VATT INSTITUTE FOR ECONOMIC RESEARCH

VATT WORKING PAPERS 161

Long-Run Effects of Selective Schools on Educational and Labor Market Outcomes

Ohto Kanninen

Mika Kortelainen

Lassi Tervonen

VATT WORKING PAPERS 161

VATT WORKING PAPERS 161

Long-Run Effects of Selective Schools on Educational and Labor Market Outcomes

Ohto Kanninen

Mika Kortelainen

Lassi Tervonen

Valtion taloudellinen tutkimuskeskus
VATT Institute for Economic Research
Helsinki 2023

VATT WORKING PAPERS 161

Ohto Kanninen:
Labour Institute for Economic Research LABORE, ohto.kanninen@labore.fi

Mika Kortelainen:
University of Turku, Finnish Institute for Health and Welfare (THL),
VATT Institute for Economic Research, mika.kortelainen@utu.fi

Lassi Tervonen:
Aalto University, Ekonominaukio 1, 02150 Espoo, lassi.tervonen@aalto.fi

VATT Working Papers:
https://doria.fi/handle/10024/147862

Valtion taloudellinen tutkimuskeskus
VATT Institute for Economic Research
Arkadiankatu 7, 00100 Helsinki, Finland

Helsinki, December 2023

Long-Run Effects of Selective Schools on Educational and Labor Market Outcomes

Ohto Kanninen, Mika Kortelainen, and Lassi Tervonen*

December 10, 2023

Abstract

This paper analyzes the effects of selective schools on students' educational and labor market outcomes. We utilize regression discontinuity design based on the centralized admission system of upper secondary schools in Finland to obtain quasi-random variation for selective high school offers and attendance. By using nationwide administrative data, we first show that the selective schools do not improve high school exit exam scores, even though there is a large jump in peer quality for students attending selective schools. Despite lacking short-term effects, we find that selective schools increase university enrollment and graduation in the long run. Yet, we do not observe positive effects on income. Importantly, our results suggest that selective high schools or better peer groups do not improve students' human capital or skills, but affect their preferences on educational choices after the secondary school.

*Kanninen: Labour Institute for Economic Research LABORE (ohto.kanninen@labore.fi); Kortelainen: University of Turku, Finnish Institute for Health and Welfare (THL), VATT Institute for Economic Research (mika.kortelainen@utu.fi); Tervonen: Aalto University, Ekonominaukio 1, 02150 Espoo (lassi.tervonen@aalto.fi). We thank Manuel Bagues, Kristof De Witte, Kristiina Huttunen, Tuomas Pekkarinen, Matti Sarvimäki, Marko Terviö, Roope Uusitalo and Hanna Virtanen for providing helpful comments.

1 Introduction

The schools individuals attend may have long-lasting effects on their lives. These effects could come, for example, in the form of better learning outcomes, non-cognitive skills or higher income in the future. In many countries, students and their parents are particularly interested in how schools that select students on the basis of earlier academic performance affect students' learning and skills. As these so-called selective schools are generally popular and admission to them can be very competitive, many seem to believe that selective schools benefit students in one way or another. However, by simply inferring that the good outcomes of selective school graduates are caused by the schools they attended, one generally ignores selection bias – these graduates could have similar outcomes even if they attended less selective schools.

Many papers have studied the effects of selective schools on test scores and other short-run educational outcomes using quasi-experimental research designs (e.g. Abdulkadiroglu, Angrist and Pathak (2014); Dobbie and Fryer Jr (2014); Clark and Del Bono (2016)). According to the recent meta-analysis by Beuermann and Jackson (2022), selective schools generally have little effects on short-run test scores. Yet, only a few papers have been able to look at longer-term outcomes like educational choices or labor market outcomes due to lack of follow-up data. In this paper we utilize nationwide administrative data in Finland to study whether selective high schools – defined here as schools in the top 10% in terms of their entrance threshold – affect various long-run educational and labor market outcomes. In addition, we replicate the representative finding of earlier literature that the effects on short-run test scores cannot be distinguished from zero.

Similarly to previous literature, a key challenge in this setting is to find a solution for identifying the effect of selective schools. Based on descriptive findings, we know that students admitted to selective schools are much more likely to attend university than those who attended other schools. In addition, those who were admitted to selective schools have higher income after the age of 25 than those who attended other schools. As these findings could be explained mostly or only by selection, we utilize exogenous variation in selective school offers at schools' entrance thresholds and use regression discontinuity design to evaluate whether selective schools causally affect individuals' outcomes.

Despite the fact that we cannot distinguish the effects on short-run test scores from zero, we do find that attending a selective school increases the probability of university enrollment and the probability of obtaining a university degree. However, we do not find positive effects on income by the age of 35. A possible mechanism behind these findings is that students at the margin earn equally well when doing something else than pursuing university education. In fact, we observe that while selective schools have a positive effect on university enrollment, they seem to impact on enrollment in universities of applied sciences (UAS) negatively. If university and UAS education have a similar effect on income for those at the margin, the disparate effects on those outcomes could explain our results. A plausible interpretation of our results is that selective high schools do not affect the skills or productivity of their students, but influence their preferences on educational and career choices. Indeed, consistent with this, we observe that selective schools increase the probability to apply to university.

We contribute to the literature that studies the effects of selective schools. Besides our study, only a few papers have looked into the effects of selective schools on long-run outcomes. Clark and Del Bono (2016) show that elite secondary school attendance increases educational attainment, but does not seem to affect labor market outcomes. As they use a different kind of identification strategy (instrumental variable approach) in a different institutional context (individuals born in the 1950s in the UK), their results hardly generalize to our setting. Perhaps closest to our study is the paper by Beuermann and Jackson (2022), who study the short- and long-run effects of selective schools in Barbados, showing that while most preferred schools do not improve test scores, they have positive effects on educational attainment, labor market outcomes, and health, especially for women. While their institutional setting and identification strategy is relatively similar to ours, we supplement their analysis and contribute to the literature by showing how selective schools affect the educational paths individuals choose. Specifically, we show that selective schools increase the probability to pursue the type of higher education that is more typical for high-SES and high-achieving individuals. Thus, we present evidence on how selective schools contribute to the type of sorting in higher education that is not based on academic ability.

More generally, our work is related to several branches of literature. It contributes to the literature studying how education affects individual's preferences. In addition to social preferences (Cappellen et al., 2020) and time preferences (Becker and Mulligan, 1997; Perez-Arce, 2017), schooling may shape individual's preferences by affecting their tastes regarding education (Oreopoulos and Salvanes, 2011). We contribute to this literature by suggesting that schooling or selective scolls may affect the type of further education individuals want to pursue. Additionally, our paper is related to economics of identity (see the seminal paper by Akerlof and Kranton (2000)) and socialization (Bisin and Verdier, 2011). Indeed, a possible explanation to our results is socialization, as selective schools seem to push students to educational paths that are typical for high-achieving and high-SES individuals. While other papers in this literature have found that the content of education can affect the identity and attitudes of individuals (Cantoni et al., 2017; Clots-Figueras and Masella, 2013; Mitrunen, 2020; Voigtlander and Voth, 2015), in our setting the syllabus is similar also in counterfactual schools, suggesting that socialization in education may also be caused by other things in the school environment, e.g. peers. For example, individuals may want to conform to social norms in their peer group (Bursztyn and Jensen, 2015). Even though we are unable to separate the effect of peers from other characteristics of selective schools, our finding is consistent with peers having larger effects on career choices and social outcomes than on test scores, a result found in the literature studying peer effects in education (see the survey by Sacerdote (2014)).

The rest of the paper is structured as follows. Section 2 provides background information on the institutional context of our analysis. Section 3 introduces the data and lays out our econometric approach, while Section 4 provides some descriptive evidence. Section 5 reports our main results and section 6 presents robustness checks and additional evidence to support our main conclusions. The last section concludes.

2 Institutional Setting

In Finland compulsory education begins the year an individual turns 7 and ends after 9 years of comprehensive school. Most of those who complete compulsory education apply to secondary education – either to general upper secondary education or to vocational education. The latter option includes many possible tracks students can choose from when applying. In this paper we focus on the general upper secondary education and for the rest of this paper we call these schools high schools. Some of these high schools have also specialized tracks (music, visualized arts, physical education etc.), but our setting allows us to study only the general track. Thus, our description of the admission system also focuses solely on it.[1]

The joint application system to secondary schools is centralized and applicants are able to rank up to five school-program combinations in their application. When oversubscribed, the admission to general track is based on comprehensive school GPA in academic subjects, ranging from 4 to 10. In these cases the entrance threshold ends up to be the GPA of the student who gets the last seat. Thus, the entrance thresholds are not known in advance, making exact treatment manipulation difficult, especially near the cutoff. Besides using these entrance thresholds as a source of exogenous variation of selective school offers, we utilize them in our definition of selective schools, as we define selective schools as the schools with thresholds among the yearly top 10% of general track thresholds.

Applicants are allocated to schools based on a deferred acceptance (DA) algorithm. Thus, each applicant is considered to her preferred program, and each program temporarily admits applicants up to its capacity. After this, rejected

[1] The description of the student selection system is based on Huttunen et al. (2023).

applicants are considered to their next most preferred program and compared to applicants who are temporarily admitted to it. Again, programs temporarily admit applicants up to their capacity and reject the excess of applicants. The algorithm continues until every applicant is allocated to a program or rejected by all of the programs shes has applied to. After this stage, admitted applicants receive offers while rejected applicants are put on a waiting list, ranked by their GPA. If an admitted applicant has not accepted the offer in two weeks, the seat is offered to the highest-ranked individual in the waiting list.

The high school typically lasts three years. During the last year students participate in the Matriculation Examination, which is an externally graded, standardized high school exit exam (HSEE).[2] Since our study focuses on individuals who started their high school studies in 1991–1999, most individuals in our sample participated in exit exams when there were three compulsory exams (mother tongue, second national language, and foreign language). Besides these exams, students had to choose at least one of the two optional exams (mathematics, humanities, and natural sciences) and they could take part in one or more additional exams. However, since 2005, only mother tongue has been compulsory. Still, during all this time, at least four exams had to be passed to graduate. Also, since 2006, the humanities and natural sciences exam was divided to separate exams for each subject (religion, ethics, psychology, philosophy, history, social studies, physics, chemistry, biology, geography, health education). There are seven possible grades in the Matriculation Examination, and good grades make it easier to gain access into higher education. The grades are, from worst to best, I (= fail, 5%), A (11%), B (20%), C (24%), M (20%), E (15%), and L (5%). The approximate share of examinees who get each grade are presented in parentheses.[3] For estimation, we give these grades numerical counterparts, from 0 to 6.

After high school, many individuals continue to university or to university of applied sciences (UAS). These are both institutions of higher education, but they differ in terms of educational content, as universities of applied sciences have a more practical approach than universities. They also differ in terms of student population, as it is relatively more common for high-achieving individuals to study in universities than in universities of applied sciences. Moreover, when HSEE GPA is kept fixed, high-SES individuals have higher university enrollment rate and lower UAS enrollment rate than low-SES individuals at almost every level of HSEE GPA (Tervonen, 2023).

[2]Since 1994, it has been possible to take one's exams during three consecutive semesters, so it is possible to start participating in exit exams already before the third year.

[3]Before 1996, grade E was not used, and top 20% of examinees got L.

3 Data and Methods

3.1 Data

We use individual-level administrative data from the Finnish National Agency for Education (EDUFI) and Statistics Finland. We observe the schools students apply to, how they rank them, their comprehensive school GPA, and the offers they receive from schools from the Joint Application Register. We use this application data for years 1991–1999, as we want to study long-run outcomes and follow individuals until they turn 35.

Besides the Joint Application Register, we use Student Register data, from which we observe enrollment in both secondary and tertiary education. We use this data from 1995 to 2018.[4] In addition, we obtain degrees completed from the Register on Degrees and Examinations and income from the FOLK module of Statistics Finland. We are also able to link parents to their children, and therefore we observe parental education as well as parental income. Furthermore, we observe applications to universities and UAS from Centralized Application Register.

We are unable to distinguish between different high school tracks before 1998. Thus, for years 1991–1997 we choose to include only schools that do not have specialized tracks in 1998. We further restrict our sample to top 10% of general high school tracks. Additionally, we follow Abdulkadiroglu, Angrist and Pathak (2014) and restrict our estimation sample to so-called sharp samples, in which the (assumed) offers are sharp at the threshold. This means that the sharp sample of school s consists of applicants to school s who rank s first or who were rejected from schools they rank above s. We pool these sharp samples, so there may be multiple observations of a single individual, as individuals may be observed in multiple sharp samples. We end up with an estimation sample of 30,165 observations.

3.2 Estimation of Thresholds

We do not observe the entrance thresholds of schools directly. Furthermore, school- and year-specific thresholds are not sharp in the sense that sometimes applicants are able to get an offer even though someone who has higher GPA does not get one. Thus, we have to estimate where the threshold of each school-year combination most likely lies. We do this by running a regression for every possible threshold, and choose the one that does best job in explaining the observed offers. Formally, for every possible threshold j we estimate

$$Y_{ist} = \beta_0 + \beta_1 GPA_{it} + \beta_2 Above_{ijst} + \epsilon_{ijst}, \tag{1}$$

[4] As the standard time for high school completion is 3 years, we do not observe the first possible year of enrollment for the application year 1991.

where Y_{ist} is the offer of individual i to school s, GPA_{it} is the comprehensive school GPA of i in year t, and $Above_{ijst}$ is a dummy indicating whether or not individual is above (or at) the threshold j. For every school we choose the threshold that generates the highest R^2. However, if the highest $R^2 < 0.5$, we drop the school-year combination in question. Also, we drop school-year combinations that do not have anyone below the threshold.

3.3 Empirical Strategy

To estimate the reduced-form effect of crossing the threshold of a selective school s on various outcomes of individual i in application year t we use the equation

$$Y_{ist} = \rho Z_{ist} + (1 - Z_{ist})f(r_{ist}) + Z_{ist}f(r_{ist}) + \lambda_{st} + \epsilon_{ist}, \tag{2}$$

where Z_{ist} is an indicator for crossing the threshold, r_{ist} is the running variable, $f(r_{ist})$ is a linear function controlling for the running variable, and λ_{st} is the school-year fixed effect. We pool the data so that all thresholds are stacked into a single threshold centered at 0. We apply triangular kernel weights in our regressions.

We estimate MSE-optimal bandwidths for every school-year combination by using the approach of Calonico, Cattaneo and Titiunik (2014). However, in our main analysis we use a fixed bandwidth, since different outcomes may have different optimal bandwidths and therefore different number of observations. We choose the fixed bandwidth, 0.5 grades, based on the mean of the school-year specific optimal bandwidths for our main (long-run) outcomes. As a robustness check, we also present the results using 19 other bandwidths ranging from 0.05 to 1 (5 to 100).

4 Descriptive Evidence

To motivate our main analysis and the choice of outcome variables, we first present descriptive evidence on education and earnings of selective school students. From Figure 1a one can see that at the age of 20 students admitted to selective schools are about six times more likely to have ever attended university than those who were rejected. At 35, the difference is 40 percentage points. Also, as can be seen in Figure 1b, those who were admitted have higher income than those were rejected after the age of 25. Before that admitted students have lower income than those who were rejected, which is probably because many of the former are still studying before they turn 25. Thus, selective school students have very different outcomes than those who did not go to selective schools.

However, these observations probably reflect also some other differences between the two groups than barely admission to selective schools. Hence, to get some idea of the characteristics of selective school students and others, Table 1 presents descriptive statistics for individuals who were admitted to selective schools, who were rejected from selective schools, and for everyone who have participated in the joint application in 1991–1999, including also those who applied to vocational schools.[5] We see that those admitted to selective schools have higher baseline GPA, are more likely to be female, are more likely to have university-educated parents, and have higher family income than those who were rejected or the average applicant. Also, almost all of them (97%) participate in high school exit exams, while only 3/4 of rejected applicants and only about half of all applicants do. Conditional on participating, those who were admitted to selective schools have better exit exam outcomes, are more likely to take the advanced math exam, and take more advanced language exams than the average applicant or those who were rejected. We also observe that they are more likely to have applied to university, have higher university enrollment, and are more likely to have graduated from university with bachelor's or master's degree than the two other groups by the age of 35. However, application, enrollment, and graduation rates for universities of applied sciences (UAS) are around the same for admitted and rejected applicants.

The descriptive statistics regarding long-run educational attainment are not surprising in the sense that selective school students have also high baseline GPA. Indeed, both applying to university and university enrollment by the age of 35 are positively associated with comprehensive school GPA, as can be seen from Figure 2. The same is true for UAS up to about 75th percentile in terms of applying – after that, the application rate declines. Similarly, the UAS enrollment rate goes up until about 80th percentile, but declines after that. Thus, university education is more common choice (in terms of applications and enrollment) than UAS only in the top end of grade distribution. Moreover, the enrollment gap is higher than the application gap for this group. A possible explanation for this is that students at the top end of the GPA distribution seem to apply to UAS as a backup option, but eventually choose to enroll in universities.

While the differences in educational attainment among selective school students and other students are large, it is plausible that the differences in Table 1 are at least partly driven by selection of high-achieving students into selective schools. Thus, the next section studies whether or not selective schools causally affect the outcomes of their students.

[5] It should be noted that while our sample consists of 30,165 observations, the number of individuals is 28,042, as some applicants are considered to multiple selective schools or apply in multiple years.

5 Main Results

5.1 First Stage

Before turning to the effects of selective schools on short- and long-run outcomes, we present how much crossing the admission cutoff affects the probability of getting an offer from selective school, i.e., the first stage. This effect is presented separately for the specific selective school in question as well as for any selective school in Figure 3. This division reflects the fact that even if one is not admitted to the school she is considered for, she may be admitted to another selective school. The first-stage estimates are also presented in Table 2. Based on these results we can conclude that crossing the threshold increases the probability of admission to a specific (preferred) selective school as well as to any selective school, though the effect on latter outcome is lower as expected.

There are several possible reasons why the estimate for specific offer does not equal one. First, some applicants may be able to get the offer even if they are below the threshold. Second, some applicants above the threshold may be able to get an offer from a school they have ranked higher, even though they scored below the threshold of that school. Third, some applicants above the threshold may be able to get an offer from some other school some other way.

Additionally, Figure 3 shows how the probability to take the high school exit exam (HSEE) jumps at the threshold. While these are smaller than the jumps in offer rates, they show that crossing the threshold clearly increases the probability to actually study in a selective school. The estimates for these effects are also presented in Figure 2.

Besides the first-stage estimates, Table 2 presents the estimates for the effects of crossing the admission cutoff on peer group characteristics. According to these estimates, by crossing the threshold one gets a peer group with higher comprehensive school GPA, higher proportion of female students, higher parental education, and higher family income.

5.2 Short-Run Outcomes

We estimate the impact of selective schools on standardized and externally evaluated high school exit exam grades, specifically for Finnish (mother tongue), English, mathematics (both basic and advanced syllabus), and exit exam GPA. The results are presented in Figure 4 and in Table 3. The estimates for Finnish and English are negative, while the other are positive. However, most of the estimates do not significantly differ from zero and are also quite small in magnitude. Only the estimate for Finnish is statistically significant at the 10% level. Thus, the general view is that we cannot reject the null hypothesis of no effect on exit exam grades, which is in line with the previous literature (Beuermann and Jackson, 2022).

As a validity check, we examine whether those crossing the selective school threshold are more likely to participate in any of the high school exit exams. Also, as both basic and advanced exams are offered in mathematics and most languages, we estimate whether selective schools affect the probability to take the exam in advanced math or the number of advanced language exams taken. All of these estimates are positive, but not statistically significant, as can be seen from Table A.1. Thus, the effects on exam grades do not seem to be driven by selection to participation or to more advanced exams.

5.3 Long-Run Outcomes

We now turn to the effects on educational and labor market outcomes. Even though we did not find evidence on the effects on short-run test scores, it could be so that the benefits of selective schools are realized later through other outcomes. For example, the selective school environment may affect students' views in a way that they start seeing university studies as a self-evident default choice. In addition, they may form networks with other high-achieving students during high school, which could lead to labor market gains later.

The results for the effects on long-run outcomes at the age of 35 are presented in Figure 5 and Table 4. The enrollment outcomes here are dummy variables indicating if one has ever enrolled in any university or UAS. According to these estimates it seems that access to selective schools increases the probability of university enrollment but decreases the probability to enroll in UAS. The effect on the probability to graduate from university is also positive and significant, while we do not observe any clear effect on the probability to graduate from UAS.

Despite the positive effect on educational attainment, we do not observe a positive effect on income. One explanation for this puzzle could be that selective schools change the preferences of their students in a way that makes them more likely to attend university, but do not affect their human capital (measured by income and test scores). Indeed, as presented in Figure 6 and in Table 5, we find that selective schools increase the probability of applying to university by the age of 35, but the effect on applications to UAS is not significantly different from zero.

Additionally, we estimate the effects on long-run outcomes for every age from 19 to 35 to construct outcome trajectories. With the help of these trajectories we can observe if the age when outcomes are measured matters.

The application trajectory for universities is presented in Figure 7a and for UAS in Figure 7b. All of the university estimates are positive and after 25 also significant at the 5% level, while the UAS estimates are insignificant and closer to zero (though negative), as can be seen from Figure 7b. These patterns support the main results presented earlier. In addition, the enrollment trajectories presented in

Figures 7c and 7d also support the main results, as all of the university estimates are positive and many of them are significant at the 5% or at the 10% level, while all of the UAS estimates are negative and most are significant at the 5% or 10% level.

Our main results suggested that selective schools increase the probability of obtaining a university degree. Figures 8a and 8b support this claim, but the positive effects seem to come relatively late. Figure 8c also supports the main result of effect on UAS degree, as most of the point estimates are close to zero and insignificant. The same holds for the effect on income percentile, as can be seen from the income trajectory presented in Figure 8d. This trajectory does not have a clear pattern, and most of the estimates do not differ significantly from zero.

It should be noted that one mechanism behind our results could be that the income at 31–35 does not reflect the true income potential of individuals in this study. For example, it could be that because selective school students study more, they also start their working careers later. Finnish university students graduate relatively late, so it is possible that the income benefits are realized only after one has turned 35. In addition, according to our results the positive effects on university degrees obtained are realized after the age of 30, so this could be the driving force behind the insignificant effects on income.

Still, it could be that the marginal student is indifferent between attending a university and a UAS in terms of future income, and because those below the threshold are somewhat more likely to attend UAS, we do not observe a positive effect on income at the age of 31–35. Thus, for marginal students the choice between universities and UAS could be just a matter of taste.

5.4 Mechanisms

We consider several mechanisms that could be behind our somewhat puzzling results. First, in selective schools students come from the upper end of the baseline grade point average (GPA) distribution. Therefore students who are able to get a seat from a selective school have a better peer group in terms of baseline GPA than those who are rejected from the same school. Thus, if better peers have a positive effect on individual's outcomes, then selective schools could have a positive effect as well. Second, students in selective schools have often more educated parents and higher family income than students in other schools. This means that the peer group one gets in selective schools is not only better in terms of GPA but also has more educated parents and higher family income on average. If students with highly educated or high-income parents are more likely to become highly educated themselves, attending a selective school means that one studies with peers who are more prone to become highly educated. If exposure to a peer group like this affects one's own tendency to aspire higher education, selective schools could have a positive effect on educational attainment. Third, selective high school students are more likely to be female. This could boost the outcomes of students in these schools, as higher proportion of female students may improve students' cognitive outcomes (Lavy and Schlosser, 2011).

To examine these possible mechanisms, we split the selective school sample by the size of the jump in the peer group characteristics that occur at the threshold. Thus, we study if the results are similar for groups where the jump is above or below the median change in terms of certain characteristic.

Maybe the most evident change is that on average students above the threshold have higher baseline GPA. Hence, it could be so that the effects are different for those who get much better peer group in terms of GPA when they are admitted to selective school than for those whose peer group quality barely changes. Columns (1) and (2) in Table 6 present the short- and long-run results for these subsamples: those year-school combinations where the jump in peer quality is above median jump and those where the jump is below median jump. Columns (3) and (4) in Table 6 present the same results, but here the peer characteristic studied is the proportion of female students. We see that the long-run effects seem to be mostly driven by school-year combinations in which the jump in peer quality is higher, but it is less clear if the effects are driven by changes in gender makeup.

As mentioned earlier, also the parental characteristics – the share of university-educated parents and average family income – of peer group changes at the selective school threshold. We also split the sample by the size of the change in these characteristics, and the results for these subsamples are presented in Table 7. Interestingly, the effects are much stronger for the school-year combinations in which the jump in family income is higher, while the effects are more ambivalent when considering changes in parental education.

Thus, the effects seem to be driven by schools where the change in peer group GPA and family income is large relative to the counterfactual school. However, we cannot conclude that the changes in peer group composition are behind the observed effects. Besides the changes in peer group, the results could be driven by something related to the selective institutions themselves. For instance, it could be that the teachers or guidance counselors in these institutions are more likely to encourage students to apply to universities. Unfortunately, we are unable to study these kind of mechanisms.

6 Validity Checks and Robustness

Our RDD strategy relies on the assumption that applicants are not able to manipulate the running variable. The standard way to check this is to examine whether there are more applicants just above the threshold than just below it (McCrary, 2008). However, this test is not applicable in a setting like ours as Zimmerman (2014) notes, because there exists discontinuities in the distribution of GPA, the running variable of our study. This can be seen by looking at Figure 9. There are GPAs that no one has, and there are also "spikes", i.e. some GPAs are very common. This fact becomes even more clear by looking at the same thing near the standardized admission threshold, as in Figure 10.

However, to provide evidence of no manipulation, we check whether pre-determined covariates are balanced at the admission cutoff. We do not find any evidence of manipulation based on Table 8, as there is no evidence of discontinuities of at the threshold.

Furthermore, it is important to check whether our estimated effects are robust to the choice of the bandwidth. So far, we have used a fixed bandwidth of 0.5 grades for all outcomes. Here, we run the same regressions as before, but using 20 different bandwidths for each outcome. The results are presented in Figures A.1-A.3. Overall, our results are robust to the choice of bandwidth.

Additionally, we check whether the results are robust to the choice of our estimation sample. Thus, in addition to top 10% schools in the treatment group, we present our results using top 5%, top 20%, and top 50% as alternative definitions of selective schools. These results presented in Figures A.4–A.6 show that the effects are mostly similar for alternative samples. However, the positive effects on university applications, university enrollment and university degrees seem to be larger the more selective the sample is, thus strengthening our results.

7 Conclusions

This paper provides new evidence on the effects of selective schools on various short- and long-run outcomes. While we do not find evidence on effects of selective schools on short-run test scores, we do find effects on later educational attainment. Specifically, we find that selective schools increase the probability of applying to university, the probability of university enrollment, and the probability of obtaining a university degree, while the same effects are negative or insignificant for UAS. Additionally, we do not find evidence of effects on income. A possible explanation to our results is that selective high schools or better peer groups do not improve students' human capital, but affect their preferences regarding educational choices after the secondary school.

References

Abdulkadiroglu, Atila, Joshua Angrist, and Parag Pathak. 2014. "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools." *Econometrica*, 82(1): 137–196.

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *The Quarterly Journal of Economics*, 115(3): 715–753.

Becker, Gary S., and Casey B. Mulligan. 1997. "The Endogenous Determination of Time Preference." *The Quarterly Journal of Economics*, 112(3): 729–758.

Beuermann, Diether W., and C. Kirabo Jackson. 2022. "The Short and Long-Run Effects of Attending The Schools that Parents Prefer." *The Journal of Human Resources*, 57(3): 725–746.

Bisin, Alberto, and Thierry Verdier. 2011. "The Economics of Cultural Transmission and Socialization." *Handbook of Social Economics*, 1: 339–416.

Bursztyn, Leonardo, and Robert Jensen. 2015. "How Does Peer Pressure Affect Educational Investments?" *The Quarterly Journal of Economics*, 130(3): 1329–1367.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, 82(6): 2295–2326.

Cantoni, Davide, Yuyu Chen, David Y. Yang, Noam Yuchtman, and Y. Jane Zhang. 2017. "Curriculum and Ideology." *Journal of Political Economy*, 125(2): 338–392.

Cappellen, Alexander, John List, Anya Samek, and Bertil Tungodden. 2020. "The Effect of Early-Childhood Education on Social Preferences." *Journal of Political Economy*, 128(7): 2739–2758.

Clark, Damon, and Emilia Del Bono. 2016. "The Long-Run Effects of Attending an Elite School: Evidence from the United Kingdom." *American Economic Journal: Applied Economics*, 8(1): 150–176.

Clots-Figueras, Irma, and Paolo Masella. 2013. "Education, Language and Identity." *The Economic Journal*, 123: 332–357.

Dobbie, Will, and Roland G. Fryer Jr. 2014. "The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools." *American Economic Journal: Applied Economics*, 6(3): 58–75.

Huttunen, Kristiina, Tuomas Pekkarinen, Roope Uusitalo, and Hanna Virtanen. 2023. "Lost Boys? Secondary Education and Crime." *Journal of Public Economics*, 218.

Lavy, Victor, and Analía Schlosser. 2011. "Mechanisms and Impacts of Gender Peer Effects at School." *American Economic Journal: Applied Economics*, 3(2): 1–33.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics*, 142(2): 698–714.

Mitrunen, Matti. 2020. "Can You Make an American? Compulsory Patriotism and Assimilation of Immigrants." *Working Paper*.

Oreopoulos, Philip, and Kjell G. Salvanes. 2011. "Priceless: The Nonpecuniary Benefits of Schooling." *The Journal of Economic Perspectives*, 25(1): 159–184.

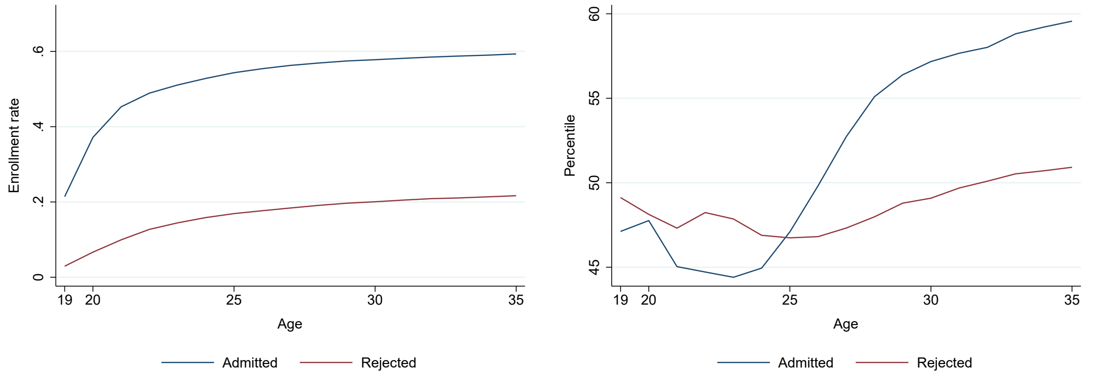
Perez-Arce, Francisco. 2017. "The effect of education on time preferences." *Economics of Education Review*, 56: 52–64.

Sacerdote, Bruce. 2014. "Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward?" *Annual Review of Economics*, 6: 253–272.

Tervonen, Lassi. 2023. "Socioeconomic Differences in Responses to Failure: Evidence from Entrance Exams." *Working Paper*.

Voigtlander, Nico, and Hans-Joachim Voth. 2015. "Nazi Indoctrination and Anti-Semitic Beliefs in Germany." *The Proceedings of the National Academy of Sciences*, 112(26): 7931–7936.

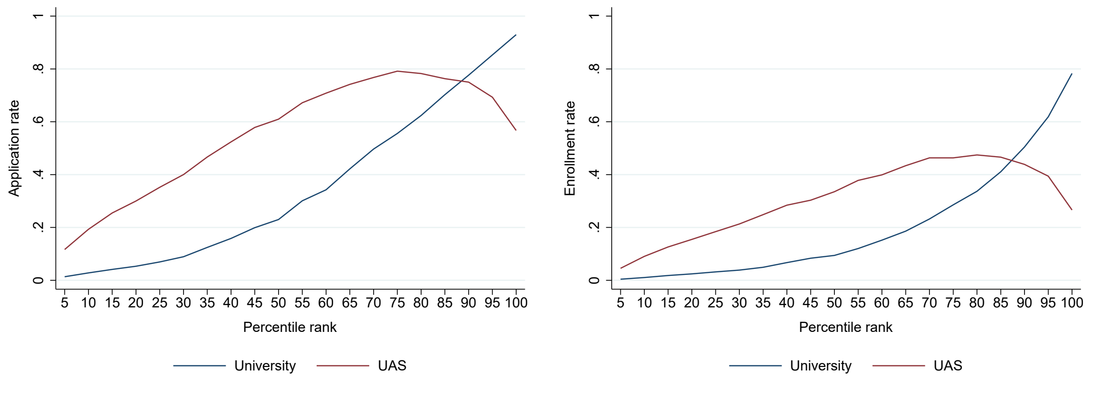
Zimmerman, Seth. 2014. "The Returns to College Admission for Academically Marginal Students." *Journal of Labor Economics*, 32(4): 711–754.



(a) University Enrollment (b) Income Percentile

Figure 1: University enrollment and income percentile by age and admission status

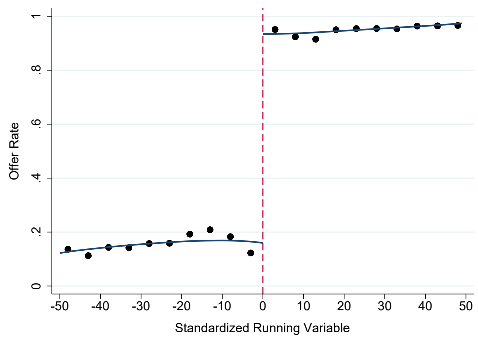
Notes: Panel (a) presents the university enrollment rates from the age of 19 to 35 for admitted and rejected selective high school applicants. Similarly, panel (b) presents the income percentiles of these two groups. These income percentiles are relative to individual's own birth cohort.



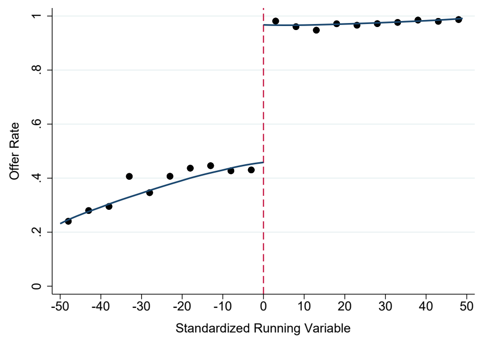
(a) Applications by GPA (b) Enrollment by GPA

Figure 2: Applications and enrollment by GPA (full sample)

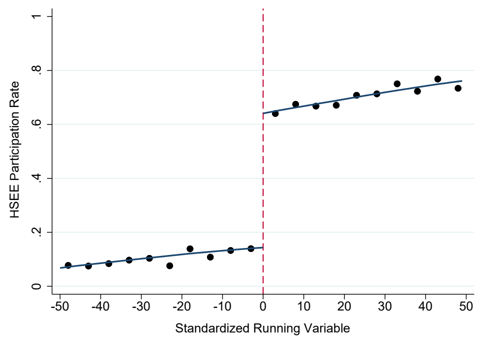
Notes: These figures present the university and UAS (university of applied sciences) application and enrollment rates (by the age of 35) by the percentile of 9th grade GPA.



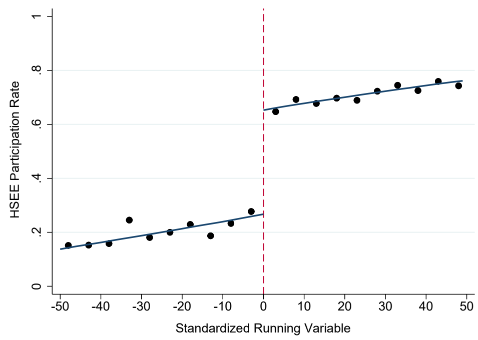
(a) Offer from a specific selective school



(b) Offer from any selective school



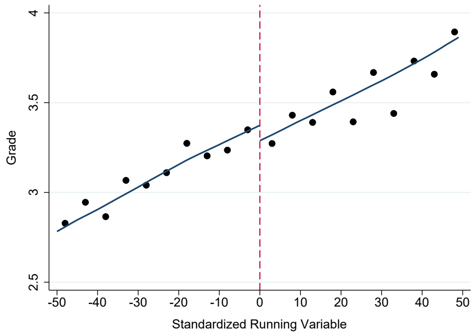
(c) HSEE in a specific selective school



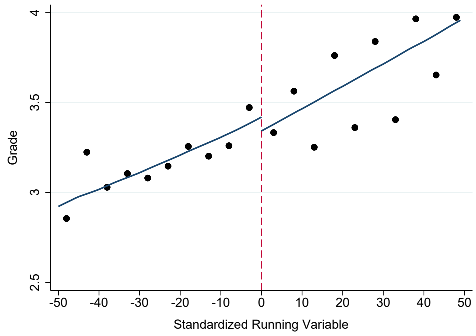
(d) HSEE in any selective school

Figure 3: First stage

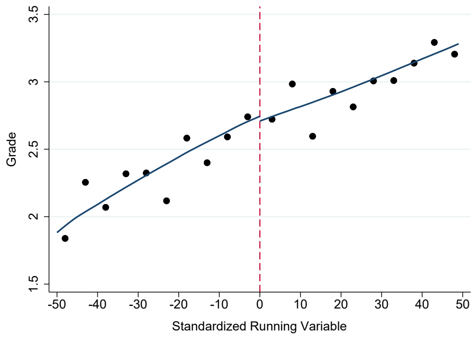
Notes: These figures show how the probability of getting a selective school offer and the probability of taking at least one high school exit exam (HSEE) in a selective school jump at the admission cutoff. The results are presented for the specific school in question (panels (a) and (c)) and for any selective school (panels (b) and (d)).



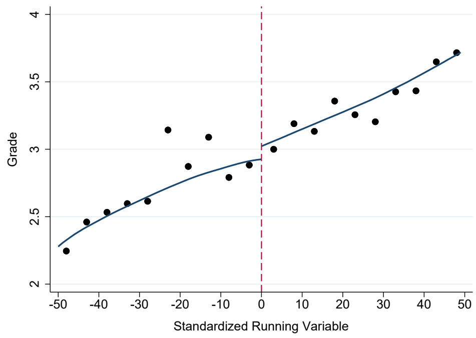
(a) Finnish



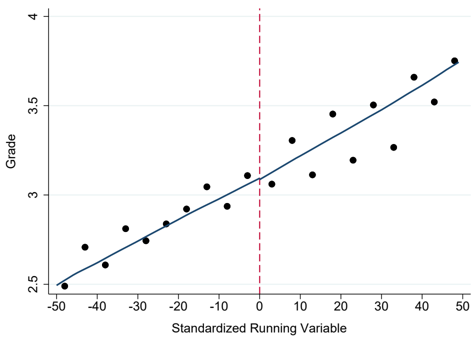
(b) English



(c) Math (adv.)



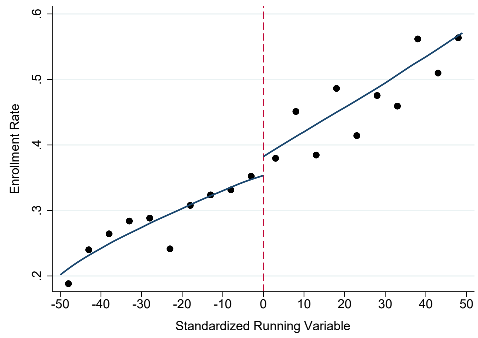
(d) Math (basic)



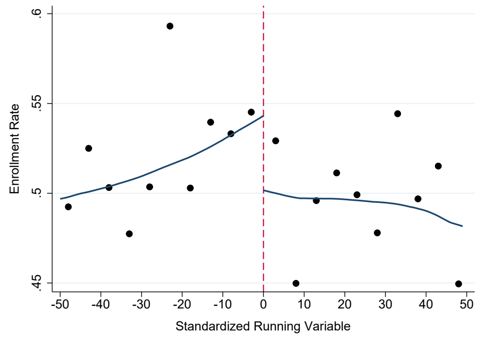
(e) GPA

Figure 4: The effect of crossing the admission cutoff on standardized test scores

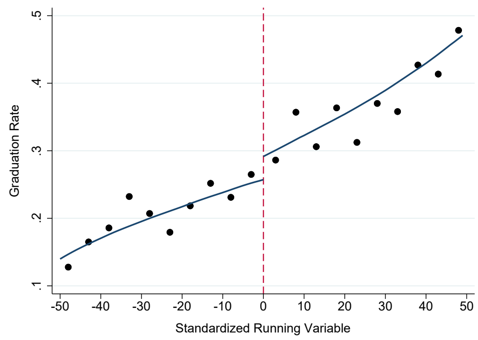
Notes: These figures present how crossing the admission cutoff affects standardized and externally evaluated high school exit exam grades.



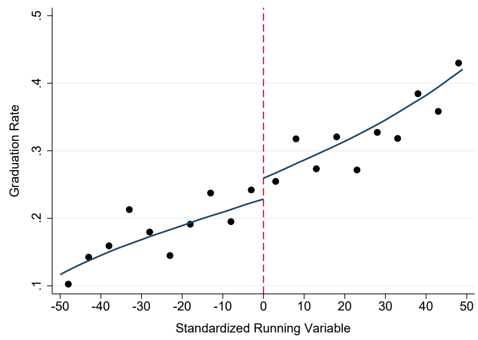
(a) University enrollment



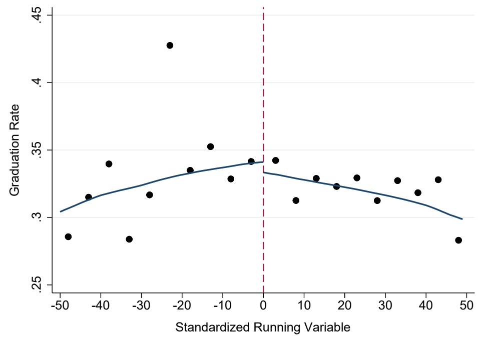
(b) UAS enrollment



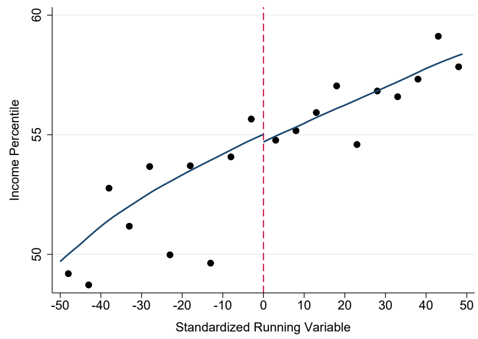
(c) Bachelor's degree



(d) Master's degree



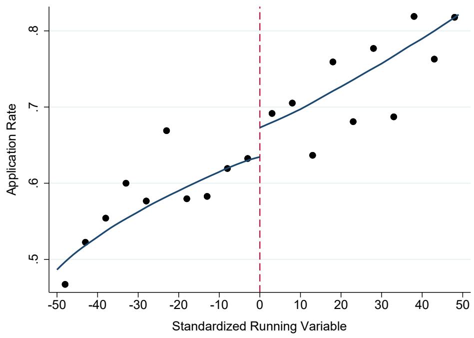
(e) UAS degree



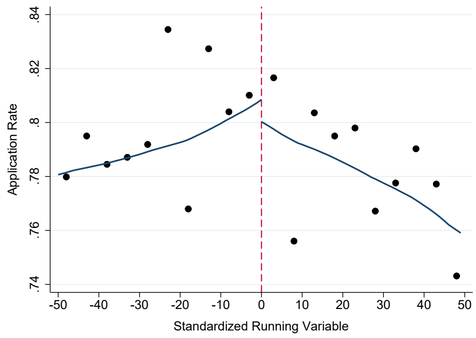
(f) Income rank

Figure 5: The effect of crossing the admission cutoff on enrollment, graduation, and income rank

Notes: These figures present how crossing the admission cutoff affects university and UAS enrollment, graduating from university with a bachelor's or master's degree, graduating from UAS, and income percentile rank at the age of 31–35.



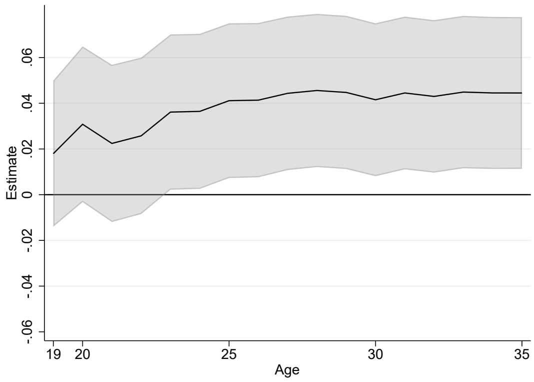
(a) University application



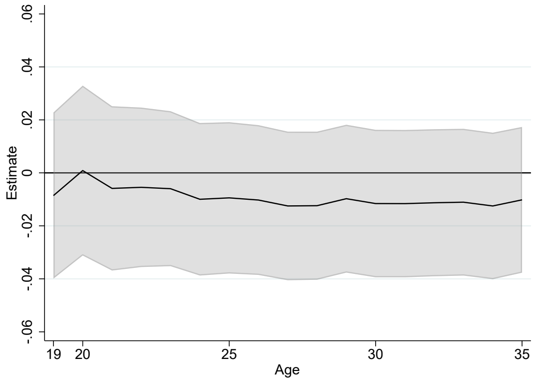
(b) UAS application

Figure 6: The effect of crossing the admission cutoff on applying to university and UAS

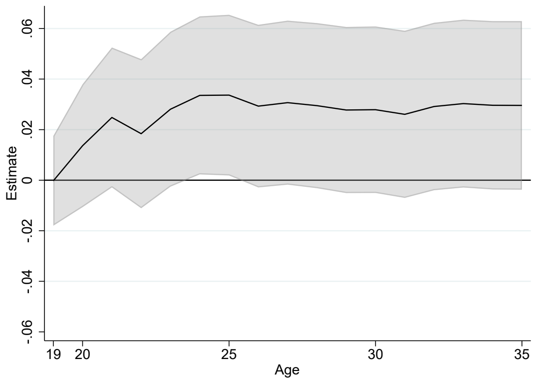
Notes: These figures present how crossing the admission cutoff affects the probability to apply to university and to UAS.



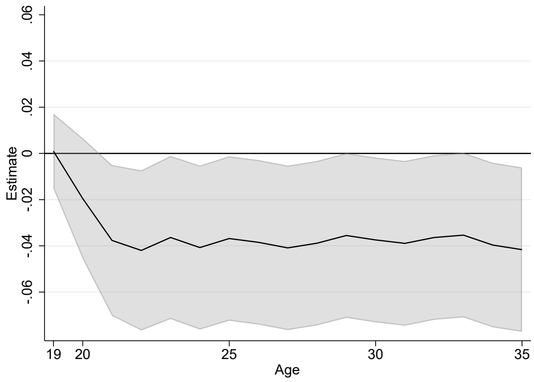
(a) University Application



(b) UAS Application



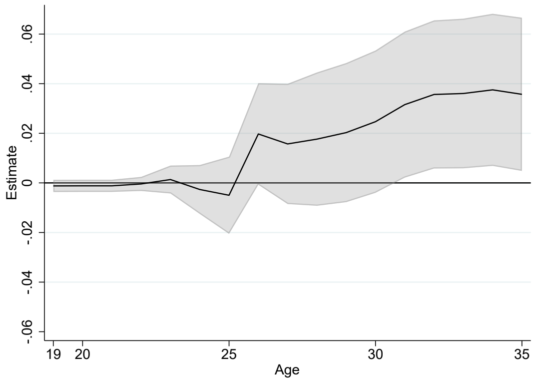
(c) University Enrollment



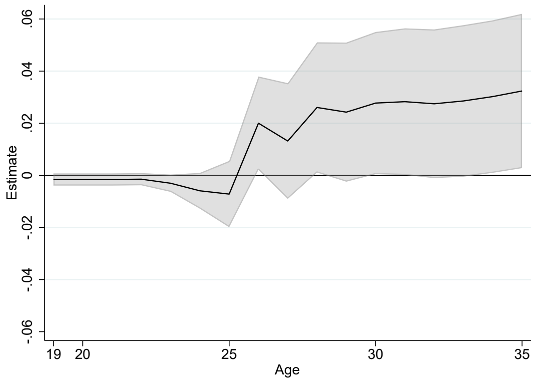
(d) UAS Enrollment

Figure 7: The effects on applying and enrollment by age

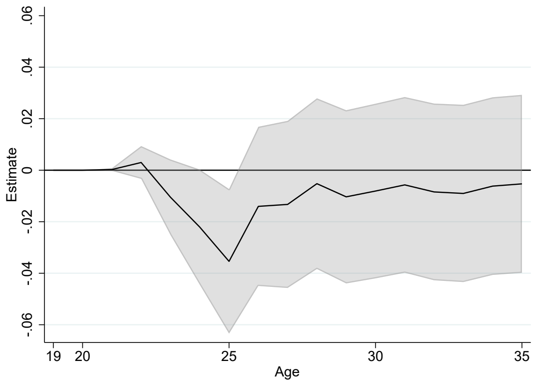
Notes: These figures present the effects of crossing the standardized admission cutoff on ever applying to university (panel (a)), ever applying to university of applied sciences (panel (b)), ever enrolling in university (panel (c)), and ever enrolling in university of applied sciences (panel (d)). The point estimates and their 95% confidence intervals are presented for ages 19–35.



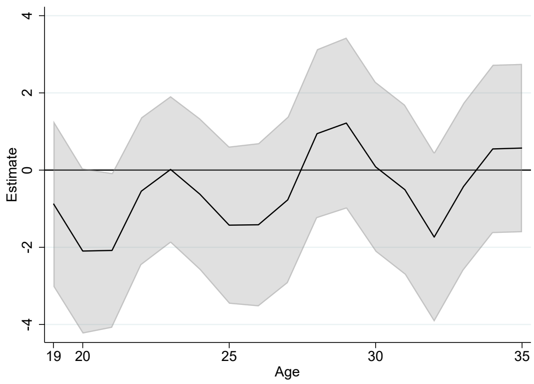
(a) Bachelor's Degree



(b) Master's Degree



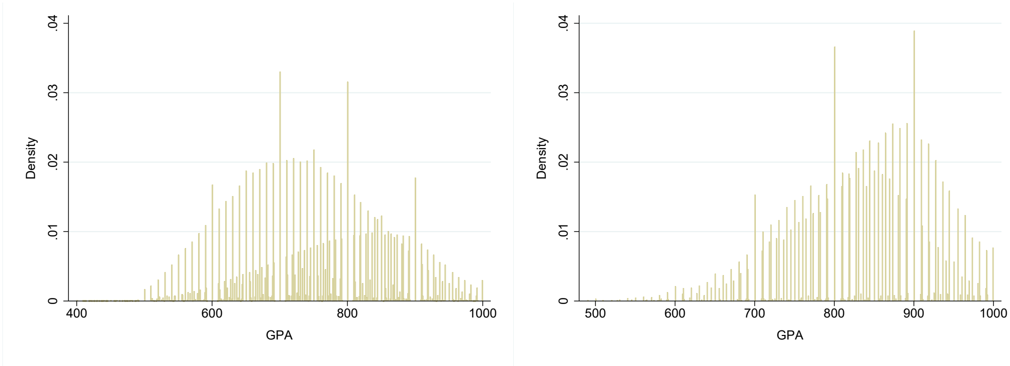
(c) UAS Degree



(d) Income Percentile

Figure 8: The effects on educational attainment and income by age

Notes: These figures present the effects of crossing the standardized admission cutoff on graduating with a bachelor's degree (panel (a)), a master's degree (panel (b)), or a UAS degree (panel (c)). Additionally, panel (d) presents the effects on individual's income percentile relative to her own birth cohort. The point estimates and their 95% confidence intervals are presented for ages 19–35.



(a) All applicants

(b) Selective school applicants

Figure 9: GPA distribution of the applicants

Notes: These figures present the 9th grade GPA distribution of all applicants (panel (a)) and of selective school applicants (panel (b)).

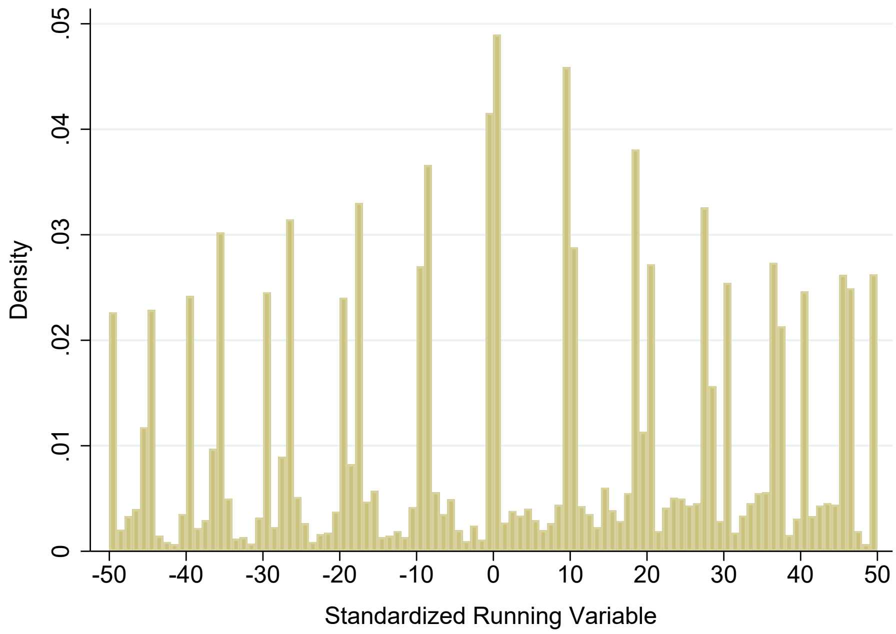


Figure 10: GPA distribution near the standardized entrance threshold

Notes: These figures present the 9th grade GPA distribution of applicants who are 0.5 grades (50) or less away from the standardized admission cutoff.

Table 1: Descriptive statistics

	(1) Admitted		(2) Rejected		(3) Full Sample	
	Mean	SD	Mean	SD	Mean	SD
A: Background						
Baseline GPA	87.08	(9.03)	54.30	(18.75)	52.14	(29.53)
Female	0.63	(0.48)	0.50	(0.50)	0.53	(0.50)
University-Ed. Parents	0.27	(0.38)	0.17	(0.32)	0.08	(0.24)
Family Income	66.18	(27.98)	59.59	(29.93)	50.94	(28.55)
B: Short-Run Outcomes						
Exit Exam Participation	0.97	(0.17)	0.75	(0.43)	0.50	(0.50)
Exit Exam GPA	62.36	(18.88)	40.43	(18.21)	50.31	(21.42)
Advanced Math	0.47	(0.50)	0.19	(0.39)	0.35	(0.48)
No. Advanced Languages	1.16	(0.41)	1.05	(0.31)	1.08	(0.34)
C: Long-Run Outcomes						
Applied to University	0.82	(0.39)	0.48	(0.50)	0.37	(0.48)
Applied to UAS	0.74	(0.44)	0.71	(0.45)	0.47	(0.50)
University Enrollment	0.59	(0.49)	0.22	(0.41)	0.22	(0.42)
UAS Enrollment	0.43	(0.50)	0.45	(0.50)	0.31	(0.46)
Bachelor's Degree at 35	0.49	(0.50)	0.15	(0.36)	0.18	(0.38)
Master's Degree at 35	0.45	(0.50)	0.13	(0.34)	0.16	(0.37)
UAS Degree at 35	0.27	(0.44)	0.28	(0.45)	0.20	(0.40)
Income Percentile at 31–35	58.38	(27.51)	50.02	(27.11)	50.27	(26.23)
Observations	16,910		11,132		792,108	

Notes: This table shows the descriptive statistics for admitted and rejected selective school applicants. Additionally, the same statistics are shown for full sample, which consists of all individuals participating in the joint application.

Table 2: First stage and peer group characteristics

	Estimate (1)	Bandwidth (2)	N (3)
A: First Stage			
Offer (Specific)	0.784***	50	12,968
	(0.010)		
Offer (Any)	0.508***	50	12,968
	(0.013)		
HSEE Participation (Specific)	0.488***	50	12,968
	(0.013)		
HSEE Participation (Any)	0.373***	50	12,968
	(0.015)		
B: Peer Group Characteristics			
Mean Rank	9.144***	50	12,710
	(0.268)		
Proportion Female	0.049***	50	12,710
	(0.003)		
Parental Education	0.062***	50	12,710
	(0.002)		
Family Income	3.451***	50	12,710
	(0.159)		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: The panel A of this table presents the effect of crossing the admission cutoff on getting a selective school and on taking at least one high school exit exam (HSEE) in a selective school. The estimates are presented separately for a specific selective school and for any selective school. Panel B of this table presents the effects on peer group characteristics.

Table 3: Short-run outcomes

	(1) Finnish	(2) English	(3) Math (adv.)	(4) Math (basic)	(5) GPA
Estimate	-0.071*	-0.034	-0.002	0.106	0.006
	(0.043)	(0.052)	(0.109)	(0.083)	(0.032)
Observations	11,352	10,718	3,585	5,129	11,995
Bandwidth	50	50	50	50	50

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table presents the effect of crossing the admission cutoff on high school exit exam grades. The estimates are presented in standard deviations. ”Math (adv.)” refers to an advanced curriculum, while ”Math (basic)” refers to a basic curriculum in mathematics.

Table 4: Long-run outcomes

	Estimate (1)	Bandwidth (2)	N (3)
University Enrollment	0.030*	50	12,968
	(0.017)		
UAS Enrollment	-0.042**	50	12,968
	(0.018)		
Bachelor's Degree	0.036**	50	12,968
	(0.016)		
Master's Degree	0.032**	50	12,968
	(0.015)		
UAS Degree	-0.011	50	12,968
	(0.017)		
Income Percentile (31–35)	-0.288	50	12,486
	(1.025)		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table presents the effect of crossing the admission cutoff on having enrolled in university or in UAS (university of applied sciences), as well as having obtained a bachelor's, master's or UAS degree by the age of 35. Additionally, it presents the effect on individual's average income percentile at the age of 31–35.

Table 5: Applications by 35

	Estimate (1)	Bandwidth (2)	N (3)
Applied to University	0.045**	50	12,968
	(0.017)		
Applied to UAS	-0.010	50	12,968
	(0.014)		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table presents the effect of crossing the admission cutoff on having ever applied to university or to UAS (university of applied sciences) by the age of 35.

Table 6: Effects by the size of changes in peer group characteristics

	GPA		Female Prop.	
	(1)	(2)	(3)	(4)
	Above	Below	Above	Below
A: Short-Run Outcomes				
Finnish	-0.043	-0.108*	-0.059	-0.070
	(0.063)	(0.060)	(0.059)	(0.063)
English	-0.028	-0.040	-0.098	0.033
	(0.075)	(0.072)	(0.072)	(0.076)
Math (adv.)	0.134	-0.110	0.035	0.004
	(0.157)	(0.152)	(0.157)	(0.152)
Math (basic)	0.142	0.052	0.349***	-0.191
	(0.122)	(0.113)	(0.120)	(0.122)
HSEE GPA	0.022	-0.040	0.042	-0.048
	(0.049)	(0.047)	(0.046)	(0.049)
B: Long-Run Outcomes				
Application (Uni.)	0.046*	0.046**	0.034	0.058**
	(0.024)	(0.023)	(0.023)	(0.025)
Application (Poly.)	-0.015	-0.005	-0.018	-0.002
	(0.021)	(0.0174)	(0.020)	(0.020)
Enrollment (Uni.)	0.049**	0.009	0.012	0.050**
	(0.026)	(0.023)	(0.024)	(0.024)
Enrollment (Poly.)	-0.081***	-0.005	-0.055**	-0.028
	(0.024)	(0.025)	(0.025)	(0.026)
Bachelor's Degree	0.062***	0.010	0.036	0.038*
	(0.022)	(0.021)	(0.022)	(0.023)
Master's Degree	0.063***	0.003	0.037*	0.029
	(0.025)	(0.018)	(0.021)	(0.022)
Polytechnic Degree	-0.022	-0.000	-0.020	0.000
	(0.025)	(0.024)	(0.024)	(0.025)
Income Percentile	-2.007	1.127	-0.238	-0.350
	(1.475)	(1.415)	(1.396)	(1.510)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table presents the effects on short- and long-run outcomes by the size of the change in the peer group that is caused by crossing the admission cutoff. The effects are presented for two groups of schools, divided by whether the changes are above or below the median change.

Table 7: Effects by the size of changes in parental characteristics

	(1)		(2)	
	Parental Education		Family Income	
	Above	Below	Above	Below
A: Short-Run Outcomes				
Finnish	-0.035	-0.098*	0.047	-0.182***
	(0.063)	(0.059)	(0.063)	(0.061)
English	-0.011	-0.057	0.037	-0.101
	(0.072)	(0.075)	(0.075)	(0.072)
Math (adv.)	0.071	-0.086	0.108	-0.084
	(0.147)	(0.162)	(0.152)	(0.156)
Math (basic)	0.121	0.076	0.191	0.011
	(0.123)	(0.112)	(0.118)	(0.117)
HSEE GPA	0.016	-0.027	0.109**	-0.106**
	(0.048)	(0.047)	(0.048)	(0.047)
B: Long-Run Outcomes				
Application (Uni.)	0.050**	0.043*	0.064***	0.025
	(0.024)	(0.024)	(0.024)	(0.024)
Application (Poly.)	-0.008	-0.013	-0.012	-0.009
	(0.021)	(0.018)	(0.020)	(0.019)
Enrollment (Uni.)	0.039	0.021	0.051**	0.009
	(0.025)	(0.023)	(0.025)	(0.023)
Enrollment (Poly.)	-0.062***	-0.022	-0.081***	-0.004
	(0.026)	(0.025)	(0.026)	(0.025)
Bachelor's Degree	0.035	0.038*	0.061***	0.010
	(0.024)	(0.021)	(0.023)	(0.022)
Master's Degree	0.025	0.041**	0.055**	0.009
	(0.023)	(0.020)	(0.022)	(0.021)
Polytechnic Degree	-0.008	-0.013	-0.010	-0.011
	(0.025)	(0.024)	(0.024)	(0.025)
Income Percentile	-1.163	0.457	-1.265	0.542
	(1.484)	(1.415)	(1.476)	(1.418)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table presents the effects on short- and long-run outcomes by the size of the change in the parental characteristics of the peer group that is caused by crossing the admission cutoff. The effects are presented for two groups of schools, divided by whether the changes are above or below the median change.

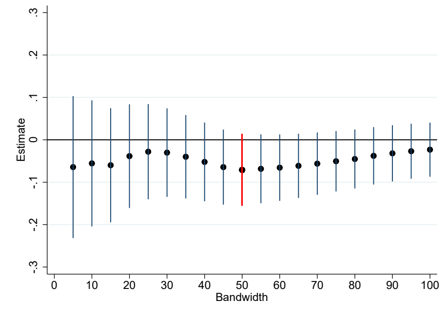
Table 8: Covariate balance

	Estimate	Bandwidth	N
	(1)	(2)	(3)
Female	-0.007	50	12,967
	(0.018)		
Parental Education	0.020	50	12,961
	(0.012)		
Family Income	1.342	50	12,935
	(1.031)		
Information on Mother	-0.001	50	12,968
	(0.001)		
Information on Father	-0.001	50	12,968
	(0.005)		

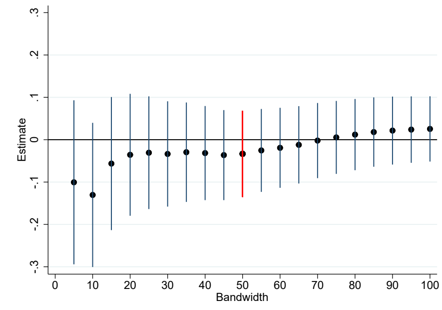
Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table presents the effects of crossing the admission cutoff on predetermined characteristics and whether parental information is observed in the data. "Parental Education" equals 0 if individual has no university-educated parents, 0.5 if one of her parents are university-educated, and 1 if two (or one in the case of single parents) of her parents are university-educated.

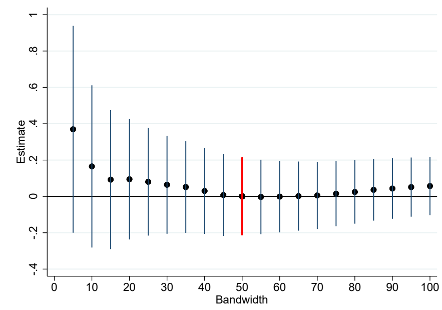
Appendix



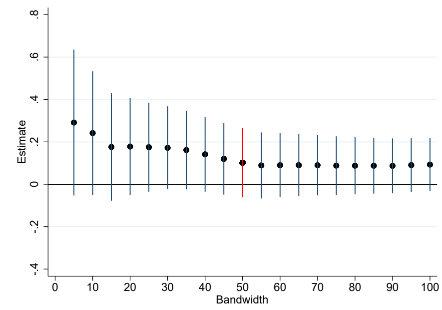
(a) Finnish



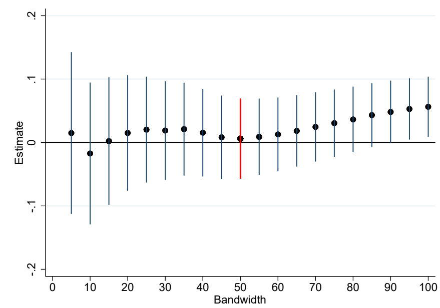
(b) English



(c) Advanced Math



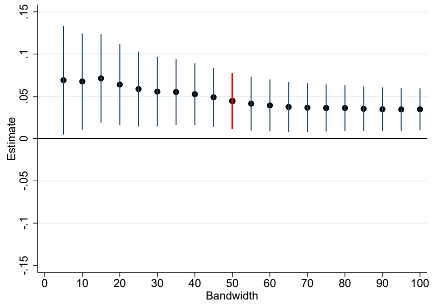
(d) Basic Math



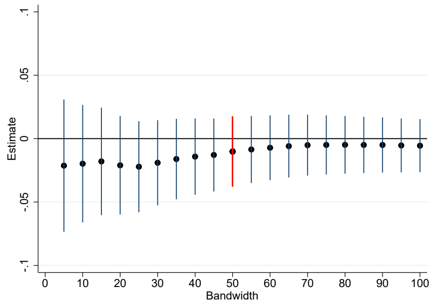
(e) GPA

Figure A.1: Robustness: Short-run estimates

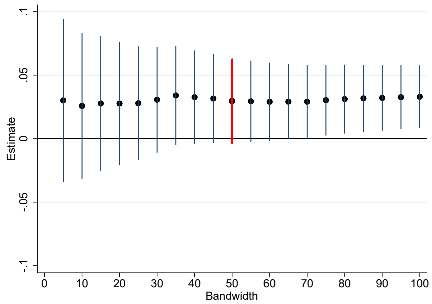
Notes: These figures present the effects of crossing the standardized admission cutoff on high school exit exam grades. The point estimates and confidence intervals are shown using 20 different bandwidths, from 5 to 100 (i.e. from 0.05 grades to 1 grade, while the grade scale goes from 4 to 10). The point estimate with the red confidence interval is the one used in the main results.



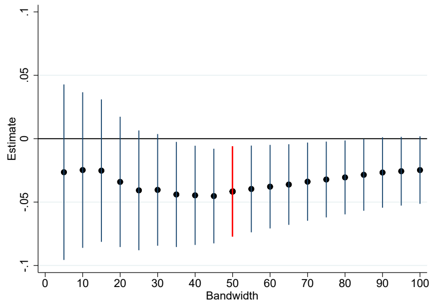
(a) University Application



(b) UAS Application



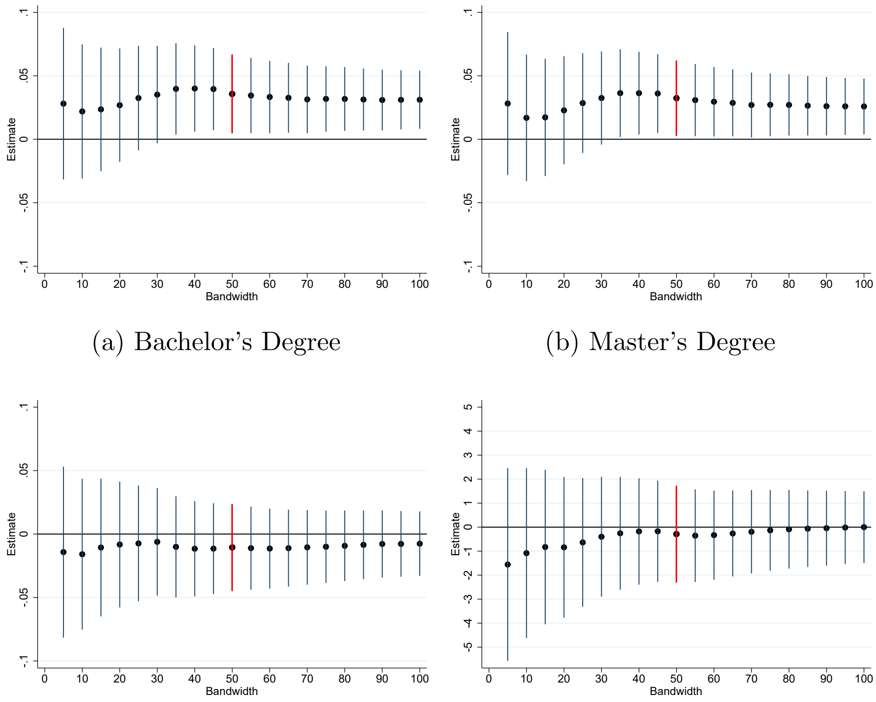
(c) University Enrollment



(d) UAS Enrollment

Figure A.2: Robustness: Application and enrollment estimates

Notes: These figures present the effects of crossing the standardized admission cutoff on (a) ever applying to university, (b) ever applying to university of applied sciences (UAS), (c) ever enrolling in university, and (d) ever enrolling in university of applied sciences by the age of 35. The point estimates and confidence intervals are shown using 20 different bandwidths from 5 to 100 (i.e. from 0.05 grades to 1 grade, while the grade scale goes from 4 to 10). The point estimate with the red confidence interval is the one used in the main results.



(a) Bachelor's Degree

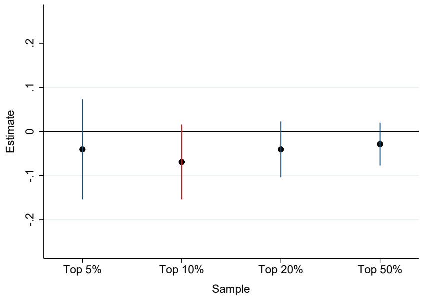
(b) Master's Degree

(c) UAS Degree

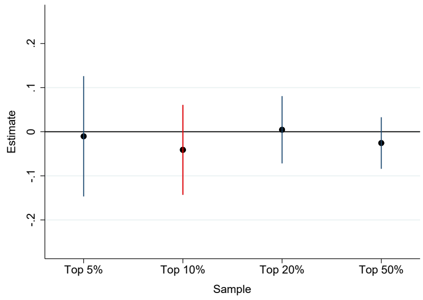
(d) Income Percentile

Figure A.3: Robustness: Degree and income estimates

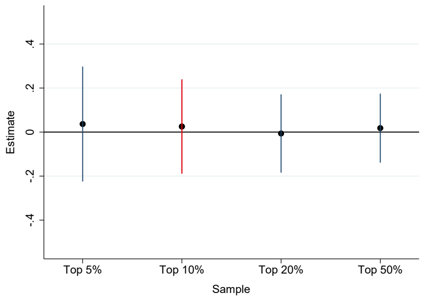
Notes: These figures present the effects of crossing the standardized admission cutoff on graduating with a (a) bachelor's degree, (b) master's degree, and a (c) university of applied sciences (UAS) degree by the age of 35. Also, panel (d) presents the effect on individual's income percentile relative to her birth cohort at the age of 35. The point estimates and confidence intervals are shown using 20 different bandwidths from 5 to 100 (i.e. from 0.05 grades to 1 grade, while the grade scale goes from 4 to 10). The point estimate with the red confidence interval is the one used in the main results.



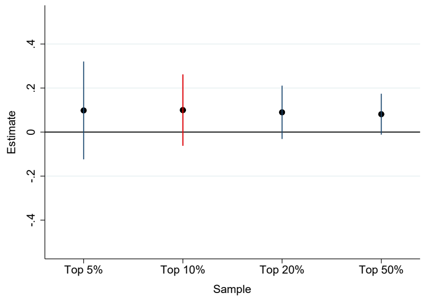
(a) Finnish



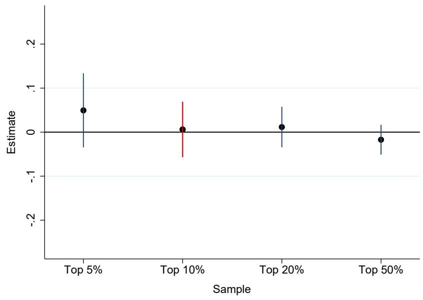
(b) English



(c) Advanced Math



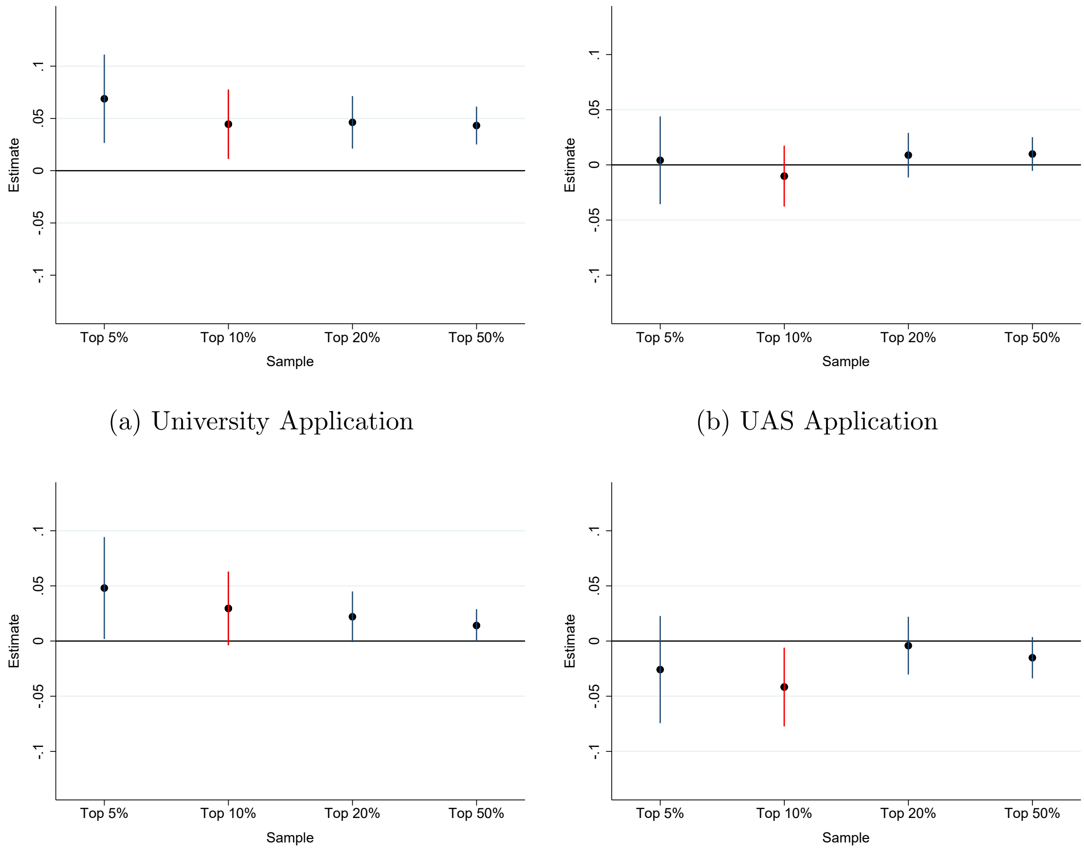
(d) Basic Math



(e) GPA

Figure A.4: Short-run estimates, alternative samples

Notes: These figures present the effects of crossing the standardized admission cutoff on high school exit exam grades. The point estimates and 95% confidence intervals are presented for four different samples: top 5%, top 10%, top 20%, and top 50% of schools. The point estimate with the red confidence interval is the one used in the main results.



(a) University Application

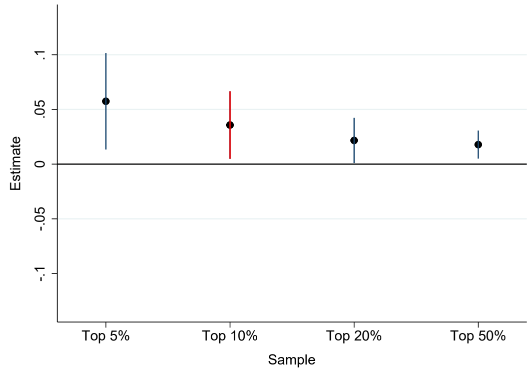
(b) UAS Application

(c) University Enrollment

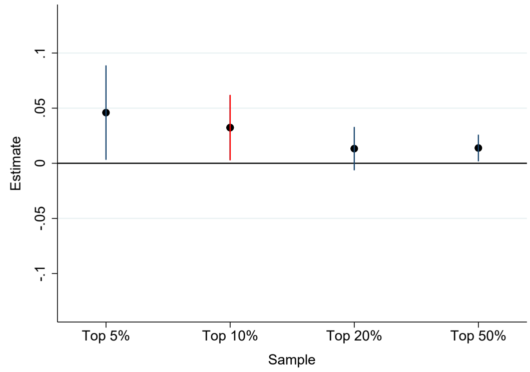
(d) UAS Enrollment

Figure A.5: Application and enrollment estimates, alternative samples

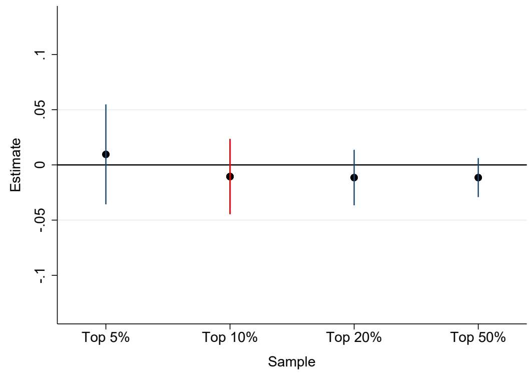
Notes: These figures present the effects of crossing the standardized admission cutoff on the probability to have (a) applied to university, (b) applied to university of applied sciences (UAS), (c) graduated from university, and (d) graduated from university of applied sciences (UAS) by the age of 35. The point estimates and 95% confidence intervals are presented for four different samples: top 5%, top 10%, top 20%, and top 50% of schools. The point estimate with the red confidence interval is the one used in the main results.



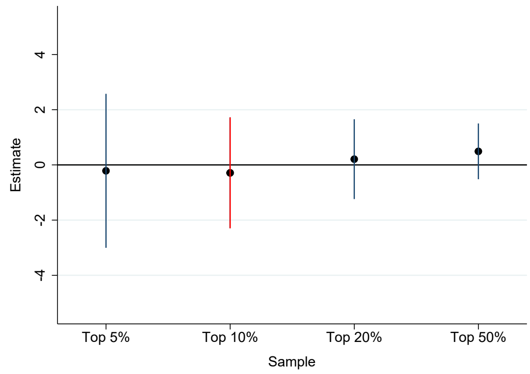
(a) Bachelor's Degree



(b) Master's Degree



(c) UAS Degree



(d) Income Percentile

Figure A.6: Degree and income estimates, alternative samples

Notes: These figures present the effects of crossing the standardized admission cutoff on graduating with a (a) bachelor's degree, (b) master's degree, and a (c) university of applied sciences (UAS) degree by the age of 35. Also, panel (d) presents the effect on individual's income percentile relative to her birth cohort at the age of 35. The point estimates and 95% confidence intervals are presented for four different samples: top 5%, top 10%, top 20%, and top 50% of schools. The point estimate with the red confidence interval is the one used in the main results.

Table A.1: Exit exam participation and advanced exams

	Estimate	Bandwidth	N
	(1)	(2)	(3)
Participation	0.012	50	12,968
	(0.009)		
Advanced Math	-0.002	50	12,023
	(0.017)		
No. Advanced Languages	0.008	50	12,023
	(0.012)		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table presents the effect of crossing the admission cutoff on participation in any high school exit exam, participation in advanced math exam, and on the number of advanced language exams the individual participated in.