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

Ohto Kanninen, Mika Kortelainen, and Lassi Tervonen*

February 21, 2022


#### Abstract

This paper analyzes the causal effects of selective school on students' short- and long-run educational and labor market outcomes. We utilize regression discontinuity design based on the centralized admission system of upper secondary schools in Finland to obtain quasirandom 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 of lacking short-term impacts, 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 also suggest 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.


[^0]
## 1 Introduction

Schools individuals attend may have long-lasting effects on their lives. These effects may come, for example, in the form of better learning outcomes, noncognitive 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 seems to believe that selective schools benefit students in one way or other. However, if individuals perceive observed outcomes of selective school graduates to be caused by selective schools, they generally ignore selection bias - the students in selective schools could have good outcomes regardless of the school they attend.

Recently, many papers have studied the effects of selective schools on test scores and other short-run 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 (forthcoming), selective schools generally have little effects on short-run test scores. At the same time, individuals are typically willing to study in selective schools, which suggests that they can generate some kind of benefits for their students, such as improved non-cognitive skills or labor market outcomes. Yet, relatively a few papers have been able to look at longer-term outcomes due to lack of follow-up data.

In this paper we utilize Finnish administrative data to study whether selective high schools affect various long-run educational and labor market outcomes. In addition, we replicate the finding that the effects on short-run test scores cannot be distinguished from zero. By selective schools we refer to high schools that are in the most selective $10 \%$ of schools in choosing students based on earlier academic performance in lower secondary school. 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 evidence, we know that students admitted to selective schools are much more likely to attend university than those who were rejected from selective schools or attended other schools. In addition, those who were admitted to selective schools earn much more after the age of 25 than those who attended other schools. Are these findings explained by selection or do selective schools have something to do with these differences? We use regression discontinuity design to tackle this problem, i.e., we evaluate whether selective schools have
a causal effect on educational attainment and choices as well as income.
Despite the fact that we do not observe effects on short-run test scores, we do find that attending a selective school affects positively on the probability of applying to university, the probability of university enrollment and the probability of obtaining a university degree. However, we do not find a positive effect on income by the age of 35 . A possible mechanism behind these findings is that students at the margin are equally well-off when doing something else than pursuing university education. In fact, we observe that while selective schools have positive effect on university enrollment, they seem to have negative effect on polytechnics enrollment. If those at the margin are indifferent in terms of income potential between university and polytechnic education, the disparate effects on those outcomes could explain our results.

A plausible interpretation of our results is that selective high schools do not affect students' human capital, but influence their preferences regarding educational choices. This could be due to the nature of the peer group one gets when crossing the selective school threshold, as peers in selective schools are generally better in terms of baseline GPA, have higher family income, and are more likely to have highly educated parents than the students in counterfactual schools. Moreover, it could be also something related to the selective institutions themselves, e.g. teachers or guidance counselors encouraging students to apply to universities more likely. Although we cannot separate mechanisms behind this result, our finding on preferences and career choice gives some support for the benefits of selective schools.

We contribute to the literature that studies the effects of selective schools. While effects on short-run outcomes have now been widely studied, only a few papers have looked into long-run outcomes (Clark and Del Bono, 2016; Beuermann and Jackson, forthcoming). Clark and Del Bono (2016) studies the impact of elite school attendance on long-run outcomes for individuals born in the 1950s and educated in a UK district that assigned students to either elite or non-elite secondary schools. As they use different kind of identification strategy (instrumental variable approach) in totally different institutional context, their results hardly generalize to our setting. Perhaps closest to our study is the recent paper by Beuermann and Jackson (forthcoming), who study the short- and long-run effects of selective schools in Barbados. While their institutional setting and identification strategy is relatively similar to ours, they concentrate on evaluating the causal impacts of attending a school that parents prefer. Instead of preferred schools, we study the long-run effect of selective schools and investigate mechanisms driving
those effects in more detail. Moreover, we are particularly interested in finding out whether selective schools affect career choices or students' preferences after the secondary school.

The rest of the paper is structured as follows. Section 2 provides background information about the mechanisms hypothesized to mediate the effects of selective schools on students' short- and long-run outcomes, the different types of schools and 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 Background

### 2.1 Mechanisms

We consider several mechanisms that could affect the outcomes of selective school students. 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 on average. If students with highly educated parents are more likely to become highly educated themselves (at every level of GPA), attending 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 and therefore also on income.

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).

### 2.2 Institutional Context

In Finland compulsory education begins the year a child 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 ("high school") 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.

The admission to high school general track is based on comprehensive school GPA in academic subjects. If a school has more applicants than it has seats, it chooses the best applicants in terms of GPA up to its capacity. In these cases the entrance threshold ends up to be the GPA of the student who gets the last seat. We define selective schools as the schools with an general track entrance threshold among the top $10 \%$ of entrance thresholds of general tracks.

The high school lasts three years. During the last year of high school students take the Matriculation Examination, which is a standardized high school exit exam. It is possible to take these exams in every academic subject, and at least four subjects must be passed to graduate. There are seven possible grades, 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\%). Before 1996, grade E was not used, and top $20 \%$ of exam takers got L. The share of exam takers who get each grade are presented in parentheses. For estimation, we give these grades numerical counterparts, 0-6.

## 3 Data and methods

### 3.1 Data

We use individual-level Finnish administrative data from the Finnish National Agency for Education (EDUFI) and Statistics Finland. We observe the schools students apply to, how they rank the schools they apply to, and their comprehensive school GPA from the Joint Application Register. We use this application data for the years 1991-1999, as we want to study long-run
outcomes and students turn 16 the year they apply to high schools and/or to vocational schools. This data includes among other things information on students' grades from comprehensive school diploma, the schools students apply to, and offers students receive from the upper secondary schools.

Besides Joint Application Register, we use Student Register data, from which we observe both high school and higher education enrollment. We use this data from 1995 to 2018. 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. We also 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 also observe parental education as well as parental income.

### 3.2 Estimation of thresholds

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

$$
\begin{equation*}
Y_{i s t}=\beta_{0}+\beta_{1} G P A_{i t}+\beta_{2} \text { Above }_{i j s t}+\epsilon_{i j s t}, \tag{1}
\end{equation*}
$$

where $Y_{i s}$ is the offer of individual $i$ to school $s, G P A_{i t}$ is the comprehensive school GPA of $i$ in year $t$, and $A_{\text {Above }}^{i j s t}$ 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, i.e. no thresholds.

### 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

$$
\begin{equation*}
Y_{i s t}=\rho Z_{i s t}+\left(1-Z_{i s t}\right) f_{1}\left(r_{i s t}\right)+Z_{i s t} f_{1}\left(r_{i s t}\right)+\lambda_{s t}+\epsilon_{i s t}, \tag{2}
\end{equation*}
$$



Figure 1: University Enrollment and Income Percentile by Age and Admission Status
where $Z_{i s t}$ is an indicator for crossing the threshold, $r_{i s t}$ is the running variable, $f_{1}\left(r_{i s t}\right)$ is a linear function controlling for the running variable, and $\lambda_{s t}$ is the school-year fixed effect. We pool the data so that all thresholds are stacked into a single threshold centered at 0 .

## 4 Descriptive evidence

Students who have attended selective schools have very different outcomes in life than those who did not go to selective schools. From Figure 1a one can see, for example, that at the age of 20 those who were admitted 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 from Figure 1b, those who were admitted earn more than those were rejected after the age of 25 . Before that they earn less, which is probably because many of them are studying before they turn 25 .

Table 1 presents more descriptive statistics for these two groups and for everyone who have participated in joint application system in 1991-1999. 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 those who were rejected or the average applicant. At the age of 35, we also observe that they are more likely to have applied to university, have higher university enrollment, and are more likely to have graduated (with bachelor's or master's degree) from university than the two other groups. However, application, enrollment, and graduation rates for polytechnics 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 are positively associated with comprehensive school GPA, as can be seen from Figure 2. The same is true for polytechnics up to about 75th percentile in terms of applying - after that, the application rate declines. The enrollment rate goes up until about 80th percentile, but declines after that. Thus, only the top students apply to and enroll in universities more than to polytechnics. The enrollment gap is higher than the application gap in the top end. A possible explanation for this is that students at the top of the GPA distribution seem to apply to polytechnics as a backup option, but eventually choose to enroll in universities.

While the differences among selective school students and other students are large, it is plausible that the differences in Table 1 are at least partly due to selection of high-achieving students into selective schools. The next section studies whether this is really the case, or if selective schools affect the outcomes of their students.

## 5 Main Results

### 5.1 First Stage Results

We present the first stage before the effects on short- and long-run outcomes. The effect of crossing the threshold on getting an offer from a specific selective school is presented in Figure 3a. However, even if one does not get an offer from some specific selective school that she prefers, it is possible that she gets an offer from another selective school. This can be seen from Figure

Table 1: Descriptive Statistics

|  | (1) <br> Admitted |  | (2) <br> Rejected |  | (3) |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | ample |
|  | Mean | SD |  |  | Mean | SD | Mean | SD |
| A: Background |  |  |  |  |  |  |
| Baseline GPA | 87.09 | (9.02) | 54.30 | (18.75) | 52.14 | (29.53) |
| Female | 0.63 | (0.48) | 0.51 | (0.50) | 0.53 | (0.50) |
| University-Ed. Parents | 0.27 | (0.38) | 0.18 | (0.32) | 0.09 | (0.24) |
| Family Income | 66.40 | (27.63) | 60.77 | (29.07) | 50.65 | (28.69) |
| B: Short-Run Outcomes |  |  |  |  |  |  |
| Exit Exam Participation | 0.97 | (0.17) | 0.75 | (0.43) | 0.50 | (0.50) |
| Exit Exam GPA | 62.37 | (18.87) | 40.43 | (18.22) | 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) | 1.06 | (0.32) | 1.08 | (0.34) |
| C: Long-Run Outcomes |  |  |  |  |  |  |
| Application (Universities) | 0.82 | (0.39) | 0.48 | (0.50) | 0.37 | (0.48) |
| Application (Polytechnics) | 0.74 | (0.44) | 0.71 | (0.45) | 0.47 | (0.50) |
| Enrollment (Universities) | 0.60 | (0.49) | 0.22 | (0.41) | 0.22 | (0.42) |
| Enrollment (Polytechnics) | 0.43 | (0.50) | 0.45 | (0.50) | 0.31 | (0.46) |
| Bachelor's Degree | 0.49 | (0.50) | 0.15 | (0.36) | 0.18 | (0.38) |
| Master's Degree | 0.45 | (0.50) | 0.13 | (0.34) | 0.16 | (0.37) |
| Polytechnic Degree | 0.27 | (0.44) | 0.28 | (0.45) | 0.20 | (0.40) |
| Observations |  | 907 |  | 135 |  | ,108 |



Figure 2: Applications and Enrollment by GPA (Full Sample)


Figure 3: First Stage

Table 2: First Stage and Peer Group Characteristics

|  | Estimate <br> $(1)$ | Bandwidth <br> $(2)$ | N <br> $(3)$ |
| :--- | :---: | :---: | :---: |
| A: First Stage |  |  |  |
| Offer (Specific) | $0.773^{* * *}$ | 52.25 | 30,165 |
| Offer (Any) | $0.014^{* * *}$ | 45.08 | 30,165 |
|  | $(0.0148)$ |  |  |
| B: Peer Group Characteristics |  |  |  |
| Mean Rank | $9.376^{* * *}$ | 32.40 | 26,839 |
|  | $(0.376)$ |  |  |
| Proportion Female | $0.049^{* * *}$ | 49.25 | 26,838 |
|  | $(0.003)$ |  |  |
| Parental Education | $0.053^{* * *}$ | 53.20 | 26,579 |
|  | $(0.004)$ |  |  |
| Family Income | $2.783^{* * *}$ | 54.72 | 25,192 |
|  | $(0.321)$ |  |  |

$$
\begin{aligned}
& \text { Standard errors in parentheses } \\
& * * * \mathrm{p}<0.01,{ }^{* *} \mathrm{p}<0.05,{ }^{*} \mathrm{p}<0.1
\end{aligned}
$$

3b. The first-stage estimates are presented in the first two rows of Table 2. Based on these estimates we can conclude that there is a clear discontinuity at the entrance threshold, as crossing the threshold increases the probability of getting an offer from a specific (preferred) selective school. It also increases the probability of getting an offer from any selective school, though this effect is lower than the former.

There are several possible reasons why the estimate for specific offer does not equal 1. 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. Third, some applicants above the threshold may be able to get an offer from some other school some other way. Fourth, there may be some measurement error. We return to this issue later.

Besides first-stage estimates, Table 2 presents the estimates for the effects of crossing the threshold on peer group characteristics. According to these estimates, by crossing the threshold one gets better peers in terms of comprehensive school GPA, a peer group with a 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 high school exit exam (Matriculation Examination) test scores. The effects are estimated for Finnish (mother tongue), English, mathematics (both basic and advanced syllabus), and exit exam GPA.

The estimates for the effects on these short-run outcomes are presented in Table 3. Panel A presents the effects on exit exam grades. The estimates for Finnish and English are negative, while the other estimates are positive. However, most of the estimates do not significantly differ from zero. Only the estimate for Finnish is statistically significant at the $10 \%$ level. To strenghten these results, we also present the effect on exit exam percentile rank in Panel B. This percentile rank is based on points in each exam. Here none of the estimates are statistically significant. Thus, as a whole, we cannot reject the null hypothesis of no effect on exit exam grades, which is in line with the previous literature (Beuermann and Jackson, forthcoming).

### 5.3 Long-run outcomes

We now turn to the long-run effects of selective schools, i.e. effects on educational attainment and labor market outcomes. Even though we did not find evidence on the effects on short-run test scores, there could be effects in the long run. It could be so that the effects of selective schools are realized later through other outcomes, and therefore the benefits are not captured by short-run test scores. For example, selective school students may learn skills that help them to get into university, e.g. how to excel in university entrance exams. Also, they may form networks with other high-achieving students during high school, and this could lead to labor market gains later. On the other hand, selective schools may simply change their students' preferences regarding education without really having effect on their human capital or productivity.

The long-run estimates at the age of 35 are presented in Table 4. The application and enrollment outcomes here are dummy variables indicating if one has ever enrolled in or applied to any university or polytechnic. According to these estimates it seems that access to selective schools increases the probability to apply to university, the probability of university enrollment, as well as the probability to graduate from university. It also seems to decrease enrollment in polytechnics. Thus, selective schools have a positive effect on educational attainment, and it seems that this positive effect comes through increased probability of applying.

However, despite the positive effect on educational attainment, we do not observe positive effect on income. One explanation for this puzzle could be that selective schools change the preferences of students in a way that makes them more likely to attend university, but their productivity remains unchanged. Moreover, it could be that the marginal student is indifferent between attending a university and a polytechnic in terms of future income, and because those below the threshold are somewhat more likely to attend polytechnics, we do not observe a positive effect on average income at the age of $31-35$. Another possibility that teachers or guidance counselors in selective schools may encourage students to send more applications, which then leads to higher university enrollment.

Additionally, we estimate the effects on these 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 4a and for
polytechnics in 4b. All of the university estimates are positive and after 25 also significant at the $5 \%$ level. Thus, selective schools increase applying to universities. However, it seems that applications to polytechnics remains unchanged, as can be seen from Figure 4b.

The trajectory of university enrollment is presented in Figure 4c. One can see that all of the estimates are positive and about half of them are significant at the $5 \%$ level. Also, almost all of them are significant at the $10 \%$ level. Thus, by looking at this graph it seems that selective schools have a positive effect on university enrollment. On the other hand, Figure 4 d shows the estimates of the effects on enrollment in polytechnics. All of the estimates are negative and many are significant at the $5 \%$ or $10 \%$ level. These effects are almost like mirror images, thus backing up the preference story.

Thus, it is not surprising that our results also suggest that selective schools increase the probability of obtaining a bachelor's degree (and maybe master's too). We do not observe any clear effect on polytechnic degree. Therefore, it seems that selective schools not only affect the type of higher education pursued, but also the quantity of it attained. Also, we present the income trajectory in Figure 5d. This trajectory does not have a clear pattern, and most of the estimates do not differ significantly from zero.

Hence, it seems that selective schools have effect on the type and quantity of educational attainment, but not on test score outcomes or income. In the next subsection we study some possible mechanisms behind these effects.

### 5.4 Heterogeneity of peer effects and mechanisms

The results are somewhat puzzling, as we observe increased enrollment but not positive effects on test scores or on income. To study the mechanisms behind these effects, we split the selective school sample by the size of the jump in the peer group characteristics that occur at the threshold. For one group the change is above the median jump, while for the other group the change is below the median jump. Thus, we study if the results are similar for groups that differ in terms of the size of the change in peer group characteristics.

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 7 present the shortand 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. 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. Columns (3) and (4) in Table 7 present the same results, but here the peer characteristic studied is the proportion of female students.

As mentioned earlier, also the parental characteristics of peer group changes at the selective school threshold. We also split the sample by the size of these changes, and the results for these subsamples are presented in columns (1) and (2) of Table 8. Interestingly, the effects are much stronger for the schoolyear combinations in which the jump in parental characteristics is higher.

It should be noted that one mechanism behind these effects 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. In this case the similar income at 35 as in the control group would not mean that their income stays similar later on.

## 6 Validity checks

### 6.1 Manipulation

Our RDD strategy relies on the assumption that applicants are not able to manipulate the running variable. In the following, we provide several pieces of evidence that support the credibility of this assumption.

The standard way to check this is to examine whether there are more applicants just above the threshold than just below it. However, this so-called McCrary's test (lähde) is not applicable in a setting like ours as Zimmerman (2014) notes, because the distribution of GPA is not continuous. This can be seen by looking at Figure 11. 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 12.

However, we can provide evidence of no manipulation by checking whether pre-determined covariates are balanced at the threshold. We do not find any evidence of manipulation based on Table 10, as the covariates are balanced.


Figure 4: The effects on applying and enrollment by age


Figure 5: The effects on educational attainment and income by age

Table 3: Short-run outcomes

|  | $(1)$ <br> Finnish | $(2)$ <br> English | $(3)$ <br> Math <br> (adv.) | $(4)$ <br> Math <br> (basic) | $(5)$ <br> GPA |
| :--- | :---: | :---: | :---: | :---: | :---: |
| A: Grades |  |  |  |  |  |
| Estimate | $-0.082^{*}$ | -0.080 | -0.035 | 0.076 | -0.015 |
|  | $(0.045)$ | $(0.056)$ | $(0.109)$ | $(0.077)$ | $(0.038)$ |
| Observations | 24,282 | 23,154 | 9,278 | 9,963 | 26,109 |
| Bandwidth | 49.50 | 48.61 | 53.63 | 68.24 | 42.86 |
| Year-School FE | YES | YES | YES | YES | YES |
|  |  |  |  |  |  |
| B: Percentiles |  |  |  |  |  |
| Estimate | $-1.733^{*}$ | -1.440 | -0.039 | 1.827 | -0.413 |
|  | $1.042)$ | $1.041)$ | $(1.916)$ | $(1.390)$ | $(0.736)$ |
| Observations | 24,282 | 23,154 | 9,278 | 9,962 | 26,109 |
| Bandwidth | 47.17 | 53.29 | 52.41 | 62.44 | 41.50 |
| Year-School FE | YES | YES | YES | YES | YES |

Standard errors in parentheses

$$
{ }^{* * *} \mathrm{p}<0.01,{ }^{* *} \mathrm{p}<0.05, * \mathrm{p}<0.1
$$

Table 4: Long-run outcomes (at 35)

|  | Estimate <br> (1) | Bandwidth <br> (2) | $\begin{gathered} \hline \mathrm{N} \\ (3) \end{gathered}$ |
| :---: | :---: | :---: | :---: |
| Application (University) | $\begin{gathered} 0.030^{* *} \\ (0.015) \end{gathered}$ | 75.18 | 30,165 |
| Application (Polytechnic) | $\begin{aligned} & -0.006 \\ & (0.014) \end{aligned}$ | 62.86 | 30,165 |
| Enrollment (University) | $\begin{gathered} 0.027 \\ (0.016) \end{gathered}$ | 59.16 | 30,165 |
| Enrollment (Polytechnic) | $\begin{gathered} -0.033^{* *} \\ (0.015) \end{gathered}$ | 77.21 | 30,165 |
| Bachelor's Degree | $\begin{gathered} 0.030^{* *} \\ (0.013) \end{gathered}$ | 78.17 | 30,165 |
| Master's Degree | $\begin{aligned} & 0.028^{*} \\ & (0.015) \end{aligned}$ | 55.91 | 30,165 |
| Polytechnic Degree | $\begin{aligned} & -0.009 \\ & (0.016) \end{aligned}$ | 62.57 | 30,165 |
| Income Percentile (31-35) | $\begin{aligned} & -0.182 \\ & (0.856) \\ & \hline \end{aligned}$ | 76,13 | 29,055 |



Figure 6: Income by GPA (Full Sample)


Figure 7: Income by GPA (Full Sample)

Table 5: Applications by field and treatment status

|  | $(1)$ <br> Control <br> $(\leq 0.1)$ | $(2)$ <br> Treated <br> $(\leq 0.1)$ | $(3)$ <br> Control <br> (All) | $(4)$ <br> Treated <br> (All) |
| :--- | :---: | :---: | :---: | :---: |
| A: Polytechnics |  |  |  |  |
| Natural Resources | .035 | .046 | .037 | .050 |
| Technology \& Transportation | .307 | .304 | .305 | .319 |
| Business \& Administration | .456 | .467 | .451 | .452 |
| Tourism, Hospitality \& Household | .202 | .182 | .171 | .167 |
| Health \& Social | .365 | .352 | .357 | .345 |
| Culture | .146 | .141 | .137 | .137 |
| Humanities \& Education | .085 | .086 | .096 | .077 |
| At least one application | .783 | .766 | .701 | .710 |
| B: Universities |  |  |  |  |
| Theology | - | - | .027 | .023 |
| Humanities | .322 | .308 | .306 | .310 |
| Arts, Design, Music \& Theater | .119 | .123 | .143 | .091 |
| Education | .184 | .176 | .170 | .157 |
| Sports \& Health Sciences | .041 | .045 | .037 | .039 |
| Social Sciences | .234 | .233 | .234 | .230 |
| Psychology | .046 | .062 | .049 | .055 |
| Law | .084 | .085 | .088 | .084 |
| Business | .206 | .210 | .216 | .228 |
| Natural Sciences | .272 | .274 | .237 | .347 |
| Agriculture, Forestry \& Vet. Medicine | - | - | .064 | .066 |
| Technology | .186 | .191 | .149 | .227 |
| Medicine \& Dentistry | .060 | .057 | .041 | .086 |
| Pharmacy | .017 | .030 | .021 | .029 |
| At least one application | .614 | .667 | .462 | .801 |
| Observations | 1,633 | 1,999 | 11,119 | 16,923 |

Table 6: Enrollment by field and treatment status

|  | $(1)$ <br> Control <br> $(\leq 0.1)$ | $(2)$ <br> Treated <br> $(\leq 0.1)$ | $(3)$ <br> Control <br> $($ All $)$ | $(4)$ <br> Treated <br> (All) |
| :--- | :---: | :---: | :---: | :---: |
| A: Polytechnics |  |  |  |  |
| Natural Resources | .017 | .025 | .019 | .024 |
| Technology \& Transportation | .255 | .231 | .276 | .231 |
| Business \& Administration | .349 | .369 | .347 | .378 |
| Tourism, Hospitality \& Household | .085 | .085 | .081 | .071 |
| Health \& Social | .259 | .260 | .239 | .264 |
| Culture | .097 | .101 | .090 | .102 |
| Humanities \& Education | .040 | .026 | .028 | .028 |
| Enrolled | .538 | .499 | .451 | .436 |
| Avg. income at 31-35 | 55.265 | 54.469 | 53.878 | 55.372 |
|  |  |  |  |  |
| B: Universities |  |  |  |  |
| Theology | - | - | .022 | .016 |
| Humanities | .161 | .149 | .156 | .176 |
| Arts, Design, Music \& Theater | .041 | .036 | .050 | .027 |
| Education | .094 | .106 | .099 | .092 |
| Sports \& Health Sciences | .025 | .024 | .022 | .018 |
| Social Sciences | .118 | .136 | .116 | .127 |
| Psychology | .014 | .019 | .012 | .021 |
| Law | .039 | .045 | .050 | .047 |
| Business | .157 | .174 | .174 | .168 |
| Natural Sciences | .217 | .203 | .197 | .235 |
| Agriculture, Forestry \& Vet. Medicine | .041 | - | .046 | .039 |
| Technology | .193 | .206 | .165 | .226 |
| Medicine \& Dentistry | .030 | .031 | .023 | .063 |
| Pharmacy | - | .019 | .012 | .022 |
| Enrolled | .347 | .401 | .217 | .595 |
| Avg. income at 31-35 | 58.566 | 58.016 | 54.791 | 61.920 |
| Observations |  |  |  |  |
|  | 1,633 | 1,999 | 11,119 | 16,923 |

Table 7: Mechanisms: Changes in Peer Group Characteristics

|  | GPA |  | Female Prop. |  |
| :--- | :---: | :---: | :---: | :---: |
|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
|  | Above | Below | Above | Below |
| A: Short-Run Outcomes |  |  |  |  |
| Finnish | -0.026 | $-0.110^{*}$ | -0.053 | $-0.123^{*}$ |
|  | $(0.063)$ | $(0.059)$ | $(0.056)$ | $(0.069)$ |
| English | $-0.164^{*}$ | -0.048 | $-0.163^{*}$ | -0.008 |
|  | $(0.086)$ | $(0.074)$ | $(0.086)$ | $(0.080)$ |
| Math (adv.) | 0.103 | -0.155 | -0.014 | 0.016 |
|  | $(0.147)$ | $(0.168)$ | $(0.152)$ | $(0.174)$ |
| Math (basic) | 0.181 | 0.004 | $0.427^{* * *}$ | $-0.226^{*}$ |
|  | $(0.125)$ | $(0.113)$ | $(0.120)$ | $(0.123)$ |
| HSEE GPA | 0.043 | -0.053 | 0.067 | -0.074 |
|  | $(0.051)$ | $(0.052)$ | $(0.048)$ | $(0.055)$ |
|  |  |  |  |  |
| B: Long-Run Outcomes |  |  |  |  |
| Application (Uni.) | 0.031 | $0.043^{*}$ | 0.020 | $0.051^{* *}$ |
|  | $(0.024)$ | $(0.025)$ | $(0.021)$ | $(0.021)$ |
| Application (Poly.) | -0.012 | -0.006 | -0.018 | 0.002 |
|  | $(0.0194)$ | $(0.0174)$ | $(0.019)$ | $(0.018)$ |
| Enrollment (Uni.) | 0.037 | 0.021 | 0.012 | $0.048^{* *}$ |
|  | $(0.023)$ | $(0.023)$ | $(0.025)$ | $(0.022)$ |
| Enrollment (Poly.) | $-0.086^{* * *}$ | -0.009 | $-0.055^{* *}$ | -0.028 |
|  | $(0.027)$ | $(0.024)$ | $(0.024)$ | $(0.025)$ |
| Bachelor's Degree | $0.042^{* *}$ | 0.020 | 0.031 | $0.039^{*}$ |
|  | $(0.021)$ | $(0.019)$ | $(0.022)$ | $(0.023)$ |
| Master's Degree | $0.055^{* *}$ | 0.013 | 0.027 | 0.031 |
|  | $(0.024)$ | $(0.018)$ | $(0.019)$ | $(0.022)$ |
| Polytechnic Degree | -0.011 | -0.004 | -0.020 | -0.007 |
| Income Percentile | $(0.022)$ | $(0.024)$ | $(0.024)$ | $(0.023)$ |
|  | $-2.495^{*}$ | 0.995 | -0.353 | -1.332 |
|  | $(1.387)$ | $(1.248)$ | $(1.348)$ | $(1.379)$ |

$$
\begin{aligned}
& \text { Standard errors in parentheses } \\
& { }^{* * *} \mathrm{p}<0.01,{ }^{* *} \mathrm{p}<0.05,{ }^{*} \mathrm{p}<0.1
\end{aligned}
$$

Table 8: Mechanisms: Changes in Parental Characteristics

|  | $(1)$ |  | $(2)$ |  |
| :--- | :---: | :---: | :---: | :---: |
|  | Parental Education | Family | Income |  |
|  | Above | Below | Above | Below |
| A: Short-Run Outcomes |  |  |  |  |
| Finnish | -0.017 | $-0.118^{* *}$ | 0.015 | $-0.160^{* * *}$ |
|  | $(0.063)$ | $(0.060)$ | $(0.063)$ | $(0.061)$ |
| English | 0.058 | $-0.183^{* *}$ | 0.029 | $-0.136^{*}$ |
|  | $(0.077)$ | $(0.078)$ | $(0.073)$ | $(0.078)$ |
| Math (adv.) | -0.013 | -0.069 | 0.126 | -0.131 |
|  | $(0.145)$ | $(0.157)$ | $(0.145)$ | $(0.154)$ |
| Math (basic) | 0.129 | 0.016 | 0.106 | 0.086 |
|  | $(0.114)$ | $(0.109)$ | $(0.117)$ | $(0.112)$ |
| HSEE GPA | 0.084 | $-0.103^{* *}$ | 0.084 | $-0.107^{* *}$ |
|  | $(0.052)$ | $(0.051)$ | $(0.051)$ | $(0.053)$ |
|  |  |  |  |  |
| B: Long-Run Outcomes |  |  |  |  |
| Application (Uni.) | $0.050^{* *}$ | 0.017 | $0.051^{* *}$ | 0.017 |
|  | $(0.020)$ | $(0.021)$ | $(0.022)$ | $(0.022)$ |
| Application (Poly.) | -0.016 | -0.005 | -0.002 | -0.016 |
|  | $(0.020)$ | $(0.018)$ | $(0.019)$ | $(0.017)$ |
| Enrollment (Uni.) | $0.055^{* *}$ | 0.004 | 0.035 | 0.027 |
|  | $(0.022)$ | $(0.023)$ | $(0.022)$ | $(0.023)$ |
| Enrollment (Poly.) | $-0.088^{* * *}$ | -0.005 | $-0.066^{* *}$ | -0.025 |
|  | $(0.027)$ | $(0.024)$ | $(0.027)$ | $(0.024)$ |
| Bachelor's Degree | $0.051^{* *}$ | 0.016 | $0.048^{* *}$ | 0.017 |
|  | $(0.023)$ | $(0.020)$ | $(0.021)$ | $(0.020)$ |
| Master's Degree | $0.046^{*}$ | 0.015 | $0.041^{* *}$ | 0.017 |
|  | $(0.024)$ | $(0.018)$ | $(0.019)$ | $(0.021)$ |
| Polytechnic Degree | -0.013 | -0.003 | 0.015 | -0.036 |
| Income Percentile | $(0.023)$ | $(0.025)$ | $(0.022)$ | $(0.024)$ |
|  | -0.673 | -0.683 | -2.145 | 0.593 |
|  | $(1.374)$ | $(1.289)$ | $(1.387)$ | $(1.286)$ |

$$
\begin{aligned}
& \text { Standard errors in parentheses } \\
& * * * \mathrm{p}<0.01,{ }^{* *} \mathrm{p}<0.05,,^{*} \mathrm{p}<0.1
\end{aligned}
$$

Table 9: Exit Exam Participation and Advanced Exams

|  | Estimate | Bandwidth | N |
| :--- | :---: | :---: | :---: |
|  | $(1)$ | $(2)$ | $(3)$ |
| Participation | 0.011 | 49.05 | 30,165 |
|  | $(0.010)$ |  |  |
| Advanced Math | 0.007 | 68.64 | 26,182 |
|  | $(0.015)$ |  |  |
| No. Advanced Languages | 0.006 | 60.92 | 26,182 |
|  | $(0.012)$ |  |  |

Standard errors in parentheses

$$
{ }^{* * *} \mathrm{p}<0.01,{ }^{* *} \mathrm{p}<0.05,^{*} \mathrm{p}<0.1
$$

### 6.2 Robustness

As a robustness check, we check whether those crossing the selective school threshold are more likely to participate in any of the high school exit exams. Also, we estimate whether selective schools make it more likely to choose advanced syllabus in math. We do this to see if the effects on math grades presented in Table 3 are driven by the probability to take the advanced test. Besides mathematics, it is possible to choose from different difficulty levels of test in languages. Thus, we check whether crossing the selective school threshold increases the amount of advanced language exams taken. All of these estimates are positive, but not statistically significant, as can be seen from Table 9.

We check whether our estimated effects are robust to the choice of the bandwidth. So far, we have used different bandwidths for each outcome. In this section we run the same regressions as before, but using 20 different bandwidths for each outcome. Figures 8-10 present short- and long-run effects based on different bandwidts. Overall, our results do not seem to depend on the choice of bandwidth.

## 7 Conclusions

This paper provides new evidence on the effects of selective schools on various short- and long-run educational and labor market outcomes. While we do not find evidence on effects of selective schools on short-run test scores or income


Figure 8: Robustness: Short-Run Estimates


Figure 9: Robustness: Application and Enrollment Estimates


Figure 10: Robustness: Degree and Income Estimates


Figure 11: GPA distribution of the applicants


Figure 12: GPA distribution near the standardized entrance threshold

Table 10: Covariate balance

|  | Estimate | Bandwidth | N |
| :--- | :---: | :---: | :---: |
|  | $(1)$ | $(2)$ | $(3)$ |
| Female | -0.006 | 54.88 | 30,162 |
|  | $(0.018)$ |  |  |
| Parental Education | 0.014 | 52.85 | 30,144 |
|  | $(0.012)$ |  |  |
| Family Income | 0.667 | 56.75 | 28,135 |
|  | $(1.028)$ |  |  |
| Information on Mother | -0.0003 | 72.30 | 30,165 |
|  | $(0.001)$ |  |  |
| Information on Father | 0.001 | 73.92 | 30,165 |
|  | $(0.004)$ |  |  |

$$
\begin{aligned}
& \hline \text { Standard errors in parentheses } \\
& * * * \mathrm{p}<0.01,{ }^{* *} \mathrm{p}<0.05,{ }^{*} \mathrm{p}<0.1
\end{aligned}
$$

in adulthood, we do find effects on 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. Our results also suggest 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.

Beuermann, Diether W., and C. Kirabo Jackson. forthcoming. "The Short and Long-Run Effects of Attending The Schools that Parents Prefer." The Journal of Human Resources.

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.

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.

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

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


[^0]:    *Kanninen: Labour Institute for Economic Research LABORE (ohto.kanninen@labore.fi); Kortelainen: University of Turku and VATT Institute for Economic Research (mika.kortelainen@utu.fi); Tervonen: Aalto University, Ekonominaukio 1, 02150 Espoo (lassi.tervonen@aalto.fi). We thank Manuel Bagues, Kristiina Huttunen, Tuomas Pekkarinen, Matti Sarvimäki and Marko Terviö for providing helpful comments.

