Swedish Institute for Social Research (SOFI)

Stockholm University

WORKING PAPER 2/2020

THE IMPACT OF ATTENDING AN INDEPENDENT UPPER SECONDARY SCHOOL: EVIDENCE FROM SWEDEN USING SCHOOL RANKING DATA

by

Karin Edmark & Lovisa Persson

The impact of attending an independent upper secondary school: Evidence from Sweden using school ranking data[♦]

March 24 2020

Karin Edmark[♣]
Lovisa Persson[♠]

Abstract

This paper provides a comprehensive study on how attending a Swedish Independent upper secondary school affects students' academic and short-term post-secondary outcomes. Beyond having access to population registers that measure school attendance and student outcomes, we are able to control for *student preferences* for independent provision, as stated in school application forms. The results from a CEM/VAM approach suggest a positive independent school effect on: final GPA, test results in English and Swedish, the likelihood of graduating on time, and attending post-secondary education. However, we also find a larger discrepancy between the final grade and the standardized test result among the independent school students, in a way that accords with more lenient grading practices among independent schools. Results from a difference-in-difference analysis around admission thresholds yielded no additional insights, due to imprecise estimates.

Key words: Private provision, mixed markets, voucher school reform, upper secondary education.

JEL-Classification: H44, I21, I26, I28.

[♦] This project has benefited from funding from the Swedish Research Council, Project number: 2014-01783. We are grateful for comments from seminar and conference participants at the Fifth Lisbon Research Workshop on Economics, Statistics and Econometrics of Education; the Swedish Institute for Social Research at Stockholm University (SOFI); the Research Institute of Industrial Economics (IFN); Maastricht University; Cambridge University; and GRIP seminar participants at Kristianstad University. We are particularly thankful for comments from Jan Sauermann, Anders Stenberg, Anna Sjögren, and Jonas Vlachos.

[♣] Stockholm University, IFAU, IFN and CESIfo; http://orcid.org/0000-0002-5629-5499; karin.edmark@sofi.su.se

[♠] Research Institute of Industrial Economics (IFN) and Kristianstad University; https://orcid.org/0000-0002-2988-6226; Lovisa.Persson@ifn.se

1. Introduction[1]

Upper secondary school is the last preparatory step before advancing to higher academic studies, entering the labor market, or commencing advanced vocational training. The effectiveness of upper secondary school consequently determines the quality of the academic abilities supplied to universities, the quality of the vocational abilities supplied to the labor market, as well as individual labor market prospects in general.

A widely discussed proposal to increase effectiveness in education is to allow entry for alternative providers with diverse approaches to learning and educational management, and/or let providers compete for students, hopefully in the dimension of educational quality (Friedman, 1955; Le Grand, 1991; Shleifer, 1998; and Hoxby, 2003). In the early 1990s, Sweden introduced a set of reforms that made entry with full voucher-funding possible for private providers of primary and secondary school. The result was a large expansion of the private school sector, in particular at the upper secondary school level (grades 10–12), where currently a quarter of all students attend a privately provided school; or as henceforth referred to, an *independent school*.[2]

We contribute to a growing international research literature on school vouchers by estimating the impact of attending a Swedish independent upper secondary school on a broad range of academic and short-run labor market outcomes. Our data consists of several merged official registers for the full population of students entering upper secondary school. The data enable us to control for student background characteristics in several dimensions. Most importantly, through our access to school application data, we are able to control for stated student preferences for different types of providers. Considering Sweden's journey from an almost complete public monopoly towards a very liberal school system by international standards, the case of Sweden should be of interest to the wider research community.

The previous literature evaluating the effects of the Swedish school voucher system is by large focused on primary and lower secondary education (grades 1–9). One exception is Hinnerich and Vlachos (2017), who show that upper secondary independent schools, on average, grade standardized tests 0.14 standard deviations more leniently than public schools: They show that the added value of independent schools is on average positive when regular teacher-graded test results are analyzed, but that it turns negative if externally re-graded test results are analyzed instead.[3] Böhlmark and Lindahl (2015) study the primary and lower secondary school level, and they find a small positive effect on students' academic achievements from an increasing independent market share within the municipality.[4]

[1] A snapshot of this project, which was made prior to generating the results presented in this manuscript, is available at https://osf.io/u8r43. After the registration of the snapshot, we encountered and corrected a few data errors, which have been corrected in this version. See the Appendix B, section B.5, for details.

[2] We use *independent* when referring to the Swedish version, and *voucher* as a catch-all term for international versions.

[3] Another forthcoming working paper, Edmark et al (2020) studies the impact of regional variation in the supply of independent schools over time.

[4] Other studies, that use the same type of municipality level variation, have found either no or positive effects on student's educational attainment, see Hennerdahl et al (2018), Sandström and Bergström (2005), Ahlin (2003), Björklund et al (2005).

Our study contributes to the previous literature in the following respects:

- First, we make use of information on stated preferences for independent vis-a-vis public schools as indicated by students' rankings of schools on official application forms. By controlling for student school rankings we reduce the risk that the results are biased due to differences in preferences for – or aversions against – either type of school.

- Second, we use a set of diverse estimation methods to paint a more robust empirical picture. That is, we estimate the effects of independent school attendance using i) a combination of coarsened exact matching (CEM) and value-added models (VAMs); and ii) a regression discontinuity inspired difference-in-differences analysis (RD/DID) around admission thresholds.

- Third, our analysis builds on data covering the entire student population, whereas Hinnerich and Vlachos (2017) are restricted to a subsample of roughly 10 percent of the student population for which there is information on externally regraded tests.

Our paper shares similarities with Kortelainen and Manninen (2019), who estimate a private school effect in Helsinki, Finland, using both RD around admissions thresholds and an added value approach. They find that the private school effect on matriculation exam scores is marginally positive but statistically insignificant. Unlike Kortelainen and Manninen (2019), we study a nationwide population of students, and we also have access to a broader range of student background variables. The Swedish case provides an interesting comparison to the Finnish case; the independent school sector in Sweden is larger, while at the same time, for-profits schools are allowed.

In general, the empirical evidence on the effects of voucher and charter schools on educational attainment is fairly inconclusive.[5] Several studies on U.S. data have found positive educational effects from attending charter[6] schools that adhere to the No Excuses approach (Dobbie and Fryer, 2019; Angrist et. al., 2013; Dobbie and Fryer, 2013; and Abdulkadiroglu et. al., 2011). Dobbie and Fryer (2019) also find positive effects on four-year college enrolment, and Angrist et. al. (2016) find that charter high schools in Boston (where many adopt the No Excuses approach) boost college preparedness. In a meta-study, Chabrier et. al. (2016) suggest that the "No-Excuses-effect" is driven by low performing fallback public schools in urban areas and intensive tutoring programs. Hahn et. al. (2018) show that high school students in private schools outperform high school students in public schools using data from Seoul, South Korea. On the other hand, Abdulkadiroglu et. al. (2018) suggest that participation in the Louisiana Scholarship (voucher) Program lowered student achievements. Studying the case of Chile, where a nationwide voucher system was implemented in 1981, Hsieh and Urquiola (2006) find no effects on educational achievement.

[5] The literature review provided in our paper is by no means exhaustive; we focus on more recent studies, or studies that are more relevant to our paper in terms of method or content relating to Sweden. We refer to e.g. Epple et. al. (2015) and Epple et. al. (2017) for reviews of earlier studies.

[6] In the context of the U.S. educational system, charter schools are usually schools that operate within the public school system. They are fully financed by public funds, and they are allowed to establish their own curriculum. However, they are usually more regulated than *voucher* schools, who are private schools that can be partly financed by a voucher given to all, or only some, families, depending on socio-economic status. Vouchers and charters share several core elements; student choice, public funding but private provision, curricular and organizational variety. Both voucher and charter schools are interesting comparison points for Swedish independent schools.

The results from our conditional-on-observables analysis using CEM/VAM suggest that attending an independent instead of a public upper secondary school has positive average effects on: students' final GPA, standardized test scores in English and Swedish, and also the likelihood of graduating on time. Analyses on subsamples suggest that the positive effects on final GPA are present in all parts of the ability distribution, and for students with varying socio-economic background. The positive effects on getting high test grades in English and Swedish are more pronounced in the upper part of the ability distribution, and the positive effect on graduating on time is more pronounced in the lower part. This is a reasonable pattern given that students with high abilities are on the margin of getting high grades, and students with low abilities are on the margin of graduating. Attending an independent school also has a positive effect on the probability of attending higher studies, including university studies, one year after graduation. Our results using CEM/VAM are robust to the use of different sampling and matching approaches, to bias-correction as suggested by Oster (2019), and to multiple hypothesis correction of p-values. However, our results from the RD/DID-analysis are overall too imprecise to be informative.

Our conditional-on-observables results mirror earlier results in Hinnerich and Vlachos (2017) who find a positive effect of independent school attendance on *teacher assessed achievements*. However, because of the potentially more generous grading standards in independent schools, documented in the same study, we cannot draw sharp conclusions about the *actual educational* added value of independent schools in Sweden, at least not based solely on our empirical results. Furthermore, we too find indications of more generous grading standards in independent schools. When we compare the standardized test grades with the final grades students get on the corresponding course, we find that students in independent schools are more likely to be "up-graded" on the course, compared to the test, but no more likely to be "down-graded. In the conclusions to this paper, we discuss the results in more detail.

2. Institutional overview Swedish upper secondary education[7]

Swedish students enter a 3-year long upper secondary education at age sixteen, after ten years of compulsory schooling. Upper secondary school is divided into six academic and twelve vocational tracks, and there is a 1-2 year long preparatory track for students whose grades do not qualify them to enter directly into any of the regular tracks. Although Swedish upper secondary school (*gymnasium*) is completely voluntary, virtually everybody – 99 percent – commences upper secondary studies.[8] [9] Upper secondary education can be provided either by the local governments, the municipalities, or by private entities; *independent schools*. Publicly (or municipally) provided and independently provided schools are both fully funded via school vouchers supplied by the municipalities, whose primary source of financing is the local income tax. Additional tuition fees are not allowed.

[7] For a more detailed review, see Appendix A.

[8] This number relates to the cohort finishing compulsory school in 2011, see Statistics Sweden (2017a).

[9] The graduation rate is however relatively low at 69 percent, compared to 85 percent in the US; 98 percent in Japan; and the OECD average at 86 percent (OECD, 2020). The low graduation rate is troubling, given that an upper secondary school diploma is correlated with significantly improved labor market prospects. In 2016, the unemployment rate among Swedish individuals aged 20–64 with an upper secondary degree was less than half of that of individuals who had not finished upper secondary education, even when restricting the comparison to individuals without post-secondary degrees (Statistics Sweden, 2017b).

As shown in Table 1, independent schools are on average smaller than their public counterparts, and they account for around one third of the market share in terms of school units. The academic track accounts for the larger share of students in both the independent and public schools and the vocational track share is roughly equal in both types of schools, but preparatory tracks are much less common in independent schools.[10] The fact that the preparatory track share is so small in the independent sector is one motivation for why we exclude preparatory tracks from our analysis. Finally, the number of students per teacher is slightly higher in the independent schools, even after adjusting for differences in academic, vocational and preparatory track shares.

TABLE 1. SCHOOL CHARACTERISTICS – SCHOOL YEAR 2013/14

	No. school units[a]	School size (No. students)	Academic tracks (student shares)	Vocational tracks (student shares)	Preparatory tracks (student shares)	Students per teacher, adjusted[b]
Independent	458	184	0.622	0.340	0.037	11.831
Public	882	273	0.562	0.320	0.118	11.368

[a] The definition of school units in the national School register changed in 2013. The new code is based on the division of headmaster responsibilities, rather than the physical school units. This has resulted in a large increase in administrative school units for the municipal schools: from 502 in school year 2011/12, to 766 in 2012/13 (after some schools had adopted the new system) and 882 in 2013/14 (when the new system was fully adopted. The number of independent schools was much less affected, and its numbers rather decreased over time; from 499 in 2011 to 484 in 2012 and 458 in 2013.
[b] The 0.5 percent top and bottom observations were excluded in order to eliminate the influence of extreme outliers, and the data was adjusted to account for the shares of students attending Academic, Vocational and Preparatory tracks, as these tend to have different student/teacher ratios. The raw data show a similar, but stronger, pattern of higher student/teacher ratios in independent schools.

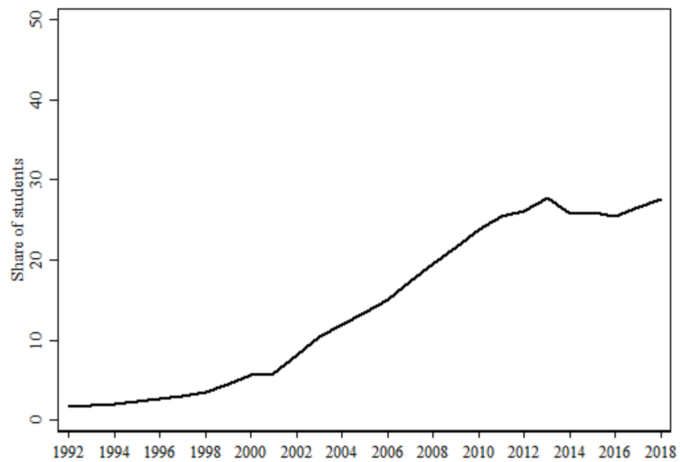
Entering upper secondary education is associated with making two choices: a choice of school and a choice of educational track. Under the current system, students choose simultaneously the school and track as one package, and are allowed to rank varying track and school combinations in their application.[11] Admission to a track and school combination is based on the grade sum, which is calculated as the sum of the grade credits of the' 16 highest graded subjects from lower secondary school (GPS9). In case of ties – i.e. several students with the same grade sum as the admission threshold – the school provider can choose from a list of allowed criteria, such as: specific subject grades, the rank of the choice, or chance. Students can choose from all independent schools in the country, and from the public schools in their home admission region. They may also apply to public schools outside of their region, but home students are then given priority in the admission process.[12]

The current regulatory framework for Swedish independent schools stems from a set of reforms implemented in the early 1990s, which greatly expanded the possibilities for independent agents to start schools and obtain full public funding.[13] As can be seen in Figure 1, the expansion of the independent market share following the school reforms was rapid, especially in the first decade of the 2000s, and a significant share of students opted for independent schools during our studied period of 2009–2013.

[10] This might be due to the fact that independent schools have only been allowed to offer these tracks since 2006.
[11] The first choice can be school A and a social science track, while the second choice could be school A and natural science track; or alternatively school B and a social science track.
[12] See Appendix A for more detailed information on admission rules that is of relevance for our RD/DID-analysis.
[13] The reforms are outlined in Propositions 1991/92:95 and 1992/93:23.

Specifically, the share of students attending an independent upper secondary school increased from 1.7 percent in 1992 to the highest share measured as of yet in 2013, at 28 percent.

FIGURE 1. INDEPENDENT UPPER SECONDARY SCHOOL MARKET SHARE



Source: Swedish National Agency for Education.

In 2013, the John Bauer group, which provided education to around 9000 students at the time, went bankrupt, which could be one of several explanations behind the stagnating independent share in the last five years.[14] [15] The bankruptcy sparked a debate about financial misconduct, and contributed to increased financial monitoring of independent providers.[16] Public oversight of the independent school sector had by then already developed from what was initially a relatively rudimentary system, to a system of more comprehensive and frequent monitoring. Since 2008, the Swedish Schools Inspectorate is responsible for the authorization of independent schools and for overseeing both public and independent schools. The Swedish Schools Inspectorate can close independent schools if severe violations are detected, but they are not authorized to close public schools.[17] [18]

The independent school reforms provided Sweden with a relatively liberal school system by international standards. For example, independent schools are allowed to be organized as for-profit organizations, such as the abovementioned JB-group. In fact, in 2013 – the year of the last cohort in our data – 85 percent of

[14] In a robustness analysis that is reported in Appendix C we find that excluding students affected by the JB bankruptcy has no qualitative impact on the results.
[15] See e.g. Sebhatu and Wennberg (2017) for an in-depth analysis of the JB-group.
[16] In 2014, the regular school inspections expanded to include also oversight of the financial situation of school providers. Starting from 2019 there are also stronger formal requirements on relevant experience and economic capability for private school providers, see https://www.skolinspektionen.se/sv/Tillstandsprovning/agar--och-ledningsprovning/.
[17] The Swedish School Inspection can instead temporarily take over the running of a municipal school. A proposal to expand the possibility to close also municipal schools is currently being investigated.
https://www.regeringen.se/pressmeddelanden/2019/11/okade-mojligheter-for-skolinspektionen-att-stanga-skolor-utreds/
[18] Chapter 3 in Angelov and Edmark (2016) describes the authorization of the independent schools in the early days of the reform, as well as the later developments. See also the National Agency for Education (Skolverket, 2004, pp. 21–22) for information on the government oversight.

upper secondary schools belonged to corporations. The remaining 15 percent were primarily organized as foundations or non-profit associations.

The government regulation concerning teaching- and instruction-related activities applies to independent and public providers alike: both are obliged to follow the same curriculum; meet the same educational goals; and use the same grading system. At the same time, school providers (or principals) – in the public as well as in the independent sector – have significant authority when making decisions that concern hiring, wage setting, allocation of resources within the school, and allocation of (a minimum total amount of) instruction time between courses and over the school year. In all essence, the same degrees of freedom offered by law thus applies to both public and independent providers, except for a few provisions regarding independent providers, including the possibility to organize as for-profit and to have a religious profile. However, both independent and public schools can profile themselves according to their offering of educational tracks, optional courses, and voluntary special instruction in sports, arts, or in other academic subjects.[19]

3. Data

The bottom line data set contains information on all individuals in Sweden that applied to upper secondary schools in 2009-2013, in the following referred to as the "application register".[20] This data set has been merged with a number of different population-wide registers held by Statistics Sweden (SCB), such as the upper secondary school attendance register, and registers containing information on students' graduation status, grades and test results, parental and student background characteristics, early work life, and post-secondary education. The sections below contain a presentation of sample restrictions, and descriptions of the variables used in the analysis.

3.1 Sample restrictions

Our sample restrictions are primarily motivated by the aim to obtain a more homogenous and comparable sample of independent and public school students. In some cases, however, observations have to be dropped because of difficulty of interpretation or suspected errors.

Starting with the original data set[21] , we limit the data set to students who start upper secondary education *immediately* after completing lower secondary school, at the age of 16, and who are *qualified* to enter a regular track in upper secondary school (i.e. who do not have to attend a preparatory track). These two restrictions reduce the sample size from 575,276 to 447,388 individuals. We then continue to make restrictions *i.–v.* below, based on information in the application register, and we thereby shrink the sample size to 296,890 individual observations. That is, we *drop* students who:

[19] For more detailed information about regulatory differences, see Section A1.1 in Appendix A.
[20] 2009 is the first year for which we observe the schools that students applied to – prior years of data show only listed track choices.
[21] The "original data set" refers to the sample size (575,276) after observations with missing observations on the following essential variables have been dropped: school ownership, educational track, and personal ID.

i. Have ranked only one school and track combination in the upper secondary application, because we want to be able to study students who have ranked both an independent and a public school.
ii. Are recorded as being admitted to more than one ranked preference, because for these students we cannot be sure which admission information is correct.
iii. Have applied to several different admission regions, or to independent schools with separate admission processes, because for these students there are several lists of ranked schools.
iv. Have applied to tracks for which they are unqualified, e.g. due to fail grades in certain subjects.
v. Were not admitted to their 1st or 2nd ranked alternative. Most students (approximately 80 percent of the applicants in the raw application data) are admitted to choice 1 or 2.

The fact that we can control for preferences for independent and public schools, as reflected in the application registers, is one of our most important contributions to the existing literature that, like us, relies on conditional-on-observables approaches. Our preferred strategy is to restrict the sample to students who have ranked both types of schools among their top two choices, which leaves us with a sample of 72 745 observations. When restricting preferences, we can be fairly certain that none of the students in the sample have any direct aversions, either rooted in political preferences or in experiences, to either of the school types, thus closing one selection channel.[22]

In addition to the *observational* sample above, we generate a separate sample to be used for an RD/DID-based analysis. This sample contains observations that are located around all competitive admission thresholds to independent and public schools. In other words, the sample will include cases where a student was marginally accepted (or not) to an independent instead of a public school, or to a public instead of an independent school. We therefore make an additional set of sample restrictions (please note, when reading what follows, that *admission group* is defined as combinations of $school \times track \times year$ based on the student's top ranked alternative). The restrictions *vi.–x.*, listed below, shrink the sample size to 12,060 individual observations. We *keep* only:

vi. Competitive admission groups, where the grade sum was actually binding in the sense that not all who applied were accepted.
vii. Admission groups that contain observations close to the admission threshold on both sides of the threshold. "Close" is defined as a ±10 grade sum unit interval; recall that the maximum grade sum is 320. (For regression specifications using a smaller data window of ±5 around the thresholds, we restrict the sample to admission groups with observations on both sides within this interval.)
viii. Individuals whose admission threshold for the lower ranked alternative was lower than that of the higher ranked alternative, and whose grades were higher or equal to than the threshold of the lower ranked alternative, so that the lower ranked school is a realistic fallback option.
ix. Individuals who have listed the same educational track for both the higher and lower ranked preference, so that we can isolate the independent/public-effect from potential track-effects.[23]
x. Individuals who apply to only non-artistic tracks, since admission to artistic tracks is not solely based on the grade sum, but also on practical admission tests.

[22] However, we will also, as a robustness check, perform estimations on the full observational sample and include dummy variables to control for the ranking of school types. These results are shown in Appendix C.
[23] We will also provide results after we relax this restriction.

In Table 2, we present an overview of the different samples and their sizes. We also show the distribution within the samples with respect to the exact rankings of the different types of schools. A more detailed exposé over sample restrictions can be found in Appendix B, Section B1.

TABLE 2. SAMPLES OVERVIEW

Observational samples	**Observations**	**RD/DID-samples**	**Observations**
Full sample	296 890	Full sample	12 060
Main sample (preference restriction)	72 745	Main sample	4 408
Independent (1) / Public (2)	34 320	Independent (1) / Public (2)	2 399
Public (1) / Independent (2)	38 425	Public (1) / Independent (2)	2 009
Supplementary samples	224 145	Supplementary samples	7 652
Independent (1) / Independent (2)	34 911	Independent (1) / Independent (2)	1 401
Public (1) / Public (2)	189 234	Public (1) / Public (2)	6 251

Note: (1) and (2) denotes the ranking of the school type.

3.2 Student background variables and other covariates

The richness of Swedish register data allows us to control for a comprehensive list of covariates on student background characteristics when estimating the effect of attending an independent school. In Table 3 we display the full list, and the averages values, of covariates for students attending independent and public schools respectively (columns 1–2).[24] The table also shows p-values (column 3) and normalized differences (Imbens and Rubin, 2015) (column 4). The sample used is the main observational sample that only includes students who have listed a mix of independent and public schools among the top two choices.

Student background characteristics in independent and public schools come across as remarkably similar according to the averages in Table 3. The (normalized) differences are less than 2 percent of the pooled standard deviations for 16 out of 20 variables. The strong balance in covariates between the samples is partly a result of restricting the sample to students who have listed both independent and public schools among the top two choices. As can be seen in Table B.6 in Appendix B, the selection into different types of schools is more pronounced when this restriction is not imposed.

Nevertheless, according to Table 3, independent school students are more likely to live in metropolitan municipalities, whereas students attending a public school are more likely to live in urban municipalities.[25] Independent school students are also somewhat more likely to have attended an independent school in grade 9. Since school choices made in lower level education are potentially important for later educational choices, we will control for 9th grade school in the empirical analysis.

[24] See also Table B.4 in Appendix B for basic summary statistics for all covariates.

[25] An urban municipality is where a majority of the population lives in urban areas, and in a rural municipality the opposite is the case. A metropolitan municipality is where at least 80 percent of the population lives in an urban area, and where the combined regional (metropolitan) population amounts to at least 500,000.

Additional controls for geography are added through the inclusion of upper secondary school municipality dummies.

TABLE 3. STUDENT BACKGROUND CHARACTERISTICS IN INDEPENDENT/PUBLIC SCHOOLS

	Independent	Public	P-value	Normalized diff.
	(1)	(2)	(3)	(4)
Household disposable income	246 095	243 240	0.177	0.010
One parent business income	0.145	0.142	0.237	0.009
One parent unemployed	0.186	0.179	0.019	0.017
One parent post-sec educ	0.550	0.554	0.276	-0.008
Both parents born in Sweden	0.730	0.729	0.739	0.002
One parent born in Sweden	0.124	0.123	0.545	0.004
No parent born in West	0.082	0.082	0.729	0.003
Born in Sweden	0.946	0.944	0.149	0.011
Born in West	0.024	0.026	0.144	-0.011
Born in non-West	0.030	0.030	0.554	-0.004
Female	0.514	0.519	0.237	-0.009
Independent9	0.186	0.172	0.000	0.037
GPS9	226.7	226.9	0.417	-0.006
High MA Test9	0.117	0.118	0.661	-0.003
High SW Test9	0.089	0.087	0.317	0.007
High EN Test9	0.224	0.216	0.015	0.018
Metropolitan municipality	0.453	0.411	0.000	0.086
Urban municipality	0.435	0.479	0.000	-0.087
Rural municipality	0.111	0.111	0.803	0.002
Regional independent share	0.254	0.248	0.000	0.078
Observations	*35,098*	*37,647*	*72,745*	*72,745*

Household income is represented per individual and in year 2016 monetary value. P-values refer to the raw differences. Missing values are replaced with imputed pooled averages. The normalized difference between samples 1 and 2 for covariate X is calculated as $(\bar{X}_1 - \bar{X}_2)/\sqrt{(S_1^2 + S_2^2)/2}$ (Imbens and Rubin, 2015).

3.3 Outcome variables

The cohorts in our data enter upper secondary education in 2009–2013 and are thus expected to graduate in 2012–2016. As 2016 is the last year recorded in our data, all post-graduation outcomes will be short-term in nature. While we are restricted to short-term outcomes, we have aimed to make use of the detailed register data to capture a broad range of the options available to students after upper secondary school. Our outcome variables thus include not only university/college educations, but also other post-secondary educations and labor income. The outcome variables are listed and categorized into three different groups in Table 4.

The outcomes in panel A are measured during, or at the end of, upper secondary school. This includes an indicator for switching school type during upper secondary school – from an independent school to a public school, or the reverse; the final 12th grade GPA, measured as the percentile rank by year among all graduating students; and the outcome “graduate on time”, i.e. after three years in upper secondary school. Students who have failed to graduate on time may either leave school without a degree (they will instead obtain a transcript of the completed courses), or complete upper secondary education by remaining in the incumbent school (or by switching to an adult education program).[26] We also define a dummy variable for remaining in upper secondary school for a 7th term, i.e. after the expected graduation.

TABLE 4. OUTCOME VARIABLES IN INDEPENDENT/PUBLIC SCHOOLS

	Independent	Public	P-value	Normalized diff.
	(1)	(2)	(3)	(4)
Panel A. Graduation and grades				
Switch independent/public	0.088	0.061	0.000	0.103
Pctile GPA12	57.143	53.217	0.000	0.140
Graduate on time	0.823	0.808	0.000	0.039
7^{th} term	0.093	0.103	0.000	-0.034
Panel B. Standardized tests				
Mathematics				
High test grade	0.059	0.050	0.000	0.038
Pass test grade	0.763	0.772	0.003	-0.021
Test grade>Course grade	0.017	0.013	0.000	0.032
Test grade<Course grade	0.301	0.271	0.000	0.066
Swedish				
High test grade	0.095	0.069	0.000	0.092
Pass test grade	0.948	0.944	0.000	0.019
Test grade>Course grade	0.098	0.099	0.614	-0.004
Test grade<Course grade	0.305	0.285	0.000	0.045
English				
High test grade	0.128	0.103	0.000	0.076
Pass test grade	0.978	0.977	0.812	0.002
Test grade>Course grade	0.108	0.111	0.195	-0.010
Test grade<Course grade	0.183	0.148	0.000	0.093
Panel C. Post-graduation[a]				
Study	0.383	0.368	0.000	0.032
Study no-prep[b]	0.312	0.295	0.000	0.037
UC≥15	0.152	0.144	0.011	0.022
Work≥50%	0.259	0.279	0.000	-0.045

The normalized difference for covariate X is calculated as $(\bar{X}_1 - \bar{X}_2)/\sqrt{(S_1^2 + S_2^2)/2}$ (Imbens and Rubin, 2015).
[a] Post-graduation outcomes are measured in the year following the graduation year, i.e. 4 years after entering upper secondary school.
[b] The pre-registered snapshot version of this table contained an error in this variable. This has been corrected, which is why the variable content for this variable differs from the same table in the snapshot.

In panel B we collect outcomes that are based on standardized tests taken in Mathematics, Swedish, and English throughout upper secondary school.[27] Our data lacks information on the exact test *scores*, but we

[26] Students have a general right to complete their upper secondary education, and have the right to retake or retest a failed course. The right to continue upper secondary education however transpires if the student is absent without valid reason for more than a month in a row.

do have information on the *grades* awarded on the tests. Based on this, we generate one outcome variable indicating whether the student was awarded a “high” grade or not, and one indicating whether the student was awarded a “pass” grade (all grades above fail) or not. In addition, we observe the final *course* grade corresponding to each test. The standardized tests are supposed to be a support and guide for the teachers’ assessments of students, but they are not strict determinants of the final course grades. In order to study how the test and course grades correspond to each other, we construct two dummy variables indicating if the test grade is higher or lower, respectively, than the course grade.

The timing and the number of tests taken varies across the educational tracks, and students in some tracks are tested in several sub courses in the same subject, resulting in multiple test observations per student. We will in our baseline estimations run the regressions on the student level averages for each outcome, such that each student gets the same weight.[28]

Panel C lists our post-graduation outcomes. We measure post-secondary school studies in the fall and create two indicator variables: the first takes on value 1 for all types of post-secondary studies, including both tertiary education (advanced and vocational training), and “complementary” types of studies such as adult complementary education, active labor market educational programs and Swedish for immigrants (see complete list in section B4.3 in Appendix B). The second dummy variable excludes the “complementary” types of studies. We also capture university studies separately by creating a dummy variable that takes on value 1 for taking university credits (UC) equivalent to 50 percent or more of a term of fulltime studies (≥15 UC). Finally, we measure labor market earnings in the form of a dummy variable for earning a “substantial amount” of labor income. We follow Forslund et. al. (2017), and define this amount as yearly earnings of at least half of the median annual work income among 45-year-olds.[29]

We recognize that studying post-graduation outcomes in the same year as graduation is probably premature, since many students choose to take a sabbatical year to work or study abroad, and we will therefore show results when measuring outcomes one year after graduation in our main results tables (4 years after entering upper secondary school). This in effect means that we are excluding the 2013 cohort from the analysis of post-graduation outcomes.[30]

Table 4 shows outcome variable averages for students attending independent and public schools respectively, as well as the normalized differences and p-values. Similar to Table 3, we use the main observational sample, which is restricted to students who have applied to a combination of independent and public schools as the two top choices. According to the raw differences in Table 4, independent school students are: more likely to switch school type, to have a higher GPA12, and somewhat more likely to graduate on time. In all test subjects, students in independent schools are more likely to receive the highest grade. The test grade is also commonly found to be more loosely connected to the final course

[27] In order to account for the fact that the exact timing and number of tests take varies across tracks, and sometimes even across schools within a track, we include fixed effects for the timing in terms of the year, school grade and term and for the course tested. These are also relevant to include due to the fact that the grading system changed during the time studied, see section B4 of Appendix B for details.

[28] We have also, as a robustness test, estimated the regressions when using each test as the level of observation. The results are overall very similar.

[29] According to Forslund et al (2017), this corresponds roughly to six months’ worth of wages for a full time employed janitor in the municipal (public) sector. A “substantial amount” is redefined as a quarter of the median income among 45-year olds, when studying outcomes in the graduation year, as the students were still in upper secondary education approximately half of that year.

[30] We also show estimates when measuring post-graduation outcomes in the same year as graduation, see Appendix C.

grade in independent schools. In particular, the course grade given is more often higher than the grade awarded on the standardized test. When looking at post-graduation outcomes, we see that the largest difference, in favor of students in public schools, is found in the propensity to work at least 50 percent one year after graduation. Furthermore, students in independent schools are somewhat more likely to register in any kind of studies in the year after graduation.

4. Empirical methods and results

4.1 Overview of the empirical estimations

We estimate the effect of attending an independent school using both conditional-on-observables approaches and quasi-experimental approaches. That is, we run added-value regressions (VAMs) that are combined with coarsened exact matching (CEM), as a way of enforcing common support, but we also estimate a regression discontinuity inspired difference-in-difference analysis (RD/DID) around admission thresholds. We argue that these different strategies, by making use of somewhat different sources of identifying variation and by relying on partly different assumptions, together provide a more comprehensive analysis than each analysis would yield on its own. Before going into the details, we write down the basic regression equation for our analysis problem as:

$$y_i = \alpha + \beta IND_i + u_i \qquad (1),$$

where y_i denotes some outcome for upper secondary student i; α is an intercept; and IND_i is a dummy variable for whether or not the students attended an upper secondary independent school instead of a public school as measured at the start of upper secondary education; and u_i is the error term.

If independent and public students were comparable in all aspects apart from what type of school they attended, the β-coefficient from equation (1) would capture the average causal effect of attending an independent – instead of a public – upper secondary school. In practice, however, independent and public students may very well differ systematically in ways that are correlated with the outcomes studied, and that may or may not be observable to the researcher.

As was explained in the introduction we deal with the selection problem by restricting the data sample to students who have expressed a relatively strong preference for both independent and public schools, more precisely, those that have listed a combination of the two school types as the top two choices in the upper secondary school applications. We then address potential remaining student selection by either conditioning on observable characteristics (combining CEM and VAM), or by further restricting the sample to students who can (more or less) plausibly be assumed to be comparable in an RD/DID-type analysis. The below sections present these respective approaches and their results in order.

4.2 Combination of CEM and VAM

4.2.1. Overview of CEM/VAM

Since admission to Swedish upper secondary school is not based on lotteries[31], we follow Dobbie and Fryer (2019), and Hinnerich and Vlachos (2017), and implement an empirical approach that combines CEM (coarsened exact matching) and VAMs (value-added-models). Abdulkadiroglu et. al. (2011) show that a conditional-on-observables approach yields test score estimates for oversubscribed Boston charter schools that are similar to the estimates obtained when leveraging charter school lotteries.[32] In a follow-up,study, Angrist et. al. (2017) conclude that the bias contained in VAMs is, in the charter school context, moderate and statistically significant. Importantly, they suggest that the bias is small enough to render observational estimates useful from a policy perspective. We take this as suggesting that VAMs may yield policy relevant estimates also in the present Swedish case; in particular as we have access to a broad set of student background variables including prior achievement and stated preferences for different types of schools.

We use CEM in order to obtain common support with respect to combinations of a set of background variables that can be considered particularly important for school choice and subsequent outcomes. We force exact matching on the following variables: gender, parents' country of birth (three dummies), GPS9 quintile, and, depending on the specification, either the county[33] where the student attends lower secondary school, or the school the student attended in 9th grade. After carrying out the matching procedure, we keep only the cells that contain both independent- and public school students.[34]

The two samples generated, one matching on upper secondary school county, and the other on 9th grade school, are then alternately used for estimating VAMs; regression models where student background variables and student's prior academic achievements are controlled for in a flexible manner (see the table notes to Tables 5.A–C for the exact covariate specification). The empirical model, which builds on equation (1), is displayed below:

$$y_{itmcp} = \alpha_t + \beta IND_i + \delta \boldsymbol{A_{it-1}} + \varphi \boldsymbol{X}_i + u_m + u_c + u_p + u_{itmcp} \tag{2}$$

Similarly to equation (1), y_{itmcp} denotes some outcome variable for student *i*, at time *t* (but note that the time when the outcome variable is measured varies), in upper secondary school municipality *m*, who attended 9th grade school *c*, and who is enrolled in track *p*. Furthermore, α_t denotes time (or cohort) fixed effects; IND_i is a dummy variable for attending private school (measured in October of the first year of upper secondary school); $\boldsymbol{A_{it-1}}$ denotes prior academic achievement; $\boldsymbol{X}_i$ denotes the remaining set of

---

[31] Lotteries may only be used in cases where several students with equal grade sum compete for the last available slot.

[32] Studying the effectiveness of charter schools in New York City, Dobbie and Fryer (2013) also show that observational estimates and lottery estimates can be qualitatively similar, although in their case the observational estimates are somewhat smaller in size. Deming (2014) present observational estimates that are similar to lottery estimates, using data on charter school lotteries in Charlotte-Mecklenburg, North Carolina.

[33] There are 21 counties in Sweden, sometimes referred to as "regions".

[34] The variables used for exact matching have been chosen to align ourselves with the previous literature, in particular with Hinnerich and Vlachos (2017), but also Dobbie and Fryer (2019). When we match on 9th grade school, we also add cohort dummies, since 9th grade school IDs cannot always be correctly linked over time.

student background characteristics that are potentially correlated with both independent school attendance and the outcome variable. Upper secondary school municipality fixed effects are included in u_m, 9th grade school fixed effects are included in u_c, track fixed effects are included in u_p, and u_{itmcp} is the error term.

Under the assumption that the included covariates and fixed effects successfully capture all systematic background differences between independent and public school students that remain in the restricted and matched samples and that are correlated with the outcome variable, the β-coefficient in equation (2) corresponds to the average treatment effect (ATE) of attending an independent school, for the sample population.

This conditional independence assumption cannot be tested. We will however construct bounds for the β-coefficient under different assumptions on the relation between unobserved and observed selection.[35] This analysis will tell us: i) what are the bounds of the β-coefficient if we assume that the selection on unobserved student characteristics is the same, half of, or double, that of the observed selection; and ii) how large would the selection on unobserved variables need to be, as a share of the observed selection, for the true β-coefficient to be zero.

4.2.2. CEM/VAM results

The VAMs are estimated on three different samples, resulting in three 3-column tables, one table for each outcome group of Table 4. Results for outcomes in group A, "Graduation and grades", are shown in Table 5.A. Column 1 shows the results for the most restricted sample, where we enforce both the preference restriction (having applied to a combination of independent and public schools for the top two options), and common support with respect to 9th grade school, gender, parents' country of birth (three dummies) and GPS9 quintile. The specification in Column 2 is our preferred specification; here we also enforce the preference restriction, but enforce common support with respect to upper secondary *school county* instead of 9th grade *school*, which has the benefit of retaining substantially more observations. In Column 3 we use the full observational sample without adding preference restrictions or preference controls. The difference between Column 3 and the remaining two columns thus shows the potential importance of utilizing information on student preferences.

Although sample sizes vary greatly as a result of alternating the restrictions discussed above, the results in Table 5.A are fairly stable across specifications. The results in the first row suggest that independent school attendance has an impact on the probability of switching school type: independent school students are more likely to switch to public schools than the other way around. The effect size of 2.3 p.p. in column 2 (our preferred specification) is quantitatively similar to the raw difference presented in Table 4. One could suspect that the effect is a result of the bankruptcy of the John Bauer (JB) group, described in Section 2.1, forcing students to switch to public schools. In a robustness analysis, which is presented in section C2 in Appendix C, we however find that the estimate for switching school is similar (0.018, compared to 0.023 in Table 5.A) also when we exclude the JB-school cases from the sample.[36] The higher

[35] This analysis will be carried out using the STATA command psacalc, see Oster (2019).

[36] More specifically, this is done in the following manner: As we lack access to school names, we cannot drop students attending JB-schools. Instead, we have dropped all observations belonging to a track×municipality×year combination where a JB-school was present.

propensity to switch from an independent to a public school rather than the reverse could thus be related to higher volatility in the independent school market in general, or to the fact that students in independent schools are more often dissatisfied with their school. The results from the analysis excluding the JB-cases are also very similar to the baseline estimates for the other outcome variables, see section C2 in Appendix C.

TABLE 5.A. GRADUATION AND GRADES – CEM/VAM

	(1)	(2)	(3)
Switch independent/public	0.0279***	0.0231***	0.0343***
Standard error	(0.0057)	(0.0049)	(0.0042)
P-value	[0.0000]	[0.0000]	[0.0000]
Observations	28837	70623	288762
Pctile GPA12	4.4302***	4.4906***	4.5410***
Standard error	(0.3474)	(0.3080)	(0.2955)
P-value	[0.0000]	[0.0000]	[0.0000]
Observations	25578	61898	254937
Graduate on time	0.0229***	0.0288***	0.0200***
Standard error	(0.0051)	(0.0041)	(0.0038)
P-value	[0.0000]	[0.0000]	[0.0000]
Observations	29440	72220	294580
7th term	-0.0133***	-0.0153***	-0.0099***
Standard error	(0.0038)	(0.0028)	(0.0024)
P-value	[0.0004]	[0.0000]	[0.0000]
Observations	29440	72220	294580
Preference restriction	YES	YES	NO
CEM on 9th grade school	YES	NO	NO
CEM on county	NO	YES	YES

Note: All regressions above include the following covariates: upper secondary school municipality dummies, prior achievement as controlled for by a cubic form of GPS9 and GPS9 quintile dummies, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade, 9th grade school dummies, track dummies, log household income, income decile dummies; and dummies indicating the following: gender, born in western country (excl. Sweden), born in non-western country, one parent post-secondary education, both parents born in Sweden, one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. Columns 1 and 2 also include a dummy indicating admission to first ranked school. Standard errors are clustered on upper secondary school. *** p<0.005, ** p<0.01, * p<0.05

Results in Table 5.A. also suggest that independent school attendance has a positive impact on the ranking of the student's GPA in 12th grade. The effect size of 4.49 percentiles in column 2 is quantitatively similar to the raw difference in Table 4. The positive independent school impact on the likelihood of graduating on time at 2.88 percentage points in column 2 is somewhat larger than the raw difference, while the negative independent school impact of staying behind for a 7th semester at 1.53 p.p. is in line with the raw difference.

Results for standardized test outcomes are shown in Table 5.B. We study four outcomes, a dummy for obtaining a pass test grade, a dummy for obtaining a high test grade, a dummy for obtaining a final course grade that is *lower* than the test grade, and finally, a dummy for obtaining a final course grade that is *higher* than the test grade. The result that stands out the most is the positive coefficient on the probability of getting a final course grade that is higher than the grade obtained on the standardized test. The coefficient is statistically significant and/or economically interesting across all samples and across all subjects. The increased probability is the highest in Mathematics at 4.62 percentage points in column 2, and the smallest in Swedish at 2.39 percentage points in column 5. The increased probability is 4.01 percentage points in English, see column 8.

A second striking result in Table 5.B is the positive coefficient on the probability of getting a high grade on the standardized tests in Swedish and in English. The probability increases with 2.04 percentage points in Swedish, and 1.55 percentage points in English. The coefficient for Mathematics is only statistically significant in the full sample estimation, but the coefficient is then relatively small at 0.5 percentage points. Remaining results that are statistically significant are also relatively small in economic terms. To summarize the findings reported in Table 5.B; attending an independent school increases the probability that you will become upgraded in your final course grade compared with the grade you received on your standardized test. This is true for all subjects. The probability of actually getting a high test grade only increases in Swedish and in English but not in Mathematics.

Finally, Table 5.C shows results for post-graduation outcomes. In column 2 the estimate of 1.99 suggests that attending an independent school has a positive impact on the probability of participating in any type of studies one year after graduation. This is the case even when excluding studies that are of a preparatory type, and the estimate then even increases to 2.44 percentage points. The effect on the probability of obtaining at least 15 university credits is also positive at 1.42 percentage points. However, the effect on the probability of earning labor income corresponding to at least a half time job is negative at -1.70 percentage points. The effects shown in Table 5.C. have the same signs as the raw differences in Table 4, and the effect sizes are quantitatively similar.

It can be noted that our results are not hugely sensitive to restricting the sample to only include individuals with preferences for both types of schools (compare full sample results in column 3 with other columns). Our results thus provide some support for the conditional on observables analysis in Hinnerich and Vlachos (2017), which does not make use of information on student preferences.

TABLE 5.B. STANDARDIZED TEST RESULTS – CEM/VAM

	Mathematics			Swedish			English		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High test grade	0.0028	0.0034	0.0050***	0.0190***	0.0204***	0.0227***	0.0161***	0.0155***	0.0182***
Standard error	(0.0028)	(0.0020)	(0.0017)	(0.0059)	(0.0046)	(0.0043)	(0.0051)	(0.0040)	(0.0036)
P-value	[0.3165]	[0.0878]	[0.0028]	[0.0014]	[0.0000]	[0.0000]	[0.0016]	[0.0001]	[0.0000]
Observations	20062	48106	193412	22052	52515	222275	20798	50202	201527
Pass test grade	-0.0090	-0.0028	0.0003	0.0065	0.0065*	0.0055*	0.0028	0.0023	0.0039***
Standard error	(0.0066)	(0.0054)	(0.0048)	(0.0034)	(0.0027)	(0.0021)	(0.0022)	(0.0016)	(0.0012)
P-value	[0.1733]	[0.6045]	[0.9443]	[0.0593]	[0.0151]	[0.0102]	[0.2056]	[0.1528]	[0.0018]
Observations	20062	48106	193412	22052	52515	222275	20798	50202	201527
Test grade>Course grade	0.0042	0.0039*	0.0044***	-0.0056	-0.0035	0.0004	0.0035	-0.0036	0.0011
Standard error	(0.0022)	(0.0017)	(0.0014)	(0.0053)	(0.0040)	(0.0032)	(0.0053)	(0.0044)	(0.0038)
P-value	[0.0629]	[0.0228]	[0.0020]	[0.2888]	[0.3735]	[0.8975]	[0.5052]	[0.4134]	[0.7788]
Observations	19322	46244	185415	20482	48724	202921	19962	48159	193922
Test grade<Course grade	0.0510***	0.0462***	0.0533***	0.0183*	0.0239***	0.0223***	0.0307***	0.0401***	0.0419***
Standard error	(0.0102)	(0.0088)	(0.0086)	(0.0089)	(0.0069)	(0.0057)	(0.0073)	(0.0058)	(0.0051)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0397]	[0.0005]	[0.0001]	[0.0000]	[0.0000]	[0.0000]
Observations	19322	46244	185415	20482	48724	202921	19962	48159	193922
Preference restriction	YES	YES	NO	YES	YES	NO	YES	YES	NO
CEM on 9th grade school	YES	NO	NO	YES	NO	NO	YES	NO	NO
CEM on county	NO	YES	YES	NO	YES	YES	NO	YES	YES

Note: Regressions are performed on individual means within each subject. All regressions above include the following covariates: upper secondary school municipality dummies, prior achievement as controlled for by a cubic form of GPS9 and GPS9 quintile dummies, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade, 9th grade school dummies, track dummies, log household income, income decile dummies; and dummies indicating the following: gender, born in western country (excl. Sweden), born in non-western country, one parent post-secondary education, both parents born in Sweden, one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. Regressions on test outcomes also include test specific dummies. Columns 1–2, 4–5, and 7–8 also include a dummy variable indicating whether the student was admitted to the first ranked choice. Standard errors are clustered on upper secondary school. *** p<0.005, ** p<0.01, * p<0.05

TABLE 5.C. POST-GRADUATION OUTCOMES – CEM/VAM

	(1)	(2)	(3)
Study	0.0261***	0.0199***	0.0165***
Standard error	(0.0068)	(0.0050)	(0.0041)
P-value	[0.0001]	[0.0001]	[0.0001]
Observations	22598	55430	230160
Study no-prep	0.0265***	0.0244***	0.0195***
Standard error	(0.0061)	(0.0046)	(0.0039)
P-value	[0.0000]	[0.0000]	[0.0000]
Observations	22598	55430	230160
UC≥15	0.0137***	0.0142***	0.0111***
Standard error	(0.0048)	(0.0034)	(0.0026)
P-value	[0.0043]	[0.0000]	[0.0000]
Observations	22598	55430	230160
Work≥50%	-0.0207***	-0.0170***	-0.0233***
Standard error	(0.0067)	(0.0049)	(0.0044)
P-value	[0.0021]	[0.0005]	[0.0000]
Observations	22585	55386	229988
Preference restriction	YES	YES	NO
CEM on 9th grade school	YES	NO	NO
CEM on county	NO	YES	YES

Note: All outcomes are measured one year after graduation. All regressions above include the following covariates: upper secondary school municipality dummies, prior achievement as controlled for by a cubic form of GPS9 and GPS9 quintile dummies, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade, 9th grade school dummies, track dummies, log household income, income decile dummies; and dummies indicating the following: gender, born in western country (excl. Sweden), born in non-western country, one parent post-secondary education, both parents born in Sweden, one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. Columns 1 and 2 also include a dummy variable indicating whether the student was admitted to the first ranked choice. Standard errors are clustered on upper secondary school. *** p<0.005, ** p<0.01, * p<0.05

4.3 Regression Discontinuity based Differences-in-Differences (RD/DID)

4.3.1 Overview of the RD/DID

Our second approach to estimating the independent school effect uses the fact that admission to an *oversubscribed* school and track combination is a deterministic function of the applicants' grade sum in 9th grade (GPS9).[37] One approach to causally identify the effect of being admitted to an independent – instead of a public school – (or the reverse) is thus to exclusively use variation among students with grade sums on the margin of admission thresholds in a Regression-Discontinuity (RD) framework. If the running variable is continuous (enough), the Local Average Treatment Effect (LATE) can be estimated by flexibly controlling for its direct influence on the outcome variable. However, since our running variable – the grade sum deviation from the admission threshold – is relatively discrete[38] [39] we instead

[37] Admission is in the form of deferred acceptance with the lower secondary final grade sum as determining factor – i.e. the order in which preferences are ranked does not matter for the probability of admission.

[38] The final grade sum increases in discrete steps that reflect getting a higher grade in one out of the total of 16 subjects. It contains distinct mass points, as is clearly visible from the figures in section B.6.1 in Appendix B. The number of mass points in

compare means among observations very close to admission thresholds[40], and combine this RD-feature with a difference-in-difference (DID) estimation.

Our RD/DID-strategy can be summarized as follows: we compare differences in outcomes between students who ended up in *different types* of schools due to having marginally different grade sums, with differences between students who were also just below of just above admission thresholds, but who listed *the same type* of school as their first and second choice.[41] The RD/DID-strategy is illustrated in Table 6.

TABLE 6: RD/DID ILLUSTRATION

	Just below Threshold (D=0)	Just above Threshold (D=1)
"Treated sample" (T=1) *(Combination of independent/public)*	$\bar{X}_{TD} = \bar{X}_{10}$	$\bar{X}_{TD} = \bar{X}_{11}$
"Untreated control sample" (T=0) *(Only independent or only public)*	$\bar{X}_{TD} = \bar{X}_{00}$	$\bar{X}_{TD} = \bar{X}_{01}$
DID-estimate	$(\bar{X}_{11} - \bar{X}_{10}) - (\bar{X}_{01} - \bar{X}_{00})$	

Using a control sample of students who only rank one type of school has some advantages. For these students, crossing the threshold is associated with having slightly higher grade sums, but it is *not* associated with attending different types of schools. By subtracting the above-below difference in these groups from the difference within the "treated" sample of students, we can thus control for potential confounding effects related to the distance to the admission threshold. An additional benefit is the possibility of differencing out two admission-effect channels, namely that students above the threshold are (1) admitted to their most preferred option, and (2) among the academically worst performing students in the class in terms of their prior achievements – two properties that are inherently linked to being close to admission thresholds and thus part of what a conventional RD would estimate.

The RD/DID-regression is implemented by estimating Equation 3:

$$y_{igt} = \alpha_{gt} + T_{it} + \varphi D_{it} + \beta(T_{it} \times D_{it}) + \delta A_{it-1} + \varphi \boldsymbol{X}_{it} + u_{igt} \qquad (3),$$

our data is furthermore fairly low – around 80 in the full RD-data sample, and closer to 40 if we focus on the data window within 50 units from the admission thresholds where the bulk of the data is located.

[39] Cattaneo et. al. (2018) discuss the issue of discrete running variables, and point out that "if the score is discrete but the number of mass points is sufficiently large, using local polynomial methods may still be appropriate. In contrast, if the number of mass points is very small, local polynomial methods will not be directly applicable". Cattaneo et. al. also point out that this is of relevance for the education literature: "despite being continuous in principle, it is common for test scores and grades to be discrete in practice". See also Kolésar and Rothe (2018) and Imbens and Wager (2018) for more examples of the recent and expanding literature on optimal bandwidth selection and how to deal with discrete running variables, or Lee and Card (2008) for an earlier reference.

[40] This part of the analysis resembles the suggestion by Cattaneo et. al. (2018) to compare averages just above and below the cutoff when the data is discrete.

[41] Section B.6.2 in Appendix B shows that the likelihood to attend an independent school increases discontinuously at the admission threshold for students with an independent school as their first choice, and a public one as their second, and the reverse holds for students with the opposite preference ordering. There is no discontinuous change for students who have listed either only independent or only public schools as the top two choices. Section B.6.1 furthermore shows no evidence of bunching at or just above the admission threshold, and section B.6.4 shows no evidence discontinuities in student background covariates at the thresholds.

where y_{igt} is the outcome variable for individual *i* in *admission group g* and year *t*. "Admission group" refers to the track×year×school combination of the student's first choice, i.e. the combination to which an admission threshold applies. α_{gt} denotes an admission group fixed effect [42]; T_{it} is a dummy variable for being in a *"treated sample"* (having applied to a combination of independent/public schools), rather than in a *"control sample"* (having listed only one type of schools); and D_{it} is an indicator for having a grade sum higher than the admission threshold. Since we have restricted the sample to only include students who list the same track as first and second choice, as was described in Section 3.1, the track fixed effects are effectively incorporated in the admission group effect. [43]

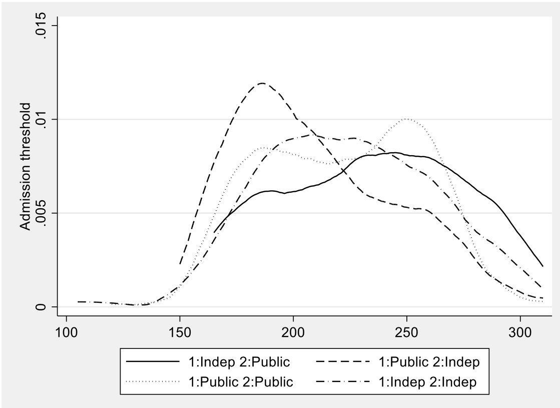
We estimate equation (3) separately for our two treatment groups. In other words, the β-coefficient will alternately estimate the impact of being above the admission threshold for students who listed an independent school as first and a public school as second preference, and the impact of being above the threshold for students with the reverse preference ordering. This means that crossing the admission threshold will be associated with a *higher* likelihood of attending an independent school for the former group, and a *lower* likelihood for the latter. Finally, the estimations include a set of student level covariates (see table notes to tables 7.A–C for the full list), $\boldsymbol{X}_{it}$, as well as the final grade sum from lower secondary education, A_{it-1}. Since we alternate between including either one of the two control groups (students who only rank independent schools or students who only rank public schools) equation (3) is estimated four times.

The specification in equation (3) rests on the assumption that the "above–below-threshold"-differences in the *control samples* are relevant counterfactual differences in case of no treatment for the *treated samples*. While this cannot be tested, Figure 2 indicates that the admission thresholds are mostly similarly distributed across the samples, although the sample of students with the preference ordering "1:Public/2:Independent" has somewhat more mass in the lower parts of the distribution. However, when we compare students with preferences "1:Independent/2:Public" with either control group, we are comparing groups that are fairly similar in terms of where thresholds are placed.

[42] We will also present results using the following alternative specifications: i) adding a basic set of student background characteristics, including 9th grade final grade sum, as covariates; ii) including application round (track×year×school) fixed effects for both the top and the second listed preference; iii) including fixed effects for the interaction admission groups (track×year×school) for the two first listed preferences. After the publication of the research plan for the project at the Open Science Framework, we also decided to add the following specifications: For the largest of the data windows, we will additionally present estimates including linear trends for the distance to the admission threshold, estimated separately above and below the threshold, as well as results when the sample includes students with different tracks as 1st and 2nd preference (and including fixed effects for these tracks).

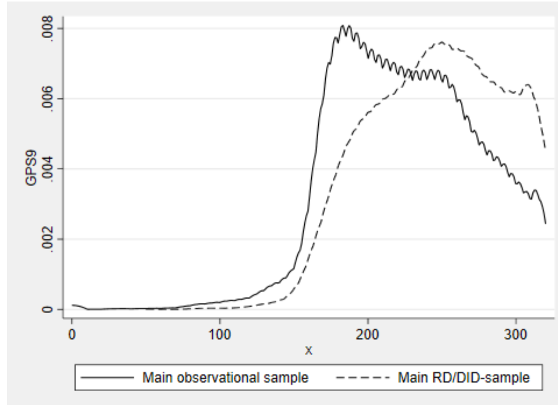
[43] In an alternative specification, which will be shown in Appendix C, we add students with different tracks listed as first and second preference in order to increase the sample size, and estimate separate fixed effects for each track preference.

FIGURE 2: KERNEL DENSITY DISTRIBUTIONS OF THE ADMISSION THRESHOLDS FOR THE FOUR RD/DID-SAMPLES



Note: This figure uses the RD/DID-samples listed in Table 2.

FIGURE 3: KERNEL DENSITY DISTRIBUTIONS OF THE GRADE 9 GRADE SUM FOR THE MAIN OBSERVATIONAL AND MAIN RD/DID-SAMPLES



Note: RD/DID-sample for the students with mixed preferences

Even though we use a different sample in the RD-analysis, compared with the CEM/VAM analysis, there is still substantial overlap in the distribution of grade sum, see Figure 3. The RD/DID sample has more density at the higher end of the distribution, which is partly a consequence of including only "competitive" or "oversubscribed" admission groups.

4.3.2 RD/DID-Regression results

In our baseline specification, "closeness" is defined as 5 points away from the admission thresholds.[44] For cohorts starting upper secondary school in 2009-2012, a 5-point increase in the grade sum corresponds to a one-step increase in the actual grade. For the last cohort in our data, 2013, the smallest unit increase is instead 2.5 points, due to a change in the grading system which increased the number of distinct pass grades from 3 to 5, see section B3 in Appendix B for details.

The results from the estimation of Equation (3) using a ±5 point window, are shown in Tables 7.A–C. In addition to showing the β-coefficient of the interaction variable *being admitted and being in the treatment group* $(T_{it} \times D_{it})$, the first row in each table shows the impact of the same interaction variable on actually *attending* an independent school in October the first year. The impact on attending an independent school is, as expected, positive when we study students who listed an independent school as first choice, and negative when a public school was the first choice. The relationship is always strongly statistically significant and around 40–50 percentage points in magnitude.[45] The IV-coefficients from using "crossing

[44] We also provide estimates for slightly larger data windows (10/20 units) in Appendix C. For these estimations, we add linear trends in the running variable, estimated separately on either side of the admission thresholds. Observations exactly at the threshold are excluded from all estimations, since the admission rules are unclear for ties at the threshold, see section A.3 in Appendix A for details.

[45] As can be seen in Appendix B, section A.6.2, the relationship is stronger for observations further away from the admission threshold, which can be explained by the possibility that students near the admission thresholds are more likely to be subject to changes in admission status after the initial admission round, when some students may have changed their mind and others admitted from waiting lists. See also section A.6.2 and A.6.3 for potential sources of error in the admission threshold that can give rise to imprecision.

the admission threshold" as an instrument for attending an independent school, is obtained by dividing the β-coefficients by these first-stage estimates, i.e. circa 0.5.[46]

Table 7.A shows the results for the graduation and grade-related outcomes. The RD/DID-estimates are, with one exception, of the same sign as their CEM/VAM-counterparts: being marginally admitted to an independent- instead of a public school (columns 1–2) is positively correlated with: switching school and graduating on time, and negatively correlated with staying behind for a 7th term. The opposite relations hold when the first choice is a public school (columns 3–4). The coefficient for the percentile rank GPA goes in the same direction as the CEM/VAM-results in columns (1), (3) and (4), but not in column (2). That said, none of the estimates in Table 7.A is statistically significant at any conventional level – the confidence intervals are very large.

In order to gain precision, we increased the data sample by i) using a larger data window around the thresholds (and including linear trends, separately estimated above and below the thresholds, in the running variable); ii) adding students with different tracks listed as the first and second preference (and including fixed effects for each of these track preferences); and iii) pooling all preference combination-samples into one joint estimation[47]. Results from these alternative estimations are shown in sections C6–C8 in Appendix C.

Overall, they give some support for the positive independent school effect on the final grade percentile rank that was found in the CEM/VAM analysis: The coefficient is of the expected sign and statistically significantly different from zero at the five percent level in 10 out of 23 specifications. The sizes of the statistically significant coefficients are reasonably in line with the CEM/VAM-results (which were a bit over 4, see Table 5.A): Most of the estimated coefficients are around 2–4, which translates into IV effect sizes of roughly 4–8 GPA percentile rank units (given that the "1st stage" estimate is around 0.5). Two of the coefficients are however very large, at around 6–8, which would suggest effects of more than 10 GPA percentile rank units. For the outcome graduate on time, out of the six coefficients that are statistically significant at the five percent level, five are of the same sign as the CEM/VAM-counterparts, and one goes in the opposite direction. For the outcomes measuring the likelihood of switching between independent and public schools, and remaining in upper secondary school after the expected graduation term, the results are in the vast majority of cases statistically insignificant from zero.

The coefficients for the outcomes based on the standardized tests in Math, Swedish and English, shown in in Table 7.B, are likewise in general very imprecisely measured. There is some indication of a positive independent school effect on the math high test grade, and this is statistically significant at the five percent level in one out of four coefficients in Table 7.B, and in 6 out of the 23 alternative specifications in sections C6–C8 of Appendix C. The pattern goes in the same direction for Swedish and English, but overall, precision is a problem, and the coefficients sometimes vary greatly between specifications. The results for the two outcomes on the discrepancy between the test and course grade are also mainly insignificant and unstable across specifications. Some of the reported estimates for the outcome for

[46] In the research plan snapshot for this paper, see https://osf.io/u8r43, we stated that we would also present IV-estimates for these specifications. This is however omitted from the draft, for the sake of brevity.

[47] In this case, we multiplied the running variable with (-1) for the subsamples where the top preference was a public school, so that being above the threshold and in a treated subsample (i.e. with mixed preferences) always predicted attending an independent school. We also added dummy variables for the four subsamples, and an interaction variable for being above the admission threshold (according to the transformed running variable) and having a public school as top preference, in order to account for the transformation of the running variable for these samples.

getting a higher grade on the standardized test than on the corresponding course, suggest that independent schools are *more strict* when setting course grades (in relation to the grades a student got on the standardized test) – i.e. the contrary to what the CEM/VAM-analysis found. However, these estimates were only statistically significant in 18 out of a total of 81 cases, and the remaining coefficients often varied a lot across specifications.

Imprecision is a problem also for the post-graduation outcomes in Table 7.C, and Appendix C sections C6–C8. The coefficients for post-graduation studies are in a few cases positive and statistically significant, but are mostly statistically insignificant from zero. The remaining outcomes are statistically insignificant in the vast majority of cases, and often have large standard errors.

We conclude that there are cases when RD/DID-estimates are in line with the CEM/VAM-results, in particular for the percentile rank GPA, but there are also example of statistically significant cases that go against the CEM/VAM-results, mainly for the outcome for getting a higher test grade than the corresponding course grade. It can be underlined, however, that the overall impression is that the results are sensitive to changes in the specification and the vast majority of coefficients have too wide confidence intervals to be informative. We must therefore not read too much into the few coefficients that come out as statistically significant. Overall, we conclude that the RD/DID-results did not provide much additional insight.

TABLE 7.A. GRADUATION AND GRADES

The coefficients in the table represent the β-coefficient in Equation (3).

Treated sample (T=1):	Independent/Public		Public/Independent	
Control sample (T=0):	Indep/Indep	Public/Public	Indep/Indep	Public/Public
	(1)	(2)	(3)	(4)
Attend Private	0.4911***	0.5015***	-0.4272***	-0.4269***
Standard error	(0.0804)	(0.0719)	(0.0667)	(0.0650)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	378	733	334	689
Number of groups	85	171	103	156
Switch independent/public	0.0240	0.0152	-0.0041	-0.0350
Standard error	(0.0693)	(0.0481)	(0.0740)	(0.0467)
P-value	[0.7296]	[0.7520]	[0.9563]	[0.4539]
Observations	364	718	322	676
Number of groups	85	171	102	155
Pctile GPA12	1.2783	-1.4700	-4.2241	-8.2873
Standard error	(4.5876)	(3.7170)	(5.2878)	(4.8837)
P-value	[0.7812]	[0.6930]	[0.4263]	[0.0918]
Observations	328	632	290	594
Number of groups	84	169	101	153
Graduate on time	0.0557	0.0117	-0.0424	-0.1073
Standard error	(0.0765)	(0.0695)	(0.0839)	(0.0693)
P-value	[0.4686]	[0.8662]	[0.6142]	[0.1237]
Observations	379	737	335	693
Number of groups	85	171	103	156
7th term	-0.0499	-0.0125	0.0103	0.0672
Standard error	(0.0635)	(0.0555)	(0.0599)	(0.0508)
P-value	[0.4348]	[0.8214]	[0.8635]	[0.1880]
Observations	379	737	335	693
Number of groups	85	171	103	156

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and standard errors are clustered on the same level. The regressions additionally include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D) The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. No trends were included. The regression sample is restricted to students with the same track preference for the top and second preference, and to observations within 5 units from the admission threshold

TABLE 7.B. STANDARDIZED TEST RESULTS *The coefficients in the table represent the β-coefficient in Equation (3).*

	Math				English				Swedish			
Treated sample (T=1):	Independent/Public		Public/Independent		Independent/Public		Public/Independent		Independent/Public		Public/Independent	
Control sample (T=0):	Indep/Indep	Public/Public	Indep/Indep	Public/Public	Indep/Indep	Public/Public	Indep/Indep	Public/Public	Indep/Indep	Public/Public	Indep/Indep	Public/Public
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Attend Private	0.5892***	0.5733***	-0.5176***	-0.4683***	0.5145***	0.5330***	-0.4522***	-0.4265***	0.4989***	0.4645***	-0.4507***	-0.4764***
Standard error	(0.0868)	(0.0811)	(0.0810)	(0.0771)	(0.0847)	(0.0772)	(0.0820)	(0.0793)	(0.0995)	(0.0936)	(0.0844)	(0.0768)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	302	584	258	540	325	592	275	542	305	572	259	526
Groups	76	151	95	139	79	152	96	138	80	164	99	152
High test grade	0.1306*	0.0599	0.0258	-0.0048	0.0499	0.0675	-0.0578	-0.0405	-0.0887	-0.0563	-0.1279*	-0.1065*
Standard error	(0.0564)	(0.0402)	(0.0274)	(0.0212)	(0.0909)	(0.0616)	(0.0877)	(0.0428)	(0.0634)	(0.0591)	(0.0545)	(0.0463)
P-value	[0.0232]	[0.1381]	[0.3498]	[0.8219]	[0.5849]	[0.2744]	[0.5113]	[0.3464]	[0.1659]	[0.3425]	[0.0209]	[0.0228]
Observations	302	586	258	542	326	594	276	544	305	574	259	528
Groups	76	151	95	139	79	153	96	139	80	165	99	152
Pass test grade	0.0979	0.0686	-0.0785	-0.0778	-0.0027	-0.0001	-0.0112	-0.0021	0.0230	0.0144	-0.0732	-0.0752
Standard error	(0.0909)	(0.0567)	(0.1028)	(0.0743)	(0.0232)	(0.0194)	(0.0378)	(0.0336)	(0.0424)	(0.0263)	(0.0607)	(0.0484)
P-value	[0.2850]	[0.2285]	[0.4471]	[0.2970]	[0.9087]	[0.9972]	[0.7674]	[0.9509]	[0.5879]	[0.5863]	[0.2309]	[0.1223]
Observations	302	586	258	542	326	594	276	544	305	574	259	528
Groups	76	151	95	139	79	153	96	139	80	165	99	152
Test grade>Course grade	0.0198	0.0215	-0.0246	-0.0412	-0.0420	-0.0159	-0.0400	-0.0211	0.0735	-0.0268	-0.0265	-0.2151***
Standard error	(0.0231)	(0.0315)	(0.0331)	(0.0264)	(0.0566)	(0.0517)	(0.0780)	(0.0589)	(0.0894)	(0.0617)	(0.0952)	(0.0705)
P-value	[0.3941]	[0.4954]	[0.4601]	[0.1205]	[0.4601]	[0.7594]	[0.6098]	[0.7206]	[0.4132]	[0.6650]	[0.7812]	[0.0027]
Observations	299	551	246	498	319	566	270	517	293	518	241	466
Groups	74	149	93	138	79	151	96	139	80	159	95	144
Test grade<Course grade	0.0128	0.0780	-0.0453	-0.0351	0.0143	-0.0295	0.1095	0.0646	0.0428	0.0142	0.1172	0.1394
Standard error	(0.0744)	(0.0577)	(0.0885)	(0.0807)	(0.0813)	(0.0621)	(0.0914)	(0.0624)	(0.1115)	(0.0816)	(0.1322)	(0.1032)
P-value	[0.8642]	[0.1787]	[0.6099]	[0.6638]	[0.8607]	[0.6354]	[0.2342]	[0.3029]	[0.7023]	[0.8621]	[0.3775]	[0.1791]
Observations	299	551	246	498	319	566	270	517	293	518	241	466
Groups	74	149	93	138	79	151	96	139	80	159	95	144

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and standard errors are clustered on the same level. The regressions additionally include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D) The regressions include the same student level covariates as Table 7.A. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. No trends were included. The regression sample is restricted to students with the same track preference for the top and second preference, and to observations within 5 units from the admission threshold.

TABLE 7.C. POST-GRADUATION OUTCOMES

The coefficients in the table represent the β-coefficient in Equation (3).

Treated sample (T=1):	Independent/Public		Public/Independent	
Control sample (T=0):	Indep/Indep	Public/Public	Indep/Indep	Public/Public
	(1)	(2)	(3)	(4)
Attend Private	0.4911***	0.5015***	-0.4272***	-0.4269***
Standard error	(0.0804)	(0.0719)	(0.0667)	(0.0650)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	378	733	334	689
Number of groups	85	171	103	156
Study	0.1741	0.0231	0.1219	-0.0005
Standard error	(0.1350)	(0.0900)	(0.1430)	(0.1152)
P-value	[0.2031]	[0.7977]	[0.3976]	[0.9966]
Observations	210	417	169	376
Groups	50	103	53	89
Study no-prep	0.0573	-0.0434	-0.0508	-0.1353
Standard error	(0.1182)	(0.0900)	(0.1301)	(0.0988)
P-value	[0.6299]	[0.6304]	[0.6976]	[0.1743]
Observations	210	417	169	376
Groups	50	103	53	89
Uni cred≥15	0.0607	-0.0182	0.0015	-0.0380
Standard error	(0.0909)	(0.0829)	(0.1044)	(0.0839)
P-value	[0.5073]	[0.8264]	[0.9883]	[0.6513]
Observations	210	417	169	376
Groups	50	103	53	89
Work ≥50%	0.1415	0.0104	0.2733	0.1050
Standard error	(0.1297)	(0.0748)	(0.1467)	(0.1116)
P-value	[0.2807]	[0.8899]	[0.0682]	[0.3494]
Observations	210	417	168	375
Groups	50	103	53	89

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and standard errors are clustered on the same level. The regressions additionally include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D) The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. No trends were included. The regression sample is restricted to students with the same track preference for the top and second preference, and to observations within 5 units from the admission threshold.

4.4 Heterogeneity analysis

We conclude that the RD/DID-estimates are overall too imprecise to draw sharp conclusions. In the heterogeneity analysis we therefore exclusively rely on the CEM/VAM approach. We estimate separate regressions for the following sub-samples:

- Academic track or vocational track
- Born in Sweden or not born in Sweden
- One parent with post-secondary education or no parent with post-secondary education
- Bottom, middle and top tercile of their cohorts' distribution of GPS9

We use our most preferred specification, i.e. where we restrict the sample to students who listed both an independent and a public school among the top two preferences, and where we match on the county where the student's upper secondary school is placed. The results are given in Tables 8.A–C. For the sake of comparison, the results for the full student sample (corresponding to column 2 in Tables 5.A–C) are reproduced in column 1 in all tables.

4.1.1 Academic vs Vocational track students

The results in columns (2) and (3) in Table 8.A suggest that academic track students gain more from independent school attendance than vocational students; the positive effect on GPA12 is larger, as is also the independent school effect on getting a high test grade in English and Swedish. However, the effect on the probability of being "upgraded" is larger for vocational track students in English, fairly similar for both groups in Math, and larger for academic track students in Swedish, see Table 8.B1–3.

Since an academic track prepares for further studies, a simple way of evaluating the success of an academic track from a post-graduation perspective is by focusing on the probability of entering higher education. Indeed, the effect on the probability on entering higher studies after graduation remains positive when studying academic tracks separately, see column (2) in Table 8.C.

For vocational students, however, judging success after graduation is somewhat harder. Even though a vocational track prepares students to enter the labor market immediately after graduation, this is clearly not the only measure of success. Another path is to continue to Higher Vocational Education, or even switch to a more academic career. Interestingly, the effect on labor market entry is fairly large and negative for vocational students, while the effect on the probability of entering higher studies, also non-preparatory studies, is positive and larger for vocational students than for academic students, although the effect on taking university credits is the same in both groups. All in all, this suggests that independent school attendance induces vocational track students to pursue further studies instead of entering the labor market. Additional regressions (available upon request) suggest that independent school attendance primarily impacts the likelihood of starting *non-university* post-secondary studies for vocational students, while the impact on starting university studies is the same for both groups.

4.4.2 Country of birth

When comparing students depending on their country of birth it is worth noting that the sample of non-Swedish born students is relatively small, see column (5) in Table 8.A. We should also remind the reader that the regression sample of students is a selected sample, since we have excluded students in preparatory tracks.

That said, we find that the positive impact of independent school attendance on the likelihood of graduating on time is more than three times larger for non-Swedish born students (8.8 percentage points) than for Swedish born students (2.6 percentage points). The impact on the GPA in 12th grade is however quite similar in magnitude. Regarding test score outcomes, the independent school effect on having a higher final course grade than test grade in math stands out as very large among non-Swedish born students, at 10 percentage points.

None of the estimates on post-graduation outcomes are statistically significant at the 5 percent level for students born outside of Sweden, although some coefficients are economically interesting. For instance, the effect on non-preparatory studies is large and positive at 4.4 percentage points, and has a relatively low p-value of 0.06.

4.4.3 Parental education level

Independent school effects relating to graduation and grades are relatively similar among students with and without a parent with a post-secondary education, see columns (6) and (7) in Table 8.A. Interestingly, students without a parent with post-secondary education, like non-Sweden born students, show larger effects on the probability of having a higher final course grade relative to the test grade, especially in Mathematics and English. The estimates on post-secondary outcomes are fairly similar in both groups, although the negative independent school effect on the probability of entering the labor market is larger among students with one highly-education parent.

4.4.4. Students with different levels of prior academic attainment

Finally, we run separate estimations for students with a GPS9 in the lower (T1), mid (T2), and upper (T3) tercile of the distribution of the full (yearly) population of students.[48]The positive coefficient on GPA in 12th grade is largest in the mid tercile group, see Column (9) in Table 8.A. The positive coefficient for graduating on time is not present in the upper part of the ability distribution, which is no surprise since most of these students are probably far above the margin when it comes to the likelihood to graduate on time. The largest positive effect is instead found in the lower achievement distribution, where the boost in grades from attending independent school may be just what it needed to reach the graduation threshold.

When looking at the standardized test results, the main difference between the GPS9-terciles is that the propensity to set a higher course grade than the standardized test grade is substantially higher for the two lower terciles than for the top tercile.[49] This pattern raises the question of whether the larger impacts on the same outcome that were reported for students with low parental education and non-Swedish born students, respectively, reflected that these groups tend to have lower grades than the student population in general. In our selected sample of students (where all students have high enough grades to avoid having to take a preparatory track), students with low educated parents do have substantially lower average GPS9 (211) than those with high educated parents (240). However, grade differences are unlikely to drive the differences with respect to country of birth: our regression sample of non-Swedish born students have almost as high average GPS9 (222) as the Swedish-born students (227).

Tables 8.B1–8.B3 also report a positive impact of independent school attendance on the likelihood of getting a high grade in English and Swedish for the top GPS9-tercile, and to some extent in Swedish for the mid tercile.

[48] The number of students in the lower tercile is lower than in the higher intervals, and this reflects that our sample is limited to students in the Academic and Vocational tracks, and excludes the students with the lowest grades who end up in the preparatory tracks.

[49] One might speculate that this reflects a stronger aversion among independent schools to set fail grades. This is supported by additional regressions (available upon request) for English and Math: the positive independent school impact on the likelihood of getting a higher course grade than test grade is larger for students who got a fail grade on the test than for students with higher test grades. For Swedish, however, the pattern is the reverse.

TABLE 8.A HETEROGENEITY – GRADUATION AND GRADES – CEM/VAM

	All	Academic	Vocational	Born Sweden	Not born Sweden	Parent post-secondary	No parent post-secondary	GPS9T1	GPS9T2	GPS9T3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Switch ind./public	0.0231***	0.0161***	0.0353***	0.0241***	-0.0138	0.0184***	0.0266***	0.0294***	0.0225***	0.0190***
Standard error	(0.0049)	(0.0047)	(0.0085)	(0.0049)	(0.0149)	(0.0044)	(0.0065)	(0.0099)	(0.0059)	(0.0044)
P-value	[0.0000]	[0.0007]	[0.0000]	[0.0000]	[0.3539]	[0.0000]	[0.0000]	[0.0030]	[0.0002]	[0.0000]
Observations	70623	47626	22661	66782	3447	38782	31356	16737	24922	28839
Pctile GPA12	4.4906***	5.4623***	2.9422***	4.4562***	5.2200***	4.5065***	4.4377***	3.2349***	5.5211***	4.1971***
Standard error	(0.3080)	(0.3699)	(0.4779)	(0.3083)	(1.1920)	(0.3298)	(0.3888)	(0.4637)	(0.4358)	(0.3912)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	61898	42061	19543	58680	2893	34387	27111	13008	21957	26822
Graduate on time	0.0288***	0.0358***	0.0173*	0.0261***	0.0877***	0.0259***	0.0335***	0.0529***	0.0389***	0.0072
Standard error	(0.0041)	(0.0048)	(0.0073)	(0.0041)	(0.0200)	(0.0048)	(0.0058)	(0.0093)	(0.0061)	(0.0046)
P-value	[0.0000]	[0.0000]	[0.0184]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.1152]
Observations	72220	48551	23325	68256	3560	39604	32112	17503	25315	29273
7^{th} term	-0.0153***	-0.0206***	-0.0049	-0.0135***	-0.0529***	-0.0136***	-0.0191***	-0.0189**	-0.0232***	-0.0077*
Standard error	(0.0028)	(0.0034)	(0.0050)	(0.0028)	(0.0154)	(0.0034)	(0.0041)	(0.0071)	(0.0045)	(0.0034)
P-value	[0.0000]	[0.0000]	[0.3256]	[0.0000]	[0.0006]	[0.0001]	[0.0000]	[0.0077]	[0.0000]	[0.0237]
Observations	72220	48551	23325	68256	3560	39604	32112	17503	25315	29273

Note: All regressions above include the following covariates: upper secondary school municipality dummies, prior achievement as controlled for by a cubic form of GPS9 and GPS9 quintile dummies, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9^{th} grade, 9^{th} grade school dummies, track dummies, log household income, income decile dummies; and dummies indicating the following: gender, born in western country (excl. Sweden), born in non-western country, one parent post-secondary education, both parents born in Sweden, one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. Columns 1 and 2 also include a dummy indicating admission to first ranked school. Standard errors are clustered on upper secondary school. *** p<0.005, ** p<0.01, * p<0.05

TABLE 8.B1. HETEROGENEITY – TEST RESULTS – MATH – CEM/VAM

	All	Academic	Vocational	Born Sweden	Not born Sweden	Parent post-secondary	No parent post-secondary	GPS9T1	GPS9T2	GPS9T3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
High test grade	0.0034	0.0035	0.0004	0.0030	-0.0004	0.0054	0.0003	0.0015**	-0.0006	0.0070
Standard error	(0.0020)	(0.0026)	(0.0018)	(0.0021)	(0.0083)	(0.0030)	(0.0017)	(0.0006)	(0.0011)	(0.0041)
P-value	[0.0878]	[0.1777]	[0.8419]	[0.1466]	[0.9592]	[0.0750]	[0.8761]	[0.0081]	[0.5642]	[0.0872]
Observations	48106	35192	12693	45380	2452	27702	20059	9435	16617	21956
Pass test grade	-0.0028	-0.0056	-0.0028	-0.0032	-0.0013	-0.0058	-0.0015	0.0114	-0.0065	-0.0049
Standard error	(0.0054)	(0.0060)	(0.0109)	(0.0055)	(0.0212)	(0.0058)	(0.0077)	(0.0123)	(0.0088)	(0.0056)
P-value	[0.6045]	[0.3519]	[0.7969]	[0.5599]	[0.9520]	[0.3156]	[0.8416]	[0.3555]	[0.4575]	[0.3828]
Observations	48106	35192	12693	45380	2452	27702	20059	9435	16617	21956
Test grade>Course grade	0.0039*	0.0055***	-0.0015	0.0040*	-0.0051	0.0070***	0.0002	-0.0018	0.0012	0.0069***
Standard error	(0.0017)	(0.0017)	(0.0040)	(0.0017)	(0.0071)	(0.0020)	(0.0024)	(0.0044)	(0.0025)	(0.0021)
P-value	[0.0228]	[0.0016]	[0.7124]	[0.0199]	[0.4675]	[0.0004]	[0.9264]	[0.6714]	[0.6459]	[0.0010]
Observations	46244	33960	12069	43633	2349	26698	19211	8940	15993	21217
Test grade<Course grade	0.0462***	0.0480***	0.0552***	0.0450***	0.1039***	0.0370***	0.0603***	0.0758***	0.0529***	0.0295**
Standard error	(0.0088)	(0.0098)	(0.0154)	(0.0088)	(0.0255)	(0.0097)	(0.0105)	(0.0140)	(0.0116)	(0.0112)
P-value	[0.0000]	[0.0000]	[0.0004]	[0.0000]	[0.0000]	[0.0001]	[0.0000]	[0.0000]	[0.0000]	[0.0084]
Observations	46244	33960	12069	43633	2349	26698	19211	8940	15993	21217

Note: Regressions are performed on individual means within each subject. All regressions above include the following covariates: upper secondary school municipality dummies, prior achievement as controlled for by a cubic form of GPS9 and GPS9 quintile dummies, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade, 9th grade school dummies, track dummies, log household income, income decile dummies; and dummies indicating the following: gender, born in western country (excl. Sweden), born in non-western country, one parent post-secondary education, both parents born in Sweden, one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. Regressions on test outcomes also include test specific dummies. Columns 1–2, 4–5, and 7–8 also include a dummy variable indicating whether the student was admitted to the first ranked choice. Standard errors are clustered on upper secondary school. *** p<0.005, ** p<0.01, * p<0.05

TABLE 8.B2. HETEROGENEITY – TEST RESULTS – SWEDISH – CEM/VAM

	All	Academic	Vocational	Born Sweden	Not born Sweden	Parent post-secondary	No parent post-secondary	GPS9T1	GPS9T2	GPS9T3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
High test grade	0.0204***	0.0254***	0.0074***	0.0202***	0.0200	0.0232***	0.0146***	0.0014	0.0117***	0.0324***
Standard error	(0.0046)	(0.0063)	(0.0025)	(0.0047)	(0.0161)	(0.0066)	(0.0033)	(0.0014)	(0.0024)	(0.0096)
P-value	[0.0000]	[0.0001]	[0.0034]	[0.0000]	[0.2144]	[0.0005]	[0.0000]	[0.3344]	[0.0000]	[0.0008]
Observations	52515	37025	15254	50148	2087	29960	22222	10620	18457	23349
Pass test grade	0.0065*	0.0064*	0.0085	0.0068*	-0.0131	0.0081***	0.0043	0.0117	0.0101*	0.0023
Standard error	(0.0027)	(0.0028)	(0.0067)	(0.0027)	(0.0186)	(0.0027)	(0.0044)	(0.0091)	(0.0044)	(0.0016)
P-value	[0.0151]	[0.0232]	[0.2047]	[0.0105]	[0.4816]	[0.0029]	[0.3289]	[0.2002]	[0.0203]	[0.1563]
Observations	52515	37025	15254	50148	2087	29960	22222	10620	18457	23349
Test grade>Course grade	-0.0035	-0.0056	0.0009	-0.0031	-0.0009	-0.0021	-0.0045	-0.0099	-0.0096	-0.0005
Standard error	(0.0040)	(0.0048)	(0.0073)	(0.0040)	(0.0248)	(0.0048)	(0.0053)	(0.0087)	(0.0064)	(0.0049)
P-value	[0.3735]	[0.2429]	[0.9022]	[0.4318]	[0.9717]	[0.6683]	[0.3988]	[0.2546]	[0.1348]	[0.9159]
Observations	48724	34476	14026	46531	1926	27799	20604	9791	17186	21662
Test grade<Course grade	0.0239***	0.0288***	0.0101	0.0237***	0.0382	0.0181*	0.0340***	0.0316**	0.0392***	0.0117
Standard error	(0.0069)	(0.0077)	(0.0115)	(0.0069)	(0.0309)	(0.0082)	(0.0084)	(0.0113)	(0.0096)	(0.0096)
P-value	[0.0005]	[0.0002]	[0.3799]	[0.0006]	[0.2164]	[0.0277]	[0.0001]	[0.0051]	[0.0000]	[0.2237]
Observations	48724	34476	14026	46531	1926	27799	20604	9791	17186	21662

Note: Regressions are performed on individual means within each subject. All regressions above include the following covariates: upper secondary school municipality dummies, prior achievement as controlled for by a cubic form of GPS9 and GPS9 quintile dummies, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade, 9th grade school dummies, track dummies, log household income, income decile dummies; and dummies indicating the following: gender, born in western country (excl. Sweden), born in non-western country, one parent post-secondary education, both parents born in Sweden, one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. Regressions on test outcomes also include test specific dummies. Columns 1–2, 4–5, and 7–8 also include a dummy variable indicating whether the student was admitted to the first ranked choice. Standard errors are clustered on upper secondary school. *** p<0.005, ** p<0.01, * p<0.05

TABLE 8.B3. HETEROGENEITY – TEST RESULTS – ENGLISH – CEM/VAM

	All	Academic	Vocational	Born Sweden	Not born Sweden	Parent post-secondary	No parent post-secondary	GPS9T1	GPS9T2	GPS9T3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
High test grade	0.0155***	0.0186***	0.0090***	0.0162***	-0.0037	0.0194***	0.0104***	0.0067**	0.0060	0.0298***
Standard error	(0.0040)	(0.0055)	(0.0025)	(0.0041)	(0.0127)	(0.0060)	(0.0030)	(0.0025)	(0.0034)	(0.0080)
P-value	[0.0001]	[0.0007]	[0.0003]	[0.0001]	[0.7700]	[0.0014]	[0.0006]	[0.0072]	[0.0747]	[0.0002]
Observations	50202	35627	14335	47375	2525	28314	21514	10576	17888	21638
Pass test grade	0.0023	0.0009	0.0056	0.0013	0.0229*	0.0011	0.0047	0.0038	0.0027	-0.0003
Standard error	(0.0016)	(0.0013)	(0.0047)	(0.0016)	(0.0114)	(0.0015)	(0.0028)	(0.0061)	(0.0025)	(0.0007)
P-value	[0.1528]	[0.4886]	[0.2283]	[0.4431]	[0.0452]	[0.4814]	[0.0981]	[0.5327]	[0.2769]	[0.6602]
Observations	50202	35627	14335	47375	2525	28314	21514	10576	17888	21638
Test grade>Course grade	-0.0036	-0.0045	-0.0025	-0.0036	0.0046	0.0003	-0.0096	-0.0175	-0.0097	0.0050
Standard error	(0.0044)	(0.0049)	(0.0085)	(0.0045)	(0.0173)	(0.0049)	(0.0059)	(0.0105)	(0.0069)	(0.0045)
P-value	[0.4134]	[0.3602]	[0.7704]	[0.4162]	[0.7876]	[0.9428]	[0.1054]	[0.0958]	[0.1601]	[0.2732]
Observations	48159	34356	13572	45451	2417	27242	20553	10034	17118	20913
Test grade<Course grade	0.0401***	0.0343***	0.0626***	0.0395***	0.0678**	0.0280***	0.0601***	0.0474***	0.0531***	0.0214*
Standard error	(0.0058)	(0.0069)	(0.0090)	(0.0059)	(0.0243)	(0.0068)	(0.0068)	(0.0084)	(0.0077)	(0.0083)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0053]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0101]
Observations	48159	34356	13572	45451	2417	27242	20553	10034	17118	20913

Note: Regressions are performed on individual means within each subject. All regressions above include the following covariates: upper secondary school municipality dummies, prior achievement as controlled for by a cubic form of GPS9 and GPS9 quintile dummies, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade, 9th grade school dummies, track dummies, log household income, income decile dummies; and dummies indicating the following: gender, born in western country (excl. Sweden), born in non-western country, one parent post-secondary education, both parents born in Sweden, one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. Regressions on test outcomes also include test specific dummies. Columns 1–2, 4–5, and 7–8 also include a dummy variable indicating whether the student was admitted to the first ranked choice. Standard errors are clustered on upper secondary school. *** p<0.005, ** p<0.01, * p<0.05

TABLE 8.C. HETEROGENIETY – POST-GRADUATION – CEM/VAM

	All	Academic	Vocational	Born Sweden	Not born Sweden	Parent post-secondary	No parent post-secondary	GPS9T1	GPS9T2	GPS9T3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Study	0.0199***	0.0168**	0.0291***	0.0215***	-0.0186	0.0217***	0.0152*	0.0147	0.0215**	0.0203**
Standard error	(0.0050)	(0.0063)	(0.0077)	(0.0051)	(0.0278)	(0.0071)	(0.0062)	(0.0079)	(0.0079)	(0.0078)
P-value	[0.0001]	[0.0081]	[0.0002]	[0.0000]	[0.5036]	[0.0022]	[0.0146]	[0.0616]	[0.0062]	[0.0091]
Observations	55430	36924	18261	52567	2588	30241	24825	13802	18935	22628
Study no-prep	0.0244***	0.0231***	0.0329***	0.0248***	0.0443	0.0258***	0.0208***	0.0220***	0.0266***	0.0229***
Standard error	(0.0046)	(0.0061)	(0.0063)	(0.0047)	(0.0238)	(0.0067)	(0.0052)	(0.0056)	(0.0075)	(0.0080)
P-value	[0.0000]	[0.0002]	[0.0000]	[0.0000]	[0.0632]	[0.0001]	[0.0001]	[0.0001]	[0.0004]	[0.0045]
Observations	55430	36924	18261	52567	2588	30241	24825	13802	18935	22628
Uni cred>=15	0.0142***	0.0160***	0.0156***	0.0149***	0.0305	0.0124*	0.0154***	0.0114***	0.0182***	0.0126
Standard error	(0.0034)	(0.0047)	(0.0036)	(0.0034)	(0.0212)	(0.0051)	(0.0040)	(0.0028)	(0.0048)	(0.0067)
P-value	[0.0000]	[0.0008]	[0.0000]	[0.0000]	[0.1517]	[0.0144]	[0.0001]	[0.0000]	[0.0001]	[0.0621]
Observations	55430	36924	18261	52567	2588	30241	24825	13802	18935	22628
Work >=50%	-0.0170***	-0.0045	-0.0423***	-0.0179***	0.0132	-0.0109	-0.0222***	-0.0237***	-0.0149	-0.0158*
Standard error	(0.0049)	(0.0054)	(0.0092)	(0.0050)	(0.0221)	(0.0058)	(0.0068)	(0.0084)	(0.0087)	(0.0064)
P-value	[0.0005]	[0.4065]	[0.0000]	[0.0004]	[0.5492]	[0.0596]	[0.0011]	[0.0047]	[0.0857]	[0.0139]
Observations	55386	36896	18245	52538	2573	30215	24809	13785	18921	22615

Note: All outcomes are measured one year after graduation. All regressions above include the following covariates: upper secondary school municipality dummies, prior achievement as controlled for by a cubic form of GPS9 and GPS9 quintile dummies, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade, 9th grade school dummies, track dummies, log household income, income decile dummies; and dummies indicating the following: gender, born in western country (excl. Sweden), born in non-western country, one parent post-secondary education, both parents born in Sweden, one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. Columns 1 and 2 also include a dummy variable indicating whether the student was admitted to the first ranked choice. Standard errors are clustered on upper secondary school. *** p<0.005, ** p<0.01, * p<0.05

5. Robustness analysis

5.1 Multiple hypothesis correction

In this section we make multiple hypothesis corrections to the p-values obtained from our most preferred empirical specification in column 2, Tables 5A–5C. We follow the procedure proposed in Hochberg (1988); see also a practical implementation in Banerjee et. al. (2015). We first rank all of the original p-values obtained from each regression from lowest to highest value, $P_1, \dots, P_m$, where m is the number of outcomes (or regressions). Next, we multiply the p-values obtained from these regressions with $(m + 1 - k)$, where k is the rank of the original p-value.

Following this procedure, and correcting for 20 regressions (the number of outcomes), we can report no substantial changes when it comes to passing rejection levels. All estimates that are statistically significant at the 0.5 percent level using original p-values are also statistically significant at the 0.5 percent level using corrected p-values, see the original and the FWER p-values in Table 9.

TABLE 9. P-VALUES CORRECTED FOR MULTIPLE HYPOTHESIS TESTING

Outcome	Original P-value	FWER P-value	Outcome	Original P-value	FWER P-value
Pctile GPA12	0.00000***	0.00000***	High test grade EN	0.00012***	0.00119***
Graduate on time	0.00000***	0.00000***	Work>=50%	0.00054***	0.00490***
Test grade<Course grade EN	0.00000***	0.00000***	Test grade<Course grade SV	0.00055***	0.00437***
7th term	0.00000***	0.00000***	Pass test grade SV	0.01506*	0.10542
Test grade<Course grade MA	0.00000***	0.00000***	Test grade>Course grade MA	0.02282*	0.13692
Study no-prep	0.00000***	0.00000***	High test grade MA	0.08778	0.43889
Switch independent/public	0.00000***	0.00003***	Pass test grade EN	0.15280	0.61119
High test grade SV	0.00001***	0.00013***	Test grade>Course grade SV	0.37351	1.12053
UC>=15	0.00003***	0.00039***	Test grade>Course grade EN	0.41345	0.82689
Study	0.00008***	0.00084***	Pass test grade MA	0.60450	0.60450

Note: *** p<0.005, ** p<0.01, * p<0.05.

5.2 Bounds on the coefficient estimates a la Oster (2019)

In this section, we analyze the robustness of our baseline CEM/VAM-estimates to unobservable variables bias following Oster (2019), who in turn builds on (among others) Altonji et. al. (2005). The idea is that information on how the covariates that are included in a regression model affect the coefficient of primary interest (β) and the R^2, can be used to evaluate the likely unobservable variable bias. Without going into details (see Oster, 2019, for these), we use the Stata-command psacalc that is provided by Oster to compute following two statistics:

i. δ(β =0); the amount of selection on unobservables that would be needed, expressed as a share of the selection on observed covariates, in order for β to go to zero.
ii. β(δ=1); the β-value we would get if the unobserved selection were of equal sign and size as the observed selection. It follows that if selection on unobservables, expressed as a share of the observed selection, is within the interval 0 and 1, then the β-coefficient of the fully specified model is bounded by our regression model's β and β(δ=1). This interval is suggested by Oster (2017) as a reasonable robustness analysis. The tables also present β(δ=0.5) and β(δ=2), i.e. β under the assumption that the unobserved selection is of the same sign and half or, and double that of, the observed selection.

In order to calculate the above statistics, we also need to make an assumption on R^2 of the fully specified model, including unobservables. We follow the recommendation by Oster and set this to 1.3 times R^2 from our regression model, or, alternatively, to one if the former should exceed one (which it never does in our case).[50]

Table 10.A shows that our main CEM/VAM-results for the graduation and post-graduation outcomes are supported by this exercise. For all of these outcomes, the degree of selection on unobservables, expressed as a share of the observed selection, that would be needed in order for the β-parameter to go to zero, $\delta(\beta = 0)$, is much larger than one. Consequently, the interval between our model estimate of β and $\beta(\delta=1)$, beta under the assumption of equal size and sign of the unobserved and observed selection, never includes zero, and this interval is additionally small in magnitude. The same holds for all of the test based outcomes in Table 10.B, except for the likelihood of receiving a Pass grade in Math, in column (2) of panel A. The independent school coefficient was however in any case not statistically significantly different from zero in the main analysis. We conclude that the robustness analysis based on Oster (2019) does *not* indicate that the main results of this paper are driven by unobserved selection.

TABLE 10.A, ESTIMATES UNDER DIFFERENT ASSUMPTIONS ON UNOBSERVABLE SELECTION, A LA OSTER (2019), GRADUATION AND POST-GRADUATION OUTCOMES

	(1)	(2)	(3)	(4)
Panel A: Graduation and grades				
	Switch ind/public	Pctile GPA12	Graduate on time	7th term
β independent school	0.0231***	4.4906***	0.0288***	-0.0153***
Standard error	(0.0049)	(0.3080)	(0.0041)	(0.0028)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	70623	61898	72220	72220
δ(β =0)	7.0100	43.3772	-8.8956	-17.0222
β(δ=0.5)	0.0223	4.5934	0.0313	-0.0163
β(δ=1)	0.0214	4.7035	0.0340	-0.0173
β(δ=2)	0.0195	4.9492	0.0401	-0.0196
Panel B: Post-graduation outcomes				
	Study	Study no-prep	UC≥15	Work ≥50%
β independent school	0.0199***	0.0244***	0.0142***	-0.0170***
Standard error	(0.0050)	(0.0046)	(0.0034)	(0.0049)
P-value	[0.0001]	[0.0000]	[0.0000]	[0.0005]
Observations	55430	55430	55430	55386
δ(β =0)	-284.1516	-26.8082	-9.9268	6.9556
β(δ=0.5)	0.0206	0.0257	0.0154	-0.0163
β(δ=1)	0.0213	0.0270	0.0167	-0.0156
β(δ=2)	0.0229	0.0300	0.0195	-0.0139

[50] Oster (2019) points out that in many instances the R^2 of a fully specified model, including unobservables, is less than one due to measurement error in the dependent variable. Setting R^2_{max} to one may in such cases give overly conservative bounds for the β-estimates. The recommendation in Oster (2019) to use 1.3 times the estimated model's R^2, is based on an analysis using randomized data from a set of published articles.

TABLE 10.B, ESTIMATES UNDER DIFFERENT ASSUMPTIONS ON UNOBSERVABLE SELECTION, A LA OSTER (2019), STANDARDIZED TEST RESULTS

	(1)	(2)	(3)	(4)
Panel A: Math				
	High test grade	Pass test grade	Test grade>Course grade	Test grade<Course grade
β independent school	0.0034	-0.0028	0.0039*	0.0462***
Standard error	(0.0020)	(0.0054)	(0.0017)	(0.0088)
P-value	[0.0878]	[0.6045]	[0.0228]	[0.0000]
Observations	48106	48106	46244	46244
δ(β =0)	3.2267	0.7193	8.1447	-23.8878
β(δ=0.5)	0.0030	-0.0009	0.0038	0.0493
β(δ=1)	0.0025	0.0012	0.0037	0.0527
β(δ=2)	0.0015	0.0060	0.0036	0.0606
Panel B: English				
	High test grade	Pass test grade	Test grade>Course grade	Test grade<Course grade
β independent school	0.0155***	0.0023	-0.0036	0.0401***
Standard error	(0.0040)	(0.0016)	(0.0044)	(0.0058)
P-value	[0.0001]	[0.1528]	[0.4134]	[0.0000]
Observations	50202	50202	48159	48159
δ(β =0)	3.9491	-3.8052	-6.8868	22.3351
β(δ=0.5)	0.0142	0.0027	-0.0040	0.0411
β(δ=1)	0.0126	0.0032	-0.0045	0.0422
β(δ=2)	0.0092	0.0042	-0.0055	0.0447
Panel C: Swedish				
	High test grade	Pass test grade	Test grade>Course grade	Test grade<Course grade
β independent school	0.0204***	0.0065*	-0.0035	0.0239***
Standard error	(0.0046)	(0.0027)	(0.0040)	(0.0069)
P-value	[0.0000]	[0.0151]	[0.3735]	[0.0005]
Observations	52515	52515	48724	48724
δ(β =0)	9.8884	-21.2779	-3.4968	10.5163
β(δ=0.5)	0.0204	0.0070	-0.0043	0.0240
β(δ=1)	0.0203	0.0075	-0.0050	0.0240
β(δ=2)	0.0200	0.0087	-0.0069	0.0242

6. Concluding discussion

Since Sweden introduced a school voucher system in 1992, the share of students who attend a publicly funded, but independently provided, upper secondary school has increased from a few percent to a quarter of all students. There are no formal limits on the amount of profits independent providers are allowed to make, but the Swedish school inspectorate has the authority to monitor independent providers and force them to closure in case of severe deficiencies.

Evaluation of the educational value added of independent schools in Sweden is a complicated matter. While Hinnerich and Vlachos (2017) document positive added value from independent schools using teacher-assessed test grades, the externally re-graded test results show the reverse; negative added value of independent schools. These results suggest that grading standards are more generous in independent schools, and that teacher-assessed achievement measures cannot be relied upon to reflect actual educational achievements.

Whereas we do not have access to re-graded tests, we estimate the added value of independent schools for a larger sample of the student population, and on a broader set of outcome variables. Similarly to Hinnerich and Vlachos (2017), we find that independent schools have positive added value with respect to teacher-assessed achievements, such as final GPA and teacher-graded results on tests, but we also find evidence that support the notion of more generous grading among independent schools. First, our results suggest that independent students' final course grades are more often "up-graded" relative to their test grade, but there is no corresponding difference in the likelihood of being "down-graded". Second, the independent school effect on test grades is positive in Swedish and English, but not in Mathematics. Since internal and external grading discrepancy has been found to be relatively small in Mathematics, Vlachos (2018) argues that the math test can be considered one of the more reliable tests. Under a strict interpretation, the fact that independent schools do not display positive added value on a test where teacher degrees-of-freedom (when grading) is small, could mean that the overall added value of independent schools is close to zero. However, there is no way for us to completely rule out that the positive effects for languages, or indeed, the positive effects on the final GPA, contain actual educational added value to some degree.

Besides Hinnerich and Vlachos (2017), other studies have also documented generous grading standards among independent schools in Sweden. Wikström and Wikström (2005) show that students from upper secondary school have a relatively high final GPA in comparison to their results on the Swedish SATs. Vlachos (2018) finds that independent schools on the compulsory level set higher final course grades, as compared with their students' results on test grades, similarly to our findings for upper secondary school students. He also finds that the discrepancy is larger for the arguably more reliable Math test than for the other tests. (This can be compared to our finding that the course grade discrepancy in Math is larger than in Swedish – but, on the other hand, similar in magnitude to English). Finally, the Swedish National Agency for Education report evidence of more generous grading among independent schools in compulsory school (Skolverket, 2019a), and that students from schools where grades appeared to be more generously set, tend to perform worse in upper secondary school compared to students with a similar grade from a school with stricter grading standards (Skolverket, 2019b).

Our results for post-graduation outcomes suggest that attending an independent school has a positive impact on the likelihood of registering for further studies. Furthermore, students attending an independent school have a higher likelihood of not only registering for university studies, but also of actually taking university credits. In light of the evidence in both our and previous studies, it is hard to know to what extent this reflects inflated grades, and to what extent it reflects actual academic preparedness. For instance, Diamond and Persson (2016) study the effect of grade inflation in Sweden and they find positive spill-over effects on students' later educational outcomes. Such impacts may work through motivational or signaling effects.

Overall, the independent school effects that we document in this study consistently shows that attending an independent school mostly benefits the individual, through higher grades and graduation rates, and a higher propensity for further studies. Whether or not this also amounts to societal benefits is more difficult to establish. When schools operate on a school market where students bring resources via vouchers, incentives to attract students by showing high achievement gains will be present, and some of our results seem to suggest, in accordance with earlier studies mentioned above, that these incentives could adversely affect the grading standards and educational measurement in Swedish upper secondary school. Antidotes, such as external grading of standardized tests, are therefore highly motived. The ongoing work by the Swedish National Agency for Education to introduce a combination of automatized and centralized grading on standardized tests, are steps in the right direction.

All of the abovementioned results are supported by our conditional-on-observables estimates using a combination of CEM and VAM. However, the RD/DID-approach was overall too imprecise, and too sensitive to specification changes, to provide much guidance. As with any non-experimental study, one could argue that our CEM/VAM approach does not provide causal estimates. Although we cannot be certain that remaining unobservable selection does not affect the results, based on our strategies of controlling for student preferences, as well as our sensitivity analysis a la Oster (2019), we find it unlikely that unobserved selection would drive the results in a substantial way.

Finally, this study, as well as the previous referenced Swedish literature, has evaluated the impact of the independent school sector without distinguishing between schools with different pedagogical approaches. For our part, this is due to practical reasons: we lack good indicators for how the schools differ, and we are furthermore, for integrity reasons, not allowed to identify specific schools in our detailed register data. US evidence however suggests that heterogeneity within the charter school sector is large and highly relevant: Schools with a "no-excuses"-approach in particular seem to perform well, whereas so called "cyber"-schools have in evaluations stood out as poor performers (Epple et. al., 2015). A relevant area for future research for the Swedish voucher school sector is thus to investigate its potential heterogeneity: What are the approaches taken by different organizations, and is there evidence that some approaches work better than others? Although such an analysis requires collection of data that allows such mapping, it has the potential of providing useful insights on best practices. It can be pointed out however, that the indicated prevalence of different grading standards poses a challenge also for such studies.

References

Abdulkadiroglu, A., J.D. Angrist, J. D., S. M. Dynarski, T. J. Kane, P. A Pathak. 2011. "Accountability and Flexibility and Public Schools: Evidence from Boston's Charters and Pilots." *The Quarterly Journal of Economics* 126: 699–748.

Abdulkadiroglu, A., Pathak, P. A., C. R. Walters. 2018. "Free to Choose: Can School Choice Reduce Student Achievement?" *American Economic Journal: Applied Economic* 10 (1): 175–206.

Ahlin, Å. 2003. "Does school competition matter? Effects of a large-scale school choice reform on student performance", Department of Economics, Uppsala University, Working Paper 2003:2.

Altonji, J. G., Elder, T. E., and C. R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113: 151–184.

Angelov, N. and K. Edmark. 2016. "När skolan själv får välja – en ESO-rapport om friskolornas etableringsmönster", ESO-rapport 2016:3.

Angrist, J. D., S. R. Cohodes, S. M. Dynarski, P. A. Pathak, and C. R. Walters. 2016. "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice." *Journal of Labor Economics* 34 (2).

Angrist, J. D., P. A. Pathak, and C. R. Walters**.** 2013. "Explaining Charter School Effectiveness." *American Economic Journal: Applied Economics* 5 (4): 1–27.

Angrist, J. D., P. D. Hull, P. A. Pathak, and C. R. Walters**.** 2017. "Leveraging lotteries for school value-added: testing and estimation." *The Quarterly Journal of Economics* 5 (4): 871–919

Banerjee, A. V., E. Duflo, R. Glennerster, and C. Kinnan. 2015. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." American Economic Journal: Applied Economics 7 (1): 22–53.

Björklund, A., Clark, M. A., Edin, P.-A., Fredriksson, P., and A. Krueger. 2005. "The education market comes to Sweden. An evaluation of Sweden's surprising school reforms". (Ed. Björklund, A), Russel Sage Foundation, New York.

Böhlmark. A., and M. Lindahl. 2015. "Independent Schools and Long-run Educational Outcomes: Evidence from Sweden's Large-scale Voucher Reform." *Economica*. 82 (327). 508-551.

Cattaneo, M. D., Idrobo, N. and R. Titiunik. 2018. "A Practical Introduction to Regression

Discontinuity Designs: Volume II" Monograph prepared for *Cambridge Elements: Quantitative and Computational Methods for Social Science,* Cambridge University Press. **Preliminary draft** available at: https://cattaneo.princeton.edu/books/Cattaneo-Idrobo-Titiunik_2018_CUP-Vol2.pdf

Chabrier, J., S. Cohodes, P. Oreopoulos. 2016 "What Can We Learn from Charter School Lotteries?" *Journal of Economic Perspectives* 30 (3): 57–84.

Deming D. J. 2014 "Using School Choice Lotteries to Test Measures of School Effectiveness". *American Economic Review: Papers and Proceedings* 104(5): 406–411.

Diamond, R. and P. Persson. 2016. "The Long-term Consequences of Teacher Discretion in Grading of High-stakes Tests". NBER Working Paper No. 22207.

Dobbie, W., and R. G. Fryer. 2019. "Charter Schools and Labor Market Outcomes." Mimeo.

Dobbie, W., and R. G. Fryer. 2013. "Getting Beneath the Veil of Effective Schools: Evidence from New York City". *American Economic Journal: Applied Economics* 5 (4): 28–60.

Edmark, K, Hussein, I. and C. Haelermans. 2020. "The Impact of Voucher Systems: Evidence from Swedish Upper Secondary Schools", forthcoming SOFI Working Paper, Stockholm University.

Epple, D., Romano, R. and R. Zimmer. 2015."Charter Schools: A Survey of Research on Their Characteristics and Effectiveness". NBER Working Papers 21256, National Bureau of Economic Research, Inc.

Epple, D., R. E. Romano, and M. Urquiola. 2017. "School Vouchers: A survey of the Economics Literature". *Journal of Economic Literature* 55 (2): 441–492.

Forslund, A., Liljeberg, L. and O. Åslund. 2017. ”Flykting- och anhöriginvandrades etablering på den svenska arbetsmarknaden", FAU-rapport 2017:4.

Friedman, M. 1955. "The Role of Government in Education." In Robert A. Solo (Ed.), *Economics and the Public Interest.* Rutgers University Press.

Le Grand, J. 1991. "Quasi-Markets and Social Policy". *The Economic Journal* 101: 1256–1267.

Hahn, Y., L.C. Wang and H. Yang. 2018. "Do greater school autonomy and accountability make a difference? Evidence from the random assignment of students into private and public high schools in Seoul". *Journal of Public Economics* 118: 15-30.

Hennerdal, P., Malmberg, B., and E.K. Andersson. (2018). Competition and school performance: Swedish school leavers from 1991–2012. *Scandinavian Journal of Educational Research* 64(1): 70–86.

Hochberg, Y. 1988. "A sharper Bonferroni procedure for multiple tests of significance". *Biometrika* 75 (4): 800–802.

Hoxby, C. 2003. "School choice and school productivity: Could school choice be the tide that lifts all boats?" In Hoxby, C. (Ed.), *The Economics of School Choice*. University of Chicago Press.

Hsieh, C. and M. Urquiola. 2006. "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program". *Journal of Public Economics* 90: 1477–1503.

Imbens. G. W. and D. B. Rubin. 2015. *Causal Inference for Statistics. Social and Biomedical Sciences.* Cambridge University Press.

Imbens, G. W. and S. Wager. 2019. "Optimized Regression Discontinuity Designs". *Review of Economics and Statistics* 101(2): 264-278.

Kolesár, M., and C. Rothe. 2018. "Inference in Regression Discontinuity Designs with a Discrete Running Variable". *American Economic Review* 108(8):2277–2304.

Kortelainen, M. and K. Manninen. 2019. "Effectiveness of Private and Public High Schools: Evidence from Finland". *CESifo Economic Studies* 424–445. Doi: 10.1093/cesifo/ifz014

Lee, D.S. and D. Card. 2008. "Regression discontinuity inference with specification error". *Journal of Econometrics* 142: 655–674.

OECD. 2020. Secondary graduation rate (indicator). doi: 10.1787/b858e05b-en (Accessed on 29 January 2020).

Oster, E. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence". *Journal of Business & Economic Statistics* 37(2): 187–204. DOI: 10.1080/07350015.2016.1227711

Sandström, M., and F. Bergström. 2005. "School vouchers in practise: competition will not hurt you". *Journal of Public Economics* 89:351–380.

Sebhatu, A. and K. Wennberg. 2017. "Institutional Pressure and Failure Dynamics in the Swedish Voucher School Sector". *Scandinavian Journal of Public Administration* 21 (3): 109–136.

Shleifer, A. 1998. "State versus private ownership". *Journal of Economic Perspectives* 12 (4): 133–150.

Skolverket. 2004. “Skarpare rutiner vid tillståndsprövning och inspektion av fristående skolor. Avrapportering av ett regeringsuppdrag.” Dnr 01-2003:1995

Skolverket. 2019a. “Analyser av likvärdig betygssättning mellan elevgrupper och skolor. Jämförelser mellan betyg och nationella prov i årskurs 9.” Rapport 475.

Skolverket. 2019b. “Grundskolors systematiska avvikelser i betygsresultat och elevers gymnasieresultat.” Promemoria, Dnr: 2018:01481.

Statistics Sweden (SCB). 2017a. “3 av 10 tar inte examen från gymnasieskolan.” Artikel publicerad 2017-10-04. https://www.scb.se/hitta-statistik/artiklar/2017/3-av-10-tar-inte-examen-fran-gymnasieskolan/

Statistics Sweden (SCB). 2017b. ”Arbetsmarknaden för personer med låg utbildning 2005–2016”. Statistiska meddelanden AM 110 SM 1704.

Hinnerich Tyrefors, B. and J. Vlachos. 2017. “The impact of upper-secondary voucher school attendance on student achievement. Swedish evidence using external and internal evaluations” *Labour Economics* 47(C): 1-14.

Vlachos, J. 2018. “Trust-Based Evaluation in a Market-Oriented School System” IFN Working Paper No. 1217.

Wikström, C. and M. Wikström. 2005. “Grade inflation and school competition: an empirical analysis based on the Swedish upper secondary schools”. *Economics of Education Review* 24 (3): 309–322.

Appendix A: Institutional overview of the Swedish upper secondary education market

This appendix provides a more detailed description of the Swedish institutional setting than is provided in the main article. The information provided here sometimes overlaps the shorter exposition of the article.

A1. Public and private provision

In Sweden, upper secondary school can be provided either by the local governments, the municipalities, or by private entities, so called *independent schools*. There is also a small number of schools run by mid-level regional jurisdictions; predominantly in nursing/care or agriculture. Municipal and independent schools are both fully funded via school vouchers provided by the municipality, which are primarily financed by the local income tax. The voucher level is to be determined using the same criterion that determines the funding to the municipality's own schools for the track in question.[51] Additional tuition fees are not allowed.

Table A1 shows some descriptive statistics on the geographic distribution and characteristics of the public and independent schools, as of school year 2013/14, which is the year in which the last cohort in our study entered upper secondary school. The table shows separate statistics for independent and public schools in Rural, Urban and Metropolitan municipalities, respectively. [52] [53] As can be seen in the first row of the table, the independent school share is substantially larger (40 percent) in the metropolitan municipalities than in the urban (21 percent) and rural (8 percent) municipalities. Nationwide, 26 percent of upper secondary students attended an independent school in year 2013. A bit more than a third, or 458, of all school units[54] are independent entities, and the independent schools are generally smaller, with an average number of 184 students compared to 273 in the public schools. Schools in rural municipalities are on average smaller.

---

[51] In cases where a municipality does not offer the tracks provided by the private school, such that exists is no municipality criterion, the voucher is instead to follow a national guideline ("Riksprislistan"). https://www.skolinspektionen.se/sv/Tillstandsprovning/Starta-fristaende-skola/Bidrag-till-fristaende-skolor/

[52] The classification of municipalities is based on municipal urbanization rate and is constructed by The Swedish Agency for Growth Policy Analysis (*Tillväxtanalys*). In short, municipalities in the three large-city areas (Stockholm, Gothenburg and Malmö) are classified as Metropolitan; detached municipalities with a predominantly urban population are classified as Urban; and the remaining municipalities with large rural populations are classified as Rural. A more detailed description is given in section A.3.1 in the appendix.

[53] Note that the geographic categorization is based on the municipality of location for the schools.

[54] It can be noted that the school unit concept used in the national registers changed in 2012-2013, and this resulted in an increase in the number of school units, in particular for the public schools. The number of independent schools was therefore approximately equal to the number of public schools in 2011, before the change of measurement (see also table note b for a comment on this).

TABLE A1. GEOGRAPHICAL DISTRIBUTION AND SCHOOL CHARACTERISTICS

	Rural	**Urban**	**Metropolitan**	**Total**
Market share independent school (student shares)	0.077	0.209	0.404	0.260
Number of school units[a]				
Independent	32	233	193	458
Public	172	531	179	882
School size (average number of students.)				
Independent	82	164	226	184
Public	184	273	360	273
Academic tracks (student shares)				
Independent	0.498	0.534	0.707	0.622
Public	0.433	0.540	0.676	0.562
Vocational tracks (student shares)				
Independent	0.457	0.428	0.257	0.340
Public	0.413	0.357	0.189	0.320
Preparatory tracks (student shares)				
Independent	0.045	0.038	0.037	0.037
Public	0.154	0.103	0.135	0.118
Number of students per teacher, adjusted for share of Voc/Ac/Prep tracks[b]				
Independent	9.602	11.675	13.407	11.831
Public	9.138	11.211	12.943	11.368

Note: Data refers to school year 2013/14. The full sample is used; i.e. before substantial sample restrictions are made. When making a classification of rural, urban or metropolitan, the three-type classification scheme made by The Swedish Agency for Growth Policy Analysis (Tillväxtanalys) is applied on a school municipality level.
[a] It can be noted that the definition of school units in the national School register changed in 2013. The new code is based on the division of headmaster responsibilities, rather than the physical school units. This means that an entity which prior to the change counted as one school in the register, may with the new classification count as several school units, each with a separate code. According to information received by e-mail from Statistics Sweden, this started to affect the number of units in the School register already in 2012/2013, as some schools started to use the new definition when submitting information for the School Register already then. The change, which in particular has affected the number of municipal schools, is clearly visible in the data: the number of municipal upper secondary schools was 502 in the fall of 2011, 766 in 2012, as, as seen in the table, 882 in 2013. The number of independent schools was rather decreasing during the same time period (likely du to schools closing rather than being an effect of the changing school unit definition) 499 in 2011, 484 in 2012, and 458 in 2013.
[b] The data is adjusted to account for varying student teacher ratios over Academic, Vocational and Preparatory track types. The raw student teacher ratio shows a similar pattern of consistently higher numbers with independent schools. The 0.5 percent top and bottom observations were excluded in order to eliminate the influence of extreme outliers.

Regarding the types of tracks offered, Table A1 shows that most students attend an academic track: this is about twice as common as being in a vocational track in both the independent and public schools. The predominance of the academic tracks is particularly high in the metropolitan areas, whereas the vocational track share is almost on par with the academic ditto in the rural municipalities. The preparatory tracks, which offer shorter catch up courses for students whose grades from compulsory school do not qualify them to enter any academic nor vocational track, are much more common in the public schools; the share of public school students who are in a preparatory track is 11.8 percent in the public schools and 3.7 percent in the independent schools.[55] A possible explanation

[55] It can be noted that the preparatory tracks are generally 1 year of length, after which the students are expected to continue to a regular track. The share of students attending a preparatory track out of all students in grade 1 in upper secondary school in 2013 was 17.4 percent, i.e. a higher number than in the table, which is based on students in all grades in upper secondary school. As most, but not all, students attend a school in the home municipality, the numbers differ slightly compared to measures based on where students reside.

for this might be that independent schools were not allowed to offer the preparatory track prior to 2006.
The table finally shows that independent schools have a slightly higher number of students per teacher (11.8) than the public schools (11.4), also after adjusting for the shares of students attending Academic, Vocational and Preparatory tracks.[56] (The unadjusted averages are 12.7 for the independent and 10.9 for the public schools). The student/teacher ratio is overall lower in more rural areas.

A2. How independent are the independent schools?

The regulation of the independent school sector was initially quite rudimentary, but has over time "caught up", and today, much of the regulation for the public schools also applies to the independent schools. This development includes the authorization of new schools, which today requires more information about the prospective providers; and the monitoring, which have become more frequent and have expanded in scope.[57] Since 2008, the Swedish Schools Inspectorate, a government agency, is responsible for the authorization of the independent schools and for the oversight of both the public and the independent schools.

As of today, the independent providers are furthermore obliged to follow the same curriculum; meet the same educational goals; and use the same grading system as public schools.[58] Headmasters and teachers in both types of schools are required to have the proper educational degree for the position – something which prior to 2002 held only for public schools – although exceptions can be made if the position can otherwise not be filled. In 2018, independent schools had a lower share of certified teachers (73 percent) than the municipal schools (85 percent).[59]

At the same time, school providers (or principals) – in both the public and independent sectors – have significant decision power within the regulatory framework when it comes to: specific hiring decisions, wage setting, and the allocation of resources within the school. The national curriculum regulates the minimum amount of total instruction time[60], but providers are free to decide how to allocate the instruction time between courses[61] and over the school year, with some restrictions; the school year shall start in August and end in July, shall comprise 40 weeks, and instruction shall be scheduled to the weekdays.[62] Instruction is by definition teacher-led. So called "distance learning" – taking instructions from a teacher via digital channels, or exclusively via educational software – is only allowed in language courses, and then only if the teacher position cannot be filled. On the other hand, schools can of course also choose to offer more teacher-led instruction time than is required by law.

[56] The adjustment was done by predicting the measure based on linear regression while inserting the same overall average track type share for all schools.
[57] Chapter 3 in Angelov and Edmark (2016) describes the authorization of the independent schools in the early days of the reform, as well as the later developments. See also the National Agency for Education (Skolverket, 2004, pp. 21–22) for information on the monitoring.
[58] An exception is made for the Waldorf schools who are allowed to use a different grading system.
[59] Skolverkets databas: Gymnasieskolan - Personalstatistik med behörighet - per ämne och kategori 2018/19.
[60] Chapter 16 §18, The Education Act (Skollag 2010:800).
[61] It is regulated what courses shall be provided in each track, and the amount of credits that is connected to each course, but it is up to the provider to decide on the instruction time for each course, see the Upper Secondary School Ordinance (Gymnasieförordning 2010:2039), Chapter 4 §22, and the Education Act, Chapter 16 §17–20, or see the web page of the National Agency for Education: https://www.skolverket.se/regler-och-ansvar/ansvar-i-skolfragor/scheman-och-larotider#h-Skoldagenslangdiolikaskolformer.
[62] The Upper Secondary School Ordinance (Gymnasieförordning 2010:2039), §1–2 Chapter 3.

The curriculum regulates what mandatory courses must be offered within in each nationally regulated track. However, schools decide what optional courses to offer, and can even choose to design their own courses from scratch.[63] Optional courses are counted against the required amount of course credits. Schools can also build profiles by offering special instruction in sports, arts, or in certain subjects such Math, languages, etc.[64]

Whereas the above regulation is similar for both types of providers, only independent schools are allowed to have a religious profile, albeit only for the non-instructional part of the school day.[65] Moreover, as of September 2019, independent schools are not bound by the Public Access to Information and Secrecy Act; i.e. their records are not public, as opposed to records in municipal schools. A proposal to eliminate this distinction is currently under governmental review.

Even though the educational regulation is overall very similar for the two types of providers, the organizational form may naturally itself have consequences for the running and management of schools. Bloom et. al. (2015) suggest that schools that are publicly funded but have more autonomy vis-à-vis the government, for example in terms of being run by private/non-government entities, have higher management scores.[66] These findings are based on surveys with school providers in several countries, including the Swedish independent school sector.

With regards to the organizational form, independent schools can be in the form of either non-profit or for-profit organizations, and in fact a large majority of the upper secondary independent schools are organized as corporations. In 2013 – the year the last cohort of our data entered upper secondary school – 85 percent of these schools belonged to corporations. The remaining 15 percent were primarily organized as foundations or non-profit associations.[67] The independent schools are furthermore often part of larger corporate groups: in 2013, more than a third of all independent upper secondary school students attended a school belonging to one of the 10 largest independent school providers, who altogether ran 153 upper secondary schools.[68] The public schools are in contrast provided by the most local tier of public government, the municipalities. It can be underlined that these do also differ largely in size; from a couple of thousand to several hundred thousand inhabitants.

To summarize, most areas of freedom by law applies to both municipal and independent providers. There are few formal provisions regarding independent providers, but existing ones include: the possibility of organizing as for-profit (including giving out dividends to owners) and adding a religious profile. Beyond these stated differences, both independent and municipal schools can profile themselves according to their offering of: nationally regulated tracks, optional courses, and voluntary special instruction in sports, arts, or in other academic subjects.

[63] New courses have to be approved by The National Agency for Education, according to the Upper Secondary School Ordinance (Gymnasieförordning 2010:2039) Chapter 1 §6.

[64] In order to do so, they need permission from the National Agency for education (if a municipal school) or the Swedish Schools Inspectorate (if an independent school).

[65] More specifically, religious activities may be added to the school day, provided that they take place outside of the instruction time and as long as participation is voluntary. In contrast, public schools shall be fully non-confessional. See Chapter 1 §6-7 of the Education Act (Skollag 2010:800).

[66] This means getting a higher score in a survey that is designed to capture management quality.

[67] See the Swedish School Register.

[68] See table 4.3 in Skolverket (2014).

A3. School and educational track choices

Entering upper secondary education is associated with making two choices: a choice of school and a choice of educational track. Under the current system, students choose simultaneously the school and track as one package, and are allowed to rank varying combinations of tracks and school in their application. The application process starts early on in the spring term, when students submit their applications to the local admission agencies. Admission offers are sent out in the summer, around July 1st, after which students have a few weeks to respond.[69,70] After that, in August/September, students may be accepted from waiting lists.

The track choice amounts to choosing among six academic and 12 vocational tracks. In addition, there are six small vocational tracks in specialized fields such as railway; airplane and shipping technicians; traditional Sami industries[71]; and professional dancing. Students whose grades do not qualify them to any of these regular tracks, are referred to a set of preparatory tracks. The aim of the preparatory tracks, which are normally a year or shorter and have individually tailored curriculums, is to qualify the students for the regular tracks. In practice however, the vast majority of the students who enter the preparatory tracks do not proceed to finishing a regular track within reasonable time: among the 2011 cohort of 9th graders, only a quarter of the students who went to preparatory tracks had completed a regular track five years later (recall that the regular tracks are normally 3 years long). The corresponding figure for the students qualifying for a regular track immediately after compulsory school was close to 80 percent.[72]

The admission criteria to upper secondary education are regulated in Chapter 7 of the Upper Secondary School Regulation (Gymnasieförordningen). The regulation states that if the number of applicants exceeds the number of available slots, admission shall be based on the grade sum, which is calculated as the sum of the grade credits for the students' 16 highest graded subjects from lower secondary graduation.[73] The regulation however leaves room for two deviations from this purely grade based admission procedure. First, ability tests are allowed to select students to the artistic track, and may also be used for specialized tracks with permission to have a special profile in for example arts or sports.[74,75] Second, a small number of slots are to be left open for applicants who cannot be judged solely on their grades due to "special circumstances" or due to being from a different grading system (such as Waldorf or foreign schools). If these slots are not filled, they are to be added to the regular grade based admission process. In case of ties – i.e. several students with the same grade sum as the admission threshold – it is up to the school provider to choose from a list of allowed criteria, such as selection based on specific subject grades; the rank of the choice; or chance.[76]

---

[69] The Upper Secondary School Regulation (Gymnasieförordningen 2010:2039, Chapter 7) states that the final admission decision shall, "if possible" be made prior to July 1st.
[70] The exact duration of the response period is determined by the local agencies, but the Swedish Association of Local Authorities and Regions ("Handböcker för gymnasieantagning 2009-13") recommends that students are given 3 weeks to respond.
[71] Sami craft, reindeer raising, nature guiding.
[72] Statistics Sweden (2017).
[73] From 2011 on, students could add credit for an additional class if they took an elective modern language class.
[74] School providers need to obtain special permission from the National Agency for Education (for the public schools) or by the Swedish School Inspectorate (for the independent schools) to use ability testing for selection to special profile tracks.
[75] See chapter 5 in the Upper Secondary School Regulation (Gymnasieförordningen 2010:2039).
[76] See the Handbooks of the Swedish Association of Local Authorities and Regions: "Handböcker för gymnasieantagning 2009-13".

Students can choose from all voucher schools in the country, and from the municipal schools in their home “admission region”.[77] [78] An admission region is one or a group of municipalities within which the resident students have equal access to all publicly operated upper secondary schools. They are formed to give students access to educational tracks that are not provided by the home municipality’s school, and/or to increase students’ school choice options. In some parts of the country, especially the more densely populated areas, there has been a trend of forming larger admission regions, thus expanding the school choice options of the students further. Such large regions were for example formed in the Gothenburg area in 2002[79]; in Stockholm county in 2008[80]; and in Southern Sweden in starting from December 2010[81]. Students may also apply to municipal schools in other admission regions, but home region students are given priority in the admission process.[82] Independent schools, on the other hand, can give no home advantage for residents, but are to evaluate applications from students from all parts of the country equally.[83]

Even though today’s admission to upper secondary school, within an admission region, is based purely on the final grades from lower secondary school, this has not always been the case for the municipal schools. Prior to the 2000s, the normal procedure was to base the educational *track* admission on the grades, but to then assign students to the public schools based on some other criterion, for example proximity to the student’s home (Molin, 2019). In the year 2000, two of the largest municipalities, Stockholm and Malmö, allowed students to be admitted to all school and track combinations based on the grades. Other regions have since followed, and according to Sund (2018), today all municipalities apply such purely grade-based admission systems. Whether or not proximity was used to determine school admission during our period of analysis is of relevance for our RD-based analysis. For many of the mid-sized and smaller municipalities, we have however not been able to collect information on the exact year in which the proximity principle for school placement was replaced with a pure grade principle. It can be pointed out that in many of the smaller and mid-sized municipalities, there was only one public school offering each track anyway, which means that assignment to schools within a given track was a non-issue. In Section B6 of Appendix B we provide results suggesting that this is not a concern for our empirical analysis.

The independent schools may manage their own admissions, or may have the admission procedure managed by the admission region. In the former case, the application shall be sent in directly to the school. This means that students who apply to both an independent school with its own admission process, and to a school handled by an admission agency, will have submitted multiple applications, and may receive separate admission offers from the different schools/agencies. The same holds for students who apply to schools belonging to different admission agencies. (As can be seen in section 3.1 of the main article, this has consequences for how we define the regression sample).

[77] Students may also apply to municipal schools outside of their admission regions, but are then not given priority in the admission process.

[78] What constitutes an admission region differs across the country, but normally a set of adjacent municipalities form an admission region. The Stockholm admission region, which roughly comprises the 20-some municipalities in Stockholm county, is one of the larger of the admission regions. See www.antagningskanslier.se for a list of the current admission regions.

[79] Email conversation with the Gothenburg upper secondary school admission agency.

[80] The 2008 regional admission excluded the Social Science and Science tracks, which were added from 2011 on (Power point presentation on “Dnr: KSL/13/0097” from the Greater Stockholm area (Kommunförbundet Stockholms län), and Sund (2018).

[81] Based on email conversation with the Malmö upper secondary school admission agency.

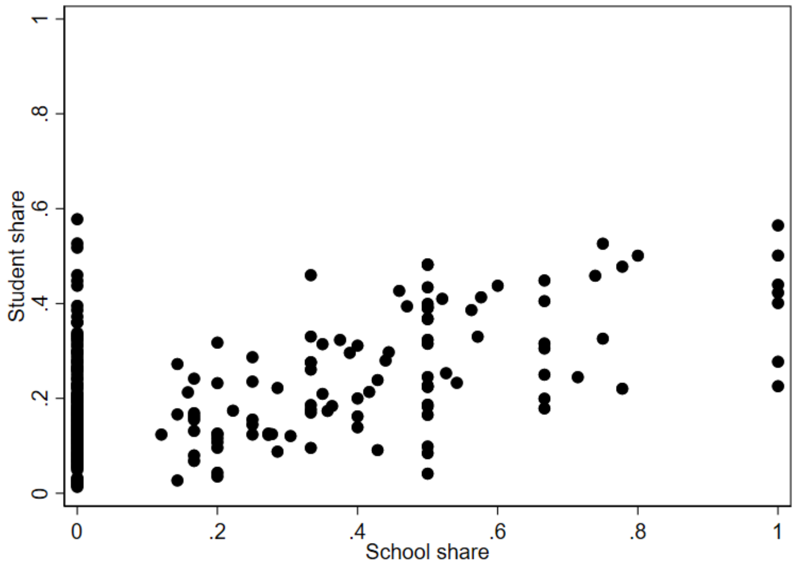
[82] See the Education Act (Skollag 2010:800), Chapter 16 §43-44.

[83] Chapter 15, §33, of the Education Act (Skollag 2010:800).

About 30 municipalities (out of a total of 290) have no upper secondary school, in which case students are given access to schools in adjacent municipalities in their admission region or to independent schools over the whole country. About 120 municipalities have a municipal upper secondary school but no independent upper secondary school. About 16 percent of students attend an upper secondary school in another municipality than their resident municipality, and this predominantly reflects students attending independent schools.

In Figure A1 we show the relationship between the independent school market share measured as the share of students residing within a municipality that attend an independent school (y-axis), and measured as the share of schools within a municipality that are independent (x-axis). The figure indicates a positive correlation, although it is far from a 1:1-relation. This is likely to in part reflect that independent schools are on average smaller than the municipal counterparts (see Table A1), and in part that it is relatively common to attend a school outside of the home municipality. The latter is underlined by the fact that the mean share of resident students that attend an upper secondary school is 0.173 in municipalities where there are no independent schools (or no schools at all), compared to 0.239 in municipalities where there is at least one independent school.

FIGURE A1. MUNICIPALITY LEVEL INDEPENDENT SCHOOL AND STUDENT SHARES



Note: The student market share is measured on the *resident municipality* level, while the school market share is measured on the *school municipality* level. Municipalities without any upper secondary schools are included as 0 independent school student share.

References to Appendix A

Angelov, N. and K. Edmark. 2016. “När skolan själv får välja – en ESO-rapport om friskolornas etableringsmönster”, ESO-rapport 2016:3.

Molin, E. 2019. “Through the Ranks – Essays on Inequality, Status and School Choice”, PhD Dissertation, Stockholm School of Economics.

Skolverket. 2004. “Skarpare rutiner vid tillståndsprövning och inspektion av fristående skolor. Avrapportering av ett regeringsuppdrag.” Dnr 01-2003:1995

Skolverket. 2014. ”Privata aktörer inom förskola och skola.” Rapport 410.

Statistics Sweden (SCB). 2017. “3 av 10 tar inte examen från gymnasieskolan.” Artikel publicerad 2017-10-04. https://www.scb.se/hitta-statistik/artiklar/2017/3-av-10-tar-inte-examen-fran-gymnasieskolan/

Sund, K. 2018. ”Långsiktiga effekter av utökade valmöjligheter till Gymnasieskolan – från närhetsprincip till betygsprincip”, Riksrevisionen rapport 2018:28.

Appendix B: Data. Sample restrictions, variables and general description.

Appendix B includes detailed information on the data sample restrictions we make, and on the included variables.

B1. Data Samples

B1.1 Observational sample for CEM/VAM-analysis

Table B1 lists the data restrictions that we apply to the Observational sample used for the CEM/VAM-analysis, and how they affect the sample size.

TABLE B1. GENERATING THE OBSERVATIONAL SAMPLE, FOR THE CEM/VAM-ANALYSIS

Sample restriction, comment	**Nr individuals**	**Sample restriction motivation**
Use Gymn_elev.dta	1180316	The raw data includes all upper secondary school students in 2009-16.
AterPnr.dta, Re-used personal ID	1179549	Some individuals have to be dropped because of issues related to their personal ID that could cause erroneous matching of individuals across different public registers.
FelPnr.dta, Erroneous personal ID	1142567	
Drop students in tracks that cannot be identified based on track code	1133454	Observations with missing information on educational track, and/or with missing information on school ownership (private or public), are dropped.
Drop if information on Indep/Public ownership is missing	1133344	
Keep year 2009–2013	821680	We study only cohorts starting in 2009-13, because for later cohorts no outcomes are available in our data. [84]
Keep only grade 1	575276	We keep in the sample individuals who start upper secondary education straight after finishing lower secondary education (no gap year), and who enter upper secondary school at the common age of 16. Some students are observed as starting grade 1 in upper secondary school several years; for example if they change tracks, or initially take a preparatory year. In such cases we only keep in the sample the first observed instance that a student enters upper secondary education. It can be noted that these sample restrictions also imply that we exclude students who enter Swedish upper secondary education in grade 2 or 3, but who did not attend Swedish upper secondary education in grade 1.
Keep only students age 16	494781	
Keep only students without gap year between lower and upper secondary school	452301	
Drop students in preparatory tracks	418916	Students in the preparatory track are excluded from the sample.

[84] While there is an application register also prior to 2009, it does not contain information on schools applied to.

Keep only students with one admission in application data	391514	Some students are recorded as being admitted to more than one or their ranked alternatives. These students are dropped from the data sample, as we cannot know which admission information is correct.
Drop students with missing info on ranked track 1 or 2	337140	Drop students with missing information on Independent/Public ownership or with missing information on track codes, for the listed preferences 1 and 2. This also excludes students who submitted only one school and track choice.
Drop students with missing info on Indep/Public school for ranked preferences 1 and/or 2	332980	
Drop students who are not accepted to either rank 1 or 2	317256	We keep only students admitted to their 1st or 2nd ranked preference in the sample. Doing so has the benefit of making students in the sample more comparable in the sense that they were all admitted to one of their top two choices.
Drop students with multiple applications	300830	Some students have applied to several admission agencies or to several schools with separate admission forms. In these cases, there is no way to infer how students rank the alternatives on the different applications forms; if the student prefers the school by track combination on one of the submitted forms over the listed school by track on another form or not. Students submitting multiple application forms are thus excluded from the sample.
Drop students who are not eligible for tracks ranked 1 and/or 2	296890	Individuals who are not eligible for either of their top two listed tracks, i.e. do not have sufficient grades in the core subjects, are dropped from the sample. [85]
Main sample: students applying to both independent and public school as rank 1-2		
Nr of student observations for sample applying to both independent and public as rank 1 and 2	**72745**	Include only students that have ranked a combination of public and private schools as their top two preferences, i.e. either an independent school as first ranked choice and a public as second, or the other way around.
Number of observations by preference order:		
1: Independent 2: Public	*34320*	
1: Public 2: Independent	*38425*	
1: Independent 2: Independent	*34911*	
1: Public 2: Public	*189234*	

[85] According to data from 2011 published by Statistics Sweden, 13 percent of students are non-eligible for a regular track, and only a quarter of the non-eligible students eventually complete upper secondary school within five years. Almost 40 percent of the non-eligible students are never accepted to a regular track, and equally many are eventually accepted to a regular track but never complete their studies

B1.2 RD/DID-sample

Table B2 lists the restrictions that are made for the sample used in the RD/DID-analysis, and how they affect the sample size.

TABLE B2. GENERATING THE RD/DID-SAMPLE

Sample restriction, comment	Nr individuals	Sample restriction motivation
Use UpperSecAppl0916.dta	882206	
Keep year 2009–2013	574648	We study only cohorts starting in 2009-13, because for later cohorts no outcomes are available in our data.
AterPnr.dta, Re-used personal ID	574227	Some individuals have to be dropped because of issues related to their personal ID that could cause erroneous matching of individuals across different public registers.
FelPnr.dta, Erroneous personal ID	570097	
Drop if missing info on admission group	541887	Drop students for whom we cannot observe school and track applied to.
Keep only applications with first priority	481272	Only students in application group=0, i.e. to first priority applicants (i.e. those with first priority to the education slots in the school and track) are included.
Keep only 1st and 2nd ranked preferences	467298	Keep only the top two ranked preferences.
Keep first observed application	457420	For students who are in application register several years, keep first observed instance.
Keep only students age 16	406154	Keep only students who turn 16 the year of application.
Keep only students without gap year between lower and upper secondary school	405391	Keep only students without gap year between lower and upper secondary school.
Drop preparatory track	400632	Drop applications to preparatory tracks.
Drop students in tracks that cannot be identified based on track code	398079	Drop observations for tracks that cannot be identified based on track code.
Drop if missing Public/Indep	396287	Drop observations with missing information on Indep/Public provision
Keep only students with one admission in application data	363767	Some students are recorded as being admitted to more than one or their ranked alternatives. These students are dropped from the data sample, as we cannot know which admission information is correct.
Drop students with multiple applications	341159	Some students have applied to several admission agencies or to several schools with separate admission forms. In these cases, there is no way to infer how students rank the alternatives on the different applications forms; if the student prefers the school by track combination on one of the submitted forms over the listed school by track on another form or not. Students submitting multiple application forms are thus excluded from the sample.
Drop students who are not eligible for tracks ranked 1 and/or 2	326629	Individuals who are not eligible for either of their top two listed tracks, i.e. do not have sufficient grades in the core subjects, are dropped from the sample.
Keep students who have ranked at least preference 1 and 2	261878	Some individuals have only listed one preference. They are dropped as they are not useful for the RDD
Keep if admission threshold 1 > admission threshold 2 *and* the student grade sum ≥ threshold 2	251264	Only students who have, in their applications to upper secondary school, ranked their listed preferences for track and school combinations in the following manner are included: the admission threshold of the first ranked preference must be higher than that of the second ranked preference. And the admission

		thresholds mush be such that the student is, based on her grade sum, admitted to either of the two alternatives. The RD-sample thus consists of two groups of students: one with a grade sum above the admission threshold to the first two ranked preferences, and one with a grade sum below the first – but above the second – ranked preference.
Keep if the same track is listed as preference 1 and 2	48643	Keep only students with the same educational track at both sides of the binding admission threshold. This is imposed in order to ensure that the only aspect changing at the threshold is whether or not the student is admitted to an independent or a public school, and not the track admitted to.
Drop Arts track	44571	Exclude students applying to the Arts track (Estetiska programmet) as admission was based on practical test in addition to grade sum. As we do not have access to the result on this test, we cannot model the admission decision for this group.
Keep if at least one student in admission group was *not* accepted	44519	Include only admission groups that were competitive, in the sense that not all who applied to the track and school in question were accepted.
Keep if at least one student in admission group *was* accepted	44395	Keep if at least one student in admission group was accepted
1: Independent 2: Public	*9415*	
1: Public 2: Independent	*8534*	
1: Independent 2: Independent	*5974*	
1: Public 2: Public	*20472*	
Number of observations in the RD-sample for the main subsamples:		
Keep if observations near threshold on both sides of threshold	**12060**	We require that there is one student observation near (defined as 10 grade sum units on a scale of 320) the admission threshold on both sides of the threshold for each admission round (i.e. for each yearly observation of first ranked school and track), and within each preference ordering.
Number of observations by preference order:		
1: Independent 2: Public	*2399*	
1: Public 2: Independent	*2009*	
1: Independent 2: Independent	*1401*	
1: Public 2: Public	*6251*	

B2. Data details: Tracks applied to as 1st and 2nd preference by the students in the RD/DID-sample

Table B3 shows the distribution of observations for alternative RD/DID-samples over different educational tracks, i.e. with respect to the tracks the students in the samples applied to. As the table shows, many of the observations are for students applying to academic tracks such as Natural Science, Social Science, Technology and Business Management and Economics.

TABLE B3. DISTRIBUTION OF TRACK OPTION 1 AND 2 FOR RD/DID-SAMPLES, EXCLUDING OBSERVATIONS AT THE ADMISSION THRESHOLD. DATA IS RESTRICTED TO STUDENTS APPLYING TO THE SAME TRACK AS 1ST AND 2ND PREFERENCE.

Data Window:	**5**				**20**				**Full window**			
Preference Ordering:	**In/Pu**	**Pu/In**	**In/In**	**Pu/Pu**	**In/Pu**	**Pu/In**	**In/In**	**Pu/Pu**	**In/Pu**	**Pu/In**	**In/In**	**Pu/Pu**
Panel A: Tracks for students starting upper secondary education in 2009-10												
BF (Child Recreation)	2	3	0	2	7	4	0	2	9	5	0	8
BP (Building and Construction)	0	3	0	6	0	15	0	14	0	20	0	24
EC (Electrical Engineering)	12	6	4	2	35	18	9	5	57	32	20	8
EN (Energy)	0	2	4	0	0	4	11	0	0	4	21	0
ES (Arts)	0	0	0	0	0	0	0	0	0	0	0	0
FP (Vehicle Engineering)	0	0	0	10	6	0	0	24	9	0	0	32
HP (Business and Administration	3	2	2	1	12	5	6	3	32	10	7	4
HR (Hotel, Restaurant and Catering)	3	0	0	0	9	0	0	0	19	0	0	0
HV (Handicraft)	15	5	0	5	32	18	0	17	64	36	0	39
IB (International Baccalaureate)	0	0	0	0	0	0	0	0	0	0	0	0
IP (Industrial Technology)	0	0	0	0	0	0	0	0	0	0	0	0
LP (Food)	0	0	0	0	0	0	0	0	0	0	0	0
MP (Media)	3	4	2	0	8	15	2	0	20	29	2	0
NP (Natural Resource Use)	0	0	0	0	0	0	0	0	0	0	0	0
NV (Natural Science)	34	12	13	55	155	33	42	232	225	79	48	800
OP (Health and Social Care)	0	3	0	0	0	10	0	0	0	17	0	0
SP (Social Science)	23	26	14	71	76	63	61	214	195	181	164	869
TE (Technology)	17	42	44	36	82	142	118	118	187	422	232	252
Panel B: Tracks for students starting upper secondary education in 2011-13												
BA (Building and Construction)	0	22	0	9	0	57	0	24	0	91	0	37
EE (Electricity and Energy)	14	20	9	16	36	73	21	60	58	112	37	99
EK (Business Management and Economics)	58	40	36	70	210	130	126	246	410	268	261	614
FT (Vehicle and Transport)	4	0	0	8	8	0	0	24	12	0	0	45
HA (Business and Administration)	0	1	1	0	0	7	2	0	0	12	3	0
HT (Hotel and Tourism)	2	0	0	0	6	0	0	0	13	0	0	0
HU (Humanities)	0	0	0	7	0	0	0	27	0	0	0	53
IN (Industrial Technology)	1	0	3	0	2	0	9	0	3	0	20	0
NA (Natural Science)	60	19	28	154	238	85	91	521	505	216	162	1398
NB (Natural Resource Use)	0	0	0	0	0	0	0	0	0	0	0	0
RL (Restaurant Management and Food)	0	0	0	2	0	0	0	3	0	0	0	3
SA (Social Science)	50	49	45	182	190	147	154	664	453	336	320	1688
VF (HVAC and Property Maintenance)	0	10	8	0	0	26	16	0	0	49	31	0
VO (Health and Social Care)	0	0	0	0	0	0	0	0	0	0	0	0
All, excluding observations at the threshold	**301**	**269**	**213**	**636**	**1112**	**852**	**668**	**2198**	**2271**	**1919**	**1328**	**5973**
All, including observations at the threshold	**429**	**359**	**286**	**914**	**1240**	**942**	**741**	**2476**	**2399**	**2009**	**1401**	**6251**

Table notes: Preference ordering In/In refers to students having listed an independent school as first and second option; In/Pu students having listed an independent as first and public as second preference, etcetera.

B3. Data Variables: Covariates

The covariates are obtained from the following registers (named in Swedish) from Statistics Sweden: Registret över totalbefolkningen (RTB); Inkomst- och taxeringsregistret (IoT); Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier (LISA); Skolverkets elevregister; Universitets- och högskoleregistret; Utbildningsregistret; Geografidatabasen; Komvux; Folkhögskolan; and Befolkningens studiedeltagande. These registers in turn are based on information from various administrative sources.

The data set consists of student level observations from the merged registers for upper secondary school applications and school attendance, for cohorts applying to and starting upper secondary school in 2009–13. The application and admittance information is observed in the summer (July-August, depending on cohort) and attendance is observed in October in the same year; i.e. the fall term of the first grade in upper secondary education.

Based on the registers, we generate the below described covariates. Where there are missing covariate values, we impute mean values and include dummy variables in the regression to control for the imputation. Summary statistics for all covariates, based on the full observational sample (see Table 2 of the main article for an overview of the samples) are shown in Table B4.

TABLE B4. DESCRIPTIVE STATISTICS: COVARIATES FOR THE FULL OBSERVATIONAL SAMPLE

Variable	Obs	Mean	Std	Min	Max
Household individual disp inc[a]	296,890	243169	282081	-4157372	75600000
One parent business income	296,890	0.14	0.34	0	1
One parent unemployed	296,890	0.17	0.37	0	1
One parent post-sec educ	296,890	0.55	0.50	0	1
Both parents born in Sweden	296,890	0.76	0.43	0	1
Only one parent born in Sweden	296,890	0.11	0.31	0	1
No parent born in West	296,890	0.07	0.25	0	1
Born in Sweden	296,877	0.95	0.22	0	1
Born in West	296,890	0.02	0.15	0	1
Born in non-West	296,890	0.03	0.16	0	1
Female	296,890	0.50	0.50	0	1
Private grade 9	296,890	0.14	0.35	0	1
GPS grade 9	296,890	228.34	47.57	0	320
High test grade Maths	296,890	0.12	0.32	0	1
High test grade Swe	296,890	0.09	0.28	0	1
High test grade English	296,890	0.21	0.41	0	1
Metropolitan municipality	296,890	0.33	0.47	0	1
Urban municipality	296,831	0.51	0.50	0	1
Rural municipality	296,890	0.16	0.36	0	1
Regional independent share	296,890	0.23	0.08	0.05	0.36

[a] Household income is given in year 2016 monetary value.

Household disposable income

The variable household disposable income contains labor and capital income, and taxable and non-taxable benefits, and comes from the Income and taxation register of Statistics Sweden (Inkomst- och Taxeringsregistret IoT) We use the individualized household disposable income *per consumption unit*. This measure takes into account that residing in a household comes with economics of scale benefits, and that the consumption needs differ between older and younger individuals, and lets the weights assigned to different household members reflect this. For example, an adult in a single household has a weight equal to one; cohabiting individuals are each assigned weights of less than one; and children are assigned lower weights than adults.

The distribution of the household income variable is, as expected, highly positively skewed. While the median household family member is endowed with SEK 217,000, the maximum household family member is endowed with SEK 37 million. We do not drop outliers, but instead we include a log transformation of household income in all estimations. The 152 observations that are either negative or zero values are replaced with a 0 after log transformation. However, a dummy to signify negative or zero values of household income is also included. We also include income deciles as covariates.

Final grades from lower secondary education (Final grade sum GPS9)

During the period under study, admission to upper secondary education was based on the students' "grade sums" from lower secondary school. Students starting lower secondary school prior to 2011, were graded on a 4-level scale: Fail; Pass; Pass with distinction; and Pass with special distinction. Each of these levels gave grade credits of: 0, 10, 15 and 20, respectively. The grade sum is defined as the sum of the grade credits of the students' best 16 subjects, and thus ranges from 0 (fail in all subjects) to 320 (highest grade in 16 subjects). For students starting lower secondary school from 2011, a different underlying grade scale was used: instead if 4 grade categories, the new system had a six-level grading scale, from A to F, with A being the highest grade, E being the lowest pass grade, and F fail. The credits attached to the grades were in this case: A:20; B:17.5; C:15; D:12.5; E:10; and F:0. This meant that the grade sum was still ranging from 0–320, but at 2.5-unit intervals instead of 5-unit intervals.

Female

We use a dummy variable defined as one if the student is female, zero if the student is male, and missing if gender information is missing.

Variables based on the students' country of birth

We generate three dummy variables indicating if the student herself is born in i) Sweden; ii) a Western country other than Sweden, and iii) a non-Western county. We define Western countries as countries in Europe, North America and Oceania.

Private school grade 9 The variable comes from the grade 9 graduation register. It takes value one if the student attended an independently provided school in grade 9, and zero if the student attended a publicly provided school. The variable is missing if information Public/Independent provider is missing. This variable is not included in regressions, instead we include all 9th grade schools as dummies.

Standardized test grade variables in Math, English and Swedish: Dummy variables for high and pass grades

We construct the three indicator variables for receiving high test grades on the national standardized tests in Mathematics, Swedish and English taken in lower secondary school. The variables are set to one if the student received the highest possible grade on the test in question ("MVG" under the pre-2011-reform grading system, and "A" under the system implemented in 2011).We also construct three indicators for receiving any pass grade on the same tests. These variables take the value one if the student was awarded any grade other than fail ("IG" under the pre-2011 system and F from 2011 on.)

Indicator variables for Metropolitan, Urban and Rural municipality

The classification of municipalities is constructed by The Swedish Agency for Growth Policy Analysis (*Tillväxtanalys*) The classification is based on the urbanization rate, i.e. the share of the population living in urban area. Municipalities are defined as metropolitan if there are at least 500,000 inhabitants residing within the municipality and the surrounding municipalities and if at least 80 percent of the municipal population lives in urban areas. The remaining (smaller) municipalities where a majority of the population lives in urban areas are classified as urban, municipalities where a majority of the population lives in rural areas are classified as rural. For instance, the municipality of Stockholm is a metropolitan municipality along with Gothenburg, Malmö and their surrounding municipalities. Detached cities like Linköping, Norrköping, Uppsala and Kiruna are classified as urban municipalities. Examples of rural municipalities are Älvsbyn, Arvidsjaur, and Robertsfors, among the municipalities in northern Sweden, and Hässleholm, Simrishamn, and Alvesta in southern Sweden. There are 290 municipalities in Sweden; 29 of them are classified as metropolitan, 131 are classified as urban, and 130 are classified as rural. In 2012, 32 percent of the total Swedish population lived in metropolitan municipalities, 50 percent in urban municipalities, and 17 percent in rural municipalities.

Academic track

Based on the student level information on the educational track of attendance measured in the fall of the first year of upper secondary education, we generate an indicator variable for attending an Academic track. The reference category is attending a Vocational track (note that preparatory track students are not included in our analysis data sample). Students for whom the track of attendance could not be identified, due to missing or uninformative track codes, were assigned missing values.

Variables for parental income, unemployment and country of birth

We generate a set of dummy variables for the parental background in terms of country of birth, highest level of completed education, business income and unemployment. We divide country of birth into Sweden; Western countries except Sweden (defined as Europe, North America and Oceania); and non-Western countries (all remaining countries). Business income is based on active and passive income from private firms, but not from closely nor widely held corporations.[86] The dummy variable generated for this variables indicates that at least one parent has positive business income. Our variable for unemployment is based on Statistics Sweden's employment indicator[87]. If defines an individual as unemployed if /s/he has an amount of yearly labor earnings lower than the basic amount. The basic amount is a figure that is used in Swedish regulations in order to determine benefit levels etcetera, and is adjusted yearly to account for inflation. The basic amount in 2013 was 44 500 SEK, or roughly 4 450 €.

Table B5 displays the exact classification of these dummy variable based on parental characteristics. The aim of the table is to clarify how we define missing values for these variables.

[86] The variable includes the following types of incomes from privately held firms (in Swedish): Inkomst av aktiv enskild näringsverksamhet + Inkomst av aktiv näringsverksamhet för delägare i handelsbolag + Inkomst av passiv enskild näringsverksamhet + Inkomst av passiv näringsverksamhet för delägare i handelsbolag.

[87] The variable "Förvärvsarbetande" from the register Inkomst- och taxering (IoT).

TABLE B5. DEFINITION OF DUMMY VARIABLES FOR PARENTAL BACKGROUND

Both parents born in Sweden (77 percent)	
Both parents born in Sweden	1
All other combinations, and one missing	0
Both parents missing value	.
Only one parent born in Sweden (12 percent)	
One parent born in Sweden, the other not	1
One parent born in Sweden, the other missing value	1
No parent born in Sweden	0
One parent born outside Sweden, the other missing value[a]	0
Both parents missing value	.
No parent born west (7 percent)	
Both parents born non-west	1
One parent born non-west, the other missing value[b]	1
Both parents born in west	0
One parent born in non-west, the other west	0
Both parents missing value	.
(at least) One parent has post-upper-secondary education	
Both parents have post-upper secondary	1
One parent has post-upper secondary, the other not	1
One parent has post-upper secondary, the other missing	1
No parent has post-upper secondary	0
One parent has no post-upper secondary, the other missing[c]	.
Both parents missing value	.
(at least) One parent has positive income from private business	
Both parents have income from private business	1
One parent has income from private business, the other not	1
One parent has income from private business, the other missing	1
No parent has income from private business	0
One parent has no income from private business, the other missing	.
Both parents missing value	.
(at least) One parent is unemployed	
Both parents unemployed	1
One parent unemployed, the other not	1
One parent unemployed, the other missing	1
No parent unemployed	0
One parent not unemployed, the other missing	.
Both parents missing value	.

[a]Most of these missing values pertain to students who are themselves born outside of Sweden. It is therefore reasonable to assume that the other parent whose value is missing, is also born outside Sweden.

[b] In most cases, when one parent is born non-west, and the other parent has missing value, the child is also born non-west. We therefore assume that the parent with missing value is born non-west.

[c] These values are set to missing because we do not know the education level of the parent with missing information, and can make no plausible assumption regarding it (missing values for this variable are more common when the child is born outside Sweden, but we cannot, based on this, infer whether the education level for the parent with missing information level is high or low.)

Table B6 finally shows the averages values for the covariates for the Full Observational samples (see Table 2 of the main article for sample definitions), as well as the normalized differences and p-values for the raw differences, for students attending independent and municipal schools, respectively, in the fall or the first year of upper secondary school.

TABLE B6. STUDENT BACKGROUND CHARACTERISTICS IN INDEPENDENT/PUBLIC SCHOOLS FOR THE FULL OBSERVATIONAL SAMPLE

	Full sample			
Variables	Indep.	Municip.	Norm. diff.	P-value
	(1)	(2)	(3)	(4)
Household disposable income	254 973	239 438	0.046	0.000
One parent business income	0.149	0.141	0.024	0.000
One parent unemployed	0.187	0.162	0.067	0.000
One parent post-sec educ	0.556	0.552	0.007	0.088
Both parents born in Sweden	0.730	0.767	-0.086	0.000
One parent born in Sweden	0.127	0.107	0.062	0.000
No parent born in West	0.080	0.066	0.052	0.000
Born in Sweden	0.947	0.951	-0.018	0.000
Born in West	0.024	0.022	0.015	0.000
Born in non-West	0.029	0.028	0.010	0.022
Female	0.518	0.492	0.052	0.000
Independent9	0.212	0.117	0.257	0.000
GPS9	227.0	228.8	-0.037	0.000
High MA Test9	0.116	0.125	-0.028	0.000
High SW Test9	0.092	0.089	0.008	0.059
High EN Test9	0.233	0.209	0.057	0.000
Metropolitan municipality	0.498	0.278	0.464	0.000
Urban municipality	0.401	0.550	-0.302	0.000
Rural municipality	0.101	0.172	-0.208	0.000
Vote share center-right	0.483	0.469	0.147	0.000
Regional independent share	0.262	0.225	0.471	0.000
Observations	*71,310*	*225,580*		

B4. Outcome variables

Below follows a more detailed and technical description of outcome variables than the shorter summary version that is available in the main paper. Summary statistics for the outcome variables, for the full observational sample (see Table 2 in the main article for sample definitions), are given in Table B7.

TABLE B7. DESCRIPTIVE STATISTICS OUTCOME VARIABLES FOR THE FULL OBSERVATIONAL SAMPLE

Variable	Obs	Mean	Std.	Min	Max
Switch school type	291,017	0.04	0.20	0	1
Pctile GPA12	256,866	55.25	27.49	3.54	99.93
Graduate on time	296,890	0.83	0.38	0	1
7th term	296,890	0.09	0.29	0	1
Post-sec. studies	231,251	0.38	0.49	0	1
Post-sec. studies no-comp [b]	231,251	0.31	0.46	0	1
UC≥15	231,251	0.16	0.36	0	1
Work ≥50%	231,068	0.28	0.45	0	1
National tests, Mathematics [c]					
High test grade	319,724	0,05	0,23	0	1
Pass test grade	319,724	0.78	0.42	0	1
Test grade>Course grade	294,446	0.01	0.11	0	1
Test grade<Course grade	294,446	0.28	0.45	0	1
National tests, Swedish [d]					
High test grade	308,238	0,08	0,27	0	1
Pass test grade	308,238	0.95	0.22	0	1
Test grade>Course grade	280,723	0.10	0.29	0	1
Test grade<Course grade	280,723	0.29	0.46	0	1
National tests, English [e]					
High test grade	305,508	0.11	0.31	0	1
Pass test grade	305,508	0.98	0.15	0	1
Test grade>Course grade	288,213	0.11	0.31	0	1
Test grade<Course grade	288,213	0.16	0.37	0	1

[a] KMX=Adult education (Komvux), LM= Active Labor market programs (Arbetsmarknadsutbildning), SFI=Swedish for Immigrants.
[b] Ibid.
[c] Note that students may take more than one tests per subject, so some students have multiple test observations within a subject. This explains the higher number of observations for the test variables.
[d] Ibid.
[e] Ibid.

B4.1 Intermediate outcomes

Switching school type: This dummy variable equals one if the student is still enrolled in the same type of school (independent or municipal) in the fall of grade 1 and grade 3, respectively, of upper secondary school. The variable is obtained from the *School register*.[88] Cases where information on Independent/Public provision is missing in either year are set to missing. This means that students who dropped out and therefore have no school information in year 3, are treated as missing observations in this variable.

Standardized test grades (Sw, En, Ma): Our data contains information on the grade received by the students on the national (standardized) tests taken in upper secondary school, for the subjects Swedish, English and Mathematics. Standardized tests are used to guide teachers' grading of students, but there is no requirement that the course grade correspond to the test grade.[89] The upper secondary test data that we have access to is available from the fall term of 2011 – previously, there was no comprehensive collection of the test data but only of subsamples. This means that tests for the students in the earlier cohorts of our data are only observed if they were taken during the later years of upper secondary education.

Students in upper secondary school take one or several Math/Swedish/English courses, with varying difficulty, depending on the educational track they attend. During the time period of our analysis, the national tests were mandatory for the initial and final course in each subject.[90] For many students in the vocational tracks, this still meant taking one test per subject, as there was only one course per subject. In addition to the mandatory tests, the National Agency for Education provided some tests for courses that were not mandatory to test in any educational track.[91] These tests could be used by the schools to guide the grading of students. There can be circumstances where individual students voluntarily take courses that are not mandatory for their educational track, and thus also participate in the course test. Our data includes both the mandatory and non-mandatory tests, and we have no indicator variable for whether or not a test was mandatory. For each course test, the National Agency for Education provided two possible test occasions each school year – one in the fall term and one in the spring term. The motivation for this is that schools have a lot of flexibility in terms of the scheduling of courses, so for some schools it makes more sense to test students in the fall, and for others in the spring, since the tests should be taken by the end of the course. Our data shows that it is by far most common to participate in the spring tests. In 2011, a reform affecting the curriculum and grading system was implemented. This means that students entering upper secondary school after the reform were subject to slightly different courses, and to a different grading system. To summarize, the following holds for the test data:

- Students in different educational tracks take different course tests, and different numbers of tests.
- Schools can choose whether to schedule the tests either in the fall or the spring term.

[88] Although we do know whether the student has switched between *types* of schools, we cannot be sure to observe all school switches that take place between schools within each school type, for the following reasons: i) school units that go through reconstructions (such as mergers) can be assigned a new School ID; ii) the definitions of school units in the School register changed in years 2012–2013, and this means that some school units cannot be linked over time In particular, the difficulty in linking schools over time stems from the fact that many schools that were defined as one school in the previous system, were under the new system recorded as several school units, each with a specific new code.

[89] https://www.skolverket.se/for-dig-som-ar.../elev-eller-foralder/betyg-och-nationella-prov/nationella-prov

[90] From January 2018, only the final subject course in each educational track is mandatory.

[91] Two such examples, according to the National Agency for Education, are Matematik 2a and Matematik 2c, see ”PM - Nationella prov i gymnasieskolan våren 2018”, Diarienummer: 5.1.1 – 2018:01623.

- The course structure and the grading system changes over time, such that students in different cohorts took different tests and were graded according to different scales.

These are data issues that need to be addressed in our analysis. We do this by dummy variables for: the test year, the test term, grade when the test was taken and the course tested.

Tables B8–9 show the classifications used, and how they reflect the shares of students receiving each grade under the two grading systems. As can be seen in the tables, the grade distribution varies a lot across the subjects as well as across course tests within subjects. In order to generate outcome variables that can be used for the entire period, we generate two dummy variables for receiving a high or pass grade, respectively, in the following manner:

- High grade = test grade MVG or test grade A
- Pass grade = test grade MVG, VG or G or test grade A–E
- In addition, we generate two dummy variables that indicate if the student received a higher, or lower, respectively, grade on the course test than the actual grade received for the course.

TABLE B8. DISTRIBUTION (%) OF TEST GRADES UNDER THE PRE-2011 CURRICULUM, BASED ON THE RAW TEST DATA

Courses in:	Math				Swedish	English	
	MAA	MAB	MAC	MAD	SVB	ENA	ENB
Pass with Special distinction (MVG)	2.4	2.3	7.7	19.2	11.9	7.6	12.8
Pass with distinction (VG)	10.8	12.9	19.2	23.6	38.2	34.2	45.0
Pass (G)	42.4	39.7	45.4	39.9	42.1	49.6	38.5
Fail (IG)	44.4	45.1	27.7	17.3	7.8	8.6	3.6

TABLE B9. DISTRIBUTION (%) OF TEST GRADES UNDER THE POST-2011 CURRICULUM, BASED ON THE RAW TEST DATA. (A IS THE HIGHEST GRADE, E IS THE LOWEST PASS GRADE, AND F IS FAIL)

Courses in:	Math							Swedish		English	
	MATMAT01A	MATMAT01B	MATMAT02A	MATMAT02B	MATMAT03B	MATMAT03C	MATMAT04	SVESVE01	SVESVE03	ENGENG05	ENGENG06
A	0.9	2.6	0.8	0.8	1.7	9.1	9.2	5.1	7.9	11.8	10.6
B	2.3	6.5	1.5	2.5	4.6	13.3	12.0	15.4	16.2	18.7	18.2
C	7.3	14.7	5.7	10.0	12.8	20.5	19.3	26.6	22.3	30.2	31.6
D	14.7	22.3	9.2	15.1	16.7	17.7	16.9	26.8	24.2	20.2	22.9
E	44.0	36.9	31.2	33.2	33.7	23.0	23.8	21.8	21.1	15.4	14.3
F	30.9	17.1	51.7	38.5	30.5	16.4	18.9	4.3	8.3	3.7	2.5

Finally, we make the following restrictions to the upper secondary test data:

- We drop tests that were not taken on a regular test date, and/or cases where another test than the regular test was used. These cases, which altogether make up less than 2% of our total raw test data, can be cases where a student was sick at the regular test date, and was therefore given a separate test (for example a test from an earlier year) at a later date. It can also be cases where a replacement test was used because of suspicions that the regular test had been leaked to students beforehand.
- In addition to the course tests in Math, English and Swedish, there are separate tests for courses in "Swedish as a second language". We drop these observations from or analysis sample, and this causes 1.7% of observations to be dropped (after we have restricted the data to the regular tests and regular test occasions, as described above).

- Finally, some students are recorded as taking the same *course* test at more than one occasion (this could happen if a student changes track and therefore needs to retake a course). These cases are dropped, and this eliminates less than one percent of the data (after the above restrictions were done). In addition, there is a small number of students who are observed as taking course tests under both the pre- and post-2011 curricula. This could also happen due to gap years or delays due to track changes.[92] For these students, we keep only the test taken under the first curriculum.

B4.2 Graduation outcomes

Graduation with complete grades: we construct a dummy variable that takes value one if the individual graduates on time with a complete set of grades. The variable takes value zero if the student is not observed in the graduation register (meaning that the student either dropped out or is still in school). It also takes value zero if the individual leaves upper secondary school within three years with *incomplete* grades, except for students who received a grade transcript of at least 2500 course credits. The reason is that this type of transcript, which was introduced in 2011, was according to Statistics Sweden often given to students who would in the years prior to 2010 count as graduates, see footnote 14 for more information. The variable is obtained from the *Graduation register*.

Percentile rank of Final GPA when graduating from upper secondary education (From the graduation register): This variable is defined as the percentile rank of final GPA, calculated separately for each graduation year among all graduating students (not just the regression sample).

The final GPA is calculated in the following manner: Students in upper secondary education take one or several sub-courses in each track, and the number of sub-courses (and the subjects taken) varies across the educational tracks. (Students in math-heavy tracks take several math courses, etc.) Each sub-course gives a number of course credits, and students are graded in each sub-course. For the calculation of the GPA, the grades are translated into grade credits, with the highest grade equaling 20 credits, and the lowest pass grade giving 10 credits.[93] The final GPA is calculated as the average grade credit over all sub-courses, weighted by the course credits.[94]

We calculate the percentile score based on the available GPA-information for all students, including those who finished upper secondary school with incomplete grades, and therefore obtained a grade transcript ("samlat betygsdokument") instead of a proper graduation certificate. The reason for including also the students with a transcript, is that the GPA includes valuable information also for many of these students, and can thus be used to rank students.[95] We however exclude students

[92] Note that the new curriculum was introduced only for the incoming students – students already attending school under the old curriculum continued under the old regulation throughout upper secondary school.
[93] For students entering upper secondary education up to and including 2010: IG:0, G:10, VG:15, VG:20, and for students entering upper secondary school starting from 2011: F:0, E:10, D:12.5, C:15, B: 17.5, A:20.
[94] Under grade system for students entering 2011 and onwards additional course credits of max 2.5 could be added to the GPA for certain courses in modern languages, English and Math, when they apply to certain University programs. These extra credits are however not included in the GPA of our data (as they vary with the University track/program applied to).
[95] According to Statistics Sweden (information in email conversation), students with a grade transcript who started upper secondary education before 2011 have all been assigned a final GPA-value of zero in the data, even though they may have had a nonzero GPA. For students entering from 2011 on, the requirements for obtaining a proper final GPA-certificate instead of a transcript were increased, and this resulted in more students ending up with a transcript instead of a certificate than previously. The graduation register data for these students contain two types of transcripts: one for students with at least 2500 course credits and one for students with fewer course credits. According to Statistics Sweden, many of the students in the former category would likely have received a proper final GPA-certificate under the previous system. This is supported by the fact that the data shows that most of the students with this type of transcript have non-zero, and on average relatively high, final GPA-values, whereas students with the latter type of transcript often have zero or very low GPA.

graduating from the IB-program, because for these students, according to Statistics Sweden, the GPA is in the data set to zero also when it is really non-zero.

B4.3 Post-graduation outcomes

Post-secondary education status: The register contains information on all individuals in education, including both the regular education system (primary, secondary and tertiary) and alternative types of education such as “folkhögskola”, adult education (Komvux), Swedish for immigrants (SFI), active labor market policy education (arbetsmarknadsutbildning) etc.

We use the register information as measured for the fall term, and measure this both for the fall term following the expected graduation of the student (if graduating on time, i.e. three years after graduating from grade 9), and four years after graduating from grade 9, i.e. giving the student one additional year.

Based on this information, we construct two indicators variables for post-secondary education:

i) An indicator for being registered in any type of post-secondary education, including also categories that are more “repeat/complementary” education, such as adult education, active labor market education, and Swedish for immigrants. The full list of education categories covered by variable is:
   a. KOMVUX (adult education)
   b. Tekniskt basår vid univ/högskola (Technical preparatory university/college year)
   c. Grundläggande Högskoleutbildning (Basic college/university education)
   d. Forskarutbildning (Post-graduate PhD education)
   e. Kvalificerad yrkesutbildning/yrkeshögskoleutbildning (Qualified vocational studies)
   f. Folkhögskola (”Folk high school”)
   g. Studiemedel för utlandsstudier (Studies abroad that qualify for Swedish study grants)
   h. Övriga med studiemedel (Other studies that qualify for Swedish study grants)
   i. Arbetsmarknadsutbildning (Labor market education)
   j. Kompletterande utbildning/konst- och kulturutbildningar (Complementary educations in arts/culture)
   k. Utbildning i svenska för invandrare (SFI) (Swedish for immigrants)
   l. Uppdragsutbildning i universitet/högskola (Commissioned education university/college)

ii) An indicator for any type of post-secondary education that *excludes* categories that are more of “repeat/complementary” education. This variable is similar to the above, apart from that it excludes the categories adult education (KOMVUX), active labor market education (Arbetsmarknadsutbildning) and Swedish for immigrants (Utbildning i svenska för invandrare SFI).

In addition, we construct a dummy variable for still being in upper secondary education measured in the fall after the individual was expected to graduate, based on the information in the education register.

Higher education credits (ECTS credits): We measure ECTS credits during the fall term, and generate a dummy variable for taking credits amounting to at least half-time equivalent studies, i.e. ≥ 15 ECTS credits (full time studies amount to 30 ECTS credits). The reference category for this variable consists of individuals who are subscribed to courses but take fewer than 15 credits and individuals who are not subscribed to any higher education courses.

Labor income: This variable is measured by yearly labor income, i.e. the sum of employment and active entrepreneurship (personal firm) income.[96] We follow Forslund et. al. (2017), and define a dummy variable for having labor income amounting to at least half of the median annual work income among 45-year olds. When studying labor earnings in the graduation year we however instead use a quarter of the median income among 45-year olds, as the students graduating in that year were still in upper secondary education approximately half of that year.

Table B10 shows how the average levels of the outcome variables for students who attended an independent or municipal school, respectively, in the fall of the first year of upper secondary school, for the Full observational and Main observational samples (See Table 2 in the main article for sample definitions).

[96] CSFVI sammanräknad förvärvsinkomst = Variabeln utgörs av summan av inkomst av tjänst och inkomst av näringsverksamhet: CSFVI = TTJ (ruta 1 på Skatteuträkningsbilagan) + NRV (ruta 2 på Skatteuträkningsbilagan)

TABLE B10. OUTCOMES IN INDEPENDENT/PUBLIC SCHOOLS FOR THE SAMPLES USED IN THE CEM/VAM-ANALYSIS

	Full sample				**Main sample**			
Variables	Indep.	Municip.	Norm. diff.	P-value	Indep.	Municip.	Norm. diff.	P-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Graduation and grades								
Switch school type	0.079	0.032	0.205	0.000	0.088	0.061	0.103	0.000
Pctile GPA12	57.231	54.640	0.093	0.000	57.143	53.217	0.140	0.000
Graduate on time	0.816	0.830	-0.038	0.000	0.823	0.808	0.039	0.000
7th term	0.097	0.092	0.016	0.000	0.093	0.103	-0.034	0.000
B. Standardized tests								
Mathematics								
High Test	0.055	0.054	0.006	0.116	0.059	0.050	0.038	0.000
Pass Test	0.763	0.783	-0.046	0.000	0.763	0.772	-0.021	0.003
Test grade>Course grade	0.017	0.012	0.037	0.000	0.017	0.013	0.032	0.000
Test grade<Course grade	0.300	0.277	0.050	0.000	0.301	0.271	0.066	0.000
Swedish								
High Test	0.098	0.076	0.077	0.000	0.095	0.069	0.092	0.000
Pass Test	0.950	0.947	0.011	0.007	0.948	0.944	0.019	0.009
Test grade>Course grade	0.099	0.094	0.016	0.000	0.098	0.099	-0.004	0.614
Test grade<Course grade	0.305	0.289	0.034	0.000	0.305	0.285	0.045	0.000
English								
High Test	0.135	0.101	0.105	0.000	0.128	0.103	0.076	0.000
Pass Test	0.979	0.978	0.010	0.019	0.978	0.977	0.002	0.812
Test grade>Course grade	0.110	0.103	0.022	0.000	0.108	0.111	-0.010	0.195
Test grade<Course grade	0.181	0.152	0.080	0.000	0.183	0.148	0.093	0.000
C. *Post-graduation*[a]								
Post-sec. studies	0.382	0.383	-0.001	0.891	0.383	0.368	0.032	0.000
Post-sec. studies no-comp[b]	0.309	0.313	-0.009	0.075	0.312	0.295	0.037	0.000
UC≥15	0.149	0.160	-0.032	0.000	0.152	0.144	0.022	0.011
Work≥50%	0.254	0.288	-0.076	0.000	0.259	0.279	-0.045	0.000

The normalized difference between samples 1 and 2 for covariate X is calculated as $\frac{\bar{X}_1-\bar{X}_2}{\sqrt{(S_1^2+S_2^2)/2}}$ (Imbens, 2015).

[a] Post-graduation outcomes are measured in the year following the graduation year, i.e. 4 years after entering upper secondary school.
[b] The pre-registered snapshot version of this table contained an error in this variable. This has been corrected, which is why the variable content for this variable differs from the same table in the snapshot.

B.5 Additional Data Details

B5.1 Errors corrected after the registration of the snapshot version of the project

After the registration of the snapshot, we corrected the following errors in the data. The data used for the analysis in this article is thus slightly different than the one of the snapshot.

- A small set of the parents of the students in our data have potentially erroneous personal id numbers. These observations were retained in the snapshot version, but are now dropped from the data.
- The snapshot version of the CEM-analysis included school dummy variables that were in a few cases miss-classified. This has been corrected.
- In the snapshot version, we unnecessarily dropped a few schools that we suspected had erroneous school-identifiers. These schools are now added to the data.
- The two outcome variables indicating that the national test grade was higher or lower than the corresponding subject course grade, were miss-classified in the snapshot version of the article. The same holds for the outcome variables that measures if the students were registered in a post-secondary education, excluding adult complementary education, active labor market educational programs and Swedish for immigrants, in the fall after the expected graduation from upper secondary school, and a year later, respectively.

B5.2 Missing information on school codes in the application data

The data on applications and admissions to upper secondary school has missing or erroneous information in school codes for relatively large shares of the data for the early cohorts. When we restrict the data to the types of observations used in the analysis of this paper[97], the share of observations with missing or erroneous school code is about 10% overall for the period 2009–2013, and is concentrated in the two first years of the period, when the share is around 20%. The reason for the large shares with missing information, according to Statistics Sweden, is that prior to 2012 information on the school code was not a mandatory piece of information to submit. From 2012, when it did become mandatory, the share of observations with missing school codes is very small.

B6. Data details related to the RD/DID-analysis

B6.1 Density of the data around the admission thresholds

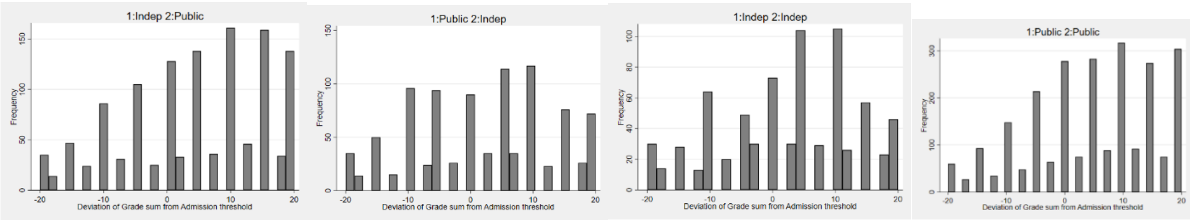
A relevant issue for the validity of the RD is that students are not able to manipulate their grade sum in order to end up marginally on the right side of the admission threshold, for example by working harder to get higher grades.[98] Even though our analysis is not a pure RD, it is interesting to study the distribution of the data. Figure B1 therefore shows the distribution of the deviation of the students' grade sum from the admission threshold. In order to see clearly the distribution of the observations around the threshold, the figures in panel A zoom in on the 20 grade sum units around the admission threshold, while panel B shows the full distribution.

[97] That is, we keep students' top two ranked alternatives; students with first priority as applicants (i.e. residing in the admission regions of the schools); and students who were qualified (i.e. had sufficient grades from lower secondary school to be eligible) for the tracks in question and who applied right after finishing lower secondary school.

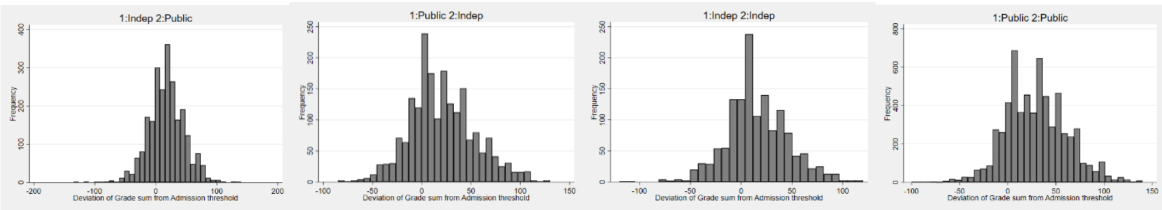
[98] It can however be pointed out that, whereas it is plausible that students will have some idea of where the admission threshold for a certain track and school combination will be (based on previous years' admissions), it is highly unlikely that they will be able to predict this with certainty or precision, as the admission threshold is a function of the number of slots available and the grade sum of all applicants.

FIGURE B1. DISTRIBUTION OF THE DISTANCE TO THE ADMISSION THRESHOLD (NOTE THE DIFFERENT VALUES ON THE Y-AXES)

PANEL A: DATA WINDOW 20 UNITS AROUND THE ADMISSION THRESHOLDS



PANEL B: FULL DATA WINDOW



Two main pieces of information are to be taken from the figures: First, they show no indication of a bunching at or just to the right of the admission threshold.[99] Second, the figures clearly show the discrete nature of the running variable. It can be noted that a "one step" increase in the running variable reflects getting a higher grade in one out of the 16 subject grades that make up the grade sum. Getting a higher subject grade means increasing the grade sum with between 5 and 10 credits, depending on the grade value, in year 2009-2012, and with 2.5 or 10 units in year 2013 due to a change in the grading system. This gives rise to the pattern that every other data bin is much lower than the rest. As can be seen in the below figures, for the full data window, bulk of observations for the running variable fall within approximately ±50 from the cutoff value zero.

B6.2 Visual presentation of the predictive power of the admission thresholds

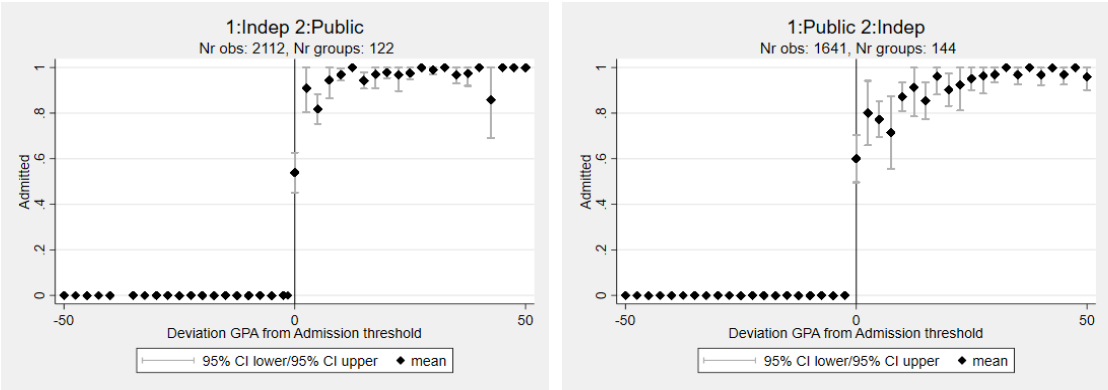
The admission threshold is inferred from our data on admissions and applications; we measure the threshold as the grade sum of the admitted student with the lowest grade sum.[100] This procedure likely gives rise to some measurement errors in the threshold values, and this is discussed further below in this section. Still, Figure B2 indicates that this is not a major issue: the inferred admission thresholds are in general good predictors of the probability of being admitted to the first preferred option. The figures show how the likelihood of admission to the most preferred alternative changes with the distance to the admission threshold, and are limited to a window of 50 on each side of the admission threshold in order to make it easier to distinguish the pattern near the admission threshold.

[99] Note also that, since the data is discrete and the admission threshold is inferred from the grade sum of the last admitted student, it would not be surprising to have a higher density just *at* the threshold, since there is, by definition, always at least one student with this grade sum value.

[100] Specifically, the admission threshold is measured as the lowest grade sum among those admitted to a track and school in a given year, among those who are qualified to the track and who apply as first prioritized applicants, i.e. reside in the application region that the school belongs to. (Note that for independent schools, the entire country forms the application region.)

FIGURE B2. DISTANCE TO THE ADMISSION THRESHOLD AND ADMISSION TO THE TOP CHOICE

PANEL A: SAMPLES OF STUDENTS WITH MIXED PREFERENCES FOR TYPE OF SCHOOL (INDEPENDENT OR PUBLIC) AS FIRST AND SECOND PREFERENCE.



PANEL B: SAMPLES OF STUDENTS WITH SAME TYPE OF SCHOOL (INDEPENDENT OR PUBLIC) AS FIRST AND SECOND PREFERENCE.

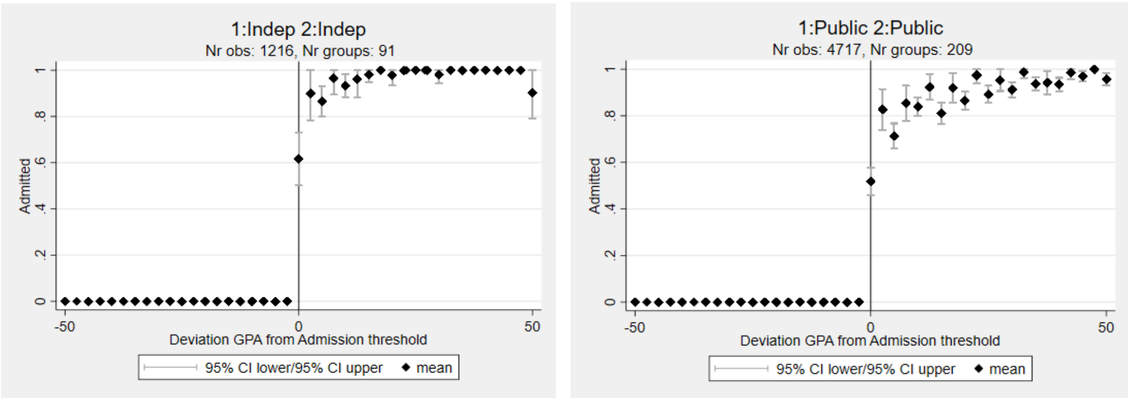


Figure notes: 95 percent confidence intervals, based on the students' t-distribution, are shown for each bin. For some bins, with a very small number of observations, the confidence interval exceeded ± 1, in which case the confidence interval extends to outside of the shown graphs. "Nr obs" denotes the total number of obs used to generate the figure, and "Nr groups" the number of admission groups, i.e. school×track×year level admission thresholds.

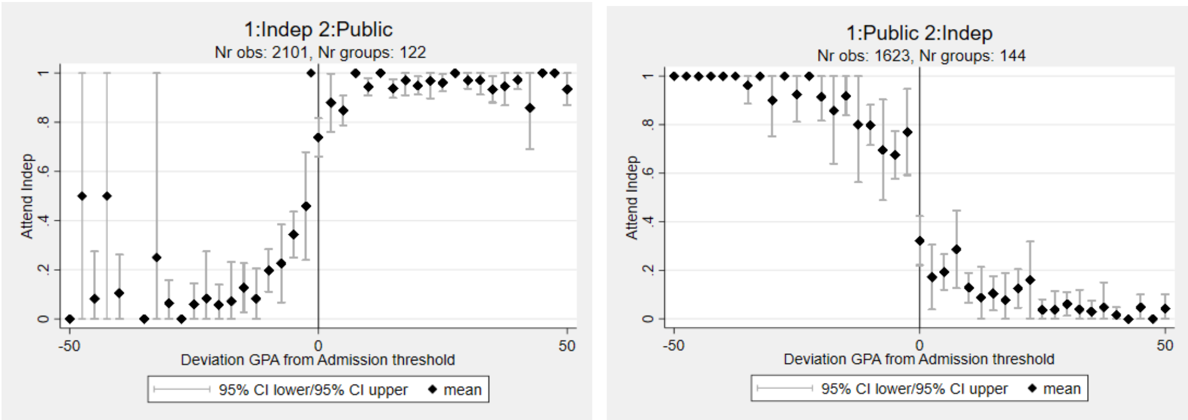
As can be seen in the figures, there is a clear discontinuous increase in the probability of being admitted to the most preferred option at the admission threshold for all subsamples. To the left of the threshold the likelihood of admission is zero for all observations, and this follows from the fact that our inferred thresholds are defined as the lowest grade sum observed among the admitted students so that there is by construction no admitted student with a lower grade sum in the data. On the right hand side of the admission threshold, the share of admitted students ranges between approximately 70 and 100 percent over the distribution of the binned data, with more of the lower values just above the threshold. At the admission threshold, the share of admitted students is 50–60 percent. A likely explanation for the lower admission share at the threshold can be found in the fact that when several students are tied at the threshold, the administration agencies could choose among a set of additional criteria, such as random allocation, grades in specific subjects, etcetera.[101] We will therefore drop the observations located at the admission threshold (value 0 of the running variable in the above figures)

[101] The Swedish Association of Local Authorities and Regions (SKL) publishes yearly handbooks with information on the current regulation and guidelines to the regional agencies regarding admission to upper secondary education. We are grateful that we were given access to the handbooks covering our data period from the SKL. The handbook for the current year can be found online: https://webbutik.skl.se/sv/artiklar/handbok-for-gymnasieantagning-2019-2020.html.

throughout the regression analysis. For additional potential explanations for why the share of admitted students is lower than 100% above the threshold, see section B6.3. We conclude from the above figures that the admission probability does indeed increase substantially at the admission threshold, although not from 0 to 100 percent.

FIGURE B3. DISTANCE TO THE ADMISSION THRESHOLD AND ATTENDING AN INDEPENDENT SCHOOL

PANEL A: SAMPLES OF STUDENTS WITH MIXED PREFERENCES FOR TYPE OF SCHOOL (INDEPENDENT OR PUBLIC) AS FIRST AND SECOND PREFERENCE.



PANEL B: SAMPLES OF STUDENTS WITH SAME TYPE OF SCHOOL (INDEPENDENT OR PUBLIC) AS FIRST AND SECOND PREFERENCE.

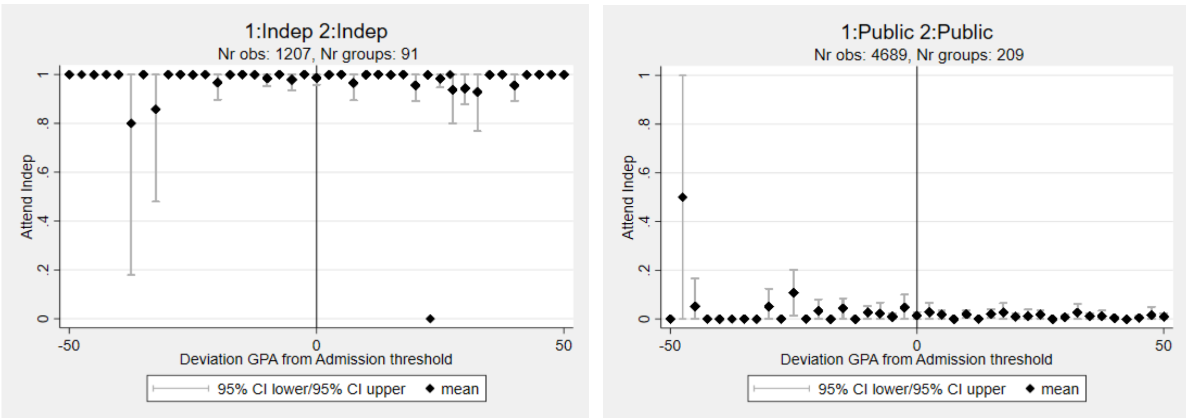


Figure notes: 95 percent confidence intervals, based on the students' t-distribution, are shown for each bin. For some bins, with a very small number of observations, the confidence interval exceeded ± 1, in which case the confidence interval extends to outside of the shown graphs. "Nr obs" denotes the total number of obs used to generate the figure, and "Nr groups" the number of admission groups, i.e. school×track×year level admission thresholds.

For the purpose of this paper is it also important to study how well the initial admissions translate into later school attendance, and in particular to the probability of attending an independent school – the treatment variable of our interest. Figure B3 therefore shows the probability of attending an independent school measured in October each year, over the distribution of the assignment variable; the distance to the admission threshold. Note that as the y-axis variable is now defined as "attending independent", we expect to see a *decrease* at the admission threshold for the students with a public school as most preferred option and independent as second, and an increase for students with the reverse preference ordering. For the samples of students listing independent schools as both options, or public, we naturally expect no change in the likelihood to attend an independent school at the admission threshold. Panel B show that this is indeed the case.

Figure B3 indicates that the discontinuous changes in the admission indicator above carries over to the likelihood of independent school attendance, albeit to a smaller magnitude. In other words, students with an independent school as first preference and public as second best option, are more likely to attend an independent school if they are to the right than to the left of the admission threshold, and the change looks discontinuous at the admission threshold. The reverse holds for students with the opposite preference ordering. The smaller magnitude of the change in probability at the admission thresholds, compared to Figure B3 above, is expected, since the independent school attendance is measured in October, after students have had time to change their mind about their school choices and potentially change schools, giving room for others to be admitted from the waiting lists.

B6.3 Additional information on the measurement of the admission threshold

We define our proxy variable for the admission threshold based on our information on upper secondary school applications and admissions. More precisely, we measure the admission thresholds as the lowest grade sum among those admitted to an educational track and school in a given year in the application data. Educational track is measured as the detailed track codes given in the application/admission data, i.e. the variable "stvkod". We take into account that students are given priority in the admission if they belong to the admission region of the school, and so we generate separate admission thresholds for each "priority group" (priority is indicated by the variable "grupp"). (Note that for independent schools the entire country form one application region.) We also make use of the variable that indicates whether students were qualified (in terms of the lower secondary school final grades) to enter each respective track applied to, the variable "beh", and we base the admission thresholds on those qualified to be admitted.

If we had perfect information on the admission status; preferences; and student grade sum, we should be able to perfectly infer the actual estimation thresholds. However, as will become clear in the analysis below, it is likely that our data contain some errors, which spill over to errors in the inferred admission thresholds. These errors may explain why the observed admission probability in our data is a bit lower than 100 percent above the estimated admission threshold. In addition to the possibility of random input errors, there is a chance that some of the admission data reflects late admissions, i.e. after some students made changes to their initial listings and others were accepted from waiting lists.[102] Both of these possibilities (errors in data and late admissions) could give rise to the lower than 100 percent admission rate above the observed admission threshold. Moreover, we cannot rule out that the regional admission agencies/schools sometimes make errors in the admission process. We have no evidence that this is the case, but it is still one possible source of error.

As commented in the main paper, schools can get permission by the National Agency for Education to base admission on ability tests, in addition to grades, for special educational tracks for high ability

[102] The reason for this is that, even though the admission data that the regional admission agencies have submitted to Statistics Sweden shall reflect the admission status from the first application round in early July each year, it cannot be ruled out that it in some cases reflect the admission status of students later in the summer or even fall. According to the instructions given by Statistics Sweden to the local admission agencies/schools, they are to send in the data of the first round of admission, such that the data reflects the admission status as of early July. However, the deadline for submitting the said data to Statistics Sweden for the time period studied here was mid-August, and some agencies/schools may have submitted the data later than that. (Reference: Email correspondence with and documentation from Statistics Sweden.)

students.[103] This applies to a low number of exceptional cases, but may explain some of the indicated errors in the figures.

For the public schools, there is an additional complicating matter, namely that some municipalities may have used different criteria for public school placement during the period under study. More specifically, we know that all municipalities currently apply grade based admission to schools (and tracks), and we know that this applied to a large set of students in 2009–13, but we cannot rule out that some municipalities were using the proximity principle to determine school placement during this period. The municipalities/admission regions with only one public school per track cause us no problems – there was only one school to be placed in for a given track – but in other cases it may matter.

How will this impact the estimations? If admission to a municipal school within an admission region with several public schools was *not* based on the student grade sum, but was instead determined by the proximity principle, then our inferred admission thresholds would be incorrectly specified. One way to test if this seems to be the case, is to run separate estimations for how well our inferred thresholds explain admission for the sample of municipalities that we know applied purely grade-based school admission, and for the sample for which we lack information, respectively. If the predictive power of the admission threshold does not differ between these two samples, then this can be seen as indicative of either that the proximity principle was indeed abandoned in all municipalities at this time – or that our inferred admission threshold nevertheless provides a good approximation of the true threshold.

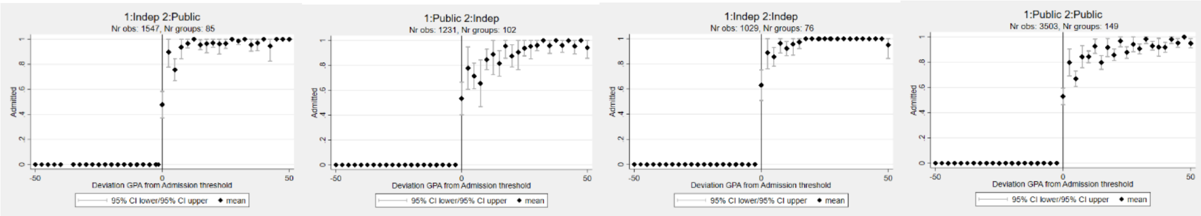
Below, separate admission figures are shown for the municipalities for which we lack information on whether grade-based school level admission was implemented during our sample period, denoted “No information sample”, and the municipalities for which we know that school placement was based on grades, denoted the “Information sample”. It can be noted that several municipalities in the no-information group, 78%, have only one school unit offering each educational track application code. (The corresponding figure for the municipalities in the “Information sample” is 70%.) This means that, unless they have agreed to form a larger admission region during the time period studied (something which is also by large unknown to us), admission to an educational track effectively determines school placement, such that admission to the track and school was in any case determined by the grades. However, as we do not know if some of them were in fact part of larger admission regions, we still include them in the no-information category.

The figures below indicate that our inferred admission threshold is highly predictable for admission for both samples. The figures show no sign of worse predictive power for the “No Information-sample” – if anything, the pattern is the reverse. This can be due either to the municipalities in the “No Information”-sample actually having implemented grade-based school admissions at this point in time, or to there just being one school for each track within the admission region anyway (i.e. admission to a track automatically meant admission to a school, and admission to a track within an admission region was for sure based on the grade sum). Or, it might be that other causes for measurement error in the inferred admission threshold blur out any differences between the samples due to the admission system. In any case, we interpret the below figures as evidence that the inferred admission threshold does a sufficiently good job at predicting admission for both samples, and we will therefore use the full sample – the combination of the two – for the analysis.

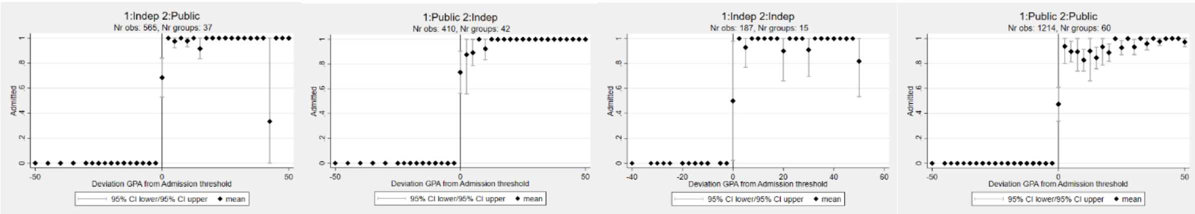
[103] See Chapter 5 in the Upper Secondary School Regulation (Gymnasieförordningen 2010:2039).

FIGURE B4. RD ADMISSION FIGURE, SEPARATE SAMPLES

PANEL A: THE “INFORMATION SAMPLE”



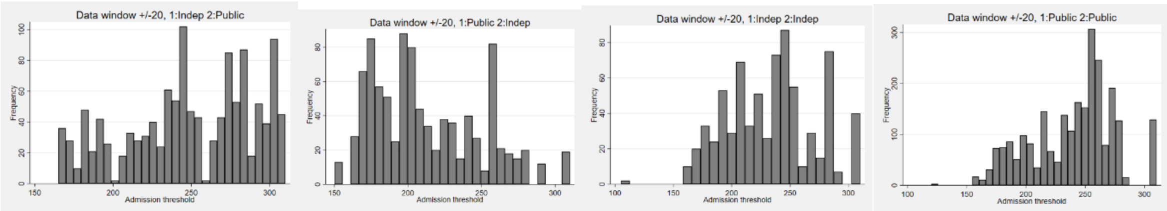
PANEL B: THE “NO INFORMATION SAMPLE”



B6.4 Additional RDD Figures

FIGURE B5. DISTRIBUTION OF THE INFERRED ADMISSION THRESHOLDS

PANEL A: DATA WINDOW ±20 (NOTE DIFFERENT RANGES OF Y-AXES)



PANEL B: FULL DATA WINDOW

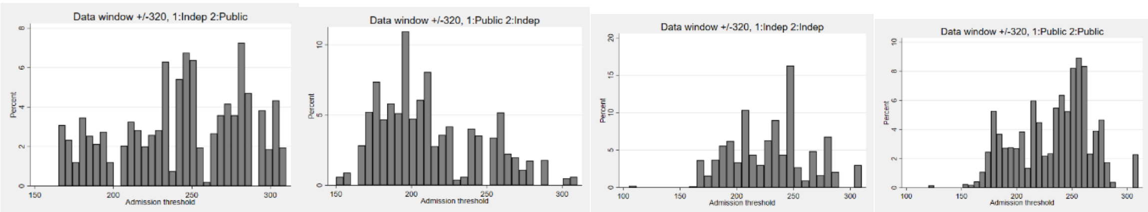
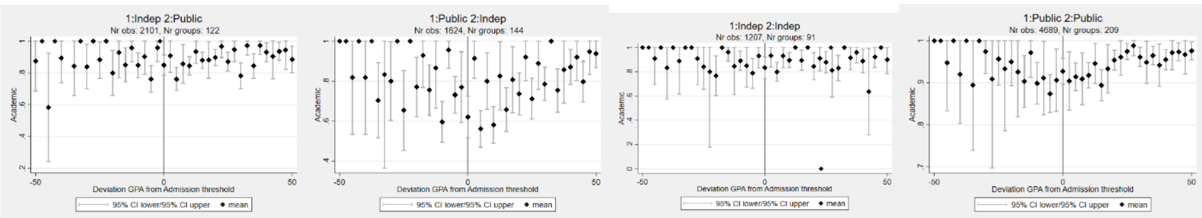


FIGURE B6. RD-FIGURES FOR ATTENDING ACADEMIC OR VOCATIONAL TRACK (IN OCTOBER OF THE FIRST GRADE OF UPPER SECONDARY SCHOOL) FOR THE FOUR SETS OF PREFERENCE-ORDER SAMPLES

PANEL A: ATTEND ACADEMIC



PANEL B: Attend Vocational

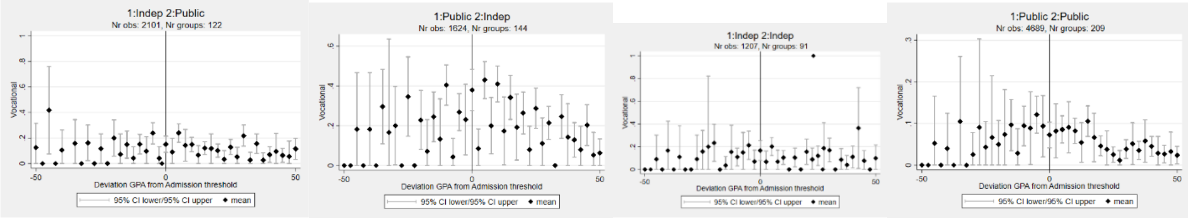
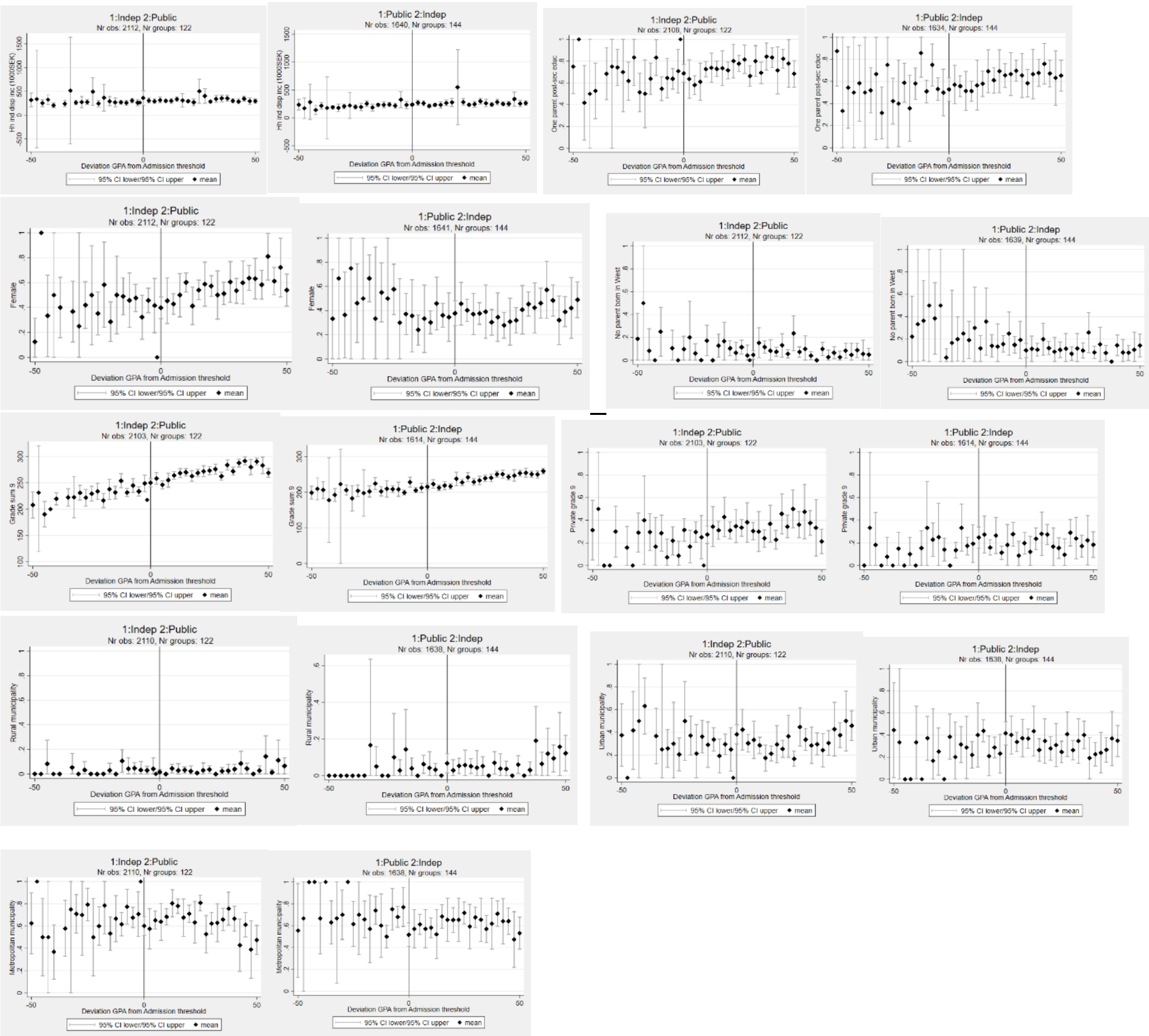


Figure B7. RD-figures for covariates for the two sets of mixed-preference samples; Independent/Public, and Public/Independent



References Appendix B

Imbens. G. W. and D. B. Rubin. 2015. *Causal Inference for Statistics. Social and Biomedical Sciences.* Cambridge University Press.

Appendix C: Supplementary Results

C.1 CEM/VAM – adding covariates

In the tables below we add control variables in a stepwise manner. All regressions are run on our preferred sample, after preference restrictions and CEM on school region has been applied. Column 7 in Table C1 thus corresponds to Column 2 in Table 5A.

TABLE C1. STEPWISE ADDING OF CONTROLS – A. GRADUATION AND GRADES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Switch ind./public	0.0272***	0.0291***	0.0270***	0.0239***	0.0236***	0.0236***	0.0231***	0.0165*
Standard error	(0.0045)	(0.0051)	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0064)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0105]
Observations	70623	70623	70623	70623	70623	70623	70623	70623
Pctile GPA12	3.9522***	2.9007**	4.3683***	4.3927***	4.4080***	4.4510***	4.4906***	4.2250***
Standard error	(1.0900)	(1.0623)	(0.3874)	(0.3800)	(0.3115)	(0.3082)	(0.3080)	(0.3862)
P-value	[0.0003]	[0.0064]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	61898	61898	61898	61898	61898	61898	61898	61898
Graduate on time	0.0151*	0.0161*	0.0261***	0.0268***	0.0278***	0.0284***	0.0288***	0.0359***
Standard error	(0.0064)	(0.0067)	(0.0047)	(0.0045)	(0.0042)	(0.0041)	(0.0041)	(0.0059)
P-value	[0.0189]	[0.0155]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	72220	72220	72220	72220	72220	72220	72220	72220
7th term	-0.0103***	-0.0104***	-0.0144***	-0.0139***	-0.0149***	-0.0151***	-0.0153***	-0.0162***
Standard error	(0.0034)	(0.0035)	(0.0031)	(0.0030)	(0.0029)	(0.0028)	(0.0028)	(0.0043)
P-value	[0.0027]	[0.0031]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0001]
Observations	72220	72220	72220	72220	72220	72220	72220	72220
School municipality	NO	YES	YES	YES	YES	YES	YES	YES
Prior Achievement	NO	NO	YES	YES	YES	YES	YES	YES
9th grade school	NO	NO	NO	YES	YES	YES	YES	YES
Track FE	NO	NO	NO	NO	YES	YES	YES	YES
Additional controls	NO	NO	NO	NO	NO	YES	YES	YES
Admitted to 1st	NO	NO	NO	NO	NO	NO	YES	YES
Independent rank 1	NO	NO	NO	NO	NO	NO	NO	YES

Note: *Prior Achievement* as controlled for by a cubic form of 9th grade sum and quintile dummy variables, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade. The following covariates are included in *additional controls*: log household income per member and income decile dummies, and dummies indicating the following: gender, born in western country, born in non-western country, one parent post-secondary education, both parents born in Sweden, only one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. ***p<0.005, ** p<0.01, * p<0.05

TABLE C2. STEPWISE ADDING OF CONTROLS – STANDARDIZED TESTS RESULTS - MATHEMATICS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High test grade	0.0052	0.0021	0.0028	0.0029	0.0034	0.0033	0.0034	0.0031
Standard error	(0.0039)	(0.0037)	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0025)
P-value	[0.1912]	[0.5750]	[0.1719]	[0.1567]	[0.0859]	[0.0908]	[0.0878]	[0.2137]
Observations	48106	48106	48106	48106	48106	48106	48106	48106
Pass test grade	-0.0117	-0.0176*	-0.0052	-0.0028	-0.0024	-0.0029	-0.0028	-0.0019
Standard error	(0.0094)	(0.0088)	(0.0057)	(0.0056)	(0.0055)	(0.0054)	(0.0054)	(0.0071)
P-value	[0.2150]	[0.0464]	[0.3640]	[0.6121]	[0.6592]	[0.5885]	[0.6045]	[0.7836]
Observations	48106	48106	48106	48106	48106	48106	48106	48106
Test grade>Course grade	0.0043*	0.0044*	0.0042*	0.0044**	0.0040*	0.0039*	0.0039*	0.0061**
Standard error	(0.0018)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0023)
P-value	[0.0149]	[0.0102]	[0.0149]	[0.0094]	[0.0204]	[0.0224]	[0.0228]	[0.0075]
Observations	46244	46244	46244	46244	46244	46244	46244	46244
Test grade<Course grade	0.0376***	0.0424***	0.0466***	0.0454***	0.0454***	0.0460***	0.0462***	0.0510***
Standard error	(0.0092)	(0.0087)	(0.0090)	(0.0088)	(0.0088)	(0.0088)	(0.0088)	(0.0099)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	46244	46244	46244	46244	46244	46244	46244	46244
School municipality	NO	YES	YES	YES	YES	YES	YES	YES
Prior Achievement	NO	NO	YES	YES	YES	YES	YES	YES
9th grade school	NO	NO	NO	YES	YES	YES	YES	YES
Track FE	NO	NO	NO	NO	YES	YES	YES	YES
Additional controls	NO	NO	NO	NO	NO	YES	YES	YES
Admitted to 1st	NO	NO	NO	NO	NO	NO	YES	YES
Independent rank 1	NO	NO	NO	NO	NO	NO	NO	YES

Note: *Prior Achievement* as controlled for by a cubic form of 9th grade sum and quintile dummy variables, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade. The following covariates are included in *additional controls*: log household income per member and income decile dummies, and dummies indicating the following: gender, born in western country, born in non-western country, one parent post-secondary education, both parents born in Sweden, only one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. ***p<0.005, ** p<0.01, * p<0.05

TABLE C3. STEPWISE ADDING OF CONTROLS – STANDARDIZED TESTS RESULTS - ENGLISH

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High test grade	0.0161	0.0060	0.0132***	0.0153***	0.0162***	0.0155***	0.0155***	0.0136**
Standard error	(0.0096)	(0.0091)	(0.0041)	(0.0039)	(0.0040)	(0.0040)	(0.0040)	(0.0050)
P-value	[0.0944]	[0.5087]	[0.0014]	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0068]
Observations	50202	50202	50202	50202	50202	50202	50202	50202
Pass test grade	-0.0008	-0.0018	0.0015	0.0021	0.0026	0.0024	0.0023	0.0020
Standard error	(0.0022)	(0.0021)	(0.0017)	(0.0017)	(0.0016)	(0.0016)	(0.0016)	(0.0025)
P-value	[0.7105]	[0.3883]	[0.3819]	[0.2270]	[0.1101]	[0.1466]	[0.1528]	[0.4243]
Observations	50202	50202	50202	50202	50202	50202	50202	50202
Test grade>Course grade	0.0020	0.0017	-0.0025	-0.0024	-0.0027	-0.0035	-0.0036	-0.0027
Standard error	(0.0049)	(0.0049)	(0.0044)	(0.0044)	(0.0044)	(0.0044)	(0.0044)	(0.0057)
P-value	[0.6782]	[0.7324]	[0.5684]	[0.5850]	[0.5425]	[0.4189]	[0.4134]	[0.6409]
Observations	48159	48159	48159	48159	48159	48159	48159	48159
Test grade<Course grade	0.0334***	0.0357***	0.0384***	0.0397***	0.0389***	0.0402***	0.0401***	0.0413***
Standard error	(0.0061)	(0.0059)	(0.0060)	(0.0059)	(0.0059)	(0.0058)	(0.0058)	(0.0068)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	48159	48159	48159	48159	48159	48159	48159	48159
School municipality	NO	YES	YES	YES	YES	YES	YES	YES
Prior Achievement	NO	NO	YES	YES	YES	YES	YES	YES
9th grade school	NO	NO	NO	YES	YES	YES	YES	YES
Track FE	NO	NO	NO	NO	YES	YES	YES	YES
Additional controls	NO	NO	NO	NO	NO	YES	YES	YES
Admitted to 1st	NO	NO	NO	NO	NO	NO	YES	YES
Independent rank 1	NO	NO	NO	NO	NO	NO	NO	YES

Note: *Prior Achievement* as controlled for by a cubic form of 9th grade sum and quintile dummy variables, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade. The following covariates are included in *additional controls*: log household income per member and income decile dummies, and dummies indicating the following: gender, born in western country, born in non-western country, one parent post-secondary education, both parents born in Sweden, only one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. ***p<0.005, ** p<0.01, * p<0.05

TABLE C4. STEPWISE ADDING OF CONTROLS – STANDARDIZED TEST RESULTS - SWEDISH

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High test grade	0.0234***	0.0182*	0.0195***	0.0192***	0.0194***	0.0204***	0.0204***	0.0196***
Standard error	(0.0077)	(0.0074)	(0.0047)	(0.0046)	(0.0046)	(0.0046)	(0.0046)	(0.0052)
P-value	[0.0024]	[0.0142]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0002]
Observations	52515	52515	52515	52515	52515	52515	52515	52515
Pass test grade	0.0054	0.0020	0.0065*	0.0063*	0.0059*	0.0064*	0.0065*	0.0050
Standard error	(0.0039)	(0.0039)	(0.0026)	(0.0026)	(0.0026)	(0.0027)	(0.0027)	(0.0040)
P-value	[0.1679]	[0.6161]	[0.0147]	[0.0164]	[0.0262]	[0.0162]	[0.0151]	[0.2154]
Observations	52515	52515	52515	52515	52515	52515	52515	52515
Test grade>Course grade	-0.0016	-0.0012	-0.0021	-0.0026	-0.0031	-0.0035	-0.0035	-0.0024
Standard error	(0.0042)	(0.0043)	(0.0039)	(0.0040)	(0.0040)	(0.0040)	(0.0040)	(0.0057)
P-value	[0.7072]	[0.7732]	[0.5885]	[0.5159]	[0.4411]	[0.3741]	[0.3735]	[0.6703]
Observations	48724	48724	48724	48724	48724	48724	48724	48724
Test grade<Course grade	0.0230***	0.0213***	0.0223***	0.0226***	0.0234***	0.0239***	0.0239***	0.0297***
Standard error	(0.0067)	(0.0067)	(0.0068)	(0.0069)	(0.0069)	(0.0069)	(0.0069)	(0.0082)
P-value	[0.0007]	[0.0015]	[0.0010]	[0.0011]	[0.0007]	[0.0006]	[0.0005]	[0.0003]
Observations	48724	48724	48724	48724	48724	48724	48724	48724
School municipality	NO	YES	YES	YES	YES	YES	YES	YES
Prior Achievement	NO	NO	YES	YES	YES	YES	YES	YES
9th grade school	NO	NO	NO	YES	YES	YES	YES	YES
Track FE	NO	NO	NO	NO	YES	YES	YES	YES
Additional controls	NO	NO	NO	NO	NO	YES	YES	YES
Admitted to 1st	NO	NO	NO	NO	NO	NO	YES	YES
Independent rank 1	NO	NO	NO	NO	NO	NO	NO	YES

Note: *Prior Achievement* as controlled for by a cubic form of 9th grade sum and quintile dummy variables, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade. The following covariates are included in *additional controls*: log household income per member and income decile dummies, and dummies indicating the following: gender, born in western country, born in non-western country, one parent post-secondary education, both parents born in Sweden, only one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. ***p<0.005, ** p<0.01, * p<0.05

TABLE C5. STEPWISE ADDING OF CONTROLS – POST GRADUATION OUTCOMES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Study	0.0162	0.0026	0.0127	0.0177*	0.0190***	0.0197***	0.0199***	0.0117
Standard error	(0.0154)	(0.0148)	(0.0077)	(0.0070)	(0.0053)	(0.0050)	(0.0050)	(0.0071)
P-value	[0.2931]	[0.8591]	[0.0982]	[0.0120]	[0.0004]	[0.0001]	[0.0001]	[0.0976]
Observations	55430	55430	55430	55430	55430	55430	55430	55430
Study no-prep	0.0177	0.0049	0.0173*	0.0222***	0.0239***	0.0241***	0.0244***	0.0143*
Standard error	(0.0168)	(0.0163)	(0.0068)	(0.0063)	(0.0048)	(0.0046)	(0.0046)	(0.0065)
P-value	[0.2937]	[0.7633]	[0.0111]	[0.0004]	[0.0000]	[0.0000]	[0.0000]	[0.0286]
Observations	55430	55430	55430	55430	55430	55430	55430	55430
UC≥15	0.0078	0.0027	0.0106*	0.0125***	0.0130***	0.0138***	0.0142***	0.0141***
Standard error	(0.0103)	(0.0099)	(0.0042)	(0.0041)	(0.0035)	(0.0034)	(0.0034)	(0.0045)
P-value	[0.4462]	[0.7840]	[0.0129]	[0.0025]	[0.0002]	[0.0001]	[0.0000]	[0.0018]
Observations	55430	55430	55430	55430	55430	55430	55430	55430
Work≥50%	-0.0207*	-0.0118	-0.0136	-0.0175*	-0.0177***	-0.0175***	-0.0170***	-0.0138*
Standard error	(0.0105)	(0.0098)	(0.0081)	(0.0078)	(0.0053)	(0.0049)	(0.0049)	(0.0064)
P-value	[0.0493]	[0.2304]	[0.0958]	[0.0251]	[0.0009]	[0.0004]	[0.0005]	[0.0312]
Observations	55386	55386	55386	55386	55386	55386	55386	55386
School municipality	NO	YES	YES	YES	YES	YES	YES	YES
Prior Achievement	NO	NO	YES	YES	YES	YES	YES	YES
9th grade school	NO	NO	NO	YES	YES	YES	YES	YES
Track FE	NO	NO	NO	NO	YES	YES	YES	YES
Additional controls	NO	NO	NO	NO	NO	YES	YES	YES
Admitted to 1st	NO	NO	NO	NO	NO	NO	YES	YES
Independent rank 1	NO	NO	NO	NO	NO	NO	NO	YES

Note: All outcomes are measured one year after graduation. *Prior Achievement* as controlled for by a cubic form of 9th grade sum and quintile dummy variables, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade. The following covariates are included in *additional controls*: log household income per member and income decile dummies, and dummies indicating the following: gender, born in western country, born in non-western country, one parent post-secondary education, both parents born in Sweden, only one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. ***p<0.005, ** p<0.01, * p<0.05

C.2 CEM/VAM – excluding JB-students

Tables C6–8 show, in the columns denoted "No JB" the results from our preferred CEM/VAM-specification, after excluding all students attending upper secondary school in track×municipality×year-combinations where schools belonging to the John Bauer-group (JB) were present. In order to facilitate the comparison, the results for the full sample (including the JB-cases) are repeated in the columns denoted "All".

TABLE C6. GRADUATION AND GRADES, AND POST-GRADUATION OUTCOMES

Graduation and grades	All	No JB	Post-graduation	All	No JB
	(1)	(2)		(3)	(4)
Switch independent/public	0.0231***	0.0182***	Study	0.0199***	0.0233***
Standard error	(0.0049)	(0.0046)	Standard error	(0.0050)	(0.0062)
P-value	[0.0000]	[0.0001]	P-value	[0.0001]	[0.0002]
Observations	70623	53222	Observations	55430	37705
Pctile GPA12	4.4906***	4.5446***	Study no-prep	0.0244***	0.0272***
Standard error	(0.3080)	(0.3278)	Standard error	(0.0046)	(0.0058)
P-value	[0.0000]	[0.0000]	P-value	[0.0000]	[0.0000]
Observations	61898	46710	Observations	55430	37705
Graduate on time	0.0288***	0.0279***	Uni cred≥15	0.0142***	0.0129***
Standard error	(0.0041)	(0.0049)	Standard error	(0.0034)	(0.0041)
P-value	[0.0000]	[0.0000]	P-value	[0.0000]	[0.0017]
Observations	72220	54424	Observations	55430	37705
7^{th} term	-0.0153***	-0.0171***	Work ≥50%	-0.0170***	-0.0168***
Standard error	(0.0028)	(0.0034)	Standard error	(0.0049)	(0.0059)
P-value	[0.0000]	[0.0000]	P-value	[0.0005]	[0.0044]
Observations	72220	54424	Observations	55386	37673

All regressions above include the following covariates: upper secondary school municipality dummies, prior achievement as controlled for by a cubic form of GPS9 and GPS9 quintile dummies, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9^{th} grade, 9^{th} grade school dummies, track dummies, log household income, income decile dummies; and dummies indicating the following: gender, born in western country (excl. Sweden), born in non-western country, one parent post-secondary education, both parents born in Sweden, one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, cohort dummies, and a dummy indicating admission to first ranked school. Standard errors are clustered on upper secondary school. *** p<0.005, ** p<0.01, * p<0.05

TABLE C7. TEST OUTCOMES

	MATH		SWEDISH		ENGLISH	
	All	No JB	All	No JB	All	No JB
	(1)	(2)	(3)	(4)	(5)	(6)
High test grade	0.0034	0.0033	0.0204***	0.0232***	0.0155***	0.0141***
Standard error	(0.0020)	(0.0024)	(0.0046)	(0.0055)	(0.0040)	(0.0045)
P-value	[0.0878]	[0.1644]	[0.0000]	[0.0000]	[0.0001]	[0.0018]
Observations	48106	36733	52515	39921	50202	38085
Pass test grade	-0.0028	0.0080	0.0065*	0.0070*	0.0023	0.0028
Standard error	(0.0054)	(0.0057)	(0.0027)	(0.0029)	(0.0016)	(0.0019)
P-value	[0.6045]	[0.1619]	[0.0151]	[0.0157]	[0.1528]	[0.1344]
Observations	48106	36733	52515	39921	50202	38085
Test grade>Course grade	0.0039*	0.0061***	-0.0035	-0.0033	-0.0036	0.0005
Standard error	(0.0017)	(0.0019)	(0.0040)	(0.0044)	(0.0044)	(0.0048)
P-value	[0.0228]	[0.0012]	[0.3735]	[0.4522]	[0.4134]	[0.9174]
Observations	46244	35261	48724	36954	48159	36481
Test grade<Course grade	0.0462***	0.0265***	0.0239***	0.0212**	0.0401***	0.0380***
Standard error	(0.0088)	(0.0092)	(0.0069)	(0.0077)	(0.0058)	(0.0066)
P-value	[0.0000]	[0.0041]	[0.0005]	[0.0058]	[0.0000]	[0.0000]
Observations	46244	35261	48724	36954	48159	36481

Note: All regressions above include the following covariates: upper secondary school municipality dummies, prior achievement as controlled for by a cubic form of GPS9 and GPS9 quintile dummies, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9^{th} grade, 9^{th} grade school dummies, track dummies, log household income, income decile dummies; and dummies indicating the following: gender, born in western country (excl. Sweden), born in non-western country, one parent post-secondary education, both parents born in Sweden, one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, cohort dummies, and a dummy indicating admission to first ranked school. Standard errors are clustered on upper secondary school. *** p<0.005, ** p<0.01, * p<0.05

C.3 CEM/VAM – post-graduation outcomes in graduation year

TABLE C8. POST-GRADUATION OUTCOMES IN SAME YEAR AS GRADUATION (T+3)

	(1)	(2)	(3)
Study t+3	0.0060	0.0081	0.0113***
Standard error	(0.0056)	(0.0044)	(0.0037)
P-value	[0.2811]	[0.0639]	[0.0021]
Observations	29440	72220	294580
Study no-prep t+3	0.0144**	0.0160***	0.0187***
Standard error	(0.0053)	(0.0042)	(0.0035)
P-value	[0.0065]	[0.0001]	[0.0000]
Observations	29440	72220	294580
UC≥15 t+3	0.0021	0.0040	0.0061***
Standard error	(0.0033)	(0.0024)	(0.0020)
P-value	[0.5170]	[0.1025]	[0.0025]
Observations	29440	72220	294580
Work≥25% t+3	-0.0090	-0.0099*	-0.0163***
Standard error	(0.0058)	(0.0045)	(0.0039)
P-value	[0.1255]	[0.0263]	[0.0000]
Observations	29433	72188	294441
Preference restriction	YES	YES	NO
CEM on 9th grade school	YES	NO	NO
CEM on region	NO	YES	YES

Note: All outcomes are measured <u>the same year as graduation.</u> All regressions above include the following covariates: upper secondary school municipality dummies, prior achievement as controlled for by a cubic form of GPS9 and GPS9 quintile dummies, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade, 9th grade school dummies, track dummies, log household income, income decile dummies; and dummies indicating the following: gender, born in western country (excl. Sweden), born in non-western country, one parent post-secondary education, both parents born in Sweden, one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, and cohort. Columns 1 and 2 also include a dummy variable indicating whether the student was admitted to the first ranked choice. Standard errors are clustered on upper secondary school.

*** p<0.005, ** p<0.01, * p<0.05

C.4 CEM/VAM – full sample with preference controls

In this section we present results when using the full observational sample and controlling for preferences by including two dummy variables, one that takes on value 1 if the student is accepted to its first choice, and one dummy that takes on value one if the student ranked an independent school as their first choice.

TABLE C9. GRADUATION AND POST-GRADUATION OUTCOMES - FULL SAMPLE

Graduation and grades		Post-graduation	
	(1)		(2)
Switch independent/public	0.0380***	Study	0.0060
Standard error	(0.0063)	Standard error	(0.0060)
P-value	[0.0000]	P-value	[0.3136]
Observations	288762	Observations	230160
Pctile GPA12	4.0575***	Study no-prep	0.0054
Standard error	(0.3491)	Standard error	(0.0054)
P-value	[0.0000]	P-value	[0.3178]
Observations	254937	Observations	230160
Graduate on time	0.0246***	Uni cred≥15	0.0074
Standard error	(0.0054)	Standard error	(0.0039)
P-value	[0.0000]	P-value	[0.0584]
Observations	294580	Observations	230160
7^{th} term	-0.0093*	Work ≥50%	-0.0197***
Standard error	(0.0039)	Standard error	(0.0058)
P-value	[0.0173]	P-value	[0.0007]
Observations	294580	Observations	229988

Note: CEM is performed on school county. All post-graduation outcomes are measured one year after graduation. All regressions above include the following covariates: upper secondary school municipality dummies, prior achievement as controlled for by a cubic form of GPS9 and GPS9 quintile dummies, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9^{th} grade, 9^{th} grade school dummies, track dummies, log household income, income decile dummies; and dummies indicating the following: gender, born in western country (excl. Sweden), born in non-western country, one parent post-secondary education, both parents born in Sweden, one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, cohort, admission to first ranked school, and ranking independent school as first choice. Standard errors are clustered on upper secondary school. *** p<0.005, ** p<0.01, * p<0.05

TABLE C10. STANDARDIZED TEST OUTCOMES - FULL SAMPLE

	Mathematics	Swedish	English
High test grade	0.0041	0.0230***	0.0142***
Standard error	(0.0021)	(0.0046)	(0.0043)
P-value	[0.0541]	[0.0000]	[0.0010]
Observations	193412	222275	201527
Pass test grade	-0.0001	0.0054	0.0021
Standard error	(0.0065)	(0.0034)	(0.0022)
P-value	[0.9887]	[0.1086]	[0.3495]
Observations	193412	222275	201527
Test grade>Course grade	0.0057**	0.0009	0.0039
Standard error	(0.0021)	(0.0048)	(0.0050)
P-value	[0.0054]	[0.8430]	[0.4399]
Observations	185415	202921	193922
Test grade<Course grade	0.0622***	0.0321***	0.0423***
Standard error	(0.0097)	(0.0071)	(0.0062)
P-value	[0.0000]	[0.0000]	[0.0000]
Observations	185415	202921	193922

Note: CEM is performed on school region. Regressions are performed on individual means within each subject. All regressions above include the following covariates: upper secondary school municipality dummies, prior achievement as controlled for by a cubic form of GPS9 and GPS9 quintile dummies, as well as 6 dummies representing pass/high test result in Math/Swe/Eng in 9th grade, 9th grade school dummies, track dummies, log household income, income decile dummies; and dummies indicating the following: gender, born in western country (excl. Sweden), born in non-western country, one parent post-secondary education, both parents born in Sweden, one parent born in Sweden, both parents born in non-western country, negative or zero household income, one parent is self-employed, one parent is unemployed, cohort, admission to first ranked school, and ranking independent school as first choice. Standard errors are clustered on upper secondary school municipality. *** p<0.005, ** p<0.01, * p<0.05

C.5 RDD/DID - post-graduation outcomes in graduation year

TABLE C11. POST-GRADUATION OUTCOMES IN T+3

Treated sample (T=1)	*Independent/Public*		*Public/Independent*	
Control sample (T=0)	*Indep/Indep*	*Public/Public*	*Indep/Indep*	*Public/Public*
	(1)	(2)	(3)	(4)
Attend Private	0.4911***	0.5015***	-0.4272***	-0.4269***
Standard error	(0.0804)	(0.0719)	(0.0667)	(0.0650)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	378	733	334	689
Number of groups	85	171	103	156
Study t+3	-0.0576	0.0251	-0.0825	-0.0571
Standard error	(0.0970)	(0.0745)	(0.0956)	(0.0750)
P-value	[0.5545]	[0.7367]	[0.3905]	[0.4482]
Observations	379	737	335	693
Number of groups	85	171	103	156
Study t+3 non-prep	-0.0716	0.0345	-0.1020	-0.0587
Standard error	(0.0868)	(0.0676)	(0.0865)	(0.0697)
P-value	[0.4119]	[0.6106]	[0.2410]	[0.4009]
Observations	379	737	335	693
Number of groups	85	171	103	156
Uni cred≥15 t+3	-0.0843	-0.0198	-0.0439	-0.0121
Standard error	(0.0583)	(0.0496)	(0.0488)	(0.0448)
P-value	[0.1518]	[0.6901]	[0.3701]	[0.7877]
Observations	379	737	335	693
Number of groups	85	171	103	156
Work ≥25% t+3	0.1243	-0.0286	0.1774	0.0226
Standard error	(0.0886)	(0.0655)	(0.1061)	(0.0769)
P-value	[0.1641]	[0.6631]	[0.0976]	[0.7689]
Observations	379	737	334	692
Number of groups	85	171	103	156

Regression details: Bandwidth:5, FE: Admission group of top choice, Standard errors clustered on Admission group of top choice, Student covariates included, no trend variables included.

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and standard errors are clustered on the same level. The regressions additionally include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D). Specifications with linear trends include the running variable separately and interacted with the dummy variable for being above the admission threshold. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included.

C.6 Alternative RDD/DID-specifications: Increasing the data window

TABLE C12. GRADUATION AND GRADES

Treated sample (T=1)	Independent/Public				Public/Independent			
Control sample (T=0)	Independent/Independent		Public/Public		Independent/Independent		Public/Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attend Private	0.5967***	0.6854***	0.6196***	0.6991***	-0.5575***	-0.6362***	-0.5472***	-0.6278***
Standard error	(0.0428)	(0.0324)	(0.0373)	(0.0295)	(0.0385)	(0.0318)	(0.0369)	(0.0302)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	1042	1772	4297	7590	1510	2505	4105	7114
Groups	169	169	777	777	325	325	664	664
Switch independent/public	-0.0320	-0.0200	-0.0275	-0.0273	-0.0259	-0.0353	-0.0167	-0.0191
Standard error	(0.0341)	(0.0251)	(0.0218)	(0.0159)	(0.0314)	(0.0224)	(0.0242)	(0.0183)
P-value	[0.3491]	[0.4261]	[0.2065]	[0.0875]	[0.4095]	[0.1160]	[0.4911]	[0.2986]
Observations	1017	1731	4209	7447	1483	2461	4023	6985
Groups	169	169	777	777	324	324	663	663
Pctile GPA12	2.5907	3.4239	3.7023*	3.1545*	0.4660	-1.0900	-0.7225	-3.3388*
Standard error	(2.7460)	(2.3737)	(1.8205)	(1.4209)	(2.5081)	(2.0062)	(2.0700)	(1.6279)
P-value	[0.3468]	[0.1511]	[0.0423]	[0.0267]	[0.8527]	[0.5873]	[0.7272]	[0.0407]
Observations	924	1579	3642	6516	1306	2181	3452	6069
Groups	168	168	760	772	323	323	650	658
Graduate on time	0.0446	0.0213	-0.0051	-0.0320	0.0836*	0.0317	-0.0148	-0.0514
Standard error	(0.0485)	(0.0350)	(0.0349)	(0.0258)	(0.0416)	(0.0325)	(0.0348)	(0.0291)
P-value	[0.3596]	[0.5439]	[0.8831]	[0.2152]	[0.0453]	[0.3296]	[0.6717]	[0.0780]
Observations	1046	1780	4321	7625	1521	2524	4134	7156
Groups	169	169	777	777	325	325	664	664
7^{th} term	-0.0487	-0.0358	0.0134	0.0039	-0.0575	-0.0187	0.0058	0.0398
Standard error	(0.0338)	(0.0249)	(0.0271)	(0.0221)	(0.0310)	(0.0231)	(0.0263)	(0.0205)
P-value	[0.1517]	[0.1514]	[0.6220]	[0.8608]	[0.0649]	[0.4184]	[0.8258]	[0.0527]
Observations	1046	1780	4321	7625	1521	2524	4134	7156
Groups	169	169	777	777	325	325	664	664
Bandwidth	10	20	10	20	10	20	10	20
Trend	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and fixed effects for the track choices of the two top choices. The regressions include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D). Specifications with linear trends include the running variable separately and interacted with the dummy variable for being above the admission threshold. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. The regression sample is restricted to students with the same track preference for the top and second preference. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

TABLE C13. POST-GRADUATION OUTCOMES

Treated sample (T=1)	Independent/Public				Public/Independent			
Control sample (T=0)	Independent/Independent		Public/Public		Independent/Independent		Public/Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attend Private	0.5967***	0.6854***	0.6196***	0.6991***	-0.5575***	-0.6362***	-0.5472***	-0.6278***
Standard error	(0.0428)	(0.0324)	(0.0373)	(0.0295)	(0.0385)	(0.0318)	(0.0369)	(0.0302)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	1042	1772	4297	7590	1510	2505	4105	7114
Groups	169	169	777	777	325	325	664	664
Study	0.2288**	0.1685*	0.1368***	0.0717	0.0672	0.0606	0.0713	0.0647
Standard error	(0.0859)	(0.0693)	(0.0448)	(0.0377)	(0.0737)	(0.0581)	(0.0518)	(0.0411)
P-value	[0.0090]	[0.0168]	[0.0024]	[0.0581]	[0.3630]	[0.2984]	[0.1697]	[0.1165]
Observations	550	932	2690	4624	828	1365	2568	4316
Groups	99	99	501	501	192	192	431	431
Study no-prep	0.0885	0.0748	0.0932*	0.0465	-0.0368	-0.0156	-0.0048	0.0017
Standard error	(0.0787)	(0.0642)	(0.0435)	(0.0335)	(0.0622)	(0.0510)	(0.0435)	(0.0357)
P-value	[0.2637]	[0.2468]	[0.0326]	[0.1664]	[0.5545]	[0.7606]	[0.9117]	[0.9621]
Observations	550	932	2690	4624	828	1365	2568	4316
Groups	99	99	501	501	192	192	431	431
Uni cred≥15	0.0043	0.0258	0.0189	0.0136	-0.0178	0.0104	0.0102	0.0145
Standard error	(0.0578)	(0.0441)	(0.0330)	(0.0258)	(0.0404)	(0.0370)	(0.0336)	(0.0294)
P-value	[0.9406]	[0.5590]	[0.5671]	[0.5971]	[0.6602]	[0.7790]	[0.7607]	[0.6224]
Observations	550	932	2690	4624	828	1365	2568	4316
Groups	99	99	501	501	192	192	431	431
Work ≥50%	-0.0651	-0.0068	-0.0123	-0.0135	-0.0098	-0.0048	0.0191	-0.0070
Standard error	(0.0800)	(0.0589)	(0.0454)	(0.0380)	(0.0701)	(0.0543)	(0.0506)	(0.0432)
P-value	[0.4179]	[0.9084]	[0.7868]	[0.7231]	[0.8887]	[0.9290]	[0.7066]	[0.8718]
Observations	550	932	2690	4623	826	1362	2567	4313
Groups	99	99	501	501	192	192	431	431
Bandwidth	10	20	10	20	10	20	10	20
Trend	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and fixed effects for the track choices of the two top choices. The regressions include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D). Specifications with linear trends include the running variable separately and interacted with the dummy variable for being above the admission threshold. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. The regression sample is restricted to students with the same track preference for the top and second preference. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

TABLE C14. STANDARDIZED TEST OUTCOMES: MATH

T=1;T=0-samples:	Indep/Public;Indep/Indep		Indep/Public;Public/Public		Public/Indep;Indep/Indep		Public/Indep;Public/Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attend Private	0.5960***	0.6826***	0.6031***	0.6759***	-0.5548***	-0.6545***	-0.5613***	-0.6546***
Standard error	(0.0483)	(0.0379)	(0.0449)	(0.0356)	(0.0457)	(0.0352)	(0.0423)	(0.0321)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	833	1438	3017	5487	1161	1949	2857	5095
Groups	154	157	643	669	295	301	549	570
High test grade	0.0457	0.0497	0.0396*	0.0472***	0.0014	-0.0135	0.0074	-0.0062
Standard error	(0.0308)	(0.0288)	(0.0183)	(0.0161)	(0.0174)	(0.0147)	(0.0096)	(0.0086)
P-value	[0.1391]	[0.0867]	[0.0313]	[0.0034]	[0.9381]	[0.3596]	[0.4401]	[0.4719]
Observations	835	1442	3030	5505	1168	1961	2874	5118
Groups	154	157	643	669	295	301	549	570
Pass test grade	0.0832	0.0183	0.0419	-0.0014	-0.0027	0.0319	-0.0462	0.0066
Standard error	(0.0535)	(0.0405)	(0.0350)	(0.0284)	(0.0466)	(0.0332)	(0.0396)	(0.0289)
P-value	[0.1220]	[0.6510]	[0.2313]	[0.9596]	[0.9538]	[0.3377]	[0.2437]	[0.8199]
Observations	835	1442	3030	5505	1168	1961	2874	5118
Groups	154	157	643	669	295	301	549	570
Test grade>Course grade	0.0587*	0.0332*	0.0199	0.0251***	0.0090	0.0027	-0.0159	-0.0103
Standard error	(0.0227)	(0.0141)	(0.0127)	(0.0087)	(0.0187)	(0.0158)	(0.0149)	(0.0110)
P-value	[0.0105]	[0.0197]	[0.1181]	[0.0041]	[0.6293]	[0.8663]	[0.2855]	[0.3466]
Observations	826	1427	2895	5253	1131	1907	2725	4845
Groups	151	154	637	667	293	298	545	569
Test grade<Course grade	-0.0330	0.0304	0.0390	0.0267	-0.0445	-0.0351	-0.0006	-0.0420
Standard error	(0.0564)	(0.0416)	(0.0340)	(0.0277)	(0.0500)	(0.0366)	(0.0419)	(0.0306)
P-value	[0.5601]	[0.4664]	[0.2508]	[0.3358]	[0.3751]	[0.3389]	[0.9876]	[0.1706]
Observations	826	1427	2895	5253	1131	1907	2725	4845
Groups	151	154	637	667	293	298	545	569
Bandwidth	10	20	10	20	10	20	10	20
Trend	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and fixed effects for the track choices of the two top choices. The regressions include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D). Specifications with linear trends include the running variable separately and interacted with the dummy variable for being above the admission threshold. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. The regression sample is restricted to students with the same track preference for the top and second preference. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

TABLE C15. STANDARDIZED TEST OUTCOMES: ENGLISH

T=1;T=0-samples:	Indep/Public;Indep/Indep		Indep/Public;Public/Public		Public/Indep;Indep/Indep		Public/Indep;Public/Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attend Private	0.5820***	0.6683***	0.5904***	0.6697***	-0.5428***	-0.6377***	-0.5679***	-0.6470***
Standard error	(0.0528)	(0.0408)	(0.0458)	(0.0362)	(0.0473)	(0.0365)	(0.0420)	(0.0333)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	869	1482	3188	5744	1201	2002	3003	5324
Groups	155	160	675	699	296	302	572	590
High test grade	0.0612	0.0459	0.0635*	0.0604*	-0.0362	-0.0418	-0.0045	-0.0167
Standard error	(0.0457)	(0.0404)	(0.0270)	(0.0265)	(0.0332)	(0.0282)	(0.0237)	(0.0204)
P-value	[0.1821]	[0.2574]	[0.0191]	[0.0229]	[0.2767]	[0.1391]	[0.8497]	[0.4142]
Observations	872	1486	3201	5761	1209	2014	3020	5346
Groups	155	160	675	699	296	302	572	590
Pass test grade	0.0169	0.0022	0.0059	-0.0018	0.0157	-0.0055	0.0156	-0.0057
Standard error	(0.0129)	(0.0104)	(0.0128)	(0.0086)	(0.0171)	(0.0125)	(0.0161)	(0.0131)
P-value	[0.1934]	[0.8358]	[0.6447]	[0.8379]	[0.3584]	[0.6610]	[0.3326]	[0.6633]
Observations	872	1486	3201	5761	1209	2014	3020	5346
Groups	155	160	675	699	296	302	572	590
Test grade>Course grade	-0.0459	-0.0674	0.0061	0.0062	-0.1021*	-0.0703*	-0.0464	-0.0290
Standard error	(0.0396)	(0.0407)	(0.0298)	(0.0253)	(0.0434)	(0.0337)	(0.0334)	(0.0273)
P-value	[0.2481]	[0.0995]	[0.8387]	[0.8062]	[0.0194]	[0.0378]	[0.1648]	[0.2891]
Observations	855	1455	3079	5529	1167	1940	2897	5107
Groups	155	160	671	695	294	299	569	587
Test grade<Course grade	0.0422	0.0713	-0.0261	0.0025	0.0519	0.0377	0.0128	0.0075
Standard error	(0.0496)	(0.0423)	(0.0294)	(0.0248)	(0.0432)	(0.0334)	(0.0353)	(0.0274)
P-value	[0.3955]	[0.0934]	[0.3745]	[0.9187]	[0.2308]	[0.2599]	[0.7178]	[0.7845]
Observations	855	1455	3079	5529	1167	1940	2897	5107
Groups	155	160	671	695	294	299	569	587
Bandwidth	10	20	10	20	10	20	10	20
Trend	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and fixed effects for the track choices of the two top choices. The regressions include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D). Specifications with linear trends include the running variable separately and interacted with the dummy variable for being above the admission threshold. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. The regression sample is restricted to students with the same track preference for the top and second preference. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

TABLE C16. STANDARDIZED TEST OUTCOMES: SWEDISH

T=1;T=0-samples:	Indep/Public;Indep/Indep		Indep/Public;Public/Public		Public/Indep;Indep/Indep		Public/Indep;Public/Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attend Private	0.5643***	0.6512***	0.6001***	0.6789***	-0.5894***	-0.6583***	-0.5815***	-0.6625***
Standard error	(0.0503)	(0.0388)	(0.0447)	(0.0341)	(0.0463)	(0.0356)	(0.0422)	(0.0322)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	853	1463	3275	5924	1149	1946	3058	5454
Groups	159	162	740	750	300	309	631	641
High test grade	0.0099	0.0353	0.0132	0.0483**	-0.0356	-0.0536*	-0.0247	-0.0233
Standard error	(0.0383)	(0.0260)	(0.0233)	(0.0183)	(0.0330)	(0.0243)	(0.0176)	(0.0142)
P-value	[0.7958]	[0.1768]	[0.5725]	[0.0084]	[0.2817]	[0.0284]	[0.1599]	[0.1016]
Observations	855	1465	3286	5939	1156	1958	3073	5476
Groups	159	162	740	750	301	310	631	641
Pass test grade	0.0081	-0.0179	-0.0195	-0.0443***	0.0068	0.0011	-0.0066	-0.0239
Standard error	(0.0273)	(0.0192)	(0.0172)	(0.0137)	(0.0311)	(0.0209)	(0.0265)	(0.0199)
P-value	[0.7663]	[0.3543]	[0.2568]	[0.0012]	[0.8263]	[0.9566]	[0.8025]	[0.2302]
Observations	855	1465	3286	5939	1156	1958	3073	5476
Groups	159	162	740	750	301	310	631	641
Test grade>Course grade	0.0244	0.0128	-0.0371	-0.0191	-0.0222	0.0164	-0.0760*	-0.0340
Standard error	(0.0538)	(0.0365)	(0.0282)	(0.0210)	(0.0471)	(0.0344)	(0.0349)	(0.0255)
P-value	[0.6514]	[0.7264]	[0.1882]	[0.3624]	[0.6382]	[0.6340]	[0.0296]	[0.1835]
Observations	793	1350	3004	5388	1075	1796	2796	4939
Groups	153	158	709	729	291	300	605	621
Test grade<Course grade	0.0257	-0.0105	-0.0482	-0.0301	0.0623	-0.0337	0.0148	-0.0270
Standard error	(0.0643)	(0.0492)	(0.0374)	(0.0301)	(0.0590)	(0.0441)	(0.0450)	(0.0339)
P-value	[0.6903]	[0.8319]	[0.1979]	[0.3182]	[0.2924]	[0.4461]	[0.7417]	[0.4250]
Observations	793	1350	3004	5388	1075	1796	2796	4939
Groups	153	158	709	729	291	300	605	621
Bandwidth	10	20	10	20	10	20	10	20
Trend	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and fixed effects for the track choices of the two top choices. The regressions include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D). Specifications with linear trends include the running variable separately and interacted with the dummy variable for being above the admission threshold. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. The regression sample is restricted to students with the same track preference for the top and second preference. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

C.7 Alternative RD/DID specifications: Including students with different track preferences for the two top choices

TABLE C17. GRADUATION AND GRADES

Treated sample (T=1)	Independent/Public				Public/Independent			
Control sample (T=0)	Independent/Independent		Public/Public		Independent/Independent		Public/Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attend Private	0.4921***	0.6793***	0.5349***	0.6957***	-0.3963***	-0.6401***	-0.4226***	-0.6282***
Standard error	(0.0672)	(0.0295)	(0.0592)	(0.0297)	(0.0605)	(0.0313)	(0.0599)	(0.0298)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	651	2981	1732	7590	574	2505	1655	7114
Groups	138	239	427	777	157	325	386	664
Switch ind./public	-0.0176	-0.0506**	-0.0089	-0.0273	-0.0353	-0.0372	-0.0524	-0.0205
Standard error	(0.0470)	(0.0192)	(0.0399)	(0.0162)	(0.0507)	(0.0225)	(0.0403)	(0.0183)
P-value	[0.7092]	[0.0090]	[0.8242]	[0.0925]	[0.4883]	[0.0987]	[0.1939]	[0.2632]
Observations	634	2923	1694	7447	560	2461	1620	6985
Groups	138	239	426	777	156	324	384	663
Pctile GPA12	2.7256	4.2649*	0.7529	2.7392*	-3.7196	-1.1913	-8.0479*	-3.3496*
Standard error	(3.7797)	(1.6673)	(3.1142)	(1.3604)	(4.4285)	(1.9421)	(3.8852)	(1.5764)
P-value	[0.4721]	[0.0111]	[0.8091]	[0.0444]	[0.4022]	[0.5401]	[0.0390]	[0.0340]
Observations	568	2628	1466	6516	501	2181	1399	6069
Groups	136	239	416	772	155	323	375	658
Graduate on time	0.0986	0.0652*	0.0118	-0.0334	-0.0575	0.0323	-0.1286*	-0.0517
Standard error	(0.0623)	(0.0284)	(0.0562)	(0.0257)	(0.0693)	(0.0331)	(0.0583)	(0.0292)
P-value	[0.1159]	[0.0225]	[0.8340]	[0.1942]	[0.4081]	[0.3299]	[0.0280]	[0.0772]
Observations	652	2993	1741	7625	575	2524	1664	7156
Groups	138	239	427	777	157	325	386	664
7th term	-0.0795	-0.0669**	0.0016	0.0031	0.0022	-0.0195	0.0552	0.0385
Standard error	(0.0527)	(0.0238)	(0.0440)	(0.0221)	(0.0529)	(0.0235)	(0.0417)	(0.0206)
P-value	[0.1338]	[0.0052]	[0.9713]	[0.8888]	[0.9671]	[0.4085]	[0.1856]	[0.0617]
Observations	652	2993	1741	7625	575	2524	1664	7156
Groups	138	239	427	777	157	325	386	664
Bandwidth	5	20	5	20	5	20	5	20
Trend	No	Linear	No	Linear	No	Linear	No	Linear

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and fixed effects for the track choices of the two top choices. The regressions include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D). Specifications with linear trends include the running variable separately and interacted with the dummy variable for being above the admission threshold. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. The regression sample is *not* restricted to students with the same track preference for the top and second preference, but students with different track preferences above for the two top choices are also included. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

TABLE C18. POST-GRADUATION OUTCOMES

Treated sample (T=1)	Independent/Public				Public/Independent			
Control sample (T=0)	Independent/Independent		Public/Public		Independent/Independent		Public/Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attend Private	0.4921***	0.6793***	0.5349***	0.6957***	-0.3963***	-0.6401***	-0.4226***	-0.6282***
Standard error	(0.0672)	(0.0295)	(0.0592)	(0.0297)	(0.0605)	(0.0313)	(0.0599)	(0.0298)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	651	2981	1732	7590	574	2505	1655	7114
Groups	138	239	427	777	157	325	386	664
Study	0.1856*	0.0460	0.1439*	0.0780*	0.1376	0.0474	0.1448	0.0611
Standard error	(0.0929)	(0.0494)	(0.0661)	(0.0375)	(0.1133)	(0.0593)	(0.0865)	(0.0411)
P-value	[0.0489]	[0.3531]	[0.0304]	[0.0384]	[0.2281]	[0.4255]	[0.0953]	[0.1376]
Observations	354	1673	1067	4624	297	1365	1010	4316
Groups	84	150	268	501	89	192	245	431
Study no-prep	0.0527	0.0123	0.0405	0.0510	-0.0470	-0.0267	-0.0226	-0.0011
Standard error	(0.0938)	(0.0432)	(0.0646)	(0.0336)	(0.1039)	(0.0518)	(0.0743)	(0.0357)
P-value	[0.5760]	[0.7758]	[0.5319]	[0.1290]	[0.6522]	[0.6074]	[0.7609]	[0.9757]
Observations	354	1673	1067	4624	297	1365	1010	4316
Groups	84	150	268	501	89	192	245	431
Uni cred≥15	0.0820	0.0014	0.0054	0.0123	0.0446	0.0106	0.0262	0.0169
Standard error	(0.0700)	(0.0341)	(0.0558)	(0.0261)	(0.0789)	(0.0383)	(0.0644)	(0.0302)
P-value	[0.2452]	[0.9663]	[0.9230]	[0.6389]	[0.5732]	[0.7825]	[0.6845]	[0.5772]
Observations	354	1673	1067	4624	297	1365	1010	4316
Groups	84	150	268	501	89	192	245	431
Work ≥50%	-0.0316	-0.0014	-0.0370	-0.0287	0.0318	0.0070	0.0032	-0.0082
Standard error	(0.0926)	(0.0460)	(0.0659)	(0.0378)	(0.1191)	(0.0538)	(0.0905)	(0.0433)
P-value	[0.7338]	[0.9760]	[0.5749]	[0.4477]	[0.7903]	[0.8973]	[0.9722]	[0.8506]
Observations	354	1672	1067	4623	296	1362	1009	4313
Groups	84	150	268	501	89	192	245	431
Bandwidth	5	20	5	20	5	20	5	20
Trend	No	Linear	No	Linear	No	Linear	No	Linear

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and fixed effects for the track choices of the two top choices. The regressions include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D). Specifications with linear trends include the running variable separately and interacted with the dummy variable for being above the admission threshold. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. The regression sample is *not* restricted to students with the same track preference for the top and second preference, but students with different track preferences above for the two top choices are also included. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

TABLE C19. STANDARDIZED TEST OUTCOMES: MATH

T=1;T=0-samples:	Indep/Public;Indep/Indep		Indep/Public;Public/Public		Public/Indep;Indep/Indep		Public/Indep;Public/Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attend Private	0.5380***	0.6682***	0.5627***	0.6757***	-0.4565***	-0.6556***	-0.4558***	-0.6557***
Standard error	(0.0757)	(0.0353)	(0.0741)	(0.0357)	(0.0732)	(0.0352)	(0.0707)	(0.0323)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	512	2341	1209	5487	450	1949	1147	5095
Groups	123	222	355	669	144	301	320	570
High test grade	0.0675	0.0423**	0.0518	0.0471***	0.0151	-0.0132	0.0067	-0.0061
Standard error	(0.0408)	(0.0160)	(0.0303)	(0.0159)	(0.0293)	(0.0150)	(0.0164)	(0.0083)
P-value	[0.1007]	[0.0086]	[0.0883]	[0.0031]	[0.6070]	[0.3793]	[0.6830]	[0.4608]
Observations	512	2348	1213	5505	450	1961	1151	5118
Groups	123	222	355	669	144	301	320	570
Pass test grade	0.1356*	0.0187	0.1267*	-0.0009	-0.0005	0.0386	-0.0369	0.0105
Standard error	(0.0664)	(0.0299)	(0.0556)	(0.0283)	(0.0803)	(0.0336)	(0.0681)	(0.0290)
P-value	[0.0432]	[0.5321]	[0.0233]	[0.9733]	[0.9951]	[0.2518]	[0.5886]	[0.7167]
Observations	512	2348	1213	5505	450	1961	1151	5118
Groups	123	222	355	669	144	301	320	570
Test grade>Course grade	0.0230	0.0353**	0.0207	0.0249***	0.0045	0.0019	-0.0037	-0.0109
Standard error	(0.0185)	(0.0131)	(0.0258)	(0.0088)	(0.0232)	(0.0152)	(0.0216)	(0.0110)
P-value	[0.2161]	[0.0075]	[0.4226]	[0.0046]	[0.8458]	[0.8978]	[0.8646]	[0.3234]
Observations	508	2315	1154	5253	437	1907	1083	4845
Groups	121	218	351	667	142	298	317	569
Test grade<Course grade	0.0484	0.0487	0.0431	0.0282	0.0058	-0.0394	-0.0356	-0.0497
Standard error	(0.0546)	(0.0284)	(0.0512)	(0.0278)	(0.0767)	(0.0373)	(0.0732)	(0.0309)
P-value	[0.3771]	[0.0873]	[0.4005]	[0.3104]	[0.9397]	[0.2917]	[0.6269]	[0.1084]
Observations	508	2315	1154	5253	437	1907	1083	4845
Groups	121	218	351	667	142	298	317	569
Bandwidth	5	20	5	20	5	20	5	20
Trend	No	Linear	No	Linear	No	Linear	No	Linear

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and fixed effects for the track choices of the two top choices. The regressions include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D). Specifications with linear trends include the running variable separately and interacted with the dummy variable for being above the admission threshold. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. The regression sample is *not* restricted to students with the same track preference for the top and second preference, but students with different track preferences above for the two top choices are also included. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

TABLE C20. STANDARDIZED TEST OUTCOMES: ENGLISH

T=1;T=0-samples:	*Indep/Public;Indep/Indep*		*Indep/Public;Public/Public*		*Public/Indep;Indep/Indep*		*Public/Indep;Public/Public*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attend Private	0.5108***	0.6662***	0.5360***	0.6699***	-0.4228***	-0.6407***	-0.4301***	-0.6463***
Standard error	(0.0699)	(0.0360)	(0.0684)	(0.0363)	(0.0745)	(0.0363)	(0.0715)	(0.0332)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	554	2422	1271	5744	483	2002	1200	5324
Groups	128	227	363	699	148	302	327	590
High test grade	0.0308	0.0169	0.0502	0.0564*	-0.0701	-0.0420	-0.0470	-0.0177
Standard error	(0.0551)	(0.0281)	(0.0479)	(0.0259)	(0.0524)	(0.0291)	(0.0349)	(0.0204)
P-value	[0.5771]	[0.5495]	[0.2952]	[0.0299]	[0.1826]	[0.1500]	[0.1793]	[0.3870]
Observations	555	2429	1275	5761	484	2014	1204	5346
Groups	128	227	363	699	148	302	327	590
Pass test grade	-0.0255	-0.0007	-0.0187	-0.0019	-0.0308	-0.0054	-0.0121	-0.0075
Standard error	(0.0189)	(0.0091)	(0.0177)	(0.0087)	(0.0328)	(0.0123)	(0.0291)	(0.0132)
P-value	[0.1812]	[0.9365]	[0.2925]	[0.8296]	[0.3501]	[0.6618]	[0.6792]	[0.5689]
Observations	555	2429	1275	5761	484	2014	1204	5346
Groups	128	227	363	699	148	302	327	590
Test grade>Course grade	-0.0372	-0.0346	-0.0015	0.0045	-0.0810	-0.0712*	-0.0154	-0.0287
Standard error	(0.0473)	(0.0329)	(0.0454)	(0.0253)	(0.0646)	(0.0338)	(0.0558)	(0.0274)
P-value	[0.4334]	[0.2942]	[0.9734]	[0.8583]	[0.2116]	[0.0361]	[0.7830]	[0.2966]
Observations	533	2362	1223	5529	464	1940	1154	5107
Groups	127	226	359	695	147	299	325	587
Test grade<Course grade	0.0361	0.0434	-0.0066	0.0043	0.0688	0.0342	0.0003	0.0081
Standard error	(0.0666)	(0.0310)	(0.0495)	(0.0249)	(0.0680)	(0.0338)	(0.0570)	(0.0270)
P-value	[0.5883]	[0.1632]	[0.8937]	[0.8644]	[0.3131]	[0.3123]	[0.9960]	[0.7654]
Observations	533	2362	1223	5529	464	1940	1154	5107
Groups	127	226	359	695	147	299	325	587
Bandwidth	5	20	5	20	5	20	5	20
Trend	No	Linear	No	Linear	No	Linear	No	Linear

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and fixed effects for the track choices of the two top choices. The regressions include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D). Specifications with linear trends include the running variable separately and interacted with the dummy variable for being above the admission threshold. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. The regression sample is *not* restricted to students with the same track preference for the top and second preference, but students with different track preferences above for the two top choices are also included. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

TABLE C21. STANDARDIZED TEST OUTCOMES: SWEDISH

T=1;T=0-samples:	*Indep/Public;Indep/Indep*		*Indep/Public;Public/Public*		*Public/Indep;Indep/Indep*		*Public/Indep;Public/Public*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attend Private	0.5007***	0.6737***	0.5135***	0.6782***	-0.4680***	-0.6621***	-0.4750***	-0.6624***
Standard error	(0.0808)	(0.0330)	(0.0773)	(0.0342)	(0.0764)	(0.0360)	(0.0694)	(0.0323)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	527	2416	1299	5924	460	1946	1232	5454
Groups	129	231	404	750	149	309	368	641
High test grade	-0.0482	0.0222	-0.0165	0.0493**	-0.0653	-0.0525*	-0.0533	-0.0243
Standard error	(0.0525)	(0.0207)	(0.0428)	(0.0182)	(0.0468)	(0.0249)	(0.0285)	(0.0143)
P-value	[0.3600]	[0.2847]	[0.6992]	[0.0070]	[0.1653]	[0.0362]	[0.0620]	[0.0913]
Observations	527	2421	1303	5939	460	1958	1236	5476
Groups	129	231	405	750	149	310	368	641
Pass test grade	0.0426	-0.0014	-0.0005	-0.0445***	-0.0217	0.0035	-0.0567	-0.0259
Standard error	(0.0301)	(0.0143)	(0.0294)	(0.0138)	(0.0493)	(0.0208)	(0.0445)	(0.0199)
P-value	[0.1602]	[0.9199]	[0.9854]	[0.0013]	[0.6608]	[0.8682]	[0.2040]	[0.1932]
Observations	527	2421	1303	5939	460	1958	1236	5476
Groups	129	231	405	750	149	310	368	641
Test grade>Course grade	0.0546	0.0320	-0.0210	-0.0211	-0.0115	0.0188	-0.1882***	-0.0348
Standard error	(0.0609)	(0.0281)	(0.0518)	(0.0212)	(0.0682)	(0.0344)	(0.0536)	(0.0256)
P-value	[0.3718]	[0.2565]	[0.6856]	[0.3194]	[0.8665]	[0.5847]	[0.0005]	[0.1739]
Observations	502	2245	1199	5388	424	1796	1121	4939
Groups	128	227	391	729	144	300	352	621
Test grade<Course grade	-0.0169	-0.0244	-0.1018	-0.0316	0.1276	-0.0384	0.0143	-0.0270
Standard error	(0.0882)	(0.0388)	(0.0639)	(0.0302)	(0.1173)	(0.0438)	(0.0836)	(0.0339)
P-value	[0.8481]	[0.5309]	[0.1119]	[0.2948]	[0.2786]	[0.3807]	[0.8639]	[0.4262]
Observations	502	2245	1199	5388	424	1796	1121	4939
Groups	128	227	391	729	144	300	352	621
Bandwidth	5	20	5	20	5	20	5	20
Trend	No	Linear	No	Linear	No	Linear	No	Linear

Note: Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, and fixed effects for the track choices of the two top choices. The regressions include dummy variables for being in a "treated" or "non-treated" sample (T=1 and T=0, respectively), and for being above or below the admission threshold (D=1 and D=0). The coefficients in the table represent the interaction variable for these two: T×D). Specifications with linear trends include the running variable separately and interacted with the dummy variable for being above the admission threshold. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. The regression sample is *not* restricted to students with the same track preference for the top and second preference, but students with different track preferences above for the two top choices are also included. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

C.8 Alternative RD/DID specifications: Pooling all samples

TABLE C22. GRADUATION AND GRADES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Attend Private	0.4411***	0.5090***	0.5586***	0.6431***	0.4562***	0.5678***	0.6532***
Standard error	(0.0502)	(0.0836)	(0.0297)	(0.0240)	(0.0437)	(0.0267)	(0.0212)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	1067	1067	2806	4806	2306	5807	10095
Groups	241	655	465	465	524	903	903
Switch independent/public	0.0342	0.0484	0.0003	0.0041	0.0243	-0.0100	-0.0115
Standard error	(0.0396)	(0.0677)	(0.0225)	(0.0167)	(0.0310)	(0.0182)	(0.0136)
P-value	[0.3884]	[0.4750]	[0.9909]	[0.8071]	[0.4331]	[0.5828]	[0.3960]
Observations	1040	1040	2746	4704	2254	5692	9908
Groups	240	641	464	464	522	902	902
Pctile GPA12	5.2115	2.9474	2.7853	3.9092**	6.1476*	2.2173	3.9471***
Standard error	(3.3851)	(5.5210)	(1.8305)	(1.4792)	(2.7087)	(1.4854)	(1.1395)
P-value	[0.1250]	[0.5939]	[0.1288]	[0.0085]	[0.0237]	[0.1359]	[0.0006]
Observations	922	922	2412	4180	1967	4948	8697
Groups	237	593	459	460	511	887	897
Graduate on time	0.0916	0.0215	0.0047	0.0063	0.1143**	0.0527*	0.0571**
Standard error	(0.0505)	(0.0738)	(0.0316)	(0.0248)	(0.0421)	(0.0264)	(0.0206)
P-value	[0.0706]	[0.7712]	[0.8826]	[0.7984]	[0.0068]	[0.0465]	[0.0056]
Observations	1072	1072	2825	4830	2316	5842	10149
Groups	241	658	465	465	524	903	903
7^{th} term	-0.0672	0.0019	-0.0100	-0.0185	-0.0615	-0.0335	-0.0514***
Standard error	(0.0385)	(0.0632)	(0.0239)	(0.0190)	(0.0323)	(0.0197)	(0.0156)
P-value	[0.0821]	[0.9764]	[0.6748]	[0.3324]	[0.0576]	[0.0893]	[0.0010]
Observations	1072	1072	2825	4830	2316	5842	10149
Groups	241	658	465	465	524	903	903
Bandwidth	5	5	10	20	5	10	20
FE	Adm.gr 1	Adm.gr 1×2	Adm.gr 1	Adm.gr 1	Adm.gr 1	Adm.gr 1	Adm.gr 1
Trend	No	No	Linear	Linear	No	Linear	Linear
Same track	Yes	Yes	Yes	Yes	No	No	No

Note: The coefficients in the table represent the interaction variable that represents being on the side of the admission threshold that predicts independent school attendance, for the samples of students with mixed preferences (i.e.being above the threshold for students with independent as first and municipal as second, and being below the threshold for students with municipal as first and independent as second preference.) Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, for all cases except column (2), which includes fixed effects for the admission groups of the first and second preference. The regressions also include the following set of dummy variables: dummies for each of the four preference combinations (combinations of Independent and Municipal as first and second preference), dummies for being above or below the admission threshold separately and interacted with a dummy for having a municipal or independent school as first preference. Specifications with Linear trends include the running variable, estimated separately above and below the admission threshold, and interacted with the indicator variable for having a municipal or private school as top preference. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. Specifications denoted "Same track: Yes" are based on the sample that is restricted to students with the same track preference for the top and second preference, and those denoted "Same track: No", are not. All regressions on the latter sample contain track fixed effects for the two track options. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

TABLE C23. POST-GRADUATION OUTCOMES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Attend Private	0.4411***	0.5090***	0.5586***	0.6431***	0.4562***	0.5678***	0.6532***
Standard error	(0.0502)	(0.0836)	(0.0297)	(0.0240)	(0.0437)	(0.0267)	(0.0212)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	1067	1067	2806	4806	2306	5807	10095
Groups	241	655	465	465	524	903	903
Study	0.0646	0.0808	0.0929	0.0637	0.0024	0.0099	-0.0094
Standard error	(0.0866)	(0.1289)	(0.0544)	(0.0413)	(0.0625)	(0.0395)	(0.0317)
P-value	[0.4566]	[0.5316]	[0.0889]	[0.1243]	[0.9694]	[0.8026]	[0.7678]
Observations	586	586	1546	2594	1364	3518	5989
Groups	139	375	270	270	329	581	581
Study no-prep	0.1063	0.0630	0.0762	0.0687	0.0437	0.0306	0.0123
Standard error	(0.0763)	(0.1142)	(0.0472)	(0.0360)	(0.0550)	(0.0333)	(0.0273)
P-value	[0.1657]	[0.5823]	[0.1078]	[0.0570]	[0.4273]	[0.3587]	[0.6524]
Observations	586	586	1546	2594	1364	3518	5989
Groups	139	375	270	270	329	581	581
Uni cred≥15	0.0486	-0.0586	0.0327	0.0283	0.0193	-0.0075	-0.0067
Standard error	(0.0615)	(0.1052)	(0.0382)	(0.0318)	(0.0440)	(0.0257)	(0.0221)
P-value	[0.4307]	[0.5785]	[0.3923]	[0.3739]	[0.6607]	[0.7690]	[0.7603]
Observations	586	586	1546	2594	1364	3518	5989
Groups	139	375	270	270	329	581	581
Work ≥50%	-0.0221	-0.1086	-0.0921	-0.0513	-0.0191	-0.0192	0.0019
Standard error	(0.0840)	(0.1544)	(0.0507)	(0.0409)	(0.0631)	(0.0390)	(0.0314)
P-value	[0.7928]	[0.4829]	[0.0705]	[0.2113]	[0.7626]	[0.6222]	[0.9516]
Observations	585	585	1545	2592	1363	3516	5985
Groups	139	374	270	270	329	581	581
Bandwidth	5	5	10	20	5	10	20
FE	Adm.gr 1	Adm.gr 1×2	Adm.gr 1	Adm.gr 1	Adm.gr 1	Adm.gr 1	Adm.gr 1
Trend	No	No	Linear	Linear	No	Linear	Linear
Same track	Yes	Yes	Yes	Yes	No	No	No

Note: The coefficients in the table represent the interaction variable that represents being on the side of the admission threshold that predicts independent school attendance, for the samples of students with mixed preferences (i.e.being above the threshold for students with independent as first and municipal as second, and being below the threshold for students with municipal as first and independent as second preference.) Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, for all cases except column (2), which includes fixed effects for the admission groups of the first and second preference. The regressions also include the following set of dummy variables: dummies for each of the four preference combinations (combinations of Independent and Municipal as first and second preference), dummies for being above or below the admission threshold separately and interacted with a dummy for having a municipal or independent school as first preference. Specifications with Linear trends include the running variable, estimated separately above and below the admission threshold, and interacted with the indicator variable for having a municipal or private school as top preference. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. Specifications denoted "Same track: Yes" are based on the sample that is restricted to students with the same track preference for the top and second preference, and those denoted "Same track: No", are not. All regressions on the latter sample contain track fixed effects for the two track options. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

TABLE C24. STANDARDIZED TEST OUTCOMES: MATH

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Attend Private	0.4811***	0.6012***	0.5694***	0.6521***	0.4869***	0.5704***	0.6619***
Standard error	(0.0584)	(0.0855)	(0.0344)	(0.0266)	(0.0530)	(0.0314)	(0.0237)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	842	842	2210	3833	1659	4178	7436
Groups	215	542	421	431	443	764	792
High test grade	0.0385	0.0651	0.0205	0.0334*	0.0232	0.0122	0.0228**
Standard error	(0.0211)	(0.0386)	(0.0146)	(0.0136)	(0.0199)	(0.0104)	(0.0088)
P-value	[0.0701]	[0.0931]	[0.1595]	[0.0141]	[0.2452]	[0.2432]	[0.0096]
Observations	844	844	2225	3850	1663	4198	7466
Groups	215	543	421	431	443	764	792
Pass test grade	0.0897	0.1619	0.0521	0.0107	0.0723	0.0554*	0.0039
Standard error	(0.0549)	(0.0877)	(0.0335)	(0.0253)	(0.0453)	(0.0273)	(0.0205)
P-value	[0.1036]	[0.0662]	[0.1213]	[0.6726]	[0.1111]	[0.0427]	[0.8479]
Observations	844	844	2225	3850	1663	4198	7466
Groups	215	543	421	431	443	764	792
Test grade>Course grade	0.0376	0.0088	0.0385**	0.0239*	0.0157	0.0269*	0.0222**
Standard error	(0.0200)	(0.0311)	(0.0143)	(0.0097)	(0.0153)	(0.0113)	(0.0083)
P-value	[0.0619]	[0.7767]	[0.0075]	[0.0144]	[0.3069]	[0.0174]	[0.0080]
Observations	797	797	2122	3667	1591	4026	7160
Groups	212	518	418	428	438	757	787
Test grade<Course grade	0.0274	0.0355	-0.0055	0.0378	0.0424	0.0219	0.0474*
Standard error	(0.0576)	(0.0931)	(0.0366)	(0.0278)	(0.0453)	(0.0283)	(0.0212)
P-value	[0.6344]	[0.7031]	[0.8811]	[0.1741]	[0.3501]	[0.4401]	[0.0254]
Observations	797	797	2122	3667	1591	4026	7160
Groups	212	518	418	428	438	757	787
Bandwidth	5	5	10	20	5	10	20
FE	Adm.gr 1	Adm.gr 1×2	Adm.gr 1	Adm.gr 1	Adm.gr 1	Adm.gr 1	Adm.gr 1
Trend	No	No	Linear	Linear	No	Linear	Linear
Same track	Yes	Yes	Yes	Yes	No	No	No

Note: The coefficients in the table represent the interaction variable that represents being on the side of the admission threshold that predicts independent school attendance, for the samples of students with mixed preferences (i.e.being above the threshold for students with independent as first and municipal as second, and being below the threshold for students with municipal as first and independent as second preference.) Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, for all cases except column (2), which includes fixed effects for the admission groups of the first and second preference. The regressions also include the following set of dummy variables: dummies for each of the four preference combinations (combinations of Independent and Municipal as first and second preference), dummies for being above or below the admission threshold separately and interacted with a dummy for having a municipal or independent school as first preference. Specifications with Linear trends include the running variable, estimated separately above and below the admission threshold, and interacted with the indicator variable for having a municipal or private school as top preference. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. Specifications denoted "Same track: Yes" are based on the sample that is restricted to students with the same track preference for the top and second preference, and those denoted "Same track: No", are not. All regressions on the latter sample contain track fixed effects for the two track options. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

TABLE C25. STANDARDIZED TEST OUTCOMES: ENGLISH

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Attend Private	0.4448***	0.5552***	0.5527***	0.6384***	0.4588***	0.5723***	0.6563***
Standard error	(0.0584)	(0.0824)	(0.0353)	(0.0274)	(0.0504)	(0.0312)	(0.0243)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	867	867	2249	3892	1754	4389	7746
Groups	217	544	419	429	455	792	817
High test grade	0.0382	-0.0219	0.0292	0.0439	0.0325	0.0103	0.0186
Standard error	(0.0437)	(0.0760)	(0.0257)	(0.0226)	(0.0298)	(0.0189)	(0.0169)
P-value	[0.3831]	[0.7734]	[0.2570]	[0.0530]	[0.2754]	[0.5873]	[0.2735]
Observations	870	870	2265	3909	1759	4410	7775
Groups	218	546	419	429	455	792	817
Pass test grade	0.0017	0.0388	-0.0069	-0.0022	-0.0059	-0.0055	0.0039
Standard error	(0.0210)	(0.0326)	(0.0114)	(0.0092)	(0.0179)	(0.0107)	(0.0084)
P-value	[0.9347]	[0.2354]	[0.5441]	[0.8128]	[0.7400]	[0.6082]	[0.6422]
Observations	870	870	2265	3909	1759	4410	7775
Groups	218	546	419	429	455	792	817
Test grade>Course grade	0.0021	0.0454	0.0219	0.0010	-0.0105	0.0038	0.0027
Standard error	(0.0389)	(0.0646)	(0.0281)	(0.0260)	(0.0376)	(0.0241)	(0.0211)
P-value	[0.9576]	[0.4829]	[0.4370]	[0.9683]	[0.7799]	[0.8758]	[0.8974]
Observations	836	836	2186	3746	1687	4246	7469
Groups	218	531	418	427	452	789	813
Test grade<Course grade	-0.0324	0.0110	0.0099	0.0318	0.0228	0.0112	0.0151
Standard error	(0.0483)	(0.0679)	(0.0314)	(0.0259)	(0.0422)	(0.0254)	(0.0204)
P-value	[0.5035]	[0.8717]	[0.7520]	[0.2203]	[0.5888]	[0.6603]	[0.4607]
Observations	836	836	2186	3746	1687	4246	7469
Groups	218	531	418	427	452	789	813
Bandwidth	5	5	10	20	5	10	20
FE	Adm.gr 1	Adm.gr 1×2	Adm.gr 1	Adm.gr 1	Adm.gr 1	Adm.gr 1	Adm.gr 1
Trend	No	No	Linear	Linear	No	Linear	Linear
Same track	Yes	Yes	Yes	Yes	No	No	No

Note: The coefficients in the table represent the interaction variable that represents being on the side of the admission threshold that predicts independent school attendance, for the samples of students with mixed preferences (i.e.being above the threshold for students with independent as first and municipal as second, and being below the threshold for students with municipal as first and independent as second preference.) Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, for all cases except column (2), which includes fixed effects for the admission groups of the first and second preference. The regressions also include the following set of dummy variables: dummies for each of the four preference combinations (combinations of Independent and Municipal as first and second preference), dummies for being above or below the admission threshold separately and interacted with a dummy for having a municipal or independent school as first preference. Specifications with Linear trends include the running variable, estimated separately above and below the admission threshold, and interacted with the indicator variable for having a municipal or private school as top preference. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. Specifications denoted "Same track: Yes" are based on the sample that is restricted to students with the same track preference for the top and second preference, and those denoted "Same track: No", are not. All regressions on the latter sample contain track fixed effects for the two track options. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1

TABLE C26. STANDARDIZED TEST OUTCOMES: SWEDISH

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Attend Private	0.4746***	0.5283***	0.5738***	0.6556***	0.4758***	0.5815***	0.6694***
Standard error	(0.0607)	(0.0970)	(0.0349)	(0.0269)	(0.0518)	(0.0306)	(0.0232)
P-value	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Observations	831	831	2182	3813	1759	4424	7870
Groups	232	529	441	449	497	856	872
High test grade	0.0324	-0.0185	0.0299	0.0372*	0.0207	0.0232	0.0260*
Standard error	(0.0375)	(0.0683)	(0.0214)	(0.0162)	(0.0280)	(0.0167)	(0.0121)
P-value	[0.3881]	[0.7863]	[0.1633]	[0.0222]	[0.4616]	[0.1651]	[0.0326]
Observations	833	833	2195	3826	1763	4442	7897
Groups	232	531	441	449	497	856	872
Pass test grade	0.0541	0.0973	0.0027	-0.0066	0.0425	0.0116	0.0127
Standard error	(0.0323)	(0.0542)	(0.0202)	(0.0147)	(0.0281)	(0.0172)	(0.0126)
P-value	[0.0954]	[0.0738]	[0.8930]	[0.6542]	[0.1314]	[0.5001]	[0.3125]
Observations	833	833	2195	3826	1763	4442	7897
Groups	232	531	441	449	497	856	872
Test grade>Course grade	0.1560**	0.1421	0.0638	0.0302	0.1162***	0.0525*	0.0303
Standard error	(0.0569)	(0.0938)	(0.0349)	(0.0257)	(0.0405)	(0.0255)	(0.0188)
P-value	[0.0066]	[0.1312]	[0.0679]	[0.2399]	[0.0043]	[0.0400]	[0.1064]
Observations	759	759	1995	3428	1623	4079	7184
Groups	224	495	426	437	480	824	848
Test grade<Course grade	-0.0624	-0.2625*	-0.0071	0.0026	-0.0259	-0.0029	0.0013
Standard error	(0.0751)	(0.1191)	(0.0420)	(0.0317)	(0.0599)	(0.0329)	(0.0256)
P-value	[0.4071]	[0.0285]	[0.8650]	[0.9354]	[0.6657]	[0.9309]	[0.9600]
Observations	759	759	1995	3428	1623	4079	7184
Groups	224	495	426	437	480	824	848
Bandwidth	5	5	10	20	5	10	20
FE	Adm.gr 1	Adm.gr 1×2	Adm.gr 1	Adm.gr 1	Adm.gr 1	Adm.gr 1	Adm.gr 1
Trend	No	No	Linear	Linear	No	Linear	Linear
Same track	Yes	Yes	Yes	Yes	No	No	No

Note: The coefficients in the table represent the interaction variable that represents being on the side of the admission threshold that predicts independent school attendance, for the samples of students with mixed preferences (i.e.being above the threshold for students with independent as first and municipal as second, and being below the threshold for students with municipal as first and independent as second preference.) Regressions include fixed effects for admission group (school×year×educational programme) for the students' top preference, for all cases except column (2), which includes fixed effects for the admission groups of the first and second preference. The regressions also include the following set of dummy variables: dummies for each of the four preference combinations (combinations of Independent and Municipal as first and second preference), dummies for being above or below the admission threshold separately and interacted with a dummy for having a municipal or independent school as first preference. Specifications with Linear trends include the running variable, estimated separately above and below the admission threshold, and interacted with the indicator variable for having a municipal or private school as top preference. The regressions include the following student level covariates: Household disposable income, a dummy variable for parents having a post-secondary degree, student level dummy variables for being female, and being born in a non-Western country, final grade sum from lower secondary education, and a dummy variable for having attended an independent school in grade 9. Missing variables for the covariates were replaced with a constant, and dummies indicating whether covariate observations were missing were included. Specifications denoted "Same track: Yes" are based on the sample that is restricted to students with the same track preference for the top and second preference, and those denoted "Same track: No", are not. All regressions on the latter sample contain track fixed effects for the two track options. Standard errors are clustered on the admission group (school×year×educational programme) for the students' top preference. *** p<0.01, ** p<0.05, * p<0.1