WORKING PAPER 2024:17

Unpacking the impact of voucher schools: evidence from Sweden

Karin Edmark
Iftikhar Hussain
Carla Haelermans

IFAU
INSTITUTE FOR EVALUATION OF LABOUR MARKET AND EDUCATION POLICY

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala.

IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

More information about IFAU and the institute's publications can be found on the website www.ifau.se

ISSN 1651-1166

Unpacking the impact of voucher schools: evidence from Sweden[a]

Karin Edmark[a] Iftikhar Hussain Carla Haelermans

September 2024

ABSTRACT

Research on voucher schools has mainly focused on whether students experience improved academic results in these schools and whether they generate competitive pressure for public schools. In this paper we focus instead on the role voucher schools might play in altering the menu of options available to students, for example with respect to vocational and academic tracks for adolescents. The setting for this paper is the period of rapid expansion of voucher schools in Sweden. Exploiting fine-grained geographical information on students' home location as well as variation in exposure across siblings, we uncover new evidence demonstrating that the introduction of voucher schools induces greater vocational education participation, and not simply a substitution of public for private vocational schools. In effect, voucher school penetration leads to a switch away from academic tracks, including academic science subjects, in favor of vocational options. We then assess the impact on medium- and long-term outcomes. The results demonstrate that voucher school penetration has a negative impact on the probability that the highest qualification is in a STEM subject by age 30. The results also reveal a negative impact on long-term labor market outcomes.

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

Key words: Private provision, independent schools, voucher school reform, vocational education, upper secondary education.

[a] We acknowledge funding from the Swedish Research Council, Project Nr: 2014-01783. We are grateful for comments from the review process of the IFAU, as well as from seminar and conference participants at the Fourth Lisbon Research Workshop on Economics, Statistics and Econometrics of Education; The 73rd Annual Congress of the IIPF in Tokyo; the Research Centre for Education and the Labour Market (ROA) at Maastricht University; the Swedish Institute for Social Research at Stockholm University (SOFI); the Institute for Evaluation of Labour Market and Education Policy (IFAU); the Uppsala Centre for Labour Studies Workshop; and the Research Institute of Industrial Economics (IFN).

[b] SOFI, Stockholm University, SE-106 91 Stockholm, Sweden, IFAU, IFN and CESIfo, karin.edmark@sofi.su.se, http://orcid.org/0000-0002-5629-5499.

[c] Department of Economics, University of Sussex. iftikhar.hussain@sussex.ac.uk.

[d] Research Centre for Education and the Labour Market (ROA), School of Business and Economics, Maastricht University. http://orcid.org/0000-0002- 9202-8427; carla.haelermans@maastrichtuniversity.nl.

1 Introduction

The use of voucher school policies has long been advocated as a way of injecting market mechanisms in the compulsory education system, thereby raising productivity in this sector (Friedman, 1962, Hoxby, 2003, Manski, 1992). A growing literature has examined the effectiveness of both small scale voucher schemes in many jurisdictions across the world, as well as nationwide voucher policies in a handful of countries. Studies have addressed: whether enrolling in a voucher school leads to improved test scores; the spillover effects for students who remain in public schools; and whether aggregate test scores and graduation rates improve under large-scale voucher schemes (Epple et al., 2017; Urquiola, 2016). Among potential mechanisms, researchers have investigated differences in productivity between voucher and public schools, competitive effects and other equilibrium supply side responses (Muralidharan and Sundararaman, 2015, Neilson, 2020) as well as sorting and changes in the quality of peers students are exposed to (Bettinger et al. 2010).

In this study we investigate a novel mechanism which may lie behind voucher school impacts: we ask whether the expansion of voucher schools leads to changes in the field of study (e.g. science versus social science) or track (e.g. academic versus vocational) taken up by students.[1] We investigate the aggregate or overall effect of voucher reforms on such outcomes in Sweden, which experienced a period of rapid voucher school expansion in the 2000's. We focus on the upper secondary education system, which has features which sets it apart from the lower levels of the compulsory schooling sector. In particular, at age 16, the onset of upper secondary schooling, students must choose not only between private (voucher) and public school options, but must also decide on whether to pursue an academic or vocational track. This

[1] Epple et al. (2017) define a voucher as "a government-supplied coupon that is used to offset tuition at an eligible private school." These authors list five countries with national voucher programs, Chile, Denmark, Netherlands, New Zealand and Sweden. We use the term "voucher schools" as short-hand for independent voucher-funded schools (Epple et al., 2017).

setting thus offers an opportunity to study the role of voucher schools in shaping students' choices and subsequent short- and long-term outcomes.

The upper secondary education system in Sweden offers around 20 sub-categories within the academic and vocational tracks, and students must simultaneously choose a school and a track. Voucher schools that enter the market can choose which of these educational tracks they offer, and are not required to offer any particular mix of tracks. Furthermore, male and female students tend to choose different tracks, and therefore there may be important distinctions to be drawn in the resulting gains or losses from the entry of voucher schools by gender.

The advantage of evaluating a large-scale or nationwide scheme such as Sweden's is that it incorporates effects including, for example, supply side responses such as the entry and exit of schools and the challenge of recruiting teachers at scale. However, identification is particularly challenging because there is no straightforward counterfactual, as might be the case with a small-scale voucher system where enrolment is determined by a lottery. Exploiting variation in the rapid rollout of voucher schools across jurisdictions in a difference-in-difference framework is one possible solution. However, concerns about differential trends and variations in local policies which may confound the treatment (voucher expansion) pose live threats to identification.

Our identification strategy exploits variation in exposure to voucher schools within families across time. Exposure to voucher schools is defined as the proportion of voucher schools at the time when students normally start upper secondary school, i.e. the year when they turn 16, in a 5km radius from students' homes.[2] Our research design compares, for example, a pair of siblings residing in an area which experiences a large change in the share of voucher schools with a pair of siblings residing in a different area experiencing little change in the

[2] We also experiment with 10km and 20km radii from the student's home.

availability of voucher schools. The identifying assumption is that in the absence of voucher schools, the within-family change in take-up of vocational track schools, for example, would have been the same in these two areas.[3]

In order to probe the plausibility of this assumption we present a set of balancing tests which demonstrate that voucher school expansion is not systematically related to changes in characteristics across siblings, including prior attainment, as well as family-by-time characteristics, such as household disposable income. Furthermore, we subject our estimates to a battery of robustness tests, including comparing our results with and without a detailed set of time-varying family covariates and pre-treatment student-level characteristics such as prior grade point average; inclusion of labor market region-by-year fixed effects and municipality trends; and variation in the size of the geographical market.

Our results demonstrate that voucher school expansion raises the likelihood of enrolling in a vocational track, whether in a voucher school or a public school. This is a key finding and demonstrates that voucher school penetration does not simply imply a substitution away from the public to the private sector within a given track, but in fact leads to a change in the aggregate mix of track choices. That is to say, voucher school expansion is not neutral in its impact on the choice of academic versus vocational tracks.

Analysis of the detailed field options within academic and vocational tracks reveals that whilst take-up of science *academic* track options declines in response to the treatment, there is a rise in the group of vocational tracks in the industry and technology fields.[4] This substitution pattern is driven by males (brothers in the sibling fixed effect design). For females (sisters) the science academic track options decline but there is no significant rise in any of the vocational

[3] It should be noted that we exploit a number of Swedish registry datasets. For outcomes relating to track and field choices, as well as upper secondary GPA scores, we have precise information on home location. When we undertake analysis of longer-term outcomes such as labor market performance, we have information on the municipality of residence, and hence we exploit municipality-level variation in the availability of voucher schools. In all cases, we include sibling fixed effects

[4] Note that the groups of tracks are broad, and this group includes also non-technical professions, including e.g. farming, bus- and truck drivers, painters, bricklayers, etcetera.

subtracks.

Investigating whether an individual's highest qualification (whether at the upper secondary or university level) is in a STEM subject reveals that there is a significant negative impact on this long-term outcome (measured at age 30) for the full sample. The negative effects seem to be stronger for the female sample than the male one, although the estimates are too noisy to make conclusive statements. However, there is no significant impact on the overall rate of graduating at the upper secondary or tertiary level, for males or females. This suggests that individuals are substituting from STEM to other subjects such that there is no net impact on the overall graduation rate.

Finally, we also estimate the impact of voucher schools on longer-term labor market outcomes (also at age 30). We find that voucher expansion has a negative impact on employment status, in the full sample and for the male sub-sample, and it also has a negative impact on employment income for the male sample.

This study relates to the literature on the impact of voucher schools (see the reviews by Epple et al., 2017, Rouse and Barrow, 2009, and Urquiola, 2016). Previous studies have focused on the possible gains for students enrolling in voucher schools (Abdulkadiroglu et al., 2018, Krueger and Zhu, 2004, Muralidharan and Sundararaman, 2015); the role of sorting between public and private schools (Bettinger et al., 2010, Epple et al., 2004); potential competition effects generated by voucher reforms (Figlio and Hart, 2014, McMillan, 2005, Neilson, 2021); and aggregate effects of large-scale voucher schemes (Böhlmark and Lindahl, 2015, Hsieh and Urquiola, 2006). On the impact of targeted vouchers in a nationwide voucher setting, see also Aguirre (2022), Correa et al. (2014), Navarro-Palau (2017) and Feigenberg et al. (2019). An emerging strand of the education markets literature also investigates the role of horizontal differentiation, where the emphasis is on the variety of schooling options and potential for matching between students and schools along these lines rather than viewing schools simply as homogenous providers of education with varying levels of quality or

productivity (Bau, 2022, Gilraine et al., 2021).

Our study builds on this large literature. As far as we are aware, we are the first to provide a detailed study of the role that voucher schools play in determining students' track as well as field of study choices, such as STEM subjects. We provide novel insights into the potential importance of this channel, which may be a key mechanism underlying the impact of voucher schools.

A second literature we contribute to is concerned with the determinants and consequences of students' choice of field of study (Altonji et al., 2016, Kirkeboen et al., 2016, Dahl et al., 2023). Researchers have investigated both demand side and supply side factors, as well as the role played by individual preference heterogeneity (Wiswall and Zafar, 2018). We demonstrate how changes in the market for secondary schooling via the emergence of voucher schools can also play an important role in determining students' education choices along these important margins.

The rest of the paper is organised as follows. Section 2 describes the institutional setting and the data, section 3 lays out our empirical strategy, sections 4 and 5 present the results for the upper secondary track choices and longer-term outcomes, respectively, section 6 discusses alternative mechanisms and section 7 concludes.

2 Institutional Background and Data

2.1 Voucher School Reforms and Track Choice in Sweden

With its roots in the privatization reforms of the early 1990s, Sweden was an early adopter of voucher school policies on a large scale. These reforms, alongside the rich data that are available in the Swedish setting, offer the opportunity to glean lessons beyond the single case study of Sweden. Chile, Denmark, the Netherlands and New Zealand have all instituted reforms which make school vouchers a central part of their education systems (Urquiola, 2016).

These reforms in Sweden enabled the entry of privately operated schools, with full public funding granted through a voucher system. Additional tuition fees are not allowed. The reforms applied both to the compulsory (grades 1–9) as well as upper secondary education system (grades 10–12). These privately operated but fully publicly funded schools are sometimes referred to as 'independent schools', or 'voucher schools'. In this paper we sometimes also use the term 'private schools' when referring to such schools. We refer to regular or traditional schools operated by the municipalities as 'public schools'.[5]

Students enter upper secondary education at age 16. They are free to choose among all voucher schools in the country, and among all public schools in their home region.[6]

Importantly for our purposes, upper secondary schooling is divided into a number of academic and vocational tracks. In order to study the tracks in more detail, we have furthermore categorized the programs into five broad subgroups; three vocational and two academic. Table 1 lists our classification of sub-categories within each track. For some of the programs, especially for vocational programs that included a variety of subtracks, the classification is somewhat arbitrary (see the table note of Table 1).

[5] For a more detailed overview of the institutional setting, see for example, Böhlmark and Lindahl (2015), Edmark and Persson (2021) and Edmark et al. (2014).
[6] They may apply to municipal schools outside their home region, but home student applicants are given priority.

Table 1 Upper secondary program included in each of the five subcategories of Vocational and Academic tracks

Vocational			Academic	
Technology and Industry oriented	*Health and Care oriented*	*Trade and Administration oriented*	*Social Science oriented and Applied Arts*	*Natural Science and Technology oriented*
Construction	Child Recreation (BF)	Business and Administration	Social Science *(Social Science, Business, Humanities)*	Natural Science *(Natural Science and Technology)*
Electrical Engineering	Health Care (OP)	Hotel, Restaurant and Catering		
Energy		Food	*Media**	International Baccalaureate
Vehicle Engineering		*Handicraft**	Arts	
Industrial				
*Natural Resource Use**				

* For several smaller programs, the classification into the broader track categories is somewhat arbitrary. For example, we have chosen to classify the Media program (MP) as belonging to the Academic social science and applied arts category, although it was technically more of a vocational track (the division between vocational and academic programs was, during the studied period, rather vague for several programs, among them the Media program, see e.g. SOU 2008:27, e.g. pp 234f and 485f, and Skolverket, 2013). The main reason for why we choose this classification is that when the program was later abolished, in school year 2011/12, much of its media-related content was added as elective study paths in the academic Social Science program. (The exception is a course in printing technique which was added to a vocational program, see Prop. 2008/09:199 p 74). Other programs for which the classification is not straightforward include the Handicraft track, which includes subfields such as hairdressing, florist, carpentry, textile design, and other (silver- and goldsmith, glassblower etc), and the Natural Resource Use track, which included subfields related to agriculture, forestry, gardening. Finally, we have chosen to assign the small number of IB track students to the Natural Science and Technology tracks – also a somewhat arbitrary classification; the alternative option would be to assign it to the other strand of Academic tracks.

In addition, for students with insufficient grades to qualify for a regular track there is also a shorter preparatory track. As explained below, we exclude this track from our analysis since voucher schools rarely offer this option. Admission to a school and track is determined solely by the students' final grade from lower secondary education, in a deferred acceptance system.[7]

[7] Ability tests may be used for admission to the arts track and special profile tracks.

Apart from being privately operated, voucher schools are, on the whole, subject to the same regulations as the public schools: they are all monitored by the Swedish School Inspectorate; follow the same curriculum and educational goals; and are required to hire certified teachers (although exceptions are allowed in case of teacher scarcity). For-profit organizations may run schools, and the vast majority of upper secondary voucher schools belong to larger corporate groups. Conversions of public schools to voucher schools have occurred, but are rare. There is no cap on the total number of voucher schools, but approval can be denied if entry is determined to have substantial negative financial consequences for the municipality's ability to provide education. See Appendix B2 for further institutional details of the voucher school system.

Figure 1 charts the expansion of upper secondary voucher schools over the period 2001 to 2010. It clearly demonstrates the dramatic expansion in the number of voucher schools, from around 150 at the beginning of the period to close to 500 by the end of the decade. The share of upper secondary students enrolled in voucher schools also exhibits a large increase, though less dramatic in proportionate terms than for the total number of schools, a consequence of the fact that voucher schools are on average substantially smaller than their public counterparts.

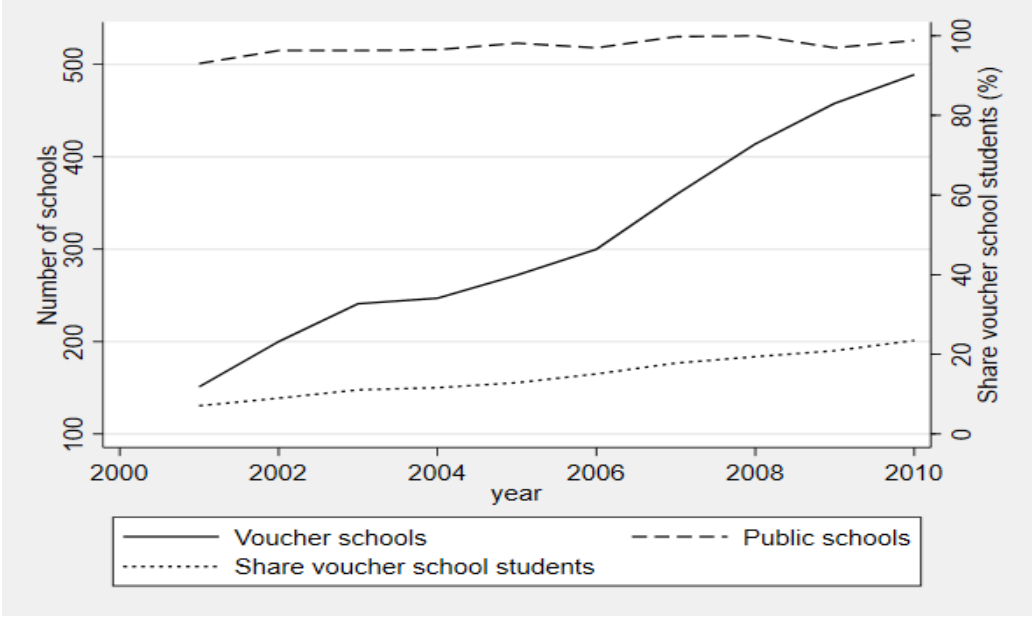
Figure 2 shows the share of grade-10 student enrolment by track type.[8] There is an upward trend in the share enrolled in the vocational tracks over most of the period under investigation, while the academic track share declines.[9] The figure also shows an increase in the preparatory track share. The preparatory track is only very rarely offered by voucher schools over this time period, and we will therefore focus on the vocational and academic tracks in this paper. Finally, note that the distribution of students across the five major tracks (the three

[8] The voucher schools are free to decide what mix of tracks to offer. Track supply in the public schools is decided by the municipalities. They shall take into account factors such as student application patterns, labour market needs and the scope for efficient resource use in their decisions, see e.g. Chapter 4 of government report SOU (2020) for an overview of the regulation

[9] It can be noted that the vocational track share has declined in the 2010's. It is believed that this is due to a reform, proposed by the government in 2008/9 and implemented in 2011, which reduced the theoretical content of the vocational tracks, which in turn reduced student's chances of qualifying for university via the vocational track.

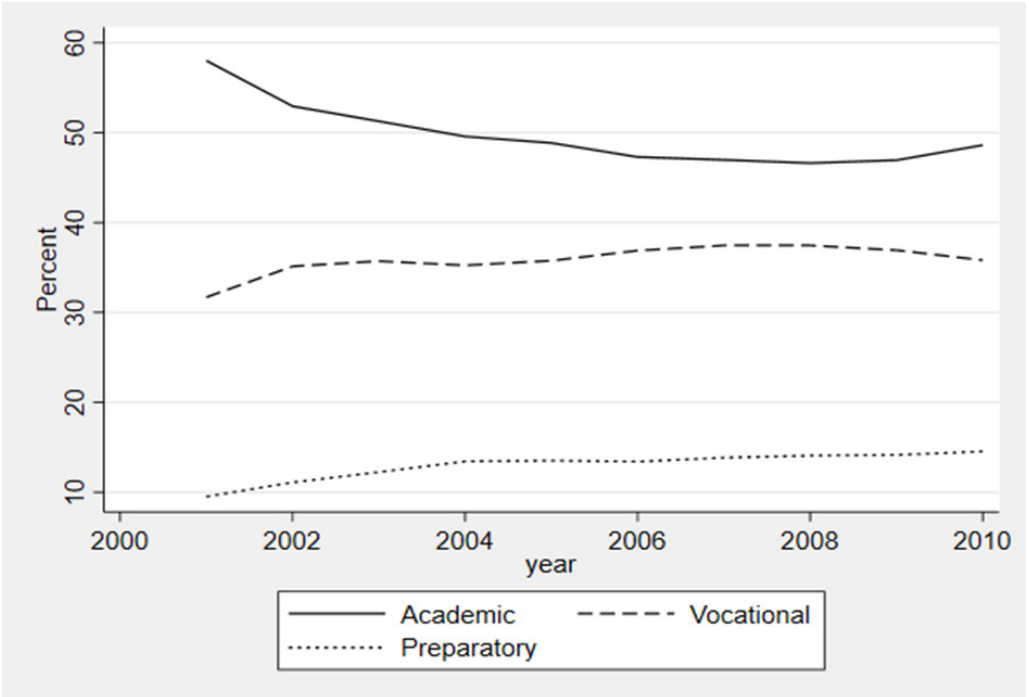
vocational tracks and two academic tracks) are broadly similar across voucher schools and traditional public schools (see Appendix B2 for further details).

Figure 1 Voucher and public upper secondary schools: number of schools and share of entry-level (grade 10) students



Notes: The voucher school status is based on information in the upper secondary school register, and the share of students in voucher schools is calculated by linking the voucher school status to the upper secondary attendance register.

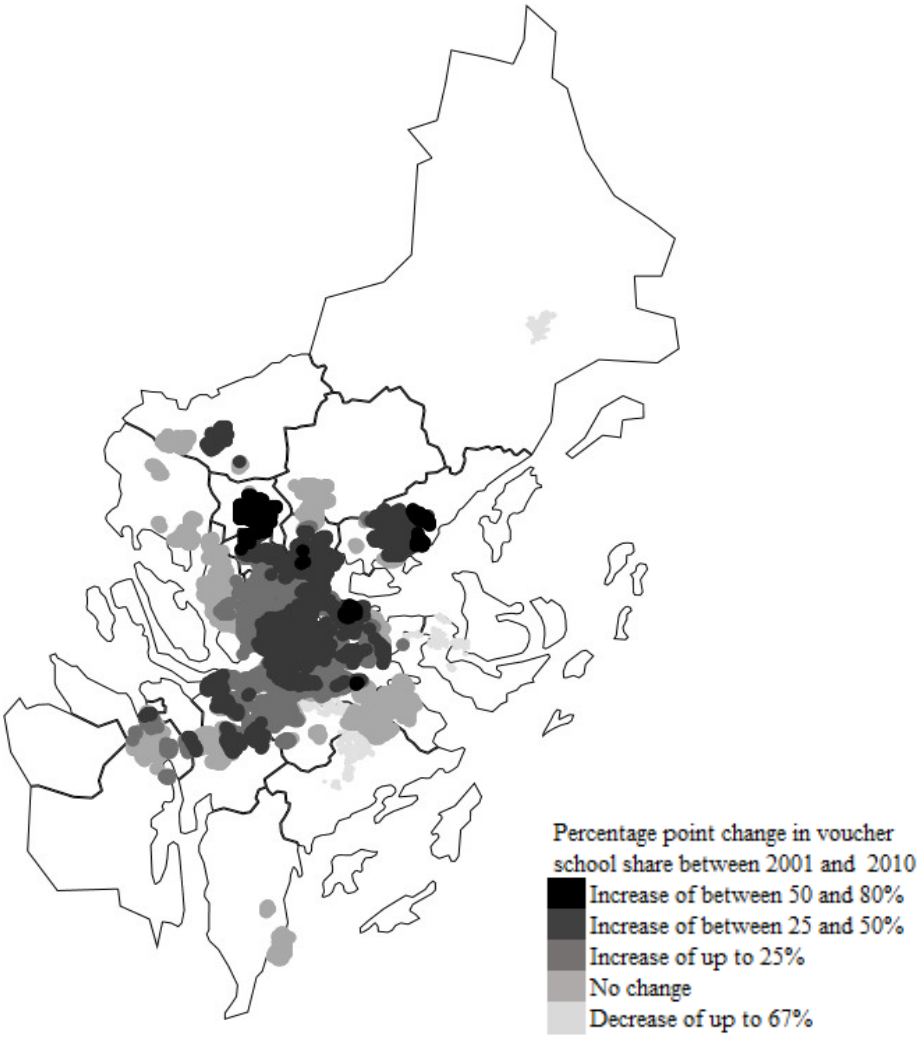
Figure 2 Figure 2 Share of entry-level students enrolled in academic, vocational and preparatory tracks



Notes: The figure is based on information on educational track of students attending the first grade of upper secondary school from the upper secondary attendance register.

As explained in greater detail in the next section, the treatment variable is the share of voucher schools within a 5km radius of the centroid of 250m-square grid cells across the whole of Sweden. As an example, Figure 3 shows a heat map depicting the growth in voucher schools over the period 2001 to 2010 for the municipalities in Stockholm county.

Figure 3. Heat map Stockholm county



Notes: The figure depicts the change in the voucher school share between 2001 and 2010, measured within a 5km radius around each grid cell, in the municipalities of Stockholm county. Each dot represents a grid cell, and darker shading represents a larger increase in voucher school share. The white areas denote grid cells where there was no student residing in the period under study, and that are as a result omitted from our regression sample.

2.2 Data

We employ various registry data held by Statistics Sweden.[10] We observe information on students' educational careers starting from the last year of lower secondary education and throughout upper secondary school. The registry data linkages also provide us information relating to individuals' longer-term outcomes in the form of university attendance status, field of study, as well as subsequent early labor market outcomes, including income and employment status. These are measured at age 30, except for university attendance, which is based on age 18-25. The data furthermore include a large set of demographic and family level background characteristics, including age, gender, parental education levels, country of birth (aggregated to larger regions), and household disposable income. We are able to link parents and children, and thus we can identify siblings. The school level data include information on number of and qualifications of the staff, and whether or not the school is a voucher school.

With respect to geographical location, we have information on the precise 250m- square grid cell in which the student resides; we observe the same information for the location of each school. The grid cells are, for confidentiality reasons, not available for very rural regions, which means that these are omitted from the analysis. Grid cells are missing for around 15% of all schools, partly due to the omission of very rural regions, and partly due to a failure to link some school addresses.[11] For the long-term education and labour market outcomes measured at age 25 or 30, we make use of another data source, which lacks the detailed grid cell geographical information; those estimations are instead based on municipal level variation in the voucher school share across siblings (discussed below). Appendix Figure A1 shows the locations of the voucher schools in our data in 2000 and 2010. It illustrates that the rapid expansion of the voucher schools over this period correlates with population density – schools opened up in all

[10] More detailed information on the data is available in the Appendix to this paper. Note that for the analysis of long-term student outcomes, we were given access to data spanning a longer period than was available in our original project data set, through the Institute for Evaluation of Labour Market and Education Policy (IFAU).
[11] See the supplementary data Appendix for details.

parts of the country but did so more rapidly in the most populous areas.

Table 2 provides summary statistics for our key variables. Panel A shows background or baseline student characteristics; panel B shows upper secondary education outcomes; and panel C summarises education track choices. Comparing the full sample with our siblings estimation sample (columns 1 and 2) reveals small positive selection in the siblings sample. For example, log household income is 10 log points higher for the siblings sample relative to the full sample. Siblings also enjoy a small advantage in upper secondary outcomes. There is little difference between the full sample and the siblings sample in upper secondary education tracks attended (panel C). Comparing the brothers sample with the sisters sample (columns 3 and 4) shows that females have substantially higher secondary grades (panel A) and upper secondary outcomes (panel B). As would be expected, there are substantial differences between females and males in upper secondary tracks (panel C). We discuss some of these differences in greater detail in the results sections below.

Table 1 Descriptive statistics

	All students	All siblings	Brothers	Sisters
Panel A: Student background characteristics				
Standardized final grade	0.28	0.36	0.21	0.51
lower secondary education	(0.74)	(0.72)	(0.69)	(0.72)
Male	0.51	0.51	1.00	0.00
	(0.50)	(0.50)	(0.00)	(0.00)
Swedish born	0.92	0.93	0.93	0.93
	(0.28)	(0.25)	(0.25)	(0.26)
Parent high education	0.54	0.58	0.58	0.57
	(0.50)	(0.49)	(0.49)	(0.50)
Log of household	12.87	12.97	12.97	12.96
disposable income	(0.48)	(0.44)	(0.44)	(0.44)
Mother employed	0.88	0.89	0.89	0.88
	(0.33)	(0.31)	(0.32)	(0.32)
Father employed	0.89	0.91	0.91	0.90
	(0.31)	(0.29)	(0.29)	(0.30)
Panel B: End of upper secondary schooling outcomes				
Graduate on time	0.76	0.80	0.78	0.82
	(0.43)	(0.40)	(0.42)	(0.38)
Percentile of final GPA	0.52	0.55	0.49	0.61
	(0.29)	(0.29)	(0.28)	(0.28)
Panel C: Educational track				
Vocational Industry/Tech	0.16	0.15	0.25	0.04
Vocational Trade/Admin	0.10	0.09	0.06	0.13
Vocational Nursing/ Care	0.06	0.06	0.03	0.09
Academic STEM	0.24	0.26	0.31	0.21
Academic Soc sci / Arts	0.43	0.43	0.33	0.52
Number of students	605807	280391	83812	76539

Note: Table shows the shares or averages of the variables for first year upper secondary students 2001-10, with standard deviations in parentheses, based on the regression sample: it includes students who qualify for the regular upper secondary track without first taking a preparatory track, and who have at least one school within a 5 km radius. The siblings samples (All siblings, Brothers, Sisters) include only families who are observed as residing in the same gridcell when all siblings started upper secondary school.

2.3 *Voucher School Expansion*

Voucher school expansion is unlikely to proceed in an idiosyncratic or random fashion. In order to evaluate this proposition, we investigate the relationship between the 10-year change in voucher school availability in a neighbourhood and characteristics of students in the neighbourhood at the start of the rapid expansion process, i.e. 2001. Specifically, we run the following regression:

$$\Delta VoucherShare_{gt} = \alpha_0 + \alpha_1 C_g^{2001} + \varepsilon_{gt}\,, \qquad (1)$$

where $VoucherShare_{gt}$ is the proportion of upper secondary voucher schools in a radius of 5km from the centre of the grid cell g, and Δ signifies the long difference between 2001 and 2010. This difference is regressed on C_g^{2001}, the mean of characteristics of 16-year old students and their families in the 5km circle, also centred at g, in 2001 (the 'baseline'). These characteristics comprise log of disposable family income, proportion of students with at least one parent with a post-secondary degree, proportion of students with at least one Swedish born parent, log of the number of 16-year olds, and the log of students' grade sum at grade 9 (i.e. prior to upper secondary school entry). The mean for the dependent variable, i.e. the change in voucher school share over the 2001 to 2010 period at the grid cell level, is 0.24 (s.d. = 0.21).

Table 3 presents estimates for model (1). The results show that voucher school expansion is related negatively to 2001 income in the local neighbourhood but positively to parental education and the proportion of parents who are Swedish born. Voucher schools also expand more where initial student performance is lower as well as in more densely populated areas.[12] In sum, we see that voucher school expansion is not random, and in order to identify its

[12] The positive correlation with parental education is in line with the findings in Edmark (2019). Edmark (2019) however finds a negative correlation with the share of students who have a Swedish background. The difference with our results may be due to the fact that Edmark (2019) studies the location patterns of lower primary private schools, and covers a different time period. It can also be noted that the Table shows the associations conditional on the other included variables. If we rerun separate regressions for each of the neighbourhood variables, without

causal impact, we require an identification strategy that takes this into account.

Table 3 Correlates of Voucher School Expansion

(Dependent variable: change in voucher school share, 2001 to 2010, RHS-variables averages within 5km radius from grid cell)

	Δ Share Voucher
Log of household disposable income	-0.3264***
	(0.0087)
Parent high education	0.6782***
	(0.0174)
Parent Swedish born	0.0369**
	(0.0160)
Log of average grade sum	-0.0417
	(0.0277)
Log of population density	0.0337***
	(0.0021)
Observations	38,889
R-squared	0.1053

Notes: The regression excludes grid cells with no upper secondary school within a 5km radius, as well as grid cells where there was no student residing at any point during 2001-10. Standard errors (in parentheses) are clustered at the grid cell level. *** p<0.01, ** p<0.05, * p<0.1.

3 Empirical Strategy

Our empirical strategy exploits the rapid expansion of Swedish voucher schools. In order to account for the non-random nature of this expansion as described above, we exploit within family variation over time in the availability of voucher schools, comparing school choices and other outcomes across siblings. Thus, this family fixed effect design accounts for time-

including the other as covariates, we obtain positive associations between voucher school share increase and household income, and student grades, and a negative association with the share of students with Swedish-born parents. That the pattern changes is not surprising, given that many of these factors are correlated. The point however remains; that voucher school expansion was not random.

invariant family traits and assess whether the changing supply of voucher schools has an impact on take-up of voucher schools, track choice and other outcomes. Exposure to voucher schools is measured by the proportion of voucher schools in a 5km radius from the family home, located via the grid cell information.

Note that our empirical strategy does not rely on, for example, between sibling differences in take-up of public versus voucher schools in order to assess the impact of enrolling in a voucher school. Rather, we compare choices and outcomes for siblings exposed to greater or fewer voucher schools in their neighbourhoods. An important issue confronting any family fixed effect research design is that variation in treatment across siblings may not be exogenous, i.e. treatment may be correlated with individual level unobservable traits, which in turn drive the outcome of interest. We provide a set of balancing tests which assuage this concern. In particular, we demonstrate that the treatment of interest, availability of voucher schools in a given radius around the family's home, is in general uncorrelated with either individual sibling covariates or with time-varying family characteristics. Furthermore, we show that results are little affected whether these covariates are included or not included in the regression models.

We estimate models of the following form:

$$y_{ifgt} = \alpha + \delta VoucherShare_{fgt} + \beta_1 X_{if} + \beta_2 W_{ft} + \beta_3 V_{mt} + \beta_4 Z_{gt} + \rho_f + \mu_t + u_{ifgt}\,, \quad (2)$$

where y_{ifgt} is an outcome (such as vocational or academic track choice, GPA, or university field or major) for student i, from family f, residing in grid cell g in year t. $VoucherShare_{fgt}$ is the proportion of upper secondary voucher schools in a radius of 5km from the centre of grid cell g, for family f residing in grid cell g, in year t. ρ_f is the family (strictly speaking, mother) fixed effect. X_{if} are predetermined student characteristics, including student's prior (grade 9) grade scores, whether the student finished lower secondary schooling (ninth grade) in a voucher school, student gender, age, sibling birth-order rank (measured among the

siblings in the regression sample), whether born in Sweden and whether born in Europe/North America. W_{ft} are time-varying family covariates, including parental employment status and household disposable income. V_{mt} are municipality-by-year covariates, which include the share of students attending a voucher school in grade 9, measures of per-student municipality funding for students in public schools and whether the municipality has a left-wing local political majority. Finally, Z_{gt} captures the local demand for upper secondary schooling, defined as the log of the number of 16-year olds residing within 5 km from the centre of grid cell g in year t. All regressions include year effects, μ_t. Standard errors are clustered at the family level. The parameter of interest is δ, the impact of the availability of voucher schools on the outcome y.[13]

We estimate this model on the sample of non-moving families, so that the impact is driven by expansion of voucher schools for a given geographical area. We also report results for the full sample, i.e. including families which move home, in the robustness section below. Results are little changed when we include these mobile families. We also assess the robustness of our results with and without time-varying individual student-level covariates.[14]

3.1 Balancing Test

Our identification strategy relies on the assumption that, for example, two pairs of siblings growing up in two different neighbourhoods experiencing differential growth in voucher schools would have experienced the same changes in the outcome if the growth in voucher schools had been the same across the two neighbourhoods.

[13] In an earlier version of this study (Edmark et al., 2020) we experimented with alternative empirical strategies, which included exploiting variation in the availability of voucher schools over time, but without controlling for family fixed effects. In our view, the identification strategy described above focuses on the most credible research design. Note that this earlier version also relied on outcomes measured at age 22, rather than up to age 30 in the current study.

[14] Our primary purpose is to assess the sensitivity of our results to including such covariates. However, a recent literature has also pointed out potential pitfalls in difference-in-differences models under a conditional common trends assumption (see for example, Caetano et al. (2022) and Sant'Anna and Zhao (2020)).

Although we can knock out neighbourhood (and family) fixed effects using our research design, this identification assumption will be violated if, for instance, voucher schools selectively sort into neighbourhoods experiencing declines (improvements) in characteristics such as public school quality or family circumstances. Under these circumstances we might expect that the younger sibling has worse (better) outcomes relative to the older sibling, independently of voucher school expansion.

In order to help assess the importance of such threats to identification, we now present balancing tests in support of our identification assumptions. Specifically, we test for correlation between the treatment variable and sibling characteristics. We evaluate whether the rise in the supply of voucher schools over time is unrelated to differences in sibling characteristics, both at the individual-level (for example, lower secondary grade score), as well as with respect to family-by-time covariates measured just before upper secondary school entry (for example, family income). Passing this balancing test lends credibility to the assumption that observable and/or unobservable differences between siblings do not drive our results. If, on the other hand, the variation we exploit is confounded by selective location of voucher schools to particular neighbourhoods, then we might expect to detect a relationship between our treatment variable and student or family background characteristics.

In order to carry out this balancing test, we estimate regressions of the following form:

$$x_{ifgt} = \alpha + \delta VoucherShare_{fgt} + \rho_f + \mu_t + u_{ifgt}\,, \qquad (3)$$

where x_{ifgt} is a pre-determined student or family characteristic, such as prior (grade 9) education performance or family income, for student i, from family f, residing in grid cell g in year t, measured the year before upper secondary school entry. $VoucherShare_{fgt}$ is the treatment variable, the share of voucher schools within 5km of the student's home, ρ_f are family fixed effects and μ_t are year effects.

The results from this exercise are shown in Table 4. Panel A reports results for the full

siblings sample, panel B for the brothers sample and panel C for the sisters sample. The results from this table demonstrate that for all three sets of estimation samples, changes in the share of voucher schools across siblings over time are uncorrelated with differences in siblings' grade scores, gender (panel A), whether born in Sweden, as well as family covariates, measured the year before upper secondary enrolment, namely log of household disposable income and father's employment status. There is evidence of positive correlation with mother's employment status (panels A and C). There is also evidence of a positive correlation between the treatment variable and being enrolled in a private lower secondary voucher school (panels A and B). This is unsurprising given that upper secondary schools are likely to collocate with lower secondary voucher schools.

At the municipality-level, there is no evidence of any relationship with per capita educational expenditure, but there is a positive association with the size of the 16-year old cohort. Finally, for the full sample (panel A) there is evidence of a negative relationship between voucher school expansion and the presence of a left wing municipal majority.

Overall, this set of balancing results demonstrates that there is little evidence of correlation between the predetermined covariates and the treatment variable after we account for family and year fixed effects. However there is some evidence of correlation between the treatment variable and mother's employment status, enrolment in a voucher school prior to upper secondary enrolment, and a left wing municipal majority. In order to assess the importance of these differences, we test for the robustness of our findings to the inclusion and exclusion of observable covariates, as well as the inclusion of labor market region-by-year effects and municipality-specific linear time trends. In practice we find that our key findings are robust to these specification checks.

Table 4 Balancing Test for Siblings Sample

	(1) Coefficients	(2) Standard Error	(3) Nr. Obs.
Panel A: All siblings			
Standardized final grade lower secondary education	0.0022	(0.0147)	279 402
Male	0.0130	(0.0136)	279 402
Swedish born	0.0001	(0.0039)	279 384
Log of household disposable income	-0.0007	(0.0056)	276 811
Mother employed	0.0112*	(0.0058)	277 917
Father employed	-0.0006	(0.0054)	272 192
Enrolled in voucher school grade 9	0.0304***	(0.0052)	279 402
Municipality per student educational costs	-44.5313	(72.7267)	278 328
Municipal council has left wing political majority	-0.0158**	(0.0074)	279 402
Log of 16-year old population within 5km radius	0.0132***	(0.0018)	279 402
Panel B: Brothers			
Standardized final grade lower secondary education	-0.0125	(0.0244)	83 216
Male	—	—	—
Swedish born	0.0025	(0.0076)	83 213
Log of household disposable income	-0.0053	(0.0102)	82 464
Mother employed	-0.0079	(0.0108)	82 744
Father employed	-0.0037	(0.0098)	81 075
Enrolled in voucher school grade 9	0.0407***	(0.0095)	83 216
Municipality per student educational costs	-2.2072	(134.7319)	82 910
Municipal council has left wing political majority	-0.0221	(0.0138)	83 216
Log of 16-year old population within 5km radius	0.0106***	(0.0034)	83 216
Panel C: Sisters			
Standardized final grade lower secondary education	-0.0050	(0.0280)	76 045
Male	—	—	—
Swedish born	-0.0075	(0.0076)	76 042
Log of household disposable income	-0.0141	(0.0108)	75 352
Mother employed	0.0319***	(0.0119)	75 654
Father employed	-0.0005	(0.0106)	73 987
Enrolled in voucher school grade 9	0.0147	(0.0101)	76 045
Municipality education costs per student	-177.7841	(141.3140)	75 736
Municipal council has left wing political majority	-0.0160	(0.0144)	76 045
Log of 16-year old population within 5km radius	0.0164***	(0.0034)	76 045

Note: Each row corresponds to a separate regression, with the listed variable as dependent variable, and the voucher school share, family, sibling order and year fixed effects are RHS variables. The table shows the coefficients for the voucher school share and their standard errors. Standard errors are clustered at the family (mother id) level. *** p<0.01, ** p<0.05, * p<0.1.

4 Impact on Upper Secondary Tracks

In this section we ask how the availability of nearby voucher school options affects students' upper secondary track choices at age 16. This analysis addresses the question whether the take-up of voucher schools simply represents a switch away from public schools to voucher schools *of the same track*, or whether the introduction of voucher schools in fact leads to a different set of track choices than would have been the case in the absence of voucher school expansion. Note that the results we present may be a consequence of voucher schools providing a different mix of track options, or they may arise through spillover effects from voucher schools to public schools. Our identification strategy cannot directly test for the importance of these various mechanisms in driving the reduced form results.

In order to undertake this analysis, we ask whether enrolling in a given track, be it at a traditional public school or voucher school, is influenced by the increased availability of voucher schools. Table 5 reports the results from this analysis. In column 1 of Table 5, the outcome variable is enrolment in a vocational track (public or voucher). Columns 2 to 6 provide a more detailed analysis by breaking down the tracks into three vocational sub- categories (industry and technology; trade and administration; nursing and care) and two academic sub-categories (science; and social science and arts). Each column in each panel of Table 5 represents a separate regression, and the estimation model is as described by equation (2). The outcome, vocational track choice, is set to one if the student chooses the vocational upper secondary track and equals zero if the choice was an academic track. The table reports the estimate of the coefficient on the proportion of voucher schools within a 5km radius of the student's home (δ in equation (2)).

Panel A of Table 5 reports results for the full siblings sample. Panels B and C report results for the sample of brothers and sisters, respectively.

Table 5 Impact on Upper Secondary Track

	(1)	(2)	(3)	(4)	(5)	(6)
		Vocational			Academic	
	Vocational	Industry/Tech	Trade/Admin	Nursing/Care	Science	Social Science/Arts
Panel A: Full siblings sample						
Voucher Share	0.0239**	0.0251***	-0.0036	0.0023	-0.0350***	0.0101
	(0.0104)	(0.0089)	(0.0077)	(0.0062)	(0.0095)	(0.0119)
Observations	280,391	280,391	280,391	280,391	280,391	280,391
Number of Mother FEs	131,028	131,028	131,028	131,028	131,028	131,028
Panel B: Brothers						
Voucher Share	0.0375**	0.0527***	-0.0149	-0.0002	-0.0499***	0.0087
	(0.0191)	(0.0186)	(0.0122)	(0.0087)	(0.0182)	(0.0205)
Observations	83,812	83,812	83,812	83,812	83,812	83,812
Number of Mother FEs	40,628	40,628	40,628	40,628	40,628	40,628
Panel C: Sisters						
Voucher Share	0.0337*	0.0157	0.0122	0.0057	-0.0420**	0.0054
	(0.0196)	(0.0110)	(0.0169)	(0.0140)	(0.0170)	(0.0238)
Observations	76,539	76,539	76,539	76,539	76,539	76,539
Number of Mother FEs	37,065	37,065	37,065	37,065	37,065	37,065

Notes: "Voucher Share" measures the share of upper secondary voucher schools within a 5 km radius from the student's home. All regressions include year, sibling order and family (mother) fixed effects as well as the following student-level covariates: lower secondary school GPA (final grade sum) in level and square; male dummy; attended voucher lower secondary school; born in Sweden; and born in Europe (except Sweden) or North America. All regressions also include the following time-varying family covariates: mother employed; father employed; household disposable income in level and square. Municipality-level variables included in all regressions are: expenditure per student on compulsory education (primary and lower secondary education); the share of students attending private school in grade 9 (lower secondary); dummy for the municipality having a left-wing political local majority. Finally, regressions also include, at the grid cell level, the log of the number of age-16 youth residing within 5km from the grid cell midpoint. All time-varying covariates are measured the year before the student enters upper secondary school, except for the first year of the panel, when the current year's values are used. Missing values in covariates were replaced by the average value of the variable, and dummy variables were added indicating the missing instances. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same grid cell measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

The result in column 1, panel A, suggests that for a sibling experiencing a 10 percentage point greater exposure to voucher schools, there is a 0.2 percentage point rise in the probability of that sibling enrolling in a vocational track school (public or voucher). This effect is statistically significant at the 5% level. Although this reduced form result may appear to be relatively small, recall that voucher schools are significantly smaller than public schools, and hence even a relatively large increase in the share of voucher schools may lead to only a small change in enrolment. For example, Figure 1 demonstrates that although the share of voucher

schools rose from around 20 percent to nearly 50 percent, the share of students enrolled in voucher schools rose by just over 10 percentage points.[15]

For the brothers and sisters samples, Panels B and C, respectively, the results in column 1 reveal that the impact on the take-up of the vocational track is similar by gender.

We now turn to the detailed breakdown of choices within the vocational and academic tracks, which reveal very interesting substitution patterns. First, the results in column 5 demonstrate that there is a decline in take-up of the science academic track as a consequence of voucher school expansion, for the full sample as well as for the brothers and sisters samples (i.e. panels A through C). The results across all three sets of estimation samples are highly statistically significant.

The full sample results in Panel A suggest that there is substitution away from the science academic track in favor of the industry and technology track (column 2, Panel A). Panels B and C suggest that this substitution pattern is driven by males: there is a strong and statistically significant rise in the industry and technology track option for males, but for females, the impact appears to be much more muted and is not significant.

For the sisters sample, a 10 percentage point rise in the voucher school share leads to a 0.4 percentage point decline in the probability of enrolling in a science academic track. For brothers, a similar rise in the voucher school share leads to a 0.5 percentage point decline in the probability of enrolling in a science academic track. Whilst males substitute into the industry and technology vocational track, there is no evidence that females exhibit a similarly strong substitution pattern. Below we explore post-secondary education and labor market outcomes, and then link back to these findings related to the science and vocational track options.

Overall, the key findings in Table 5 are that voucher school penetration leads to: (i) a

[15] In Appendix Table A1 we also investigate the effect of increasing private school supply on private school enrolment. The results (Table A1, column 1) show that a 10 percentage point rise in the proportion of private schools results in a 0.8 percentage point rise in the probability of attending a private school. The table also reports the results using alternative specifications. All results are consistent with the baseline specification in column 1.

rise in the take-up of the vocational track option, for males as well as females; (ii) a fall in enrolment in the science academic track, also for males and females; and (iii) a rise in the industry and technology vocational option, but for males only.

Robustness Checks

We present a set of robustness checks for our core set of track results. We first assess the sensitivity of our results to the inclusion or exclusion of the control variables. We compare estimates with and without covariates which vary across siblings. Appendix Table A2 shows that when we exclude these covariates, the results for track choice are virtually identical to those reported in Table 5, where covariates are included.

Next, we relax the sample restriction which excludes mobile families. Appendix Table A3 includes mobile families. The results are largely robust to this sample selection choice, with the negative impact on science track consistent with the main set of results, as is the overall positive vocational track result for the full sample and the sisters sample, as well as the positive industry and technology outcome for the brothers sample. There is some evidence to suggest that the industry and technology option also experiences a rise for the sisters sample (column 2, Panel C).

In Appendix Tables A4 and A5 we include labor market region-by-year effects and municipality-specific linear time trends, respectively.[16] The addition of regional or municipal trends has the advantage of helping control for other local changes to the education system. For instance, some municipalities expanded their school admission regions over this time period, typically by replacing previous proximity-based upper secondary placement algorithms by grade-based school admission. In some regions, municipalities merged to form joint admission

[16] Note that the local labor market regions are defined by Statistics Sweden with the aim of connecting municipalities between which there is substantial commuting. Since commuting patterns change over time, Statistics Sweden generate new local labor market regions each year. We use the 2001 version. In 2001 there were 88 local labor market regions (compared to approximately 290 municipalities).

areas.[17] The results in Appendix Tables A4 and A5 generally support our main conclusions. In Table A4 the impact on vocational track choices overall is not significant, though it is significant for males at the 10% level (Table A4, column 1). Other results in this table are line with our main set of results, including the decline in the take-up of the science academic track and the rise in the vocational technology track. The results in Table A5, which includes municipality-specific linear time trends, are consistent with our baseline estimates.

Appendix Tables A6 and A7 use 10km and 20km radii from the student's home location to measure the penetration (proportion) of voucher schools, respectively. The results from these two tables suggest that our main conclusions are generally robust to these changes in the definition of the treatment variable, although the impact on the overall vocational track choice (column 1 in both tables) is no longer statistically significant. This last result is a consequence of larger standard errors and a small (e.g. 20% for the overall vocational take- up in Table A6, column 1, Panel A) decline in the estimated coefficients. These results would suggest that the narrower, 5km, distance measure we use in our baseline specification is more salient or relevant to students than the larger radii of 10km and 20km. The results for the science academic track choice in Tables A6 and A7 are in line with our baseline results.

One concern with the family fixed effect approach is that parental human capital investments may differ across children within the family. For example, there is mixed evidence on whether parents compensate or reinforce for endowment differences among their children (see for example, Bharadwaj et al., 2018). Furthermore, there is evidence to support the hypothesis that there are human capital spillovers between siblings (Altmejd et al., 2021). Although we cannot fully account for such issues with the data at our disposal, we run regression estimates of our main model without family fixed effects. Results are reported in Table A15.

[17] These changes applied to the municipality operated schools – voucher school admission at the upper secondary level was grade based with no proximity priority throughout the period. See Molin (2019) and Sund (2018) for further details of the allocation rules.

This shows that although there are some relatively minor quantitative differences in the estimates when family fixed effects are excluded, the main conclusions we draw from our analysis remain unchanged.

Overall, these checks broadly support our conclusions from the core set of results reported earlier. The evidence suggests that the impact on vocational track choices, including within sub-categories such as the choice of the science track option and the industry and technology option, largely survive the battery of checks we apply.

Heterogeneous Impact

In this section we focus on how the treatment impact varies by parental education as well as minority or national origin status. In Table 6, we divide the sample up by high or low parental education background and report a separate set of results for each subsample in panels A to D, respectively. High parental education is defined as at least one parents having completed a post-secondary education qualification.

The results suggest that there is a substantially stronger response to voucher school expansion on vocational track take-up for students with lower parental education (column 1). Substitution patterns between academic science and vocational industry / technology tracks are in line with previous results for both groups (columns 2 and 5 in Panels A and B), although the effects are somewhat stronger for students from the lower parental education background.

Turning to impact by national origin status, we proxy this by whether at least one parent is born outside Sweden (the 'minority' group) or whether both parents are Swedish- born (the 'majority' group). The results in panels C and D of Table 6 suggest that there is a substantially larger response in terms of vocational track take-up for the minority group. There are also interesting contrasts in terms of take-up of the vocational trade / administration track (negative impact for the majority group but large positive impact for the minority group). There is no evidence to suggest that the minority group increases take-up of vocational industry / technology

track. The impact on the academic science track, on the other hand, is negative and statistically significant for both sets of students.

Table 6 Upper Secondary Track: Heterogeneous Impact

	(1)	(2)	(3)	(4)	(5)	(6)
		Vocational			*Academic*	
	Vocational	Industry/Tech	Trade/Admin	Nursing/Care	Science	Social Science/Arts
Panel A: Parents high education						
Voucher Share	0.0056	0.0227**	-0.0089	-0.0083	-0.0279*	0.0195
	(0.0132)	(0.0109)	(0.0088)	(0.0071)	(0.0143)	(0.0166)
Observations	156,906	156,906	156,906	156,906	156,906	156,906
Number of Mother FEs	73,293	73,293	73,293	73,293	73,293	73,293
Panel B:Parents low education						
Voucher Share	0.0407**	0.0350**	-0.0067	0.0126	-0.0433***	0.0031
	(0.0171)	(0.0148)	(0.0137)	(0.0109)	(0.0128)	(0.0177)
Observations	111,376	111,376	111,376	111,376	111,376	111,376
Number of Mother FEs	52,306	52,306	52,306	52,306	52,306	52,306
Panel C: Both parents Swedish-born						
Voucher Share	0.0145	0.0272***	-0.0165*	0.0039	-0.0370***	0.0220
	(0.0119)	(0.0102)	(0.0087)	(0.0071)	(0.0108)	(0.0135)
Observations	210,129	210,129	210,129	210,129	210,129	210,129
Number of Mother FEs	98,878	98,878	98,878	98,878	98,878	98,878
Panel D: At least one parent born abroad						
Voucher Share	0.0587***	0.0198	0.0369**	0.0021	-0.0349*	-0.0255
	(0.0218)	(0.0180)	(0.0167)	(0.0130)	(0.0202)	(0.0255)
Observations	68,981	68,981	68,981	68,981	68,981	68,981
Number of Mother FEs	31,571	31,571	31,571	31,571	31,571	31,571

Notes: "Voucher Share" measures the share of upper secondary voucher schools within a 5 km radius from the student's home. All regressions include year, sibling order and family (mother) fixed effects as well as the same covariates as Table 5. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same grid cell measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

5 Education and Labor Market Outcomes

We now turn to the impact of voucher school expansion on students' short- and long-term outcomes. We first discuss the impact on upper secondary school education outcomes (graduate on time and GPA) and then move on to longer-term outcomes (post-secondary education and labor market outcomes).

5.1 Graduate on Time and GPA

Applying the same empirical approach as for the track choice outcomes, we assess the impact of voucher school expansion on two upper secondary school education outcomes: whether students graduate on time from upper secondary school; and their percentile rank on their GPA score. Table 7 reports results from the estimation of model (2) for these two outcomes. Percentile rank on GPA score is calculated by cohort across the full population of students. Students are defined as graduating on time if they complete upper secondary school within three years from starting. These outcomes shall not be interpreted as "pure" measures of student ability; they also reflect the different subject content of the various tracks, as well as potential differences in grading standards across both schools and tracks.[18]

Table 7, Panels A through C report the results for the full sample of siblings, the brothers sample, and the sisters sample, respectively. The results suggest that in general there does not appear to be any impact of voucher school expansion on the likelihood of students graduating on time or the GPA score. One exception is the marginally significant positive impact for females on graduating on time (column 1, panel C).

In Appendix Table A8 we examine whether the treatment effect on the graduate on time and percentile GPA outcomes varies by parental education background or minority status (as proxied by whether at least one parent is born is outside Sweden). The results suggest that there is some evidence of a positive impact on graduating on time for students whose parents are relatively highly educated, as well as for those with at least one parent born abroad (column 1, panels A and D respectively).[19]

[18] There is ample evidence of more generous grading standards among Swedish voucher schools, see e.g. Hinnerich and Vlachos (2017) and Edmark and Persson (2021).

[19] Previous studies of the Swedish voucher reforms have found mixed evidence of any impact on test scores. Early work suggested relatively large effects (Sandström and Bergström, 2005), however, Böhlmark and Lindahl (2015) find modest positive effects. Recent work by Hinnerich and Vlachos (2017) has pointed to the possibility of artificial gains through grade inflation; Edmark and Persson (2021) use a value added framework and come to similar conclusions regarding grade inflation

Table 7 Impact on Student Graduation and Final Marks, Upper Secondary School

	(1)	(2)
	Graduate on time	Pctile GPA
Panel A: Full Siblings sample		
Voucher Share	0.0152	-0.0081
	(0.0098)	(0.0052)
Observations	280,391	253,862
Number of Mother FEs	131,028	119,076
Panel B: Brothers		
Voucher Share	-0.0080	-0.0066
	(0.0185)	(0.0092)
Observations	83,812	74,878
Number of Mother FEs	40,628	36,378
Panel C: Sisters		
Voucher Share	0.0302*	-0.0157
	(0.0183)	(0.0101)
Observations	76,539	69,491
Number of Mother FEs	37,065	33,706

Notes: "Voucher Share" measures the share of upper secondary voucher schools within a 5 km radius from the student's home. All regressions include year, sibling order and family (mother) fixed effects as well as the same covariates as Table 5. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same grid cell measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

We probe the robustness of the results reported in Table 7 by experimenting with the exclusion of covariates, inclusion of mobile families, inclusion of labor market region-by-year fixed effects as well as municipality linear trends, and varying the radii of the proportion of voucher schools from the student's home. The results are reported in the Appendix Tables A9 and A10. These show that the size of the coefficient estimates in almost all cases are very similar to those reported in the baseline case in Table 7. For some of the specifications there is evidence of a positive impact on graduating on time and a negative impact on percentile GPA rank; however, there is no consistent pattern of statistical significance across the specifications.

5.2 Long-term Analysis: Post-Secondary Education and Labor Market Outcomes

We next exploit longer-term data in order to examine upper secondary, university and labor market outcomes by age 30. As explained above, by using an alternative Swedish registry data source (see the data appendix for details), we are able to complement the above short- term analysis with the long-term impact on education and labour market outcomes. This data set lacks information on locations at the grid cell level. We therefore carry out this part of the analysis by exploiting municipality level variation in the voucher school share in combination with the sibling fixed effects model. The estimation equation is as follows:

$$y_{ifmt} = \alpha + \delta VoucherShare_{fmt} + \beta_1 X_{if} + \beta_2 W_{ft} + \beta_3 V_{mt} + \rho_f + \mu_t + u_{ifmt},$$

where m now indicates municipality so that voucher share variable varies at the municipality-level. All other variables are as defined for equation (2).

Column 1 of Table 8 estimates the impact of voucher school expansion on the likelihood that an individual's highest level of completed education at age 30 (which may be at the upper secondary or university level) is in a STEM subject. The results demonstrate that there is a statistically significant negative impact for the full sample (column 1, panel A). This suggests that a 10 percentage point rise in the local share of voucher schools at the time of enrolling in upper secondary school leads to a 0.2 percentage point decline in the probability that the highest level of completed education is in a STEM subject.

Undertaking the analysis by gender, the point estimate for the impact on the probability that the highest level of completed education is in a STEM subject is negative for both males and females (column 1, panels B and C), although it is more than twice as large for females than males. However, it is not statistically significant in either the male or female sample.

Investigating the impact on tertiary-level qualification in a STEM subject, column 2

reveals negative point estimates overall, substantially larger impact for females than males, but none of these is statistically significant at conventional levels.

Table 8 Impact on Long-Term Outcomes, Municipality Level Variation

	(1) Highest qualification is in STEM	(2) STEM subject tertiary/post-secondary degree	(3) Tertiary/Post-secondary degree	(4) University credits (by age 25)	(5) Employed	(6) Log employment income
Panel A: Full Siblings sample						
Voucher Share	-0.0189**	-0.0080	0.0076	-0.0007	-0.0155**	-0.0429
	(0.0079)	(0.0063)	(0.0077)	(0.0052)	(0.0062)	(0.0289)
Observations	413,684	413,684	420,083	841,063	420,099	420,099
Number of Mother FEs	191,539	191,539	194,387	371,009	194,394	194,394
Panel B: Brothers						
Voucher Share	-0.0047	-0.0063	-0.0194	-0.0058	-0.0189*	-0.0958*
	(0.0170)	(0.0132)	(0.0143)	(0.0090)	(0.0114)	(0.0540)
Observations	119,350	119,350	121,622	258,380	121,628	121,628
Number of Mother FEs	57,247	57,247	58,313	120,926	58,316	58,316
Panel C: Sisters						
Voucher Share	-0.0112	-0.0147	0.0211	0.0092	-0.0135	-0.0480
	(0.0114)	(0.0104)	(0.0151)	(0.0100)	(0.0127)	(0.0565)
Observations	112,821	112,821	114,304	240,043	114,306	114,306
Number of Mother FEs	54,108	54,108	54,809	112,512	54,810	54,810

Notes: "Voucher Share" measures the share of upper secondary voucher schools at the municipality level. All regressions include year, sibling order and family (mother) fixed effects as well as the following student-level covariates: lower secondary school GPA (final grade sum) in level and square (separate coefficients are estimated for final GPA before and after 1998, since a new grading system was introduced for lower secondary school in that year); male dummy; attended voucher lower secondary school; born in Sweden; and born in Europe (except Sweden) or North America. All regressions also include the following time-varying family covariates: mother employed; father employed; and household disposable income in level and square (household income enters in the form of separate values for each parents' individual component of the households' joint income). Municipality-level variables included in all regressions are: expenditure per student on compulsory education (primary and lower secondary education); the share of students attending private school in grade 9 (lower secondary); dummy for the municipality having a left-wing political local majority. Finally, regressions also include, at the municipality level, the log of the number of age-16 youth residing in the municipality. All time-varying covariates are measured the year before the student enters upper secondary school, except for the first year of the panel, when the current year's values are used. Missing values in covariates were replaced by the average value of the variable, and dummy variables were added indicating the missing instances. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same municipality measured the year they turn 16. The regression outcomes are measured at age 30, and the sample includes students starting upper secondary education in 1995-2005, except for the outcome "university credits (by age25)" which includes students starting upper secondary education in 1995-2009. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

There is no evidence of a statistically significant impact on the overall likelihood of obtaining a university degree (column 3), either for the full, brothers or sisters samples. This

would suggest that women may be substituting away from STEM subjects in favor of other disciplines, so that there is no net impact on the probability of women obtaining a university degree. The final post-secondary education outcome is a dummy variable for having taken any university course credits by age 25 (column 4).[20] The results for this outcome also suggest no significant impact arising from voucher school penetration.

We next turn to two labor market outcomes, measured at age 30: employment status and log employment income. The results suggest that there is a negative impact on employment status for the full siblings sample, significant at the 5% level. A 10 percentage point rise in the share of voucher schools leads to a 0.2 percentage point fall in the probability of being employed at age 30. The point estimate is negative and economically meaningful for males and females (column 5, panels B and C), although it is statistically significant only for males.

For log income (column 6), there is a marginally significant negative impact for males: a 10 percentage point rise in the local share of voucher schools at the time of enrolling in upper secondary school leads to a 1 percent decline in income, significant at the 10% level. The point estimates for the full sample (panel A) and the female sample (panel C) are negative and economically meaningful, but are not statistically significant.

We also undertake a series of robustness tests for the baseline long-term education and labor market outcome results produced in Table 8. The appendix Table A11 excludes covariates which vary across siblings. These results are virtually unchanged relative to the baseline results. Appendix Table A12 includes mobile families. Once again the key results reported earlier are not sensitive to this variation in the choice of estimation sample. Appendix Table A13 includes labor market region-by-year effects. This specification also supports the main set of results. One difference is that the negative impact on graduating in a STEM subject at university is now statistically significant for the overall sample as well as for females. Finally, when

[20] Note that all long-term outcomes are measured at age 30 *except* university credits, which are measured at age 25.

municipality-specific linear time trends are included (appendix Table A14), estimates of the impact on highest qualification in STEM as well as labor market outcomes are no longer significant. One interpretation of these results is that our baseline set of results are not robust to the inclusion of such time trends. However, we would caution against such a strict interpretation of this robustness check. In particular, it can be argued that given the lack of 'pre-treatment' data, the estimation of pre- existing trends is especially difficult in the current scenario.

Finally, Table 9 investigates heterogeneous response to voucher school penetration by parental education background as well as by Swedish-born status of the parents. This shows that the estimated impact of voucher penetration on the probability that an individual's highest level of completed education is in a STEM subject is negative across all four groups (column 1, panels A through D), but is statistically significant only for students with highly educated parents as well as minority students (i.e. those with at least one parent born outside Sweden).

Interestingly, the results demonstrate a statistically significant negative impact on graduating from university in a STEM subject for minority students (column 2, panel D in Table 9). This accords with the earlier evidence demonstrating that for this group of students, voucher penetration leads to lower take-up of academic science track, with no evidence of substitution into the vocational industry and technology track (see the discussion above relating to the evidence on track choice in Table 6).

The estimated impact on employment outcomes is also negative across all groups (columns 5 and 6), though it is only statistically significant for students with highly educated parents as well as the majority Swedish background group of students in the case of employment status (column 5, panels A and C) and for the majority group of students for log employment income (column 6, panel C).

Table 9 Heterogeneous Impact, Long-Term Outcomes, Municipality Level Variation

	(1) Highest qualification is in STEM	(2) STEM subject tertiary/post-secondary degree	(3) Tertiary/Post-secondary degree	(4) University credits (by age 25)	(5) Employed	(6) Log employment income
Panel A: Parents high education						
Voucher Share	-0.0296**	-0.0107	0.0095	0.0064	-0.0175*	-0.0498
	(0.0121)	(0.011+D33	(0.0112)	(0.0080)	(0.0096)	(0.0445)
Observations	185,184	185,184	187,359	388,979	187,363	187,363
Number of Mother FEs	86,061	86,061	87,036	173,042	87,038	87,038
Panel B:Parents low education						
Voucher Share	-0.0045	-0.0043	0.0036	-0.0047	-0.0098	-0.0188
	(0.0111)	(0.0078)	(0.0114)	(0.0073)	(0.0087)	(0.0405)
Observations	200,171	200,171	203,814	396,214	203,822	203,822
Number of Mother FEs	92,872	92,872	94,495	175,267	94,498	94,498
Panel C: Both parents Swedish-born						
Voucher Share	-0.0094	0.0009	0.0072	-0.0031	-0.0173***	-0.0522*
	(0.0086)	(0.0070)	(0.0084)	(0.0057)	(0.0066)	(0.0305)
Observations	335,864	335,864	340,785	676,075	340,796	340,796
Number of Mother FEs	155,905	155,905	158,111	299,088	158,116	158,116
Panel D: At least one parent born abroad						
Voucher Share	-0.0523**	-0.0538***	-0.0046	0.0022	-0.0114	-0.0503
	(0.0204)	(0.0163)	(0.0208)	(0.0134)	(0.0190)	(0.0897)
Observations	69,966	69,966	71,333	148,205	71,338	71,338
Number of Mother FEs	32,054	32,054	32,646	64,841	32,648	32,648

Notes: "Voucher Share" measures the share of upper secondary voucher schools at the municipality level. All regressions include year, sibling order and family (mother) fixed effects as well as the same covariates as Table 8. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same municipality measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

6 Alternative Mechanisms: Teacher, School and Peer Quality

The impact of voucher schools on student outcomes may be mediated via a number of mechanisms. Choice of academic or vocational track is one channel. In this section we investigate three alternative channels, focusing on school, teacher and peer quality. We implement the same model employed for the track choice analysis, using characteristics of the school the student is exposed to as the outcome variable, in order to assess how voucher penetration alters this feature for students.

Turning first to school characteristics, column 1 in Table 10 reports the impact of voucher school expansion on the student-teacher ratio, the number of students per teachers in the school. The estimates in column 1 imply that private school expansion leads to a *fall* in the student-personnel ratio that students are exposed to, i.e. voucher school expansion implies that students experience smaller class sizes. This could be an effect of a shift towards more teacher intensive vocational education. The impact appears to be stronger for the female sample than the for the male one (column 1, panels B and C).

Column 2 in Table 10 reports the impact of voucher school expansion on the share of qualified teachers, measured as the proportion of teachers, out of all teachers, who are qualified for the subject they teach. The results suggest that a rise in the proportion of voucher schools leads to a fall in the proportion of certified teachers that students are exposed to (panel A, column 2). This time the results are statistically significant for males, but not for females (panels B and C).

We next turn to the quality of peers that students meet at school (measured within the type of track the student attends).[21] We use three measures of peer characteristics: the proportion of peers with at least one parent attaining post-secondary education; the proportion of peers who are Swedish born; and the average GPA of peers at grade 9. The results are reported in the final three columns of Table 10. The results suggest that although voucher school expansion leads to students experiencing peers whose parents are less likely to have a post-secondary education and are less likely to be Swedish born, there is no significant impact on the average GPA of their fellow students.

Overall we conclude that voucher school expansion leads to ambiguous effects on school and peer quality experienced by students. There is some improvement in class size, teachers are less likely to be formally qualified, and some changes in peer quality, although

[21] Prior studies documenting the impact of voucher schools on sorting in Sweden include Björklund et al. (2005) and Böhlmark et al. (2016).

there is no significant change in the peer quality as measured by average GPA.

Table 10 Mechanisms

	(1)	(2)	(3)	(4)	(5)
	School characteristics		*Average peer characteristics in track type and school*		
	Pupils per teacher	Share Qualified Teachers	Share with a High-Educated Parent	Share Swedish-born	Stand. Final Lower-Secondary GPA
Panel A: Full Siblings sample					
Voucher Share	-0.3308***	-0.0221***	-0.0103***	-0.0122***	-0.0049
	(0.0728)	(0.0039)	(0.0036)	(0.0016)	(0.0097)
Observations	272,177	276,414	262,871	277,659	277,378
Number of Mother FEs	127,339	129,225	123,062	129,819	129,691
Panel B: Brothers					
Voucher Share	-0.0786	-0.0246***	-0.0213***	-0.0142***	-0.0269
	(0.1286)	(0.0073)	(0.0067)	(0.0028)	(0.0184)
Observations	81,175	82,628	78,392	83,001	82,902
Number of Mother FEs	39,374	40,057	38,031	40,247	40,199
Panel C: Sisters					
Voucher Share	-0.4595***	-0.0109	-0.0087	-0.0099***	-0.0111
	(0.1438)	(0.0070)	(0.0068)	(0.0030)	(0.0177)
Observations	74,283	75,384	71,311	75,633	75,529
Number of Mother FEs	35,991	36,515	34,577	36,634	36,583

Notes: Teacher characteristics (column 1 and 2) are measured at the school level, and peer characteristics (column 3–5) are measured at the school and track type level, where track type follows the 5 categories of vocational and academic tracks shown in Table 5. "Voucher Share" measures the share of upper secondary voucher schools within a 5 km radius from the student's home. All regressions include year, sibling order and family (mother) fixed effects as well as the same covariates as Table 5. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same grid cell measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

7 Conclusion

The role played by voucher schools in influencing students' education choices remains an underexplored area. This study makes an attempt to fill this gap. We focus on the role of voucher schools in shaping field of study among students attending upper secondary school in Sweden, as well as subsequent short- and long-term outcomes. These are important questions and addressing them helps shed new light on how voucher schools impact the education market, as well as improving our understanding of what determines adolescents' field choices. The empirical research design exploits the rapid expansion of voucher schools in Sweden, as well

as sibling comparisons.

Our findings suggest that increased voucher school penetration increases the likelihood that students enrol in vocation tracks, whether these are provided by traditional public or voucher schools. These results demonstrate that voucher school expansion is not neutral in its impact on the takeup of academic and vocational tracks. Detailed analysis of the impact on choices within tracks reveals that students substitute away from science academic tracks and instead select vocational and technology tracks. These substitution patterns appear to be driven by males.

One way to address whether or not such shifts in choices induced by voucher schools enhance welfare for student is to investigate their impact on longer-term education and labor market outcomes. Although we find that voucher school penetration leads to a reduced probability that an individual's highest qualification is in a STEM subject, we find no evidence of any adverse impact on university graduation rates.

We do however, uncover a negative impact of voucher school expansion on labor market outcomes, namely employment status and log income. Although our empirical strategy does not allow us to pin down the precise mechanism which drives these negative labor market consequences, we do explore other potential mediating factors, including changes in measures of school and peer quality arising from voucher school expansion. Analysis of the precise mechanisms in play would be an important focus for future research. One interesting aspect is the role of the funding system for upper secondary specializations. In the Swedish setting, the vocational STEM-related programs typically have higher reimbursement levels per student. This is motivated given that they tend to require more equipment and smaller classes, but might also make them more attractive to private providers. Other aspects of interest for further study relate to the design of admissions to schools and tracks, as well as regulation relating to the entry and operation of voucher schools. Such analyses would be useful for evaluating whether the impacts found in this paper are likely to generalize to other voucher school settings, that may differ in

design from the Swedish upper secondary case.

References

Altmejd, A., Barrios-Fernández, A., Drlje, M., Goodman, J., Hurwitz, M., Kovac, D., ... & Smith, J. (2021). "O brother, where start thou? Sibling spillovers on college and major choice in four countries." *The Quarterly Journal of Economics*, 136(3), 1831-1886.

Altonji, J. G., Arcidiacono, P., and Maurel, A. (2016). "The analysis of field choice in college and graduate school: Determinants and wage effects," in *Handbook of the Economics of Education* (Vol. 5, pp. 305-396). Elsevier.

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.

Aguirre, J. (2022). "How can progressive vouchers help the poor benefit from school choice? Evidence from the Chilean voucher system," *Journal of Human Resources*, 57(3), 956-997.

Bellégo & Pape. 2021. "Dealing with Logs and Zeros in Regression Models". SSRN Scholarly Paper ID 3444996. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3444996.

Bau, N, (2022). "Estimating An Equilibrium Model of Horizontal Competition in Education," forthcoming *Journal of Political Economy*.

Bettinger, E., Kremer, M., & Saavedra, J. E. (2010). "Are educational vouchers only redistributive?" *The Economic Journal*, 120(546), F204-F228.

Bharadwaj, P., Eberhard, J. P., & Neilson, C. A. (2018). "Health at birth, parental investments, and academic outcomes." *Journal of Labor Economics*, 36(2), 349-394.

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

Böhlmark, A., Holmlund H. and M. Lindahl. (2016). "Parental choice, neighbourhood segregation or cream skimming? An analysis of school segregation after a generalized choice reform", *Journal of Population Economics* 29: 1155–1190.

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.

Caetano, C., Callaway, B., Payne, S., & Rodrigues, H. S. A. (2022). "Difference in differences with time-varying covariates". arXiv preprint arXiv:2202.02903.

Chen & Roth. 2024. "Logs with Zeros? Some Problems and Solutions" *The Quarterly Journal of Economics*, Volume 139, Issue 2, Pages 891–936, https://doi.org/10.1093/qje/qjad054

Cohn, Zack and Wardlaw, "Count (and Count-Like) Data in Finance" (2022). *Journal of Financial Economics*, Forthcoming, Available at http://dx.doi.org/10.2139/ssrn.3800339

Correa, J. A., Parro, F., & Reyes, L. (2014). "The effects of vouchers on school results: evidence from Chile's targeted voucher program". *Journal of Human Capital*, 8(4), 351-398.

Dahl, G.B, Rooth, D., & Stenberg, A. (2023). "High School Majors and Future Earnings". *American Economic Journal: Applied Economics*, 15(1), 351-82.

Edmark, K., Frölich, M., & Wondratschek, V. (2014). "Sweden's school choice reform and equality of opportunity". *Labour Economic*s, 30, 129-142.

Edmark, K., Hussain, I., and Haelermans, C. (2020). "The Impact of Voucher Schools: Evidence from Swedish Upper Secondary Schools," Working Paper 3/2020, Swedish Institute for Social Research (SOFI), Stockholm University.

Edmark, K., & Persson, L. (2021). "The impact of attending an independent upper secondary school: Evidence from Sweden using school ranking data." *Economics of Education Review*, 84, 102148.

Epple, D., Figlio, D., & Romano, R. (2004). "Competition between private and public schools: testing stratification and pricing predictions." *Journal of Public Economics*, 88(7-8), 1215- 1245.

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.

Feigenberg, B., Yan, R., & Rivkin, S. (2019). "Illusory gains from Chile's targeted school voucher experiment," *The Economic Journal*, 129(623), 2805-2832.

Friedman, M. (1962). "The Role of Government in Education," in *Capitalism and Freedom*, University of Chicago Press.

Gilraine, M., Petronijevic, U., & Singleton, J. D. (2021). "Horizontal differentiation and the policy effect of charter schools," *American Economic Journal: Economic Policy*, 13(3), 239- 276.

Hinnerich, T. 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.

Hoxby, C. M. (2003). "School choice and school productivity. Could school choice be a tide that lifts all boats?" in *The Economics of School Choice* (pp. 287-342). University of Chicago Press.

Hsieh, C. T., and Urquiola, M. (2006). "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program," *Journal of Public Economics*, 90(8-9), 1477-1503.

Kirkeboen, L. J., Leuven, E., & Mogstad, M. (2016). "Field of study, earnings, and self-selection." *The Quarterly Journal of Economics*, 131(3), 1057-1111.

Krueger, A. B. and P. Zhu. (2004). Another Look at the New York City School Voucher Experiment. American Behavioral Scientist, 47(5), 658–698.

Manski, C. F. (1992). "Educational choice (vouchers) and social mobility," *Economics of Education Review*, 11(4), 351-369.

McMillan, R. (2005). "Competition, incentives, and public school productivity". *Journal of Public Economics*, 89(5-6), 1133-1154.

Molin, E (2019). "School Choice and Student Sorting". Chapter in Ph D thesis: "Through the Ranks. Essays on Inequality, Status and School Choice," Stockholm School of Economics.

Navarro-Palau, P. (2017). "Effects of differentiated school vouchers: Evidence from a policy change and date of birth cutoffs". *Economics of Education Review*, 58, 86-107.

Neilson, C. (2021). "Targeted vouchers, competition among schools, and the academic achievement of poor students," Working Paper, Yale University.

Rouse, C. E., and Barrow, L. (2009). "School vouchers and student achievement: Recent evidence and remaining questions," *Annual Review of Economics*, 1(1), 17-42.

Sandström, M. and F. Bergström. (2005). School vouchers in practice: competition will not hurt you. Journal of Public Economics, 89, 351–80.

Sant'Anna, P. H., & Zhao, J. (2020). "Doubly robust difference-in-differences estimators". *Journal of Econometrics*, 219(1), 101-122.

Söderström, M. and R. Uutsitalo (2010). School Choice and Segregation: Evidence from an Admission Reform", *Scandinavian Journal of Economics* 112(1), 55–76.

Muralidharan, K., and Sundararaman, V. (2015). "The aggregate effect of school choice: Evidence from a two-stage experiment in India," *The Quarterly Journal of Economics*, 130(3), 1011-1066.

SOU (2020). "Gemensamt ansvar – en modell för planering och dimensionering av gymnasial utbildning", Statens offentliga utredningar SOU 2020:33.

Sund, K (2018). "Långsiktiga effekter av utökade valmöjligheter till gymnasieskolan – från närhetsprincip till betygsprincip", The National Audit Office (Riksrevisionen) 2018:28.

Urquiola, M. (2016). "Competition among schools: Traditional public and private schools," *Handbook of the Economics of Education* (Vol. 5, pp. 209-237). Elsevier.

Wiswall, M., & Zafar, B. (2018). "Preference for the workplace, investment in human capital, and gender." *The Quarterly Journal of Economics*, 133(1), 457-507.

Government bills and reports:

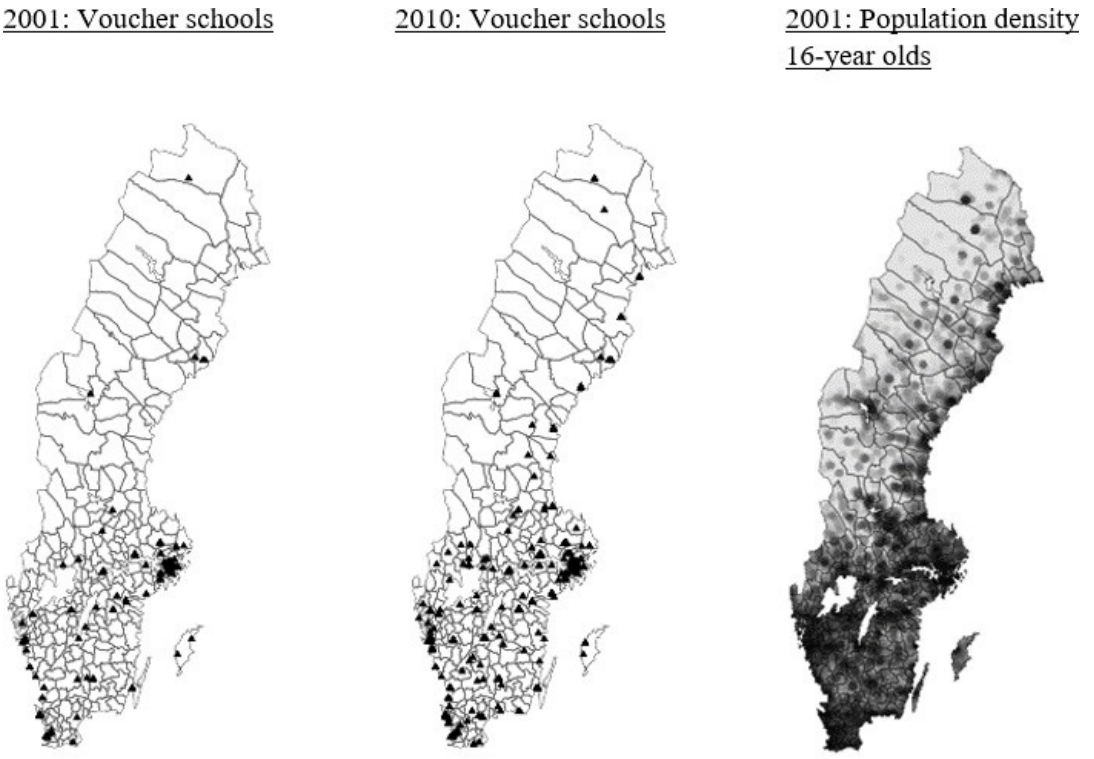
SOU 2008:27. Gymnasieskolan i en ny tid − förslag och bedömningar.

Government bill: Prop. 2008/09:199. Högre krav och kvalitet i den nya gymnasieskolan. (Information on the approval of the bill is available here: https://www.riksdagen.se/sv/dokument-och-lagar/dokument/betankande/hogre-krav-och-kvalitet-i-den-nya-gymnasieskolan_gx01ubu3/)

APPENDICES

Appendix A: Figures and tables

Figure A1 Voucher school locations in 2001 and 2010



Notes: The figures are based on grid cell coordinate information for school buildings, and for individual residential addresses.

Table A1: Impact on Voucher School Enrolment, alternative regression specifications.

	(1) Baseline specification	(2) No student covariates	(3) Including movers	(4) Labour market by year FE	(5) Municipality linear trends	(6) 10 km radius	(7) 20 km radius
Panel A: Full siblings sample							
Voucher Share	0.0764***	0.0773***	0.0643***	0.0829***	0.0889***	0.1086***	0.1125***
	(0.0093)	(0.0093)	(0.0074)	(0.0116)	(0.0114)	(0.0093)	(0.0090)
Observations	280,391	280,391	360,087	280,391	280,391	330,468	376,369
Number of Mother FEs	131,028	131,028	170,847	131,028	131,028	153,670	174,025
Panel B: Brothers							
Voucher Share	0.1019***	0.1036***	0.0847***	0.1007***	0.1111***	0.1475***	0.1315***
	(0.0169)	(0.0169)	(0.0140)	(0.0212)	(0.0210)	(0.0173)	(0.0168)
Observations	83,812	83,812	104,487	83,812	83,812	98,914	112,796
Number of Mother FEs	40,628	40,628	51,478	40,628	40,628	47,724	54,157
Panel C: Sisters							
Voucher Share	0.0436**	0.0428**	0.0407***	0.0683***	0.0462**	0.0552***	0.0548***
	(0.0173)	(0.0173)	(0.0140)	(0.0220)	(0.0211)	(0.0177)	(0.0168)
Observations	76,539	76,539	98,266	76,539	76,539	90,185	102,922
Number of Mother FEs	37,065	37,065	48,561	37,065	37,065	43,487	49,423

Notes: "Voucher Share" measures the share of upper secondary voucher schools within a 5 km radius from the student's home, except column (5)/(6)/, which measure the share of upper secondary voucher schools within a 10/20/ km radius. All regressions include year, sibling order and family (mother) fixed effects. Column (4) additionally include local labour market by year fixed effects, and column (5) includes municipality-specific linear time trends. All regressions except column (2) include the following student-level covariates: lower secondary school GPA (final grade sum) in level and square; male dummy; attended voucher lower secondary school; born in Sweden; and born in Europe (except Sweden) or North America. All regressions also include the following time-varying family covariates: mother employed; father employed; household disposable income in level and square. Municipality-level variables included in all regressions are: expenditure per student on compulsory education (primary and lower secondary education); the share of students attending private school in grade 9 (lower secondary); dummy for the municipality having a left-wing political local majority. Finally, regressions also include, at the grid cell level, the log of the number of age-16 youth residing within 5km from the grid cell midpoint (this is replaced by the number residing within 10km/20km radius for the 10km/20km specifications in columns 6 and 7). All time-varying covariates are measured the year before the student enters upper secondary school, except for the first year of the panel, when the current year's values are used. Missing values in covariates were replaced by the average value of the variable, and dummy variables were added indicating the missing instances. The sibling sample is, with the exception of column (3), restricted to non-moving households, defined as siblings who reside in the same grid cell measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Impact on Upper Secondary Track, student covariates omitted, 5 km radius.

	(1)	(2)	(3)	(4)	(5)	(6)
		Vocational			Academic	
	Vocational	Industry/Tech	Trade/Admin	Nursing/Care	Science	Social Science/Arts
Panel A: Full siblings sample						
Voucher Share	0.0235**	0.0270***	-0.0047	0.0013	-0.0328***	0.0082
	(0.0110)	(0.0095)	(0.0079)	(0.0064)	(0.0101)	(0.0122)
Observations	280,391	280,391	280,391	280,391	280,391	280,391
Number of Mother FEs	131,028	131,028	131,028	131,028	131,028	131,028
Panel B: Brothers						
Voucher Share	0.0403**	0.0539***	-0.0139	0.0002	-0.0534***	0.0091
	(0.0201)	(0.0191)	(0.0123)	(0.0087)	(0.0193)	(0.0206)
Observations	83,812	83,812	83,812	83,812	83,812	83,812
Number of Mother FEs	40,628	40,628	40,628	40,628	40,628	40,628
Panel C: Sisters						
Voucher Share	0.0350*	0.0160	0.0130	0.0059	-0.0422**	0.0041
	(0.0205)	(0.0110)	(0.0172)	(0.0142)	(0.0181)	(0.0241)
Observations	76,539	76,539	76,539	76,539	76,539	76,539
Number of Mother FEs	37,065	37,065	37,065	37,065	37,065	37,065

Notes: "Voucher Share" measures the share of upper secondary voucher schools within a 5 km radius from the student's home. All regressions include year, sibling order and family (mother) fixed effects, and the following municipality-level variables: expenditure per student on compulsory education (primary and lower secondary education); the share of students attending private school in grade 9 (lower secondary); dummy for the municipality having a left-wing political local majority. Finally, regressions also include, at the grid cell level, the log of the number of age-16 youth residing within 5km from the grid cell midpoint. All time-varying covariates are measured the year before the student enters upper secondary school, except for the first year of the panel, when the current year's values are used. Missing values in covariates were replaced by the average value of the variable, and dummy variables were added indicating the missing instances. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same grid cell measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Impact on Upper Secondary Track, movers included, 5 km radius.

	(1)	(2)	(3)	(4)	(5)	(6)
		Vocational			Academic	
	Vocational	Industry/Tech	Trade/Admin	Nursing/Care	Science	Social Science/Arts
Panel A: Full siblings sample						
Voucher Share	0.0158**	0.0118*	0.0060	-0.0020	-0.0208***	0.0056
	(0.0080)	(0.0067)	(0.0061)	(0.0047)	(0.0070)	(0.0090)
Observations	360,087	360,087	360,087	360,087	360,087	360,087
Number of Mother FEs	170,847	170,847	170,847	170,847	170,847	170,847
Panel B: Brothers						
Voucher Share	0.0203	0.0277*	-0.0120	0.0045	-0.0313**	0.0062
	(0.0154)	(0.0149)	(0.0096)	(0.0068)	(0.0138)	(0.0162)
Observations	104,487	104,487	104,487	104,487	104,487	104,487
Number of Mother FEs	51,478	51,478	51,478	51,478	51,478	51,478
Panel C: Sisters						
Voucher Share	0.0339**	0.0171**	0.0220*	-0.0054	-0.0297**	-0.0013
	(0.0149)	(0.0085)	(0.0130)	(0.0108)	(0.0122)	(0.0176)
Observations	98,266	98,266	98,266	98,266	98,266	98,266
Number of Mother FEs	48,561	48,561	48,561	48,561	48,561	48,561

Notes: "Voucher Share" measures the share of upper secondary voucher schools within a 5 km radius from the student's home. All regressions include year, sibling order and family (mother) fixed effects as well as the following student-level covariates: lower secondary school GPA (final grade sum) in level and square; male dummy; attended voucher lower secondary school; born in Sweden; and born in Europe (except Sweden) or North America. All regressions also include the following time-varying family covariates: mother employed; father employed; household disposable income in level and square. Municipality-level variables included in all regressions are: expenditure per student on compulsory education (primary and lower secondary education); the share of students attending private school in grade 9 (lower secondary); dummy for the municipality having a left-wing political local majority. Finally, regressions also include, at the grid cell level, the log of the number of age-16 youth residing within 5km from the grid cell midpoint. All time-varying covariates are measured the year before the student enters upper secondary school, except for the first year of the panel, when the current year's values are used. Missing values in covariates were replaced by the average value of the variable, and dummy variables were added indicating the missing instances. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Impact on Upper Secondary Track, including labour market by year fixed effects, 5 km radius.

	(1)	(2)	(3)	(4)	(5)	(6)
		Vocational			Academic	
	Vocational	Industry/Tech	Trade/Admin	Nursing/Care	Science	Social Science/Arts
Panel A: Full siblings sample						
Voucher Share	0.0126	0.0176*	-0.0098	0.0047	-0.0255**	0.0142
	(0.0122)	(0.0104)	(0.0091)	(0.0071)	(0.0111)	(0.0141)
Observations	280,391	280,391	280,391	280,391	280,391	280,391
Number of Mother FEs	131,028	131,028	131,028	131,028	131,028	131,028
Panel B: Brothers						
Voucher Share	0.0399*	0.0375*	-0.0029	0.0051	-0.0460**	0.0074
	(0.0223)	(0.0217)	(0.0144)	(0.0096)	(0.0214)	(0.0244)
Observations	83,812	83,812	83,812	83,812	83,812	83,812
Number of Mother FEs	40,628	40,628	40,628	40,628	40,628	40,628
Panel C: Sisters						
Voucher Share	0.0198	0.0230*	-0.0021	-0.0011	-0.0342*	0.0118
	(0.0233)	(0.0135)	(0.0199)	(0.0161)	(0.0203)	(0.0287)
Observations	76,539	76,539	76,539	76,539	76,539	76,539
Number of Mother FEs	37,065	37,065	37,065	37,065	37,065	37,065

Notes: "Voucher Share" measures the share of upper secondary voucher schools within a 5 km radius from the student's home. All regressions include year, sibling order and family (mother) fixed effects, and local labour market by year fixed effects. The regressions also include the same covariates as Table A3. All time-varying covariates are measured the year before the student enters upper secondary school, except for the first year of the panel, when the current year's values are used. Missing values in covariates were replaced by the average value of the variable, and dummy variables were added indicating the missing instances. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same grid cell measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Impact on Upper Secondary Track, including municipality specific linear trends, 5 km radius.

	(1)	(2)	(3)	(4)	(5)	(6)
			Vocational		Academic	
	Vocational	Industry/Tech	Trade/Admin	Nursing/Care	Science	Social Science/Arts
Panel A: Full siblings sample						
Voucher Share	0.0435***	0.0343***	0.0125	-0.0032	-0.0302***	-0.0109
	(0.0123)	(0.0105)	(0.0092)	(0.0072)	(0.0114)	(0.0143)
Observations	280,391	280,391	280,391	280,391	280,391	280,391
Number of Mother FEs	131,028	131,028	131,028	131,028	131,028	131,028
Panel B: Brothers						
Voucher Share	0.0670***	0.0688***	0.0101	-0.0119	-0.0581***	-0.0116
	(0.0227)	(0.0220)	(0.0147)	(0.0102)	(0.0220)	(0.0252)
Observations	83,812	83,812	83,812	83,812	83,812	83,812
Number of Mother FEs	40,628	40,628	40,628	40,628	40,628	40,628
Panel C: Sisters						
Voucher Share	0.0453*	0.0180	0.0232	0.0041	-0.0506**	0.0061
	(0.0233)	(0.0132)	(0.0201)	(0.0163)	(0.0206)	(0.0286)
Observations	76,539	76,539	76,539	76,539	76,539	76,539
Number of Mother FEs	37,065	37,065	37,065	37,065	37,065	37,065

Notes: "Voucher Share" measures the share of upper secondary voucher schools within a 5 km radius from the student's home. All regressions include year, sibling order and family (mother) fixed effects, and municipality specific linear time trends. The regressions also include the same covariates as Table A3. All time-varying covariates are measured the year before the student enters upper secondary school, except for the first year of the panel, when the current year's values are used. Missing values in covariates were replaced by the average value of the variable, and dummy variables were added indicating the missing instances. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same grid cell measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Impact on Upper Secondary Track, 10 km radius.

	(1)	(2)	(3)	(4)	(5)	(6)
		Vocational			Academic	
	Vocational	Industry/Tech	Trade/Admin	Nursing/Care	Science	Social Science/Arts
Panel A: Full siblings sample						
Voucher Share	0.0186	0.0220**	-0.0057	0.0023	-0.0540***	0.0336**
	(0.0120)	(0.0106)	(0.0087)	(0.0074)	(0.0108)	(0.0130)
Observations	330,468	330,468	330,468	330,468	330,468	330,468
Number of Mother FEs	153,670	153,670	153,670	153,670	153,670	153,670
Panel B: Brothers						
Voucher Share	0.0349	0.0442**	-0.0055	-0.0038	-0.0797***	0.0433**
	(0.0222)	(0.0213)	(0.0134)	(0.0101)	(0.0211)	(0.0219)
Observations	98,914	98,914	98,914	98,914	98,914	98,914
Number of Mother FEs	47,724	47,724	47,724	47,724	47,724	47,724
Panel C: Sisters						
Voucher Share	0.0123	0.0090	-0.0080	0.0116	-0.0533***	0.0310
	(0.0226)	(0.0124)	(0.0189)	(0.0167)	(0.0187)	(0.0258)
Observations	90,185	90,185	90,185	90,185	90,185	90,185
Number of Mother FEs	43,487	43,487	43,487	43,487	43,487	43,487

Notes: "Voucher Share" measures the share of upper secondary voucher schools within a 10 km radius from the student's home. All regressions include year, sibling order and family (mother) fixed effects, as well as the same covariates as Table A3 (with the exception that the number of 16-year olds is measured within a 10km radius from the student's residence, instead of within 5km). All time-varying covariates are measured the year before the student enters upper secondary school, except for the first year of the panel, when the current year's values are used. Missing values in covariates were replaced by the average value of the variable, and dummy variables were added indicating the missing instances. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same grid cell measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Impact on Upper Secondary Track, 20 km radius.

	(1)	(2)	(3)	(4)	(5)	(6)
		Vocational			Academic	
	Vocational	Industry/Tech	Trade/Admin	Nursing/Care	Science	Social Science/Arts
Panel A: Full siblings sample						
Voucher Share	0.0193	0.0135	0.0054	0.0004	-0.0439***	0.0232*
	(0.0123)	(0.0105)	(0.0091)	(0.0079)	(0.0105)	(0.0132)
Observations	376,369	376,369	376,369	376,369	376,369	376,369
Number of Mother FEs	174,025	174,025	174,025	174,025	174,025	174,025
Panel B: Brothers						
Voucher Share	0.0260	0.0472**	-0.0188	-0.0023	-0.0672***	0.0378*
	(0.0227)	(0.0225)	(0.0138)	(0.0110)	(0.0206)	(0.0226)
Observations	112,796	112,796	112,796	112,796	112,796	112,796
Number of Mother FEs	54,157	54,157	54,157	54,157	54,157	54,157
Panel C: Sisters						
Voucher Share	0.0019	-0.0126	0.0072	0.0075	-0.0420**	0.0359
	(0.0239)	(0.0129)	(0.0205)	(0.0181)	(0.0184)	(0.0274)
Observations	102,922	102,922	102,922	102,922	102,922	102,922
Number of Mother FEs	49,423	49,423	49,423	49,423	49,423	49,423

Notes: "Voucher Share" measures the share of upper secondary voucher schools within a 20 km radius from the student's home. All regressions include year, sibling order and family (mother) fixed effects, as well as the same covariates as Table A3 (with the exception that the number of 16-year olds is measured within a 20km radius from the student's residence, instead of within 5km). All time-varying covariates are measured the year before the student enters upper secondary school, except for the first year of the panel, when the current year's values are used. Missing values in covariates were replaced by the average value of the variable, and dummy variables were added indicating the missing instances. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same grid cell measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Impact on Graduation and Grades, 5 km radius

	(1)	(2)
	Graduate on time	Pctile GPA
Panel A: Parents high education		
Voucher Share	0.0235*	-0.0039
	(0.0129)	(0.0071)
Observations	156,906	144,764
Number of Mother FEs	73,293	67,796
Panel B:Parents low education		
Voucher Share	0.0098	-0.0096
	(0.0156)	(0.0080)
Observations	111,376	98,821
Number of Mother FEs	52,306	46,639
Panel C: Both parents Swedish-born		
Voucher Share	0.0115	-0.0068
	(0.0109)	(0.0058)
Observations	210,129	192,158
Number of Mother FEs	98,878	90,711
Panel D: At least one parent born abroad		
Voucher Share	0.0367*	-0.0077
	(0.0221)	(0.0116)
Observations	68,981	60,646
Number of Mother FEs	31,571	27,881

Notes: Panel A shows the results for the sub-sample of students who have at leastone parent with a completed post-secondary education, and panel B shows the results for studets that do not. Panel C shows the results for students with both parents born in Sweden, and Panel D shows the results for students with at least one foreign-born parent. "Voucher Share" measures the share of upper secondary voucher schools within a 5 km radius from the student's home. All regressions include year, sibling order and family (mother) fixed effects, as well as the same covariates as Table A3. All time-varying covariates are measured the year before the student enters upper secondary school, except for the first year of the panel, when the current year's values are used. Missing values in covariates were replaced by the average value of the variable, and dummy variables were added indicating the missing instances. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same grid cell measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A9: Impact on Graduating on time, Alternative estimation models.

	(1)	(2)	(3)	(4)	(5)	(6)
	No student covariates	Including movers	Labour market by year FE	Municipality linear trends	10 km radius	20 km radius
Panel A: Full Siblings sample						
Voucher Share	0.0148	0.0141*	0.0085	0.0085	0.0185*	0.0194*
	(0.0101)	(0.0078)	(0.0117)	(0.0118)	(0.0107)	(0.0109)
Observations	280,391	360,087	280,391	280,391	330,468	376,369
Number of Mother FEs	131,028	170,847	131,028	131,028	153,670	174,025
Panel B: Brothers						
Voucher Share	-0.0104	0.0027	-0.0146	-0.0205	-0.0110	0.0072
	(0.0190)	(0.0150)	(0.0221)	(0.0220)	(0.0206)	(0.0208)
Observations	83,812	104,487	83,812	83,812	98,914	112,796
Number of Mother FEs	40,628	51,478	40,628	40,628	47,724	54,157
Panel C: Sisters						
Voucher Share	0.0284	0.0292**	0.0320	0.0180	0.0196	0.0239
	(0.0187)	(0.0142)	(0.0220)	(0.0222)	(0.0194)	(0.0200)
Observations	76,539	98,266	76,539	76,539	90,185	102,922
Number of Mother FEs	37,065	48,561	37,065	37,065	43,487	49,423

Notes: See Table A1. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A10: Impact on Percentile GPA, Alternative estimation models.

	(1)	(2)	(3)	(4)	(5)	(6)
	No student covariates	Including movers	Labour market by year FE	Municipality linear trends	10 km radius	20 km radius
Panel A: Full Siblings sample						
Voucher Share	-0.0065	-0.0084**	-0.0218***	-0.0041	-0.0071	0.0003
	(0.0067)	(0.0041)	(0.0062)	(0.0062)	(0.0071)	(0.0058)
Observations	253,862	320,051	253,862	253,862	299,510	341,044
Number of Mother FEs	119,076	152,41	119,076	119,076	139,817	158,330
Panel B: Brothers						
Voucher Share	-0.0072	-0.0076	-0.0205*	-0.0046	0.0026	-0.0013
	(0.0114)	(0.0075)	(0.0111)	(0.0110)	(0.0122)	(0.0103)
Observations	74,878	91,629	74,878	74,878	88,473	100,899
Number of Mother FEs	36,378	45,193	36,378	36,378	42,781	48,548
Panel C: Sisters						
Voucher Share	-0.0132	-0.0172**	-0.0235*	-0.0079	-0.0233*	-0.0084
	(0.0124)	(0.0081)	(0.0124)	(0.0123)	(0.0129)	(0.0112)
Observations	69,491	87,66	69,491	69,491	81,931	93,394
Number of Mother FEs	33,706	43,345	33,706	33,706	39,589	44,941

Notes: See Table A1. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A11: Impact on Long Term Student Outcomes, no Student Covariates, Municipality Level Variation in Private School Share

	(1) Highest qualification is in STEM	(2) STEM subject tertiary/post-secondary degree	(3) Tertiary/Post-secondary degree	(4) University credits (by age 25)	(5) Log employment income	(6) Employed
Panel A: Full Siblings sample						
Private Share	-0.0197**	-0.0094	0.0049	-0.0042	-0.0432	-0.0156**
	(0.0086)	(0.0065)	(0.0084)	(0.0056)	(0.0292)	(0.0062)
Observations	413,684	413,684	420,083	841,063	420,099	420,099
Number of Mother FEs	191,539	191,539	194,387	371,009	194,394	194,394
Panel B: Brothers						
Private Share	-0.0087	-0.0104	-0.0229	-0.0129	-0.0944*	-0.0180
	(0.0171)	(0.0135)	(0.0152)	(0.0097)	(0.0542)	(0.0115)
Observations	119,35	119,35	121,622	258,38	121,628	121,628
Number of Mother FEs	57,247	57,247	58,313	120,926	58,316	58,316
Panel C: Sisters						
Private Share	-0.0120	-0.0157	0.0159	0.0050	-0.0545	-0.0148
	(0.0115)	(0.0105)	(0.0158)	(0.0106)	(0.0568)	(0.0127)
Observations	112,821	112,821	114,304	240,043	114,306	114,306
Number of Mother FEs	54,108	54,108	54,809	112,512	54,810	54,810

Notes: "Voucher Share" measures the share of upper secondary voucher schools at the municipality level. All regressions include year, sibling order and family (mother) fixed effects as well as the following municipality-level variables: expenditure per student on compulsory education (primary and lower secondary education); the share of students attending private school in grade 9 (lower secondary); dummy for the municipality having a left-wing political local majority and the log of the number of age-16 youth residing in the municipality. All time-varying covariates are measured the year before the student enters upper secondary school, except for the first year of the panel, when the current year's values are used. Missing values in covariates were replaced by the average value of the variable, and dummy variables were added indicating the missing instances. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same municipality measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A12: Impact on Long Term Student Outcomes, Including Movers, Municipality Level Variation in Private School Share

	(1) Highest qualification is in STEM	(2) STEM subject tertiary/post-secondary degree	(3) Tertiary/Post-secondary degree	(4) University credits (by age 25)	(5) Log employment income	(6) Employed
Panel A: Full Siblings sample						
Private Share	-0.0189***	-0.0074	0.0045	-0.0022	-0.0452	-0.0150**
	(0.0073)	(0.0058)	(0.0072)	(0.0047)	(0.0277)	(0.0059)
Observations	438,23	438,23	445,084	903,454	445,104	445,104
Number of Mother FEs	202,371	202,371	205,406	396,506	205,415	205,415
Panel B: Brothers						
Private Share	-0.0042	-0.0039	-0.0210	-0.0088	-0.1045**	-0.0207*
	(0.0161)	(0.0123)	(0.0134)	(0.0083)	(0.0520)	(0.0109)
Observations	125,413	125,413	127,816	274,484	127,824	127,824
Number of Mother FEs	60,090	60,090	61,213	128,171	61,217	61,217
Panel C: Sisters						
Private Share	-0.0100	-0.0102	0.0214	0.0027	-0.0690	-0.0179
	(0.0105)	(0.0095)	(0.0141)	(0.0091)	(0.0538)	(0.0120)
Observations	119,638	119,638	121,223	257,908	121,225	121,225
Number of Mother FEs	57,293	57,293	58,042	120,555	58,043	58,043

Notes: "Voucher Share" measures the share of upper secondary voucher schools at the municipality level. All regressions include year, sibling order and family (mother) fixed effects as well as the following student-level covariates: lower secondary school GPA (final grade sum) in level and square (separate coefficients are estimated for final GPA before and after 1998, since a new grading system was introduced for lower secondary school in that year); male dummy; attended voucher lower secondary school; born in Sweden; and born in Europe (except Sweden) or North America. All regressions also include the following time-varying family covariates: mother employed; father employed; and household disposable income in level and square (household income enters in the form of separate values for each parents' individual component of the households' joint income). Municipality-level variables included in all regressions are: expenditure per student on compulsory education (primary and lower secondary education); the share of students attending private school in grade 9 (lower secondary); dummy for the municipality having a left-wing political local majority. Finally, regressions also include, at the municipality level, the log of the number of age-16 youth residing in the municipality. All time-varying covariates are measured the year before the student enters upper secondary school, except for the first year of the panel, when the current year's values are used. Missing values in covariates were replaced by the average value of the variable, and dummy variables were added indicating the missing instances. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A13: Impact on Long Term Student Outcomes, Including Local Labour Market-by-Year Fixed Effects, Municipality Level Variation in Private School Share

	(1)	(2)	(3)	(4)	(5)	(6)
	Highest qualification is in STEM	STEM subject tertiary/post-secondary degree	Tertiary/Post-secondary degree	University credits (by age 25)	Log employment income	Employed
Panel A: Full Siblings sample						
Private Share	-0.0200**	-0.0139*	-0.0006	0.0026	-0.0647*	-0.0175**
	(0.0097)	(0.0078)	(0.0095)	(0.0063)	(0.0357)	(0.0077)
Observations	413,684	413,684	420,083	841,063	420,099	420,099
Number of Mother FEs	191,539	191,539	194,387	371,009	194,394	194,394
Panel B: Brothers						
Private Share	0.0041	-0.0163	-0.0263	0.0025	-0.1484**	-0.0252*
	(0.0209)	(0.0160)	(0.0174)	(0.0109)	(0.0661)	(0.0139)
Observations	119,350	119,350	121,622	258,380	121,628	121,628
Number of Mother FEs	57,247	57,247	58,313	120,926	58,316	58,316
Panel C: Sisters						
Private Share	-0.0159	-0.0220*	0.0100	0.0103	-0.0016	0.0002
	(0.0140)	(0.0129)	(0.0186)	(0.0123)	(0.0703)	(0.0158)
Observations	112,821	112,821	114,304	240,043	114,306	114,306
Number of Mother FEs	54,108	54,108	54,809	112,512	54,810	54,810

Notes: "Voucher Share" measures the share of upper secondary voucher schools at the municipality level. All regressions include year, sibling order and family (mother) fixed effects, and local labour market by year fixed effects, and the same covariates as Table A12. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same municipality measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A14: Impact on Long Term Student Outcomes, Including Municipality-specific Linear Trends, Municipality Level Variation in Private School Share

	(1)	(2)	(3)	(4)	(5)	(6)
	Highest qualification is in STEM	STEM subject tertiary/post-secondary degree	Tertiary/Post-secondary degree	University credits (by age 25)	Log employment income	Employed
Panel A: Full Siblings sample						
Private Share	-0.0055	0.0060	0.0132	-0.0012	-0.0027	-0.0055
	(0.0111)	(0.0090)	(0.0108)	(0.0068)	(0.0406)	(0.0087)
Observations	413,684	413,684	420,083	841,063	420,099	420,099
Number of Mother FEs	191,539	191,539	194,387	371,009	194,394	194,394
Panel B: Brothers						
Private Share	0.0209	0.0123	-0.0093	-0.0007	0.0051	0.0039
	(0.0239)	(0.0187)	(0.0200)	(0.0120)	(0.0756)	(0.0157)
Observations	119,350	119,350	121,622	258,380	121,628	121,628
Number of Mother FEs	57,247	57,247	58,313	120,926	58,316	58,316
Panel C: Sisters						
Private Share	0.0006	-0.0057	0.0138	0.0103	-0.0836	-0.0078
	(0.0164)	(0.0152)	(0.0210)	(0.0133)	(0.0792)	(0.0180)
Observations	112,821	112,821	114,304	240,043	114,306	114,306
Number of Mother FEs	54,108	54,108	54,809	112,512	54,810	54,810

Notes: "Voucher Share" measures the share of upper secondary voucher schools at the municipality level. All regressions include year, sibling order and family (mother) fixed effects, municipality-specific linear time trends, and the same covariates as Table A12. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same municipality measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A15 Impact on Upper Secondary Track, Voucher school attendance, Graduation and Final Grade, gridcell specification, sibling sample.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Vocational		Academic			
	Vocational	Industry/Tech	Trade/Admin	Nursing/Care	Science	Social Science/Arts	Graduate on time	Pctile GPA
Panel A: All siblings								
Voucher Share	0.0166*	0.0138*	0.0025	0.0002	-0.0315***	0.0151	0.0120	-0.0028
	(0.0087)	(0.0071)	(0.0061)	(0.0050)	(0.0081)	(0.0100)	(0.0078)	(0.0043)
Observations	280,391	280,391	280,391	280,391	280,391	280,391	280,391	253,862
Number of Gridcell FEs	32,908	32,908	32,908	32,908	32,908	32,908	32,908	32,086
Panel B: Brothers								
Voucher Share	0.0321*	0.0344**	-0.0065	0.0041	-0.0489***	0.0143	-0.0063	-0.0006
	(0.0173)	(0.0166)	(0.0104)	(0.0073)	(0.0163)	(0.0186)	(0.0158)	(0.0082)
Observations	83,812	83,812	83,812	83,812	83,812	83,812	83,812	74,878
Number of Gridcell FEs	20,827	20,827	20,827	20,827	20,827	20,827	20,827	19,735
Panel C: Sisters								
Voucher Share	0.0022	0.0019	0.0186	-0.0182	-0.0180	0.0132	0.0129	-0.0140
	(0.0170)	(0.0095)	(0.0144)	(0.0121)	(0.0147)	(0.0206)	(0.0156)	(0.0090)
Observations	76,539	76,539	76,539	76,539	76,539	76,539	76,539	69,491
Number of Gridcell FEs	19,742	19,742	19,742	19,742	19,742	19,742	19,742	18,749

Notes: "Voucher Share" measures the share of upper secondary voucher schools within a 5 km radius from the student's home. All regressions include year, and gridcell fixed effects as well as the following student- level covariates: lower secondary school GPA (final grade sum) in level and square; male dummy; attended voucher lower secondary school; born in Sweden; and born in Europe (except Sweden) or North America. All regressions also include the following time-varying family covariates: mother employed; father employed; household disposable income in level and square. Municipality-level variables included in all regressions are: expenditure per student on compulsory education (primary and lower secondary education); the share of students attending private school in grade 9 (lower secondary); dummy for the municipality having a left- wing political local majority. Finally, regressions also include, at the grid cell level, the log of the number of age-16 youth residing within 5km from the grid cell midpoint. All time-varying covariates are measured the year before the student enters upper secondary school, except for the first year of the panel, when the current year's values are used. Missing values in covariates were replaced by the average value of the variable, and dummy variables were added indicating the missing instances. The sibling sample is restricted to non-moving households, defined as siblings who reside in the same grid cell measured the year they turn 16. Standard errors, clustered at the family level (mother id), in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix B1: Data details

B1.1 Data sources

The analysis is based on registry data held by Statistics Sweden. The data allow linking parents and children, and contain extensive background information on e.g. gender, country of birth (grouped), household disposable income, earnings and employment, and educational attainment. The school register contains information on all upper secondary schools in Sweden. It contains information on the voucher-status of schools, as well as the municipality of location, and can be linked to more detailed geographical information.

Two alternative and separate data sets were used for the estimations: The main analysis, on outcomes relating to students' short-term outcomes and school characteristics, was carried out using data accessed through the server system for micro data research at Statistics Sweden, MONA. These data include detailed geographical information that allowed us to construct measures of voucher school presence within different radii around students' residences. They were however limited in the sense that they lacked information on student outcomes in the long run, as they reached adulthood. In order to add this to the analysis, we were graciously given access to an alternative set of data through the Institute for the Evaluation of Education and Labour Market Policy (IFAU). These data are based on the same type of Statistics Sweden registry data, and contain roughly similar information as our main data set, but for a longer period of time. They therefore allowed us to study students' educational attainment and labour market outcomes until the age of 30. The data accessed through the IFAU did however not contain geographical information at the same level of detail as the main data set, which is why the analysis of long-term outcomes was carried out using voucher school variation at the municipality by year level instead of at the finer grid cell level.

B1.2 Sample restrictions

The main raw data set covers the full school aged population during 2001-10 and their parents. The alternative data set used for the long-term analysis covers the same population but includes earlier cohorts: The analysis of the educational and labour market outcomes measured at age 30 was carried out using data on students entering upper secondary school in years 1995-2005, whereas the analysis for university credits taken by age 25 uses cohorts entering upper secondary school in 1995-2009.

The main regression analysis, for the shorter-term analysis, was carried out after implementing the following restrictions to the initial raw data:

- We included only students who finished lower secondary school with a basic qualification (sufficient number of pass grades) to a regular Academic or Vocational educational track. Students who lack this qualification cannot continue straight from lower to upper secondary regular programs, but first need to attend a preparatory program.
- We included only students with non-missing information on the type of upper secondary track attended, and if the attended school is a municipal or voucher school.
- We included only students residing in grid cells /municipalities/ with at least one upper secondary school within a radius of 5km (or 10km/20km/within the municipality, depending on the specification). The restriction is needed for us to be able to construct a variable measuring the share of voucher schools in the vicinity.
- We included only students' first instances of attending the first year of upper secondary school. That is, potential repeat years, for example due to students changing educational track, are excluded. Furthermore, only students who start grade 10 at age 15-17 are included (16 is the normal starting age.)
- The calculations for private school shares exclude schools that offer only the preparatory track (there are a few such public schools in the data), and exclude the (very few) schools for which the type of track offered was not observed. The motivations for excluding schools offering only the preparatory track is that all students in the regression data set qualify for at

least some regular educational track. (The types of tracks offered by the schools are identified based on track information for students attending grade 10.)

- The regression analysis was carried out using variation within siblings, defined as having the same mother. The regression sample was thus further restricted to students with at least one sibling in the regression data panel.

Below, we list the sample restrictions step-wise, and the number of remaining observations after carrying out each step, for the main sibling regression sample.

Sample restrictions	*Number of observations*
Students in upper secondary education grade 10 year 2001-10, keeping only the first instance a student is observed as attending grade 10.	1,171,422
Drop students younger than 15 or older than 17 (age 16 is the normal upper secondary school start age).	1,165,270
Drop observations with missing grid cell coordinates for student residence (note that grid cell information is lacking for students in locations with low population density).	959,369
Drop students who are not qualified to enter a regular upper secondary academic or vocational program based on their lower secondary final grades. (These students first need to take preparatory courses.)	841,235
Drop students with missing information on type of school (voucher or not) or type of program (type of academic or vocational).	826,056
Drop student observations with missing mother identifier.	824,730
Keep only students for whom i) the outcome variable is non-missing; ii) a sibling is observed in the regression data sample; iii) the variable for the share of voucher schools within 5 km is defined (meaning that there is at least one upper secondary school within 5 km from where the student resides); and iv) the family (based on mother id) resides in the same grid cell when all observed siblings enter upper secondary education.	Approximately 280,000 students (exact size varies depending on the number of non-missing observations in the outcome variable).

B1.3 Information on the geographic location of schools.

As mentioned above, the main analysis was carried out using grid cell level voucher-share variables, and the long-term analysis used municipality level voucher-shares due to data limitations. This section provides details on the geographical data employed to construct the voucher share variables for the main analysis.

The geographical information is in the form of 250m^2 grid cell coordinates linked to schools and to students' residential addresses. Our data set has missing grid cell coordinate information for the following cases:

i) 2001-2003: Grid cell coordinates are missing for all schools.
ii) 2004-2008: Grid cell coordinates are missing for the first three years of operation for each start-up school. This follows from the fact that a start-up school enters our data the year the first cohort of students graduates from the school.
iii) Grid cell coordinates are also missing for schools with incomplete address information in the School Register. For such schools, Statistics Sweden cannot link schools to grid cell coordinates.
iv) 250m^2 grid cell coordinates are missing for students residing in very rural areas, since for such areas Statistics Sweden only gives access to larger grids, in order to disable identification of individual schools based on the coordinate information.

We deal with the missing school grid cell coordinates by replacing missing coordinate information for a school and year with the nearest available future data for the school in question. For a few schools, there are missing observations in later years, and non-missing observations in earlier years. For these cases, the earlier year's observations are used for later years. This means that the geographical information will be measured with error for schools that change location between the unobserved years and the observed year. Changes of school locations are however rare, so this is likely a minor issue.

After the above replacements were made, there is still a relatively large number of schools with missing information on geographical coordinates: 16% of the schools in our data lack 250m^2 grid cell coordinate information. The share is 17.4% among the non-voucher schools and 14% among the voucher schools. This missing grid cell coordinates stem both from the fact that grid cells at this high level of detail are not provided by Statistics Sweden for schools in rural areas, and from the fact that some schools could not be linked to grid cells due to missing or incomplete address information in the School register.

B1.4 Variable definitions

Outcome variables: Short run variables

Educational track

The educational track is measured according to the upper secondary program the student attends in the fall of the first term of upper secondary school. We construct dummy variables for attending the Vocational track (instead of the Academic track)[22]; as well as for the five subcategories of Vocational and Academic tracks that are shown in Table A1.

Dummy variable for graduating on time

The dummy variable takes value one for students who are observed as graduating from an upper secondary program on time – i.e. three years from starting, and value zero otherwise. Students from

[22] There is also the option of attending a preparatory first year, but this is not a relevant option for the students in our regression sample, who are all qualified to enter the regular upper secondary Vocational and Academic programs.

Waldorff schools or the IB-program are treated as graduating on time, although their grade values are not shown in the register. Student with incomplete final grades (failing some of the courses) are given value zero.

Pctile GPA

The variable is defined as the year-wise percentile rank (from 0 to 0.9999) of students' final grade point average from upper secondary school. This is calculated among all students graduating in a given year (except for the few students from Waldorff and IB schools, whose grades are not included in the registers). Students graduating with incomplete grades, due to failing some courses, are treated as having a GPA of zero, i.e. are assigned the lowest rank. Students who are not observed in the final grades register, for example due to dropping out, are treated as missing.

Outcome variables: Long run variables

Highest level of completed education at age 30 is in STEM

The variable is measured at age 30, and is defined as a dummy variable which takes value one if the individuals' highest level of completed education is in a STEM field, and value 0 otherwise. STEM is defined as either of the following categories of the Swedish SUN2000 education classification system (which is based on the ISCED-system):

- 42 Biological and Environmental sciences. (*Biologi och miljövetenskap*)
- 44 Physical Sciences, Chemistry and Geology. (*Fysik, kemi och geovetenskap*)
- 46 Mathematics and Natural Sciences. (*Matematik och övrig naturvetenskap*)
- 48 Information and Communication Technologies (ICTs). (*Data*)
- 52 Engineering and Engineering Trades. (*Teknik och teknisk industri*)
- 54 Manufacturing and Processing. (*Material och tillverkning*)
- 58 Architecture and Construction. (*Samhällsbyggnad och byggnadsteknik*)

Highest level of completed education at age 30 is a post-secondary degree

The variable is measured at age 30, and is defined as a dummy variable which takes value one if the individuals' highest level of completed education is a post-secondary degree, including short post-secondary educations; and value 0 otherwise. The variable is based on the Swedish SUN2000 education classification system (which is based on the ISCED-system).

Highest level of completed education at age 30 is a post-secondary STEM degree

The variable is measured at age 30, and is defined as the interaction of the two above dummy variables; highest level of completed education is in the STEM field and is a post-secondary degree.

Having completed at least some university studies by age 25

The variable is defined as a dummy variable which takes value one if the student has taken a non- zero amount of university credits between age 18 and age 25, and is zero otherwise. The variable is generated based on register information reported to the university credit administrative system LADOK. It includes the vast majority of universities and colleges.[23]

Employed at age 30

The variable is defined as a dummy variable for being employed, based on Statistics Sweden's employment indicator (syssstat).

Log of employment income age 30

The variable is defined as the natural log of the total annual earnings from employment and self-employment (personal firms) in SEK, after zeroes were replaced with ones, in order for individuals

[23] The most notable exception is the Stockholm School of Economics, a business school with around 2000 students, which was not included until 2011.

with zero earnings not to be recorded as missing observations. This simple solution can however be criticized (see Bellégo and Pape, 2021, Cohn et al., 2022, and Chen and Roth, 2024 for discussions on the issue).It is therefore an advantage to add the binary employment variable as an alternative labour market outcome.

Outcome variables: School and peer characteristics

School pupil-teacher ratio

The variable is defined as the number of students per teacher. The variable takes into account hours worked by weighing observations by the work time, i.e. counting a teacher working half time as half a teacher etcetera.

Share of qualified teachers in the school

The variable is defined as the share of teachers in the school that have the proper education for the subjects they teach. The variable takes into account hours worked in the same manner as the above variable, i.e. by counting a teacher working for example half time as half a teacher.

Share with a high educated parent

The variable is defined as the share of a student's peers who have at least one parent with a post-secondary degree. Peers are measured as the students who attend the same grade, school and type of track as the individual, and type of track is classified according to the five categories of Table A1.

Share that are Swedish born

The variable is defined as the share of a student's peers who are born in Sweden. Peers are measured in the same manner as for the Share with a high educated parent above.

Average Standardized Lower Secondary Final GPA

The variable is defined as the average value of the standardized final GPA from lower secondary education, among the students' peers. Peers are measured in the same manner as for the Share with a high educated parent above.

B1.4 Additional data choices

Below, we list additional data choices made:

- In the small number of cases where there was duplicate and conflicting information for an individual within a year, the variable was replaced as missing, unless it could clearly be inferred which one of the observations was the true one.
- The top and bottom 0.5 % observations, defined separately for the distributions for each year of data, were dropped for the variables household disposable income and the school level pupil-teacher ratio, as there were some very large outliers in these variables.
- The following data details differ between the main regression sample and the long-run analysis regression sample, due to slightly differences in data availability:
    - Household disposable income, which is included as a covariate, was in the main regression sample generated in the following manner, in order to take account of the fact that a student with separated parents may have different observations for household income for each parent: If household disposable income was the same for mother and father, we used either value. If household disposable income was not the same for mother and father, we used the average of the two values. If either the mother's or the father's income information is missing, we used the other parent's information.
    - Household disposable income was in the long-run analysis sample available in the form of each parents' individual consumption unit component of household disposable income. This means that each of the parents have been assigned an

amount that reflects their share of the household according to their assumed consumption weights. In this case, we kept the mothers' and fathers' disposable income variables as separate covariates.

- The student background covariates that vary over time were measured at age 15, i.e. the year before students normally start upper secondary education. For the first year of the data panel for the main sample, the variables are however measured the same year as the students started upper secondary education, since earlier information was not available in our data. The same holds for some of the variables in the regression sample of the long-run analysis.
- We dealt with missing observations in student background and local covariates in the following manner: The missing observations were replaced with constant values (we chose to replace them with the average of the respective variable) and dummy variables we added to the regressions indicating the incidences of missing values.

Appendix B2: Further Details on Voucher School Supply and Growth

Figure B1 shows the national annual distribution of students (from the first grade of upper secondary schools) by each track type, for voucher and traditional public schools. This figure demonstrates that after the initial years of voucher school growth, the distribution of schools across the five major tracks (three vocational tracks and two academic tracks) are broadly similar.

Table B1 shows descriptive statistics for students enrolled in the different types of schools. This demonstrates that there is considerable variation in the 'quality' of voucher schools, as measured by the characteristics of the students in these schools. Furthermore, on average the student populations look very similar in the two types of schools. For example, standardized final grades are 0.30 in voucher schools and 0.27 in public schools; the proportion of students with a high educated parent is 0.57 in voucher schools and 0.53 in public ones.

With respect to reimbursement rates for different track options for voucher schools, Figure B2 shows that the vocational programs in general receive higher reimbursement (higher voucher levels), consistent with the higher costs associated with these courses, arising from smaller class size and more equipment.

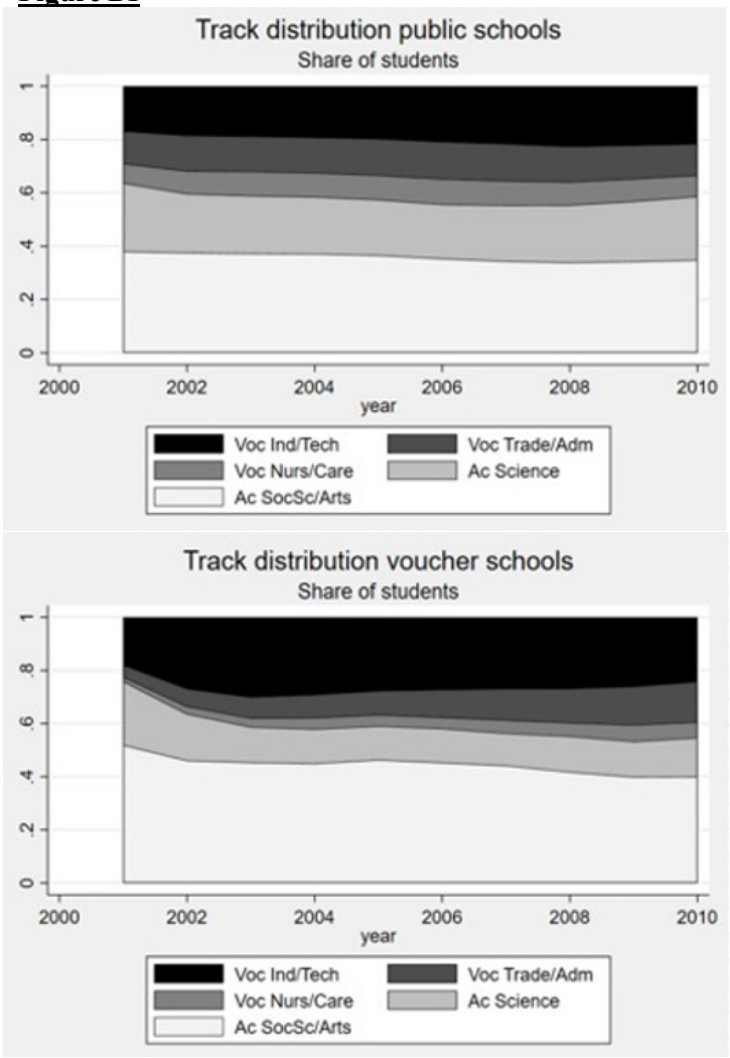
Finally, Figures B3 and B4 plot the relationship between vocational track enrolment in a municipality and the share of voucher schools, in levels and in changes. Figure B3 shows that the correlation in levels is negative. Figure B4, which shows changes in both variables between 2001 and 2010, demonstrates that there is a positive correlation, in line with our main set of results.

Table B1 Background Characteristics of Students, by Type of School

	All	Voucher	Public
Standardized final grade lower secondary education	0.28	0.30	0.27
	(0.74)	(0.76)	(0.73)
Male	0.51	0.52	0.50
	(0.50)	(0.50)	(0.50)
Swedish born	0.92	0.93	0.91
	(0.28)	(0.26)	(0.28)
Parent high education	0.54	0.57	0.53
	(0.50)	(0.50)	(0.50)
Log of household disposable income	12.87	12.91	12.86
	(0.48)	(0.51)	(0.47)
Mother employed	0.88	0.88	0.88
	(0.33)	(0.33)	(0.33)
Father employed	0.89	0.88	0.89
	(0.31)	(0.32)	(0.31)
Percent of students attending each option in 2001	100.00	9.46	90.54
Percent of students attending each option in 2010	100.00	29.49	70.51

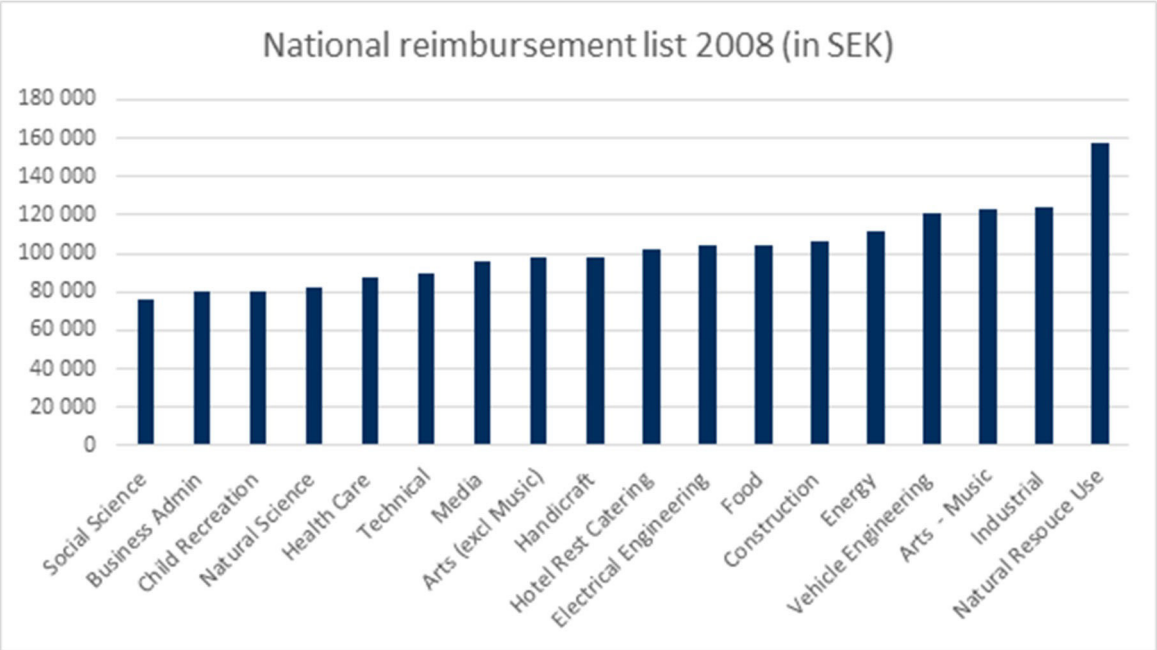
Note: Standard deviations are shown in parentheses. The sample is based on students attending the first year of an upper secondary school track. Students in rural locations, residing further than 5km away from any upper secondary school, are excluded, as are students attending the preparatory track. The preparatory track caters to students with too low grades from lower secondary school to qualify for an upper secondary track.

Figure B1



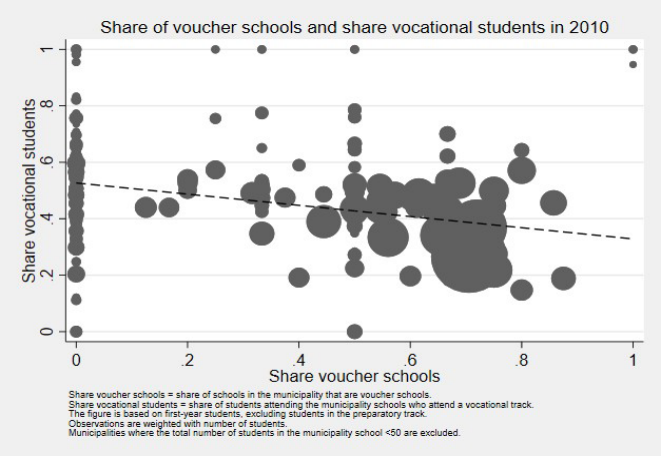
Note: The figures are based on all students attending first year of upper secondary school, excluding students in the preparatory track. The preparatory track caters to students with too low grades from lower secondary school to qualify for an upper secondary track.

Figure B2: National per student reimbursement per program



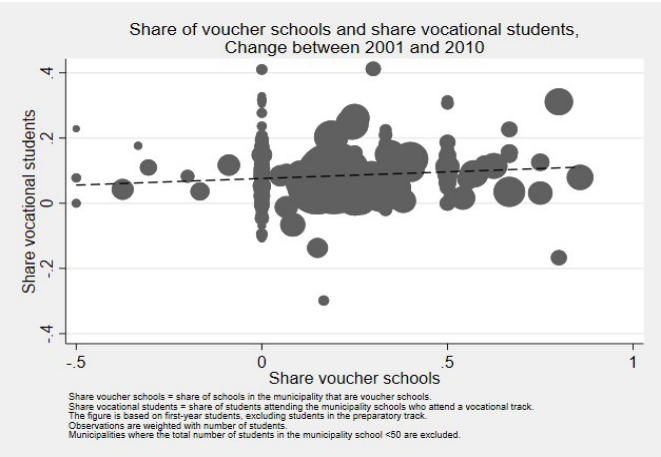
Note: this figure shows the national per student reimbursement per program as of 2008 (Swedish Krona in 2008 monetary value)

Figure B3: Relationship between vocational track enrolment and share of voucher schools, levels



Note: municipalities with no upper secondary schools, or whose upper secondary school/s/ has in total fewer than 50 students, are excluded, as such cases are likely to represent very rural locations with a very limited supply of upper secondary education within the municipality. Each dot represents a municipality, and is weighted by the student population size. Share vocational students measure the share attending a vocational track, among the students in the schools in the municipality. Share voucher schools represent the share of the schools in the municipality that are voucher schools.

Figure B4: Relationship between vocational track enrolment and share of voucher schools, changes, 2001-2010



Note: municipalities with no upper secondary schools, or whose upper secondary school/s/ has in total fewer than 50 students, are excluded, as such cases are likely to represent very rural locations with a very limited supply of upper secondary education within the municipality. Each dot represents a municipality, and is weighted by the student population size. Share vocational students measure the share attending a vocational track, among the students in the schools in the municipality. Share voucher schools represent the share of the schools in the municipality that are voucher schools.