**Working Paper No. 23-04**

# First generation elite: the role of school networks

Sarah Cattan
Institute for Fiscal Studies

Kjell G. Salvanes
Norwegian School of Economics

Emma Tominey
University of York

High school students from non-elite backgrounds are less likely to have peers with elite educated parents than their elite counterparts in Norway. We show this difference in social capital is a key driver of the high intergenerational persistence in elite education. We identify a positive elite peer effect on enrolment in elite programmes and disentangle underlying mechanisms. Exploiting a lottery in the assessment system, a causal mediation analysis shows the overall positive peer effect reflects a positive effect on application behaviour (conditional on GPA), which dominates a negative effect on student GPA. We consider implications for income mobility finding that encouraging further mixing between elite and non-elite students in high school could improve mobility across the whole distribution.

VERSION: May 2023

Suggested citation: Cattan, S., Salvanes, K. G., Tominey, E. (2023). *First generation elite: the role of school networks* (CEPEO Working Paper No. 23-04). Centre for Education Policy and Equalising Opportunities, UCL.
https://EconPapers.repec.org/RePEc:ucl:cepeow:23-04

## Highlights

- Kids from disadvantaged backgrounds are very unlikely to attend the top elite university programmes. They are also much less likely to interact with other kids whose parents achieved these elite degrees. We show that this difference in social capital is a key driver for the intergenerational persistence in education.
- Causally identify whether and why interacting with students from elite families (elite peers) during high school drives the decision to enrol in an elite degree programme for high and low SES students. To disentangle the mechanisms we use a causal decomposition method which exploits a lottery in high school assessments.
- There is a large peer effect for kids from disadvantaged backgrounds, telling us that overall interacting with elite peers does increase the likelihood of first generation elites, as it increases enrolment into the elite degree programmes. The causal mediation shows two important channels. First, grades fall from exposure to elite peers and this is driven by a teacher downgrade. Second, conditional on grades, exposure to elite peers raises applications to elite degrees through aspiration and role model channels and this effect is large enough to counteract the negative effect of grades.
- We have identified a mechanism to increase intergenerational mobility, by reducing segregation of students in high school.

## Why does this matter?

We should care about the lack of low SES students in elite degrees because it contributes to inequalities and the paper shows that providing role models / aspirations can help to encourage first generation elites. On the other side though, teacher assessments of low SES students can harm their chances and more education is needed to counteract this.

# First generation elite: the role of school networks

**Sarah Cattan**
Institute for Fiscal Studies and IZA

**Kjell G. Salvanes**
Norwegian School of Economics, CEPR, CESifo, HCEO and IZA

**Emma Tominey**
Department of Economics, University of York, HCEO and IZA

May 3, 2023

**Abstract**

High school students from non-elite backgrounds are less likely to have peers with elite educated parents than their elite counterparts in Norway. We show this difference in social capital is a key driver of the high intergenerational persistence in elite education. We identify a positive elite peer effect on enrolment in elite programmes and disentangle underlying mechanisms. Exploiting a lottery in the assessment system, a causal mediation analysis shows the overall positive peer effect reflects a positive effect on application behaviour (conditional on GPA), which dominates a negative effect on student GPA. We consider implications for income mobility finding that encouraging further mixing between elite and non-elite students in high school could improve mobility across the whole distribution

**Keywords:** Peers, Elite university, Subject choice, Social mobility; Teacher bias
**JEL codes:** I24, J24, J62

**Acknowledgements:** We have received useful comments from many including seminar audiences at the Yale University, Princeton University, University of Chicago, Columbia University, the University of Oxford, University of Turin, University of California San Diego,University of California Los Angeles, CEPEO Institute of Education UCL, University of Copenhagen, University of Sheffield participants at the NORFACE IFS workshop, York Workshop for Labour and Family Economics, LSE and Sapienza workshop on Opportunity and Mobility, Erasmus University Rotterdam, SOFIE, University of Bergen workshop on Health, Human Capital and Social Insurance, University of Essex, Royal Holloway University of London. All remaining errors are our own. The authors gratefully knowledge financial support from from the Research Council of Norway through project no. 262675, and project no. 275274, the ESRC-funded Centre for the Microeconomic Analysis of Public Policy (ES/T014334/1), the Rockwool Foundation, and the European Research Council for grants agreement no. 819752 - DEVORHBIOSHIP - ERC-2018COG and ERC-2014-CoG-646917-ROMIA.

## 1 Introduction

Recent research documents very high levels of socioeconomic segregation in elite graduate and post-graduate degrees. In the US, pupils from the top quintile of the parental income distribution have been found to be around 23 times more likely to attend an elite ('Ivy Plus') college than pupils from the bottom quintile (Chetty et al., 2020a). In Chile and the UK, graduates from fee-paying private high schools are over-represented in elite university programs by factors of 16 and 7 respectively (Barrios-Fernandez et al., 2022; Britton et al., 2021).

In Norway, the setting of this study, elite education is also highly selective and associated with top paying jobs. Elite programmes have a high school GPA cut-off around 40% higher than other degree programmes and only admit 3 to 4% of each birth cohort. As shown in Figure 1, even compared to non-elite graduate students, elite graduates are disproportionately present at the top of the income distribution at age 28-40, with many of them in leadership positions in the private and public sectors (Bütikofer et al., 2021; Kirkebøen, 2010).

Like in these other countries, elite programmes in Norway have an over-representation of students from high socio-economic backgrounds: 7 students come from the top 20% richest families for every student from the bottom 20%. Yet, many of the factors thought to be driving the socioeconomic segregation of students in elite degrees in the US, Chile and even the UK are unlikely to be at play in Norway and most European countries. There are no university tuition fees. There are no expensive private feeder high schools (Zimmerman, 2019; Barrios-Fernandez et al., 2022; Britton et al., 2021; Michelman et al., 2022). And there are no legacy enrolment policies (Chetty et al., 2020b). Instead, a centralised admission system allocates students to degrees based on their degree preference and their high school GPA. Comparable admission systems and no tuition fees are not only the standard in Nordic countries but also the preferred model in most European countries.

Why then are there still so few first generation elites in Norway? This paper argues that social capital is a key part of the answer. The component of social capital we focus on is the degree of exposure to peers from elite-educated families in high school ('elite peers' henceforth). We document that children from non-elite backgrounds are much less likely to have elite peers than children from elite backgrounds. We ask whether this lack of exposure causally hinders students from non-elite backgrounds to become first generation elite and, if so, what mechanisms drive this impact. These notions are tightly linked to the idea of economic connectedness in friendship networks, which Chetty et al. (2022) find to be positively correlated with upward mobility.[1]

A key innovation of the paper is to ask *why* being exposed to elite peers influences higher education outcomes. As mentioned above, for a student to be admitted to an elite degree, they need to have a high enough GPA *and* they need to express a preference for that program. Elite peers can potentially influence both margins through a number of mechanisms.

First, elite peers can influence GPA through at least two mechanisms. On the one hand, elite peers can have spillovers on the effort and subsequent learning of other students. These spillovers could arise, for example, if elite peers, who are likely to be high achieving and highly motivated, help their classmates learn more effectively and/or impart their aspirations for higher education onto them. On the other hand, in a system where GPA is partly based on teacher assessments like it is in Norway, the presence of elite peers could distort teacher's assessment behaviour. If teachers 'mark on a curve', a higher proportion of elite students will create downward pressure on the rank of other students since elite students are likely to be highly ranked. And if teachers are prone to implicit bias, a higher proportion of elite peers may also trigger further distortion in the way that teachers assess elite students relative to non-elite students of similar ability.[2] A priori, it is not clear whether elite peers would have a positive or negative effect on overall GPA. In this paper, we identify the effect of elite peers on overall GPA and exploit that GPA is based on scores to *both* blindly assessed and teacher assessed exams to disentangle these potentially counter-acting mechanisms from each other.

In addition to influencing high school GPA, elite peers could affect student's probability to enrol in an elite degree through a second margin - by influencing their university application decisions conditional on their GPA. Interactions with elite peers – and possibly with their parents – might change students' information set about the content of and returns to these programs and/or their beliefs about which programs are a good match for them (Lundberg 2020; Porter and Serra 2020; Michelman et al. 2022; Mani and Riley 2019). In turn, this might encourage students who have a GPA high enough for an elite degree to actually apply and/or to make better decisions regarding

[1]Another explanation is ability differences, but we remove this as a possibility by controlling flexibly for student ability.

[2]See Campbell (2015) and Doyle et al. (2023) for evidence of a teacher bias against lower income students and Papageorge et al. (2020) for evidence that teacher expectations are important for student outcomes such as college completion.

the set of programs to which they apply and increase their chances of admission. In this paper, we identify the causal impact of elite peers on elite degree enrolment conditional on GPA – i.e. via this second margin – by exploiting a unique feature of the Norwegian institutional context whereby a lottery assigns students to taking certain subjects as blindly assessed exams. We show evidence to support that this lottery provides a credible source of exogenous variation in GPA.

The paper is divided into four parts, each yielding a key finding. All the analysis is based on administrative register data tracking nine cohorts of students who finished middle school between 2002 and 2012 through their education and, for those who reach 30 or more by the last year of observation (2018), into the labour market. We link these individuals to their parents to identify their socio-economic backgrounds. Using school identifiers, we link them to their high school peers and peers' parents to create our measure of social capital - the proportion of peers' parents who have an elite education in the student's high school cohort.

The first part of the paper asks what effect elite peers have on the probability to enrol in an elite degree for students from different socio-economic backgrounds. To identify this effect, we exploit within school, between cohort variation in peer composition across nine cohorts of high school students. This approach has been extensively used since it was initially proposed by Hoxby (2000),[3] and we perform an extensive set of robustness checks to probe its validity in our context.[4]

Our first finding is that exposure to elite peers increases enrolment of the average student in an elite degree. However, this effect is three times larger for students with at least one elite educated parent (high SES) than it is for students with low educated parents (low SES).[5] Combined with the fact that high SES students are on average twice as likely to be exposed than low SES students to elite peers are, these estimates imply that the socioeconomic gradient in exposure to elite peers in high school explains 12% of the gap in elite degree enrolment between these two groups. In other words, the socioeconomic segregation of students into high schools is a significant driver of the intergenerational persistence in elite education in Norway.

The second part of the paper starts delving into the mechanisms of this effect and asks what

[3]Among others, see Angrist and Lang (2004); Lavy et al. (2011); Black et al. (2013); Cools et al. (2019).

[4]These include testing whether the variation in the exposure to elite peers across time within schools can be considered random, including school-specific linear trends, full interactions between school and cohort fixed effects, dropping schools whose variation appears non-random and controlling for a measure of school quality which varies across time which is teacher traits.

[5]We explore the sensitivity of these results to different definitions of elite and SES. A similar SES gradient emerges when SES is defined by parent income and when elite is defined by peers' parents' income or occupational prestige.

effect elite peers have on high school GPA. Using a similar identification strategy as above (this time with GPA as the outcome), we estimate the effect of elite peers on overall high school GPA and on each of its components. This unique feature of the Norwegian context helps us tease out whether elite peers have a different effect on blindly assessed exam scores (which provide the cleanest measure of underlying effort or knowledge) and on teacher assessed exam scores.

Our second finding is that exposure to elite peers is detrimental for overall GPA (and hence for chances to be admitted in an elite degree through that channel), and particularly so for low SES students. However, elite peers significantly *improve* students' scores on blindly assessed exams, which suggests that elite peers do have positive spillovers on the effort and knowledge of other students. What drives the negative effect of elite peers on GPA is a large negative effect of elite peers on teacher assessed exam scores. We show that this negative effect is partially explained by elite peers lowering the rank of other students. But, even within student cohort rank, the downgrade is larger for the low SES students, which could point to the possibility that the more elite peers there are in a cohort, the more teachers become implicitly biased against low SES students.

The third part of the paper aims to identify the effect that elite peers have on students' probability to enrol in an elite degree *over and beyond* their effect on GPA. This is challenging because GPA is endogenous with respect to individual elite degree enrolment. To estimate elite peer effects conditional on GPA, we therefore need to instrument GPA. We exploit exogenous variation in GPA resulting from a lottery inherent in the Norwegian examination system whereby schools randomise the subjects on which students are blind externally-assessed in their third year.[6]

Our third finding is that elite peers have a positive effect on the probability of enrolling in an elite degree conditional on GPA. We use these results to decompose the overall effect of elite peers on elite degree enrolment into i) an *indirect* effect of elite peers, working through their effect on GPA, and ii) a *direct* effect (conditional on GPA). This is a causal mediation analysis in the sense that it takes into account that the mediator, here high school GPA, is endogenous (Celli, 2021). We find that the positive direct effect of elite peers on students' application behaviour conditional on GPA dominates any negative indirect effect through GPA (driven by the teacher downgrade).

The fourth part of the paper turns to the labour market implications of the elite peer effects

[6] We show that student assignment to externally-assessed maths exams is both balanced on a number of students' observable characteristics and a strong determinant of GPA, making it a plausibly valid and relevant instrument for GPA.

we have uncovered. So far, we have shown that the lack of mixing between children from elite and non-elite backgrounds in high school is a key driver of the intergenerational persistence in elite education. Whether it also plays a role in driving the intergenerational persistence in income depends on the earnings returns to elite degrees across the SES distribution. The final part of the paper provides descriptive evidence using the subset of cohorts for whom we can observe earnings from age 30 and discusses the implications of this evidence for intergenerational mobility.[7]

Our fourth finding is that the earnings premium associated with elite degrees at age 30-32 is high and only slightly lower for low SES students than for high SES students. To quantify the impact that elite peers have on the earnings of students across the parental income distribution, we estimate how the intergenerational 'rank-rank' coefficient (from a regression of parents' percentile rank on their child's rank at age 30-32) varies with the degree of exposure to elite peers. We find that exposure to elite peers in high school raises mobility at the bottom of the parental income distribution, but also exacerbates intergenerational persistence at the top. A direct implication of this finding is that a policy that encourages further mixing between elite and non-elite students in high school could improve mobility across the whole distribution. We illustrate this point by way of a series of simulations which reassign low SES students in schools with low exposure to elite peers, into schools with a high exposure - and vice versa for the high SES students.

Our paper speaks to three strands of literature. First, it contributes to the intergenerational mobility literature (for example, see Corak et al. (2014) and Adermon et al. (2021)). Specifically our paper relates to a small but growing literature on the role of social capital in driving mobility (Chetty et al., 2022, 2020b; Barrios-Fernandez et al., 2022). We provide causal evidence that the segregation of children from elite families at the high school level is a driver of the persistence in elite education and in income, especially at the top of the distribution. Whilst previous studies are based on data from the US or Chile, these countries can to some extent be considered outliers in terms of their high level of income inequality intergenerational mobility as indicated by the Gatsby curve (Corak, 2013). The Norwegian context is more comparable to other European countries and

[7] Due non-classical measurement error in earnings both for fathers and for sons estimates of intergenerational mobility may be downward biased depending on which age is a good predictor for life-time earnings. Bhuller et al. (2017) finds that earnings measured in the early 30s are a good predictor of life time earnings for Norway using data across the whole life-cycle. Nybom and Stuhler (2017) finds that using income ranks in contrast to log earnings is less dependent on the exact age of measuring earnings. We use the rank of means earnings for ages around 30 for the students.

opportune to study the role of social networks in driving inequality, since other drivers, such as credit constraints and legacy university admission systems, are less likely to be salient, if at all

Second the paper relates to the large literature on the heterogeneous effect of peer characteristics on educational and economic outcomes. It is most closely linked to Bertoni et al. (2020) and Cools et al. (2019) who also focus on the effects of peers with high parental education and their heterogeneity (by socioeconomic background and by gender, respectively) and closely related to Dahl et al. (2021) and Altmejd et al. (2021) who identify sibling peer effects in the field of high school and college major respectively. Our paper exploits unique features of the institutional context to provide a rich description of the mechanisms underlying the elite peer effect on elite degree enrolment and to quantify the relative contribution of these peer effects working through GPA and over and beyond GPA.

Finally, our analysis of elite peer effects on blind and non-blind assessments speaks to the literature on the impact of teacher discretion on academic achievement and long-term outcomes. Several papers in the literature contrast these two types of assessments to provide evidence of teacher stereotypes (Lavy, 2008; Lavy and Sand, 2018; Burgess and Greaves, 2013), while other papers directly elicit teacher bias using Implicit Bias tests (Carlana, 2019; Alesina et al., 2018). To the extent we measure student ability well, our results are consistent with teachers downgrading similarly able students on the basis of their SES status and show that such behaviour can have profound consequences for social mobility. Our results also suggest a clear policy implication not only for Norway but also for others systems, including the US, where university admission is partly based on teachers discretionary evaluations of performance: increasing the weight that blindly assessed exams has in GPA can increase the chances that high-ability low SES students become first generation elite.

## 2 The Norwegian Education System

**High school** Norwegian education has been compulsory until the age of 15-16 since 1959; all students must now complete seven years of primary school followed by three years of middle school (Black et al., 2005a).[8] After completing these 10 years of education, students decide whether to

[8]The seven years of primary school includes a year of preschool education, which was made mandatory in 1996.

continue their education in high school or to drop out to join the labour force. Those who continue onto high school choose between an academic track and a vocational track. The academic track, which we focus on in this paper, lasts 3 years and is geared towards preparing students to attend higher education.

The assignment mechanism of students to high schools varies across counties and cohorts. In some counties (including all rural counties), schools have catchment areas and geographical distance determines student high school allocation.[9] Other counties have a free high school choice system where intake is centralised and based on middle school GPA. During our period of analysis, which focuses on cohorts graduating from middle school between 2002 and 2010, eight out of nineteen counties had free school choice. Because these areas tend to be the most densely populated areas, the majority of high school students in our sample had free school choice.[10]

**Higher education** Higher education institutions include universities (in Bergen, Oslo, Trondheim and Tromsø) and university colleges. Since the early 2000s, Norwegian universities offer three-year bachelor degrees and five-year combined bachelor-master degrees. 98% of university students attend a public institution, and even private institutions are funded and regulated by the Ministry of Education and Research. There are no tuition fees for attending a public higher education in Norway, and most students are eligible for financial support (part loan/part grant) from the Norwegian State Educational Loan Fund (NSELF).[11]

To pursue a higher education, students must apply for a combination of a field of study at a specific institution (e.g. law at the University of Oslo). Since the late 1990s, admission to public higher education institutions has been centralised and is based on student ranking for programmes and high school GPA, conditional on students having completed the requisite high school modules (e.g. maths at high school is required for a maths degree). Every year, the deadline for applying to programmes is mid-April, which is when students first submit their ranking of up to fifteen pro-

[9] A small number of private colleges instead require tuition fees for students - only one in our sample.

[10] Some counties have changed their assignment systems over recent years. For example, the two largest cities in Norway - Oslo and Bergen - have varied their intake systems over recent years (Bütikofer et al., 2020; Dalla-Zuanna et al., 2020). Oslo moved from a local catchment to school choice admissions based system between 2006-2009 but reverted back from 2010; whilst Bergen moved to school choice admissions from 2006 onward. We test and reject that our results are sensitive to the type of admissions system.

[11] The NSELF is a national body founded in 1947 with the task to provide student aid in the form of direct transfers or scholarships and to issue loans under conditions specified by the Norwegian state. Since the 1980s financial aid is not dependent on the student's own means or that of their parents.

grammes to a central organisation - the Norwegian Universities and Colleges Admission Service.[12] Students can adjust their ranking until July. Then offers are made sequentially where the order is determined by the students' application score derived from the student's high school GPA.

**Elite degrees** Whereas 'elite' higher education refers to highly competitive, private institutions with high tuition fees such as Ivy League colleges in the US (e.g. Chetty et al. 2020b) and 'Russell Group' universities in the UK (Britton et al., 2021), in Norway 'elite education' refers to a set of specific degrees at specific institutions that are both highly selective and associated with the best earnings outcomes. Specifically, elite degrees are five-year masters degrees in a select set of subjects at specific universities, and we follow Bütikofer et al. (2018) in defining elite programs as degrees at the master level (or above) in Economics from the Norwegian School of Economics, Engineering at the Norwegian University of Science and Technology, Engineering School in Trondheim or Norwegian University of Science and Technology and in Economics, Law or Medicine from the University of Oslo, Bergen, Trondheim and Tromsø. Not only are these elite programmes associated with high earnings, but a majority of future leaders in the private and public sectors are recruited from these institutions (Kirkebøen, 2010; Bütikofer et al., 2021). Figure 1 plots the earnings percentile distribution across education groups and confirms that the elite educated are positioned high in the income distribution.

Similarly to the US or the UK, access to elite degrees is highly competitive: only a very small proportion - around 3% per birth cohort - attend these elite degree programmes. Important differences with other contexts, however, are that there are no tuition fees for these degrees and no easier access for legacy students - because a centralised admission system allocates students based only on high school GPA (given the student's ranking of programmes). As we show in the next section however, despite these equalising features of the higher education system in Norway, there is a very strong socioeconomic gradient in the likelihood to pursue a higher education and an even stronger one in the likelihood to pursue an elite degree. This is true even conditional on previous ability, measured through national exams taken prior to high school. Our paper aims to better understand the role that high school social networks play in driving these inequalities.

[12] The Norwegian Universities and Colleges Admission Service handles the admission process to all universities and to most university colleges, and therefore to all elite degrees we consider in this paper.

**High school GPA** High school GPA is a combination of grades on three types of exams: i) teacher (internal) assessments, ii) grades on externally (blindly) assessed exams, and iii) grades on oral exams assessed by both the student's teacher and an external examiner. In each of the three years of high school, students receive a teacher assessment on all subjects. In addition, they must take several mandatory exams in May or June of each academic year. In their first year, 20 percent of students are chosen randomly to sit for a final exam in one course. In their second year, all students sit a final exam in one course, either oral or written and the subject of these exams is chosen at random at the county level of each school.

In their third year, all students take four exams: one written exam in Norwegian language, two written exams in two other subjects and a final oral exam in one other subject. It is the responsibility of the county to allocate a student to a topic for the written examination, with the exception of mandatory exams (Norwegian in the third year). As described in detail in Andersen and Lokken (2020), there are several administrative procedures in place to avoid any non-random selection of students for particular courses or type of exams, and in fact there are no incentives to do so. We make use of this lottery later in the paper (section 7). Like Andersen and Lokken (2020), we find strong support for the random assignment of exam subject within school and programme of study in our sample.

## 3 Data and descriptive statistics

### 3.1 Data source and definition of key variables

Our data comes from Norwegian register and administrative data that have been linked by Statistics Norway. We select our sample to include all students finishing middle school and entering the academic track of high school between 2002 and 2010. The linked data allows us to follow these students from middle school through to high school, onto university (if they ever enrol) and the labour market. The data links students' educational records to a rich set of information on their parents, including parental education, occupation and income. It also contains school identifiers, which allows us to identify students' peers.[13]

[13] By far most students starting high school in Norway do this within the Norwegian school system. There is no tradition of attending high school in other countries. Some families will of course move to another country during high school, and we lose track of them in the data, but this is negligible.

Data on individuals' education attainment comes from the national education database, which contains codes for the highest completed level of education.[14] We use this information to define our main outcome variable $Y_{isc}$ as an indicator for whether student $i$ entering high school $s$ in cohort $c$ enrols into an elite degree within six years of completing middle school.[15] As mentioned earlier, elite degrees are defined as the set of 5-year bachelor/masters degree in law, medicine, and STEM obtained in the best institutions of the country (see full list of degrees in section 2). Even though high school is only three years long, we define the outcome as enrolling into a degree within six years of completing middle school because it is very common in Norway to have one or two gap years between high school and university in order to travel, work or complete military service (which has been mandatory for men and women since 2015).[16]

We define the student's peer group as all students entering the same high school in the same year. On average students are exposed to 95 high school peers.[17] We construct our main treatment variable, $P_{-isc}$, as the proportion of parents who have an elite education in the student $i$'s cohort $c$ in high school $s$ (excluding student $i$'s own parents).[18] In a sensitivity analysis presented in subsection 5.4, we make use of two alternative definitions of 'elite' peers. The first is based on peers' family income and measures the proportion of elite peers as the proportion of peers' parents in the top 5% of the income distribution (of high school students' parents).[19] The second is based on occupational prestige and defines the proportion of elite peers as the proportion of peers' parents working in an elite occupation, i.e. as a lawyer, doctor or in a STEM occupation (using the occupation classification into STEM from Deming and Noray (2018)).

In all regressions we control for a set of covariates relating to the individual student or their

[14]These codes are in the NUS2000 format, which is a six-digit code containing highly detailed information on both the level and field of a person's education.

[15]We focus on enrolment in elite degrees as opposed to completion of an elite degree as our main outcome because our interest in this paper lies in how peers shape subject choice. Peers could also shape students' ability to complete the degree they enrol in, but studying this mechanism is beyond the scope of this paper.

[16]2% of the sample of students with a degree in STEM, law or medicine study for the degree abroad but, as it is not possible to link the institution, these students are excluded from our sample.

[17]This varies between 30 and 159 at the 10th and 90th percentile. We explore in sensitivity the robustness to results dropping small schools.

[18]Note that we use the same grouping of elite education for students and parents since these elite groups have been stable over time in terms of being very competitive to enter and a basis for recruitment to top positions in the labour market, paying top salaries (see Strømme and Hansen 2017).

[19]This measure of family income is constructed by summing the income of mothers and fathers at the end of middle school and deflating to 2020 prices. It is calculated within the sample of academic high school students. When compared to the overall population, families in the top 5% of the family income distribution of academic high school students represent 3% of the whole population.

parent, which are all predetermined with respect to the student entering high school. Covariates about the student include an indicator for gender; whether they were born in Norway and their middle school GPA (standardised to have a mean of 0 and standard deviation of 1 within cohort of all middle school students). Covariates for the student's parents include indicators for whether the mother and the father's highest levels of education are compulsory education, high school degree, or university/post-graduate degree; a variable measuring whether the student has zero, one or two elite educated parents and an indicator for whether the student's household is in the richest decile (based on the distribution across all cohorts in our sample of household income distribution measured at the end of middle school and deflated to 2020 prices).

Throughout the paper, we distinguish between two groups of students with different socioeconomic status, which we define based on the education of their parents. The 'low SES' group includes students with at least one parent with no further education beyond compulsory education (10 years of education) and no parent with an elite education. The 'high SES' group includes students with at least one parent with an elite education and no parent with compulsory education. In paragraph 5.4, we also present our main results for the intermediate SES group in section 5, though we focus most of the discussion on the low and high SES groups for expositional simplicity. There we also show how results compare when defining SES based on household income instead of parental education.

### 3.2 Descriptive statistics of the analytical sample

Table 1 provides summary statistics for the individuals in our analytical sample in the first column, and in the low and high SES sample (as defined above) in the second and third columns. Our sample has close to 178,000 students studying in 556 high schools spread throughout Norway. As mentioned earlier, most students in the academic track attend a higher education institution, but only one in ten pursue an elite degree, reflecting their high selectivity. This represents 3% of the cohort graduating from middle school. Among students pursuing an elite degree, close to 70% complete a 5-year STEM or Economics/Business masters degree, while 20% complete a law masters degree and 10% complete a medical degree.

The second and third columns of Table 1 compare the probability of enrolling in a higher education degree between the low and high SES groups and confirms the existence of a parental education gradient[20] The gradient is much more pronounced when it comes to enrolling in an elite degree, with the probability of enrolling being five times as large for high SES students than for low SES ones. These patterns align with Bütikofer et al. (2021), whose findings suggest that, although Norway has one of the lowest intergenerational income elasticities in the world, intergenerational education persistence is high and comparable to other countries, including the US, with much lower levels of income mobility.

Moving down the table, we see that exposure of students to elite peers is socially graded: the proportion of elite peers in their school cohort is twice as high among high SES students than it is for low SES students. This social gradient is also visible with the other definitions of elite peers above. This segregation of elite students is a result of the admission system in Norwegian, which leads students to self-select into high schools either based on ability (in municipalities with free choice) or income (in municipalities with catchment areas).

Other statistics included in Table 1 show that high school students are disproportionately female (60%) and selected on family income, as 25% of their families have income in the top 10% of the income distribution. They are also selected on ability: the average middle school GPA in the sample is 0.67 standard deviations above the GPA of the average middle school student. As mentioned in section 2, high school GPA is an average of grades on the three types of assessments taken by each student (teacher assessments, written and oral assessments) across the three years of high school. We standardise GPA within each cohort of high school students to have a mean of 0 and standard deviation of 1. As Table 1 show, high SES students perform better than low SES students.

To look at long-term earnings implications of our results (section 8), we use data on labour market earnings (before taxes and transfers) for ages 30-32. While our full sample includes students born between 1986-1993, for this part of the analysis we use the cohorts born 1986-1988 for whom we can observe earnings at least from age 30. We smooth out the transitory component of income as much as possible by calculating, for each individual, the mean income across the available years.[21] To analyse the effect of exposure to elite peers on long run outcomes, we calculate the student's

[20] Note that about 50 percent of a cohort attend vocational high school, and by far most of them are recruited from low SES backgrounds.

[21] Bhuller et al. (2017) suggest rank stability of earnings from age 30. Our last year of earnings data is measured in November 2018, so for those born in 1986 income data is available at the full age range of 30-32; whilst for those born in 1988 income is available at age 30. We deflate earnings to 2020 prices.

earnings percentile rank within each birth cohort.[22] Whilst percentile rank of students in the total sample is 58 on average, low SES students in our sample average at the 55th percentile, whilst high SES students at the 65th percentile. To estimate intergenerational mobility regressions, for each student, we calculate parents' percentile rank of income by taking the average of real household income when the child was 15-19, the ages when students make decisions about the pursuit of elite education (see Chetty et al. 2020a). The percentile rank of income is calculated across the population of parents.

## 4 Empirical strategy

We start the analysis by estimating the effect of exogenously increasing $P_{-ics}$, the proportion of elite educated parents in student $i$'s cohort $c$ in high school $s$. Our strategy identifies this effect from within school, between cohort variation in outcomes and in the proportion of elite peers in the student's cohort. We operationalize this strategy by estimating the following benchmark specification by OLS:

$$Y_{ics} = \beta_1 P_{-ics} + X_{ics}^{'}\beta_2 + \alpha_s + \rho_c + \epsilon_{ics} \tag{1}$$

where $Y_{ics}$ is an indicator for whether student $i$ in cohort $c$ and school $s$ and enrols in an elite degree within 6 years of graduating from middle school; $P_{-ics}$ measures the proportion of cohort-school peers' parents who have an elite degree, excluding student $i$; $X_{ics}$ is a vector of student $i$'s characteristics (gender, Norwegian born, middle school GPA, mother and father's education, proportion of own parents with an elite degree, and whether family income in the top decile); $\alpha_s$ is a school indicator; $\rho_c$ is a cohort effect; and $\epsilon_{ics}$ is an error term. Our benchmark model assumes a linear peer effect, but we test for non-linearities in sensitivity analyses. We cluster standard errors at the school level to account for unobserved correlation of error terms within schools and follow Hoxby (2000) in weighting regressions by school size to take account of the parent peer variables group averages, taken from groups of different sizes.

In equation (1), the parameter of interest is $\beta_1$. This parameter will be identified with year-to-year variation in exposure to elite peers within schools, which can be thought of a luck. The

[22] We use data on the full birth cohort (and not only our sample members).

standard deviation of our treatment variable is 0.056 in the raw data and reduced by less then half to 0.026, once we remove school and cohort effects. To visualise this identifying variation, we plot the time series of the proportion of parents with an elite degree across the years of entry into high school, for a randomly picked school within each decile of average cohort intake size. As Figure A2 shows, whether the school is small, medium or large there is a lot of movement from one year to the next in the exposure to elite peers.

The within school, cross-cohort variation in $P_{-ics}$ at the core of our empirical strategy arises from slightly different reasons, depending on the high school's admission system. In schools where admission is distance-based, year-to-year variation in the proportion of parents with an elite degree in a given school results from year-to-year demographic changes in the composition of families living in the area (as in Hoxby (2000)). In schools where admission is GPA-based, variation in $P_{-ics}$ comes from year-to-year variation in the parental education of students whose middle school grades are high enough to be admitted into a certain school.[23]

OLS estimates of equation (1) will be unbiased if cohort-to-cohort variation in the proportion of elite educated parents is random within schools and conditional on student characteristics (including student middle school GPA). By controlling for school fixed effects and middle school GPA, our empirical strategy allows families to select their children's high school based on their knowledge of the composition of the school and based on their children's ability (and other characteristics). However, as explained in Hoxby (2000), the strategy relies on the idea that there is some variation in adjacent cohorts' peer composition within a school that is idiosyncratic and beyond the easy management of parents and schools. That is, "even parents who make very active decisions about their child's schooling cannot perfectly predict how their child's actual cohort within a school will turn out" (Hoxby, 2000).[24]

The identifying assumption would break down if $P_{-ics}$ is correlated with unobserved, time-varying determinants of the student's achievement, conditional on the controls included in the model. This could happen if certain types of families systematically moved to areas (in catchment area municipalities) or applied to schools (in free school choice municipalities) based on their

[23]While both sources of variation should be equally valid to identify the parameters of interest, we also re-estimate our model in the two subsamples defined by the procedure for admissions to high schools. The results are almost identical in the two samples (results available upon request).

[24]Here we focus on idiosyncratic variation in cohort composition, as opposed to classroom composition, so we need not worry about schools and parents manipulating the assignment of students to classrooms.

knowledge of trends in the composition of the student body and/or in school outcomes, resulting for example from changes in neighbourhood quality (e.g. gentrification) and/or in school quality (e.g. change in school management) over time.

Our model does control for a number of time-varying student and family characteristics, including student's middle school GPA, parental education, parental income and occupation, which could already capture such trends. So the threat to identification arises only if changes in self-selection of families into certain schools were not well proxied by those variables. This could for example happen if, within a parental education and income group, parents with a greater taste for children's education achievement self-selected into schools with the reputation to be improving or to be increasingly attracting like-minded families. In this case, $P_{-ics}$ could be positively correlated with $\epsilon_{ics}$, which would lead to an over-estimate of the causal effect of elite peers on elite degree enrolment.

While our identifying assumption is untestable, we perform a number of robustness checks to gauge its likely validity. First, we re-estimate our main model in equation (1) including school-specific linear trends in order to control for trends in students' characteristics and/or school characteristics which may not be captured by the controls included in the model. That is, we estimate the following model:

$$Y_{ics} = \beta_1 P_{-ics} + X'_{ics}\beta_2 + \alpha_s + c \times D_{is} + \rho_c + \epsilon_{ics} \tag{2}$$

where $c$ is a cohort (linear) trend and $D_s$ is an indicator for whether the student is in school $s$.

A limitation of this first test is that time trends in outcomes may not be captured by the linear term well. We therefore perform a second robustness check, which pools the data for low and high SES student and estimates the model this time including a full set of interactions between school and cohort fixed effects. In this model, it is possible to identify the difference in the elite peer effect between the low and high SES group. Our third test is based on an idea proposed by Hoxby (2000) and referred to as 'drop if more than random' in this paper. This check consists in re-estimating equation (1) on the sample of schools excluding those where within school, between cohort variation in $P_{-ics}$ is greater than what would be observed if such variation was random.[25]

[25]We first regress for each school the proportion of elite peers on a constant and a quadratic in years, estimating the school-specific time trends. Next the cohorts for each school are randomly reordered five times. If the reordering

As a main concern with our identification strategy is that elite parents select into schools of improving quality, we implement a fourth check, whereby we aim to control for observable changes in school quality. We pinpoint one specific dimension of school quality: the characteristics of teachers teaching in school $s$ to cohort $c$, which we can also measure in our data. In this test, we augment our benchmark specification with three variables measuring teacher traits at the school level: the proportion of female teachers, the proportion of teachers from a professional background, the proportion from a low skilled background and the average age of teachers (as a proxy of teacher experience).

As we discuss in subsection 5.2, the estimates of coefficients $\gamma_1$ for the low and high SES samples (and of the difference between the two in the second check) in all these robustness checks are very similar to the benchmark estimates in Equation 1. This provides us with strong confidence that our identifying assumption holds in this context. Nevertheless, we also perform a series of placebo tests checking whether the within school variation in the proportion of elite educated parents is associated with changes in student birth outcomes. We pick birth outcomes because these student characteristics cannot be causally affected by peers but are likely to be correlated with the unobserved characteristics of other students selecting in the same schools. As we also discuss in subsection 5.2, the results of these placebo tests confirm that our treatment variable is unlikely to be correlated with unobservable time-varying determinants of achievement.

## 5 Elite peer effects on elite degree enrolment

### 5.1 Benchmark results

The estimates of Equation 1 for the full sample are reported in Table 2. Across all students (column 1), exposure to elite peers significantly increases average students' enrolment in elite education. A one standard deviation (SD) increase in the proportion of elite educated parents in a school-cohort leads to a 2.6 percentage point (ppts) increase in the likelihood that students in this school-cohort

of cohorts results in the original ordering, the process is repeated until the new ordering does not reflect the true order. After each random reordering, the regression of the proportion of elite peers on a constant and a quadratic in years is repeated, thereby estimating the time trends that would occur if cohorts were randomly assigned within a school. Following Hoxby (2000), if the $R^2$ of the regression using the true cohorts is 1.05 times the smallest $R^2$ of the five regressions with false assignment of cohorts, then the school is flagged as having changes in the composition of elite peers as "more than random". The benchmark estimation is then repeated on the sample of schools which have not been flagged.

enrol in an elite degree. Columns (2) and (3) of Table 2 report the estimates of $\beta_1$ in the benchmark model in the samples of low SES and high SES students and show that the effect of elite peers in one's high school cohort is three times larger for high SES students than it is for low SES students (4 ppts vs 1.3 ppts). Both effects are statistically significant from zero and different from each other. These estimated peer effects are economically significant, comparing to around one third of the size of the gender differences in enrolment (see Table A1 for the full set of results).

Combined with the summary statistics presented in Table 1, these results imply that low SES students face a double disadvantage: not only are they exposed to a smaller share of elite peers in their school cohort than high SES children are on average, but being exposed to elite peers is also less beneficial to their future educational outcomes than it is for high SES children.[26] To measure the contribution of these two sources of disadvantage to the SES gap in elite education enrolment, we perform an Kitagawa–Blinder–Oaxaca decomposition of the gap in elite education enrolment between low and high SES students. We re-estimate the benchmark model estimated on the sample pooling the low and high SES subsamples and use the estimates of the model to compute the SES gap in elite degree enrolment that is attributable to the average SES gap in the explanatory variables and to the SES gap in the coefficients associated with these variables.

According to this decomposition (Table A2), the SES gap in average exposure to elite peers in high school explains 1.5 ppts or 7.2% of the SES gap in elite degree enrolment, while the SES gap in the *effect* of such exposure explains another 1 ppt or 4.8% of the SES gap in elite degree enrolment. To get a sense of the relative importance of elite peers in explaining the SES gap in elite degree enrolment, we present the results of the decomposition for a selected set of covariates included in the model in the same table. For example, the SES gap in ability (as measured by middle school GPA) explains 5 ppts or 24% of the SES gap in elite degree enrolment.

### 5.2 Validity of identification strategy

As described in section 4, we perform a number of checks to probe the validity of our identification strategy. We first re-estimate our main specification augmented with school-specific linear trends according to the specification in Equation 2. The results of this specification, which are reported

[26]See section A1 for heterogeneity by gender which shows that the SES gaps exist within gender, but also across gender as the elite peer effect is larger for male students than females.

in column (2) of Table 3, are very similar to those from our benchmark specification (included in the first column of the table for easy comparison).

Second we allow for further flexibility in the school specific trends, by including full interactions between school and cohort fixed effects. The results, presented in Panel C) estimate a peer effect for high SES students of 0.05 and for low SES students as 0.019, displaying the same SES gradient as our benchmark specification.

Next, we perform the 'drop if more than random' check, whereby we re-estimate the benchmark model on the sample of schools for which the cross-cohort variation in the proportion of elite families across cohorts is in line with variation from a random or fictitious ordering of cohorts. The estimates of the model on this sample, reported in column (4) of Table 3, are also very similar to those obtained on the whole sample. Finally, column (5) of Table 3 controls for the potential presence of changing school quality within schools across years, by adding time-varying school average teacher traits. Again, the estimate of the elite peer coefficient is very similar to the benchmark.

Table A3 reports the results of our placebo tests. Each row reports the coefficient on the elite peer variables in Equation 1 where the dependent variable is a different birth outcome. As expected, the exposure to elite peers during high school is unrelated to outcomes measured before high school. Together with the highly robust estimates of the elite peer effect, we take this evidence as strongly supportive of our identifying assumption within our benchmark model.

Finally, we consider several reasons why our identification strategy may not be valid for the whole sample and test the robustness of our findings in different sub-samples. Specifically, we consider whether our results vary across samples defined by school size, high school admission mechanism, student birth order, and high school majors. We also consider whether our results are likely to be influenced by measurement error in our treatment variable. These additional robustness checks are all described in Appendix subsection A2.1 and reported in Table A5. Overall, our benchmark results are highly robust, which further strengthens our confidence in our identification strategy.

### 5.3 Specification checks

Our main model assumes linear elite peer effects, but one may argue that those could be non-linear and vary either with the degree of exposure to elite peers or with student ability.[27] We discuss these potential mispecification of our benchmark model here.

If elite peers were non-linear in exposure to elite peers, the social gradient in elite peer effects we reported in Table 2 could merely reflect that low SES students have lower average exposure to peers than high SES students. To test this possibility, we re-estimate our main model, this time allowing the effect of elite peers to enter the model quadratically. Figure A3 plots the densities of $P_{-ics}$ for each SES group and the marginal effect of the proportion of elite families as implied by the estimates of this specification for each group.[28] The graph shows there is common support for most of the distribution of the treatment variable across the groups. Moreover, there is little evidence of non-linearity in the effect of the proportion of elite families on students' outcomes through most of the distribution of the treatment variable.[29] This confirms that our finding of a socio-economic gradient in the elite peer effect is not driven by a mispecification of our benchmark model.

SES and student ability are positive correlated, so the SES gradient in elite peer effects we have established could reflect a student ability gradient. To assess whether that is the case, we re-estimate the benchmark model, where we now allow for the effect of elite peers on enrolment to vary with middle school GPA, our measure of student ability. Figure A4 plots out elite degree enrolment predicted by the estimates of this model for the low and high SES subsamples. Interestingly, for low ability students, the likelihood to enrol into an elite degree programme is very low for both low and high SES students. Moving along the ability distribution, the SES gradient materialises and increases with student ability. In other words, compared to high SES students of the same ability, it is the low SES students with medium to high ability who face the greatest disadvantage in terms of the benefits of being exposed to elite peers.

[27] Several papers in the related literature have shown empirical evidence of non-linearities when considering the effects of high achieving peers. For example, Feld and Zoelitz (2017), Lavy et al. (2011) and Tincani (2017)

[28] That is, we estimate: $Y_{isc} = \beta_{11} P_{-ics} + \beta_{12} P_{-ics} \times P_{-ics} + X'_{ics}\beta_2 + \alpha_s + \rho_c + \epsilon_{ics}$. The estimates of the coefficients $\beta_{11}$ and $\beta_{12}$ are then used to compute these marginal effects and are reported in column (6) of Table 3.

[29] The one exception is for the high SES group, for whom the coefficient on the quadratic term is negative and statistically significant and the non-linearity kicks in at high levels of $P_{-ics}$ where there is little support. Importantly, across the distribution of proportion of elite families, the peer effect is higher for high SES students.

### 5.4 Robustness to alternative definitions of elite and SES

**Definitions of elite** Parents who graduated from an elite education programme are elite in a way that represents a mix of knowledge and human capital, as well social standing (Barrios-Fernandez et al., 2022). Elite educated parents are likely to have information about elite institutions. They are also likely to have high levels of income and consumption and possibly to hold prestigious occupations within their community. All these could be reasons why high school students are influenced or inspired by peers' elite parents to pursue an elite education themselves.

To tease out whether our measure of elite peers (based on education) is appropriate to capture these various facets to elite, we estimate our benchmark model with two alternative definitions of elite peers: the proportion of peers' parents with an income in the top 5% of the family income distribution and the proportion of peers' parents with a prestigious occupation (lawyer, doctor or STEM) (see section 3 for exact definitions). Like our benchmark measure of elite peers, all these variables are standardised to have mean 0 and standard deviation 1.

Exposure to parents with an elite, prestigious occupation increases enrolment into an elite degree programme by 0.4 and 1.1 ppts for low and high SES students respectively (column 2). Similarly, exposure to parents with an elite level of income increases enrolment by 1.1 and 2.4 ppts for high and low SES students respectively (column 3). When including all three measures of elite in the regression together in column 4, interestingly the coefficients on exposure to elite educated parents is very robust and in general the other definitions of elite peers are no longer significant - except for the low SES exposure to elite income parents. This suggests that defining elite peers as the proportion of elite educated parents as we do in our benchmark results is likely to pick up the different mix of educational achievement and social status held by those parents.

**Definitions of SES** Our definition of household SES based on parental education is motivated by an interest in understanding why intergenerational persistence in education is so high. Our main results compare impacts of elite peers for two SES groups defined at the bottom and top of the parental education distribution, and we test the robustness of our findings to alternative definitions of SES based on parental education as well as on family income.

Figure A5 shows that the effect of elite peers for the intermediate SES group (comprising of children who are neither in the low nor the high SES groups) lies in between the effect we estimate for the low SES and high SES samples. Accordingly, if we redefine the low SES group as those with no elite educated parents, we still find a SES gradient that is, as expected, flatter than in our benchmark estimation.[30] Moving to an income-based definition of SES, Figure A6 plots the effect of exposure to elite peers across students' family income percentile rank and shows again a SES gradient in the effect of elite peers. The gradient is flatter than the benchmark SES gradient, as shown by the two lines intersecting with the y-axis at the points representing the estimates from Table 2, however our benchmark estimates lie within the confidence intervals when SES is defined by income. Overall, these results provide reassurance that our findings are not driven by the specific definitions of SES that we focus on in our benchmark results.

## 6 Elite peer effects on students' academic performance

Having established the presence of a significant elite peer effect on elite degree enrolment and a socioeconomic gradient in this effect, we turn to analysing the mechanisms underlying this effect. We start this by analysing the effect of elite peers on high school GPA, a central determinant of higher education enrolment decisions in Norway.

We estimate the elite peer effects on overall GPA by using the same identification strategy as above and estimating Equation 1, this time with high school GPA as dependent variable. The estimates of these models are reported in Panel A of Table 4 and show that an increase in the proportion of elite peers in a student's school cohort has a negative and statistically significant effect on overall GPA across the whole sample. Coefficients in the second and third column of the table reveal a strong socio-economic gradient in this effect. Specifically, exposure to elite peers have a significantly detrimental effect on the GPA of low SES students, reducing their grade by 17.1% SD, and a smaller detrimental effect on the GPA of high SES students of 4.6% SD.

As explained in section 2, overall GPA is a weighted average of blindly assessed written exams, teacher-assessed internal grades, and oral exams assessed jointly by the student's teacher and an external examiner. To explain why elite peers have a negative effect on high school GPA and a particularly negative one for low SES students, we re-estimate the model this time with each GPA

[30] The coefficients (standard error) are 0.024 (0.003) and 0.040 (0.008) for the low and high SES samples respectively.

component as dependent variable (Panel B of Table 4). First, we find that exposure to elite peers in high school *increases* grades on externally-assessed written exams for both high and low SES students. Among all components of high school GPA, externally-assessed written exams can be considered as the cleanest measure of learning or knowledge. As a result, we interpret this positive effect of elite peers on externally assessed exam scores as reflecting positive spillovers of elite peers on the learning or effort and motivation put into learning by other students in the class. The fact that all students seem to benefit from their elite peers, regardless of their ability or socio-economic background, also indicates it is unlikely teachers' teaching behaviour changes depending on the fraction of elite (high-ability) students in the cohort. It is also reassuring in suggesting that low and high SES students do interact with each other.

In contrast, exposure to elite peers *decreases* the grades of low SES students on exams assessed by teachers either fully (internal grades) or partly (oral exams). The negative effect on teacher assessment is four times as large for low SES students than it is for high SES students. For oral exams, the elite peer effect is insignificant for high SES while it is negative and statistically significant for low SES students.[31]

The SES gradient in the negative effect of elite peers on exams where teachers have some discretion over students' grade is consistent with teachers marking on a curve, i.e. that they assess student's achievement relative to others.[32] If this were the case, because elite students tend of be high achieving, an increase in elite peers in the cohort would create downward pressure on the rank of other students. This downward pressure would be more strongly felt among lower ranked students than among higher ranked students, thus creating a more negative effect of elite peers on teachers assessments of low SES students than on those of high SES students.[33]

To corroborate this conjecture, we present in Figure A7 estimates of the elite peer effect on teacher assessments, this time allowing an interaction between the peer effect and the student's

[31]In the final three panels of Figure A5 the coefficients on the effect of exposure to elite peers on overall GPA and its components are plotted for low, medium and high SES households. On the whole, the peer effect for middle SES households sits in between estimates for low and high SES households, suggesting a linear SES gradient, although the confidence intervals often overlap across samples. The exception is for the written assessment where the exposure to elite peers has the same coefficient across household socioeconomic background.

[32]In Norway, teachers are not officially supposed to mark to a curve, but may nonetheless grade students relative to others.

[33]Table A6 illustrates that the effect of adding an elite student to the school cohort lowers the rank of students calculated amongst their high school peers (which is the dependent variable of the regression). Therefore, even though the written exam score marked nationally increases with elite peer exposure, their rank within the classroom falls.

ability rank within the cohort, where the rank is calculated using the middle school GPA, ranked across all students within the same high school cohort. The figure clearly shows that the negative effect of the exposure to elite peers on the teacher assessment is driven by lower ranked students within the cohort, thus in line with the idea that the presence of elite peers in the cohort have a mechanical negative effect on the rank of other students.

However, this 'rank effect' does not fully explain why the effect of elite peers on GPA is *more negative* for low SES students than it is for high SES students. Indeed, as is clear from Figure A7, the negative effect of elite peers on teacher assessments is always larger for low SES students than it is for high SES, conditional on their middle school GPA rank. While our data does not allow us to further explain these results directly, these patterns are consistent with the existence of a systematic teacher bias against low-SES students, where this bias responds to cohort composition and gets exacerbated by the presence of more elite peers. Specifically, the higher the fraction of elite peers in the cohort, the stronger the bias in teachers' assessment against low SES students.[34]

## 7 Elite peer effects on elite degree enrolment conditional on GPA : results from a causal mediation analysis

The paper so far has established that elite peers have a positive effect on the probability that students enrol in an elite degree (section 5), but a negative effect on their high school GPA (section 6). This implies that, conditional on GPA, elite peers must have a positive effect on students' likelihood to apply to elite degrees.[35] Such positive effect could derive from elite peers or their parents by acting as role models and/or providing information about these educational routes and their returns.

[34]Table A7 replicates Table 3 but for the high school GPA outcome and shows that the benchmark estimates are robust to the validity tests. Figure A8 plots the marginal effect of exposure to elite peers implied from the quadratic specification, showing again that the SES gradient in high school GPA is present across the distribution of GPA.

[35]We do not directly observe application behaviour and so are not in a position to comment on whether elite peers have an effect on the probably that students apply for an elite degree versus on the probability that they make better strategic decisions and end up being more likely to be accepted into an elite degree. We leave this interesting question to future work.

### 7.1 IV strategy to estimate the direct effect of elite peers

In this section, we aim to quantify the effect of elite peers on the probability of enrolling in an elite degree conditional on GPA. We refer to this effect as the direct effect of elite peers - as opposed to the indirect effect of elite peers working through GPA. We will use a causal mediation analysis to separate these two effects. To estimate the direct effect of elite peers, we need to estimate the following model, which corresponds to our benchmark model augmented to include high school GPA as an explanatory variable:

$$Y_{ics} = \gamma_1 P_{-ics} + \gamma_2 GPA_{ics} + X^{'}_{ics}\gamma_3 + \alpha_s + \rho_c + \epsilon_{ics} \tag{3}$$

where $GPA_{ics}$ refers to high school GPA.

In the model above, $GPA_{ics}$ is likely endogenous because it is likely to be correlated with unobserved individual determinants of elite degree enrolment. As a result, an OLS estimation of Equation 3 would fail to recover $\gamma_2$ in an unbiased way. And because $P_{-ics}$ and $GPA_{ics}$ are correlated, it would also fail to recover $\gamma_1$ in an unbiased way. To overcome this issues, we need to instrument high school GPA with a variable that is strongly predictive of GPA (relevance), but that only affects the probability of enrolling in an elite degree through its impact on GPA. This exercise can be seen as a causal mediation analysis with endogenous mediators discussed in Celli (2021).[36]

To instrument high school GPA, we propose to exploit a unique feature of the Norwegian high school system and provide some justification for our instrument below. Specifically, we exploit a lottery which randomly allocates students to take externally assessed examinations in a specific subject in the final, third year of high school. We define our instrument as an indicator that takes the value 1 if student $i$ was randomized into taking math as an externally assessed subject in the third year of high school, and 0 otherwise.

We focus on the randomisation into taking math (as opposed to another subject) because it is likely to have the strongest first stage in our context. Indeed, whilst one could argue that the marking of maths subjects are less subjective than other subjects, mathematics has been shown to

[36]There are a few examples of mediation analysis taking account of the endogeneity of mediators through an instrumental variables strategy in the economics literature. For example, see Aklin and Bayer (2017), Attanasio et al. (2020) and Nicoletti et al. (2023)

be prone to strong teacher bias, which are therefore more strongly circumvented in blindly assessed exams (Copur-Gencturk et al., 2020).[37] Interestingly, in our context and in line with this, we find that the negative effect of elite peers on teacher assessments is particularly strong for maths.[38] Being randomly assigned into a written maths exam therefore is likely to raise GPA most strongly And as we see below, the first stage F-statistic is high particularly for low SES students.

In addition to being relevant, the instrument must also satisfy the rank condition. In other words, being assigned to a written math exam must affect elite degree enrolment decisions only through its effect on high school GPA. First, this condition could be violated if the probability of being assigned a written exam is higher for certain schools, or programmes of study for example. However, we include school and programme of study fixed effects in the regression to account for the fact that randomisation is done within school and programme of study.

Second, even conditional on these fixed effects, it may be that students with particular traits are more likely to take the written exams. For example, the students with highly educated parents may ask to be assigned the written exams and the same is true for boys versus girls. Alternatively teachers may assign students with specific characteristics to the maths tests. However this is not possible, because the randomization takes place at the county level. Teachers, parents and students are not involved with the assignment to the subject of the written exam. In any case, to reassure that the assignment to students to the written exam is random, Table A8 reports the coefficients (and standard error) of a regression of the instrumental variable on the set of covariates included in the benchmark specification, augmented by indicators for programme of study in high school (groupings of majors into social science, humanities, science and general), school and cohort fixed effects. The table shows very little significance of student characteristics in predicting the lottery assignment to take a maths exam, validating the exclusion restriction.

To operationalise this IV strategy, we estimate Equation 3 jointly with the following first stage

[37] Copur-Gencturk et al. (2020) show that teachers assessment of maths performance for two students providing a similar answer are linked to their SES, gender and ethnicity.

[38] When repeating the analysis in Table 4 but replacing the dependent variable with the teacher assessment in maths, English or Norwegian, the negative downgrade is of greatest magnitude for maths. For low SES students the coefficient (standard error) is -0.034(0.006), -0.018(0.005) and -0.013(0.004) and for high SES students is -0.009(0.009), -0.006(0.006) and -0.005(0.005) for maths, English and Norwegian respectively.

equation:

$$GPA_{ics} = \delta_1 P_{-ics} + Z_{ics}^{'}\delta_2 + X_{ics}^{'}\delta_3 + \alpha_s + \rho_c + \epsilon_{ics} \tag{4}$$

where $Z_{ics}$ denotes the instrumental variable and the notation for other terms is as before. The direct effect of exposure to elite peers on student enrolment is the conditional effect given by coefficient $\gamma_1$. The indirect effect of exposure to elite peers through the channel of high school GPA is the product of $\delta_1$ from equation 4 and $\gamma_2$ from 3 (i.e. the product of the effect of elite peers on high school grades and the effect of high school grades on elite enrolment).

### 7.2 IV and mediation analysis results

Panel A of Table 5 reports the first stage estimates of $\delta_1$ and $\delta_2$ from Equation 4, separately for the low and high SES samples. These estimates confirm that the instrument is relevant for the low SES student sample where the F-statistic on the instrumental variable is 77. The F-stat for the high SES sample is lower at 5.6, which is intuitive since high SES students perform very highly anyway and there was a lower downgrade in teacher assessments for these students. For this reason we now focus on reporting the results for low SES students.

Panel B of Table 5 contrasts the OLS estimates of the overall effect (i.e. unconditional on GPA) of elite peers on elite degree enrolment in columns 1 and 3[39], with the IV estimates of the direct effect of elite peers (parameter $\gamma_1$ of Equation 3) in columns 2 and 4. Columns 2 and 4 in Panel B of Table 5 also report the estimates of $\gamma_2$ of Equation 3, the effect of GPA on the probability of enrolling in an elite degree.

As expected, high school GPA has a strong positive and statistically significant effect on the probability of enrolling in an elite degree. The coefficient on the elite peers in the IV specification is 0.026 and statistically significant in the low SES sample, which means that an increase in exposure to elite peers by one standard deviation encourages low SES students to raise their enrolment in elite degree by 2.6 percentage points (conditional on GPA). This direct effect is consistent with elite peers (and/or their families) raising students' motivation or aspiration to pursue an elite education over and beyond any effect they may have on academic performance.

[39] Note that the OLS estimates are slightly different from those presented in Table 2 because we now also control for high school programme indicators (as in the first stage equation).

The final rows of Table 5 decompose the total peer effect on student enrolment to an elite degree from Table 2 into the direct effect $\gamma_1$ from Equation 3 and the indirect effect $(\delta_1 * \gamma_2)$ working through GPA. Interestingly the direct effect of exposure to elite peers through the mechanisms of information and role models is large enough to cancel out the negative effect coming from grades. This suggests two policy conclusions. First, programmes directly providing role models and information for high-ability, low SES students may be effective at raising enrolment into elite programmes. Second, a policy reform to increase the proportion of written maths examinations assessed blindly for low SES students would reduce the teacher bias from exposure to elite peers and raise enrolment of low SES students in elite degree programmes.[40]

## 8 How does exposure to elite peers shape intergenerational income mobility?

We conclude our analysis by considering the extent to which the elite peer effect on elite degree enrolment (and the social gradient in this effect) we have uncovered in this paper translates onto earnings and inequalities therein. If that is the case, policies aimed at reducing socioeconomic segregation in high school could be a potential lever to reduce the intergenerational persistence of earnings.

The key parameter that will determine the answer to these questions is the return to an elite degree and the gap in returns between low and high SES students. While elite degrees have been shown to have high labour market returns in other contexts, this evidence is, to our knowledge, missing for Norway. Moreover, it is not a given that the returns would be positive for low SES students - a requirement for any policy aimed at increasing the number of first generation elites to have an impact in the labour markets for this group.[41]

In this final section, we therefore aim to tackle three questions: is the earnings premium to an elite degree positive across SES backgrounds? Does exposure to elite peers in high school raise the

[40] The direct effect for high SES students also increases once we condition on high school GPA, however with such a low F-statistic we do not consider these results as reliable.

[41] Zimmerman (2019) shows that the returns to business focused elite degrees in Chile are close to zero for males and females not from private high schools, which are the types of high schools that charge high tuition and serve upper-income households and hence that are rarely attended by low SES students. On the contrary the returns are similar or even higher for low compared to high SES students on elite medical school programmes. Michelman et al. (2022) attributes the different returns across the programmes as a requirement to schmooze when moving from business programmes and into the labour market, which is not required for other programmes, such as medicine. Hastings et al. (2013) on the other hand finds large positive returns for highly selective degrees across SES.

longer-run outcome of earnings age 30-32? Does exposure to elite degrees exacerbate or mitigate the link between child and parents' earnings? To answer these questions, we use data on the earnings of the three oldest cohorts in our data (born between 1986-1988). For these cohorts it is possible to measure income for some ages between 30 and 32 years old, which has been shown to be the age at which earnings rank becomes relatively stable and predictive of earnings rank at older ages (Bhuller et al., 2017).

**Earnings premium to enrolling in an elite degree** To measure the earnings premium to enrolling in an elite degree, we estimate a Mincer style regression of earnings on an indicator for whether the student enrolled in a degree and an indicator for enrolling in an elite degree (with the category of no degree is omitted) on the set of individual level controls we included in (Equation 1), school and cohort fixed effects. We estimate this specification for low and high SES students separately, as the earnings premium could be different between the two. The results of this specification are reported in columns 1 and 2 of Table 6 where the dependent variable is the (within cohort) percentile rank of earnings age 30-32 in Panel A and an indicator for earning in the richest decile in Panel B. We find evidence of a very high average earnings premium to enrolling in an elite degree, which is only slightly smaller than they are for high SES students.[42] Although the evidence presented here is not necessarily causal, it does suggest that the returns to becoming first generation elite are likely to be positive so that increasing the probability that low SES students enrol in an elite degree could increase social mobility.

**Effect of elite peers on earnings** Having established a positive earnings premium to enrolling in an elite degree across SES backgrounds, we move on to tackling the second question.[43] We estimate the effect of being exposed to elite peers during high school on adult earnings by re-estimating our benchmark equation (1), this time with the indicator for the percentile rank (Panel A) and earnings

[42] When we estimate the Mincer equations allowing a different coefficient for studying for different elite programmes, of law, STEM or medicine, we find interesting findings that are consistent with different returns across programmes requiring different levels of schmoozing, in Zimmerman (2019). In particular, the earnings premium measured by an indicator for earning in the top percentile associated with studying or a law degree is higher for high SES, whereas for the medicine degree is higher for low SES students. See columns (5) and (6) of Table A9. It is only for the outcome at the top percentile that we find these patterns, as for the earnings percentile (columns (1) and (2) or earning in the richest decile (columns (3) and (4) the coefficients associated with the earnings premium from the different elite degrees are very similar across Low and High SES students.

[43] This is important given that Dahl et al. (2021) find that increasing integration of different genders in military training drives short-run but not longer-run outcomes such as attitudes, field of study or occupation.

in the top decile (Panel B) as outcomes. We present the estimates of this specification in columns (3) through (4) of Table 6. Being exposed to elite peers in high school increases the percentile rank but this effect is lower for low SES students than it is for high SES students (0.86 percentiles compared to 2.5). It also increases the probability of being in the richest decile at age 30-32 but only for high SES students (2.2 ppts).[44] These results are in line with our earlier evidence that elite peers have a less positive effect for low SES students than for high SES students but suggests additionally that the effect of elite peers persists into later life.[45]

**Implications for intergenerational income mobility** We have shown so far that elite peers increase the educational attainment and earnings of low SES students, but they have a stronger effect on the outcomes of high SES students. This means that while exposure to elite peers could increase the number of first generation elites and thus increase mobility at the bottom of the parental income distribution, it could also exacerbate the lack of mobility at the top of the parental income distribution. To verify this hypothesis empirically, we estimate the extent to which the degree of intergenerational income persistence across the parental income distribution varies with the child's exposure to elite peers in high school. To do so, we estimate the equation:

$$r^c_{ics} = \delta_{11} r^p_{ics} + \delta_{12}(r^p_{ics})^2 + \delta_{21} r^p_{ics} \times P^{high}_{-ics} + \delta_{22}(r^p_{ics})^2 \times P^{high}_{-ics} + X'_{ics}\beta_3 + \alpha_s + \rho_c + \epsilon_{ics} \tag{5}$$

where $r^c_{ics}$ is the child's rank in the children's earnings distribution at age 30-32, $r^p_{ics}$ is the parents' rank in the parents' earnings distribution when the child was 15-19, and $P^{high}_{-ics}$ is an indicator taking the value 1 if child $i$ had a level of exposure to elite peers during high school that was higher than the mean (c. 6%) and 0 otherwise. Note that this equation is a flexible version of the typical rank-rank regression used in the intergenerational mobility literature where we have allowed some non-linearity in the rank-rank coefficient across the parental income distribution and across the

[44]See Table A10 and Figure A9 which show respectively the validity tests for the outcome of earnings and the figure plotting the peer effect in a specification which allows for the peer variable to enter the model in a quadratic specification.

[45]The results do not imply that the only way through which elite peers affect earnings is by boosting students' probability of enrolling in an elite degree. Indeed, elite peers may have other effects on earnings over and beyond their effect on educational attainment (for example through connections that could help secure a good job). When re-estimating the model this time also controlling for whether the student has enrolled in an elite education (results available upon request), we still find a positive effect and an SES gradient of elite degrees on earnings. Understanding these mechanisms is beyond the scope of the paper but we note that these findings are an interesting avenue for future research.

level of exposure to elite peers in high school.[46]

We depict the estimates of this equation with Figure 2, which shows the predicted value of child's earnings rank, as a function of the parent's rank and the level of exposure to elite peers. For any value of the parent percentile rank, the child's percentile rank is higher in the high exposure group with above average proportion of elite peers than in the low exposure treatment. Importantly, the additional uplift in the relationship is highest at the bottom and the top of the parent income distribution. This means that whilst exposure to elite peers lifts mobility at the bottom of the parental income distribution, it also increases persistence at the top of the income distribution. Segregation of children from elite educated parents in high school (leading to low exposure at the bottom of the distribution and high exposure at the top) could therefore be a factor explaining why intergenerational persistence in income is particularly high at the top of the distribution in Norway (Pekkarinen et al., 2017) and in other contexts (Chetty et al., 2014).

**Policy implications** As discussed earlier, the high school system in Norway is such that it creates some degree of socioeconomic segregation in high schools, with children of elite educated parents being more likely to be exposed to other children of elite educated parents than children of non-elite educated parents. This segregation could be a factor exacerbating the intergenerational persistence both in education and in earnings in Norway. This implies that policies that would improve the mixing of students across socio-economic classes (i.e., to use the term of Chetty et al. (2022), their economic connectedness), could increase mobility both at the top and at the bottom of the parental income distribution.

To illustrate this point, we use our estimates to conduct several simulations through which low and high SES students are re-assigned across high schools in a way that reduces the level of socioeconomic segregation more or less strongly. These simple simulations do not intend to mimic the effect of particular interventions (e.g. busing, affirmative action or quotas), which policy makers could and have used to reduce socio-economic segregation. Rather, the objectives is to illustrate the point above and to quantify the amount of re-assignment that would be needed to discern an impact on the intergenerational persistence of income, abstracting from general equilibrium effects.

In a nutshell, the simulations take a low SES student from a school with the lowest exposure

[46] This follows a similar strategy of Pekkarinen et al. 2009, used for example in Bütikofer and Salvanes 2020 and Kaila et al. 2021.

to elite peers and moves them into a school with the highest exposure to elite peers whilst simultaneously taking a high SES student from the high exposure school and moving them to the low exposure school. The different simulations vary both the number of students within each school that are reassigned and the number of schools that are included in the reassignment. Once the students have been reassigned, we re-calculate the mean exposure to elite peers to all pupils within the schools that have seen their student body change. Using our estimates of the effect of exposure to elite peers on the percentile earnings rank from Table 6, we simulate each student's percentile rank and re-estimate the rank-rank regression above. We compare the estimate of the intergenerational persistence in income under each simulation with the actual degree of intergenerational persistence estimated in the true data (see section A3 for full details on how we conduct the simulation).

Table A11 shows that for all simulations, intergenerational mobility rises once the exposure to elite peers is re-balanced across low and high SES students, as indicated by the higher intercept and the flatter slope coefficient. The increase in intergenerational mobility is shown graphically for the most extreme simulation we considered (moving 5 students from each school in Norway) in Figure A10. After the simulation, the intercept is higher suggesting a higher earnings rank for very low SES students, and the gradient of the relationship between parent and child percentile rank is flatter.

## 9 Conclusion

Socioeconomic inequalities in elite education are high, even in Scandinavian countries where income inequality is notoriously low. The main contribution of this paper is to show that social interactions in high school plays a key role in driving inequalities in education and earnings outcomes within and across generations.

First, we show that Norway's high school admission system creates high levels of segregation of children of elite educated into the same high schools. Moreover, we show that being exposed elite peers in high school has a stronger positive effect on the probability of enrolling in an elite degree and on the earnings of high SES student than on those of low SES students. Together, these findings suggest that this segregation is responsible for reducing mobility at the bottom of the socioeconomic distribution while exacerbating persistence at the top.

These findings are important in the context of the intergenerational mobility literature. They provide causal evidence that social capital and in particular the level of economic connectedness among high school peers is a mechanisms behind the intergenerational persistence (or lack thereof) of education and income. Moreover, they support the idea that social capital is a key reason why intergenerational income persistence is so high at the top of the distribution in Norway. While our findings are for the Norwegian context, they may also be relevant to explain similar patterns in other countries, such as the US (Chetty et al., 2014).

The second contribution of our paper is to explain the mechanisms through which elite peers affect children's educational outcomes. Our findings are threefold. First, exposure to elite peers raises the test score in written and blind marked high school examinations suggesting increased learning or effort of students. This peer effect is of equal size for low and high SES students, which reassuringly suggests that the students from different backgrounds interact with students from elite families. Second, through a teacher assessment downgrade, elite peer exposure penalises the GPA of low SES students much more than for high SES students. We argue that this pattern reflects the fact that elite peers push the rank of other students down the distribution of teacher grades. Conditional on rank however, the teacher assessment downgrade remains stronger for low SES than high SES students, which could suggest that the presence of elite peers also triggers a teacher bias against low SES. Third, we tease out the final role model channel by showing through a causal mediation analysis that, conditional on GPA, students' exposure to elite peers increases their likelihood to apply to an elite degree. Because this positive peer effect dominates any negative peer effect through teacher grades, we find the overall effect of peers to be positive.

Overall, our findings suggest that considering peer interactions is very important for policymakers interested in improving the life chances of low SES students as well as intergenerational mobility. A direct implication of our findings is that policies increasing the social mixing of students from different parental education background could be beneficial to improve social mobility across the parental income distribution. In absence of such mechanism to re-allocate students across schools, our findings do suggest that increasing the reliance of university admission systems on standardised, blindly assessed tests could improve the educational and economic chances of low SES students by increasing the benefits that low SES students have from being exposed to elite peers. Moreover, role model interventions which provides information to encourage students to apply to elite degrees and help them maximise their chances to be admitted could also increase the chances that high-ability students from low SES backgrounds become first generation elite.

## References

Adermon, A., M. Lindahl, and M. Palme (2021). Dynastic human capital, inequality, and intergenerational mobility. *American Economic Review 111*(5), 1523–48.

Aklin, M. and P. Bayer (2017). How can we estimate the effectiveness of institutions? Solving the post-treatment versus omitted variable bias dilemma. *Working paper*.

Alesina, A., M. Carlana, E. La Ferrara, and P. Pinotti (2018, December). Revealing stereotypes: Evidence from immigrants in schools. Working Paper 25333, National Bureau of Economic Research.

Altmejd, A., A. Barrios-Fernández, M. Drlje, J. Goodman, M. Hurwitz, D. Kovac, C. Mulhern, C. Neilson, and J. Smith (2021, 03). O Brother, Where Start Thou? Sibling Spillovers on College and Major Choice in Four Countries*. *The Quarterly Journal of Economics 136*(3), 1831–1886.

Andersen, M. and S. Lokken (2020). The final straw: high school dropout for marginal students. Statistics Norway Discussion Paper No. 894.

Angrist, J. and K. Lang (2004). Does school integration generate peer effects? Evidence from boston's metco program. *American Economic Review 94*(5), 1613–1634.

Attanasio, O., S. Cattan, E. Fitzsimons, C. Meghir, and M. Rubio-Codina (2020). Estimating the production function for human capital: results from a randomized controlled trial in colombia. *American Economic Review 110*(1), 48–85.

Barrios-Fernandez, A., C. Neilson, and S. Zimmerman (2022). Elite universities and the intergenerational transmission of human and social capital. Unpublished manuscript, Yale University.

Bertoni, M., G. Brunello, and L. Cappellari (2020). Who benefits from privileged peers? evidence from siblings in schools. *Journal of Applied Econometrics 35*(7), 1–24.

Bhuller, M., M. Mogstad, and K. G. Salvanes (2017). Life-cycle earnings, education premiums, and internal rates of return. *Journal of Labor Economics 35*(4), 993–1030.

Black, S., P. Devereux, and K. Salvanes (2005a). Why the apple doesn't fall far: understanding the intergenerational transmission of education. *American Economic Review 95*(1).

Black, S., P. Devereux, and K. G. Salvanes (2005b). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics 120*(2), 669–700.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2013). Under Pressure? The effect of peers on outcomes of young adults. *Journal of Labor Economics 31*(1), 119–153.

Britton, J., E. Drayton, and L. van der Erve (2021). Which university degrees are best for intergenerational mobility? IFS report.

Burgess, S. and E. Greaves (2013). Test scores, subjective assessment and stereotyping of ethnic minorities. *Journal of Labor Economics 1*(1), 37–89.

Bütikofer, A., R. Ginja, F. Landaud, and K. Lokken (2020). School selectivity, peers and mental health. *Unpublished*.

Bütikofer, A., S. Jensen, and K. G. Salvanes (2018). The role of parenthood on the gender gap among top earners. *European Economic Review 109*, 103–123.

Bütikofer, A., E. Risa, and K. G. Salvanes (2021). Status traps and human capital investment. *Unpublished*.

Bütikofer, A. and K. G. Salvanes (2020). Disease control and inequality reduction: Evidence from a tuberculosis testing and vaccination campaign. *The Review of Economic Studies 87*(5), 2087–2125.

Campbell, T. (2015). Stereotyped at seven? biases in teacher judgement of pupils' ability and attainment. *Journal of Social Policy 44*(3), 517–547.

Carlana, M. (2019, 03). Implicit Stereotypes: Evidence from Teachers' Gender Bias. *The Quarterly Journal of Economics 134*(3), 1163–1224.

Celli, V. (2021). Causal mediation analysis in economics: Objectives, assumptions, models. *Journal of Economic Surveys 36*(1), 214–234.

Chetty, R., J. N. Friedman, E. Saez, N. Turner, and D. Yagan (2020a). Income segregation and intergenerational mobility across colleges in the united states. *The Quarterly Journal of Economics 135*(3), 1567–1633.

Chetty, R., J. N. Friedman, E. Saez, N. Turner, and D. Yagan (2020b, 02). Income Segregation and Intergenerational Mobility Across Colleges in the United States*. *The Quarterly Journal of Economics 135*(3), 1567–1633.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014, 09). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States *. *The Quarterly Journal of Economics 129*(4), 1553–1623.

Chetty, R., M. Jackson, and T. e. a. Kuchler (2022). Social capital I: measurement and associations with economic mobility. *Nature 608*, 108–121.

Cools, A., R. Fernández, and E. Patacchini (2019). Girls, boys, and high achievers. Technical report, National Bureau of Economic Research.

Copur-Gencturk, Y., J. R. Cimpian, S. T. Lubienski, and I. Thacker (2020). Teachers' bias against the mathematical ability of female, black, and hispanic students. *Educational Researcher 49*(1), 30–43.

Corak, M. (2013, September). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives 27*(3), 79–102.

Corak, M., M. J. Lindquist, and B. Mazumder (2014). A comparison of upward and downward intergenerational mobility in canada, sweden and the united states. *Labour Economics 30*, 185–200.

Dahl, G., D.-O. Rooth, and A. Stenberg (2021). Intergenerational and sibling peer effects by gender in high school majors. *NBER Working Paper 27618*.

Dahl, G. B., A. Kotsadam, and D.-O. Rooth (2021). Does integration change gender attitudes? the effect of randomly assigning women to traditionally male teams. *The Quarterly Journal of Economics 136*(2), 987–1030.

Dahl, G. B., D.-O. Rooth, and A. Stenberg (2023). High school majors and future earnings. *American Economic Journal: Applied Economics 15*(1), 351–82.

Dalla-Zuanna, A., K. Liu, and K. G. Salvanes (2020). Pulled-in and crowded-out: Heterogeneous outcomes of merit-based school choice. Unpublished, University of Cambridge.

Deming, D. J. and K. L. Noray (2018). Stem careers and the changing skill requirements of work. Technical report, National Bureau of Economic Research.

Doyle, L., M. J. Easterbrook, and P. R. Harris (2023). Roles of socioeconomic status, ethnicity and teacher beliefs in academic grading. *British Journal of Educational Psychology 93*(1), 91–112.

Feld, J. and U. Zoelitz (2017). Understanding peer effects - on the nature, estimation and channels of peer effects. *Journal of Labor Economics 35*(2), 34–68.

Hastings, J., C. Neilson, and S. Zimmerman (2013). Are some degrees worth more than others? evidence from college admission cutoffs in chile. National Bureau of Economic Research.

Hoxby, C. (2000). Peer effects in the classroom: Learning from gender and race variation. Technical report, National Bureau of Economic Research.

Kaila, M., E. Nix, and K. Riukula (2021). Disparate impacts of job loss by parental income and implications for intergenerational mobility. Working Paper 53, Federal Reserve Bank of Minneapolis.

Kirkebøen, L. J. (2010). Forskjeller i livsløpsinntekt mellom utdanningsgrupper.

Lavy, V. (2008). Do gender stereotypes reduce girls' or boys' human capital outcomes? evidence from a natural experiment. *Journal of Public Economics 92*(10), 2083–2105.

Lavy, V., M. D. Paserman, and A. Schlosser (2011, 08). Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers In the Classroom. *The Economic Journal 122*(559), 208–237.

Lavy, V. and E. Sand (2018). On the origins of gender gaps in human capital: Short- and long-term consequences of teachers' biases. *Journal of Public Economics 167*, 263–279.

Lundberg, S. (2020). Educational gender gaps. *Southern Economic Journal 87*(2), 416–439.

Mani, A. and E. Riley (2019). Social networks, role models, peer effects aspirations. Technical report. WIDER Working Paper 2019/120.

Michelman, V., J. Price, and S. D. Zimmerman (2022). Old boys' clubs and upward mobility among the educational elite. *The Quarterly Journal of Economics 137*(2), 845–909.

Nicoletti, C., K. G. Salvanes, and E. Tominey (2023). Mothers working during preschool years and child skills: does income compensate? *Journal of Labor Economics 41*(2), 000–000.

Nybom, M. and J. Stuhler (2017). Biases in standard measures of intergenerational income dependence. *Journal of Human Resources 52*(3), 800–825.

Papageorge, N. W., S. Gershenson, and K. M. Kang (2020). Teacher expectations matter. *Review of Economics and Statistics 102*(2), 234–251.

Pekkarinen, T., K. Salvanes, and M. Sarvimaki (2017). The evolution of social mobility: Norway during the twentieth century. *The Scandinavian Journal of Economics 119*(1), 5–33.

Pekkarinen, T., R. Uusitalo, and S. Kerr (2009). School tracking and intergenerational income mobility: Evidence from the finnish comprehensive school reform. *Journal of Public Economics 93*(7–8), 965–973.

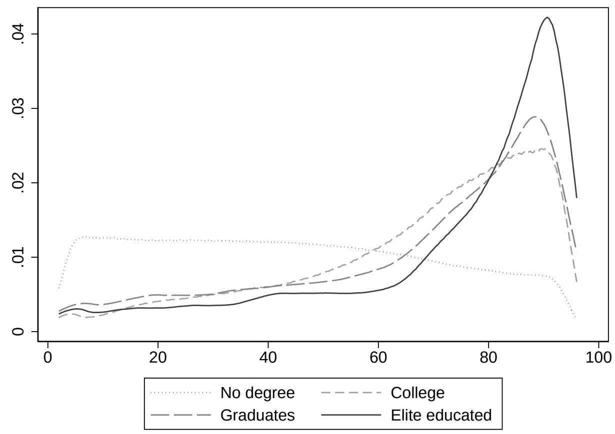
Porter, C. and D. Serra (2020). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics 12*(3), 226–54.

Strømme, T. B. and M. N. Hansen (2017). Closure in the elite professions: the field of law and medicine in an egalitarian context. *Journal of Education and Work 30*(2), 168–185.

Tincani, M. (2017). Heterogeneous peer effects in the classroom. HCEO Working Paper 2017-006.

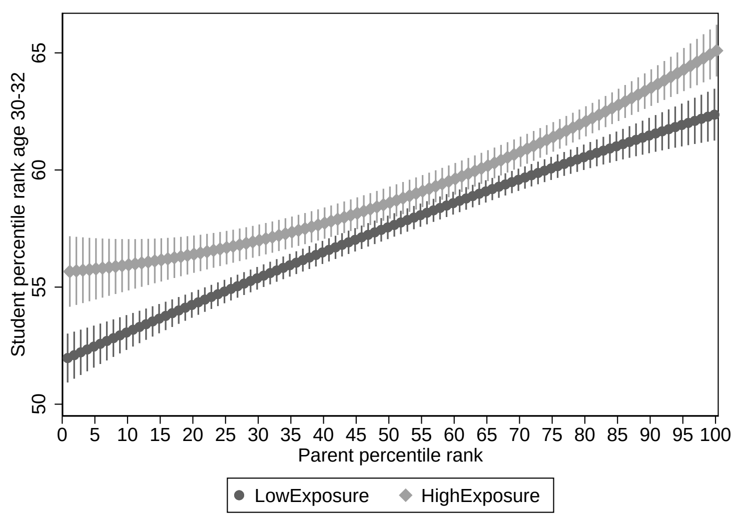
Zimmerman, S. D. (2019). Elite colleges and upward mobility to top jobs and top incomes. *American Economic Review 109*(1), 1–47.

**Figure 1:** Density of earnings percentiles by education level




Notes: This graph plots the density of earnings percentiles across educational groups. Sample is the population of Norway aged 28-40 between 1993-2001. The percentile rank of earnings is calculated within each birth cohort.

**Figure 2:** Intergenerational mobility: estimating the percentile rank-rank correlation across exposure to elite peers




Notes: This graph plots the fitted values from an intergenerational mobility rank-rank regression allowing for the interaction between exposure to elite peers and the parent percentile rank to be quadratic. High (low) exposure is defined as above (below) mean proportion of elite peers in the high school cohort.

**Table 1:** Summary statistics of the sample

| | Full sample | Low SES sample | High SES sample |
|---|---|---|---|
| | Mean (sd) | Mean (sd) | Mean (sd) |
| Enrolls in higher education | 0.904 | 0.861 | 0.956 |
| Enrolls in elite degree | 0.102 | 0.053 | 0.260 |
| % of parent with elite degree | 0.061 | 0.047 | 0.100 |
| | (0.056) | (0.047) | (0.068) |
| **Covariates** | | | |
| Female | 0.601 | 0.651 | 0.527 |
| Born in Norway | 0.873 | 0.836 | 0.852 |
| Middle school GPA (std) | 0.676 | 0.496 | 0.921 |
| | (0.634) | (0.639) | (0.591) |
| *Mother's highest education level* | | | |
| Compulsory education | 0.516 | 0.932 | 0.161 |
| High school degree | 0.126 | 0.068 | 0.144 |
| University degree | 0.358 | 0.000 | 0.695 |
| *Father's highest education level* | | | |
| Compulsory education | 0.578 | 0.916 | 0.073 |
| High school degree | 0.139 | 0.084 | 0.042 |
| University degree | 0.282 | 0.000 | 0.885 |
| % of own parents with an elite degree | 0.066 | 0.000 | 0.580 |
| | (0.194) | () | (0.183) |
| Family income in the top decile | 0.214 | 0.123 | 0.485 |
| | (0.309) | (0.244) | (0.352) |
| % of peer parents in top 5% of sample | 0.191 | 0.17 | 0.240 |
| | (0.100) | (0.095) | (0.110) |
| % of peer parents in elite occupation | 0.017 | 0.014 | 0.023 |
| **Mechanisms** | | | |
| High school GPA (std) | 0.013 | -0.252 | 0.494 |
| | (0.999) | (0.951) | (1.000) |
| **Long-run** | | | |
| Student in top decile of earnings 30-32* | 0.141 | 0.104 | 0.230 |
| Student percentile rank 30-32 | 58.494 | 55.181 | 64.667 |
| | (26.728) | (25.926) | (28.068) |
| N | 177,219 | 58,328 | 20,018 |

Notes: Sample of students ending middle school and entering high school between 2002-2010. The table presents means and standard deviations (in parentheses) of the main variables used in the analysis. Elite degree status defined as enrolment into Economics/Business, Engineering, Law of Medicine at a top institution (see section 3). Elite occupation takes the value 1 for STEM occupations, lawyers or doctors. High school GPA is standardized within cohort to have mean 0 and standard deviation 1. Middle school GPA is standardized to have mean 0 and standard deviation 1 within the cohort. *Measured for oldest 5 cohorts where sample size is 59,043; 20,454; 6,765 for the total sample; low SES and high SES respectively.

**Table 2:** Effect of elite peers on the probability of enrolling in an elite degree

| | (1)<br>Full sample | (2)<br>Low SES | (3)<br>High SES |
|---|---|---|---|
| Proportion of parents with elite degree (std) | 0.026*** | 0.013*** | 0.040*** |
| | (0.003) | (0.003) | (0.008) |
| *Number of students* | *177,219* | *58,328* | *20,018* |
| *Number of schools* | *556* | *524* | *459* |

Notes: OLS estimates of a regression of an indicator for whether the student is enrolled in an elite degree within 6 years of starting high school on: the proportion of parents with elite degree in the student's school's cohort, student's gender, middle school GPA, an indicator for whether the student was born in Norway, mother and father's highest education level, a variable measuring the number of student's own parents who have an elite education, and an indicator for whether the student's family income is in the top decile of the overall income distribution. Regressions include cohort and school fixed effects. Column (1) reports the coefficient on the proportion of parents with an elite degree estimated in the full sample, column (2) and column (3) report the same coefficient estimated in the low SES and high SES samples, respectively. The low SES sample is defined as the group of students who have at least one parent with no more than the compulsory level of education, but no parent with an elite education. The high SES sample is defined as the group of students who have at least one parent with an elite degree, but no parent with a compulsory level of education. Regressions are weighted by school size to take account of the parent peer variables group averages, taken from groups of different sizes. Standard errors clustered at the school level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

**Table 3:** Validity of the empirical strategy

| | (1) Benchmark | (2) School-specific linear trends | (3) School-cohort fixed effects interacted | (4) 'Drop if more than random' | (5) Including teacher traits | (6) Quadratic specification |
|---|---|---|---|---|---|---|
| **A - Low SES students sample** | | | | | | |
| Proportion of parents with elite degree (std) | 0.013*** | 0.013*** | | 0.010** | 0.015*** | 0.014*** |
| | (0.003) | (0.003) | | (0.004) | (0.004) | (0.003) |
| Proportion of parents with elite degree squared | | | | | | -0.001 |
| | | | | | | (0.001) |
| *Number of pupils* | *58,610* | *58,610* | | *28,181* | *37,270* | *58,610* |
| *Number of schools* | *524* | *524* | | *284* | *390* | *524* |
| **B - High SES students sample** | | | | | | |
| Proportion of parents with elite degree (std) | 0.040*** | 0.047*** | | 0.038*** | 0.053*** | 0.058*** |
| | (0.008) | (0.008) | | (0.013) | (0.009) | (0.010) |
| Proportion of parents with elite degree squared | | | | | | -0.008** |
| | | | | | | (0.004) |
| *Number of pupils* | *20,018* | *20,018* | | *8,420* | *12,737* | *20,018* |
| *Number of schools* | *459* | *459* | | *240* | *349* | *459* |
| **C - Low and High SES students sample** | | | | | | |
| Proportion of parents with elite degree (std) | | | 0.050*** | | | |
| | | | (0.004) | | | |
| Indicator for low SES | | | 0.041*** | | | |
| | | | (0.014) | | | |
| Proportion of parents with elite degree * low | | | -0.031*** | | | |
| | | | (0.003) | | | |
| *Number of pupils* | | | *78,540* | | | |

Notes: OLS estimates of the coefficient on the variable measuring the proportion of elite educated parents in the student's youth cohort in different specifications in the low SES sample (Panel A), in the high SES sample (Panel B) and pooled sample of low and high SES (Panel C). Column (1) the benchmark specification (equation 1 and Table 2). Column (2) is benchmark specification controlling also for school-specific linear trends. Column (3) includes fully interacted fixed effects for the school and cohort estimating on the pooled sample of low and high SES students. Column (4) the benchmark specification estimated on the subsample of schools where variation in the elite peer variable evolves over time in a random way. Specifically, we drop the schools where the $R^2$ from a school-level regression of the proportion of elite educated peers on a quadratic in year is 1.05 times the $R^2$ from five regressions where cohorts are randomly re-ordered for each. See section 4. Column (5) the benchmark specification including additionally average traits of teachers within schools across cohorts: the proportion of females, the proportion of teachers from a professional or low skilled background and average age. The teacher background is defined by the occupation of their father. Column (6) refers to the benchmark specification augmented with a quadratic term in the elite peer variable. Regressions are weighted by school size. Standard errors clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

**Table 4:** Elite peer effect on overall GPA and its components

| | (1) Full sample | (2) Low SES | (3) High SES |
|---|---|---|---|
| **A - Dep. var: overall GPA** | | | |
| | -0.118*** | -0.171*** | -0.046*** |
| | (0.013) | (0.016) | (0.012) |
| *Number of observations* | *177,219* | *58,610* | *20,018* |
| | | | |
| **B - Dep. var: components of GPA** | | | |
| Externally assessed written exam grades | 0.025*** | 0.030** | 0.030* |
| | (0.009) | (0.012) | (0.016) |
| *Number of observations* | *177,219* | *58,610* | *20,018* |
| | | | |
| Teacher-assessed internal grades | -0.110*** | -0.162*** | -0.040*** |
| | (0.013) | (0.016) | (0.012) |
| *Number of observations* | *177,219* | *58,610* | *20,018* |
| | | | |
| Semi-externally assessed oral exam grades | -0.036*** | -0.064*** | -0.012 |
| | (0.008) | (0.011) | (0.014) |
| *Number of pupils* | *149,488* | *49,414* | *17,189* |

Notes: OLS estimates of the effect of the proportion of parents with an elite degree in the student's school cohort in the benchmark model where the dependent variable is now a measure of academic performance. See notes to Table 2 for detailed list of controls. The measures of academic performance are: overall high school GPA (row 1), average performance on externally assessed written exams across all three years of high school (row 2), average performance on teacher assessed grades across all three years of high school (row 3), and average performance on oral exams marked by an external examiner and the student's teachers across all three years of high school (row 4). All measures of performance are standardized to have mean 0 and standard deviation 1 within cohort. Column (1) reports the coefficient on the proportion of parents with an elite degree estimated in the full sample, column (2) and column (3) report the same coefficient estimated in the low SES and high SES samples, respectively. Standard errors clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

**Table 5:** IV estimates and decomposition of the total effect of elite peers on elite degree enrolment

| | Low SES | | High SES | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | OLS | IV | OLS | IV |
| **A - Dependent variable: GPA** | | | | |
| Proportion of parents with elite degree (std) | | -0.039*** | | -0.010 |
| | | (0.007) | | (0.010) |
| Student took written math exam (IV) | | 0.095*** | | 0.043*** |
| | | (0.011) | | (0.018) |
| *F stat* | | *77.471* | | *5.567* |
| **B - Dependent variable: elite degree enrolment** | | | | |
| Proportion of parents with elite degree (std) | 0.011*** | 0.026*** | 0.032*** | 0.049*** |
| | (0.004) | (0.004) | (0.009) | (0.019) |
| Overall high school GPA | | 0.385*** | | 1.771** |
| | | (0.053) | | (0.714) |
| **C - Decomposition** | | | | |
| Direct effect | | 0.026 | | 0.049 |
| Indirect effect | | -0.015 | | -0.018 |
| Total effect | | 0.011 | | 0.031 |
| *Number of pupils* | *58,586* | *58,586* | *19,968* | *19,968* |
| *Number of schools* | *500* | *500* | *409* | *409* |

Notes: Data source, Norwegian administrative data. Sample of students ending middle school and entering high school between 2002-2010. Columns (1) and (3) report OLS estimation. Columns (2) and (4) report two-stage least squares estimation, where the IV for high school GPA is lottery to take written exam in maths in year 3 of high school. Dependent variable is indicator for studying for an elite (graduate) degree. Model controls the same as Table 2 including school, cohort and additionally high school program fixed effects. Regressions are weighted by school size to take account of the parent peer variables group averages, taken from groups of different sizes. Standard errors clustered at the school level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

**Table 6:** Long-term implications for earnings

| | (1) Low SES | (2) High SES | (3) Low SES | (4) High SES |
|---|---|---|---|---|
| **A - Dependent variable: Earnings percentile** | | | | |
| Student ever enrolled in degree | 10.185*** | 14.369*** | | |
| | (0.455) | (1.574) | | |
| Student ever enrolled in elite degree | 26.701*** | 30.521*** | | |
| | (0.828) | (1.639) | | |
| Proportion of parents with elite degree | | | 0.816*** | 2.462*** |
| | | | (0.344) | (0.615) |
| **B - Dependent variable: Richest decile** | | | | |
| Student ever enrolled in degree | 0.028*** | 0.082*** | | |
| | (0.005) | (0.024) | | |
| Student ever enrolled in elite degree | 0.250*** | 0.284*** | | |
| | (0.010) | (0.025) | | |
| Proportion of parents with elite degree | | | 0.004 | 0.022*** |
| | | | (0.004) | (0.009) |
| *Number of pupils* | *20,454* | *6,765* | *20,454* | *6,765* |
| *Number of schools* | | | *457* | *372* |

Notes: Columns (1) and (2) run a Mincer-style regression of earnings on an indicator for degree and an elite degree. The omitted category is no degree. The regressions include a gender dummy and year of birth dummy variables as controls. Columns (3) and (4) estimate the benchmark specification (including all controls and fixed effects) but with the dependent variable changed to be the earnings percentile of the student (Panel A) and an indicator for being in the top decile of the earnings distribution (Panel B). The low SES sample in columns (1) and (3) is defined as the group of students who have at least one parent with no more than the compulsory level of education, but no parent with an elite education. The high SES sample in columns (2) and (4) is defined as the group of students who have at least one parent with an elite education, but no parent with a compulsory level of education. Sample of birth cohorts 1986-1988. Income is deflated to 2020. For the cohorts 1986; 1987 and 1988 income is measured ages 30-32; 30-31 and 30 respectively (see section 3). Standard errors clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

## Online Appendix

## A1 Gender heterogeneity

There are very well documented differences in college major choice across genders which contribute towards the gender pay gap. We re-estimate the benchmark model from Equation 1 separately by gender to understand if the SES gap also exists within gender. The results in Table A12 suggest that the effect of exposure to elite educated peers during high school on elite enrolment is larger for males than females (3.9ppt compared to 1.8ppt in the full sample). Within each gender, the SES gradient is still present and the peer effect is considerably larger for low SES males or females compared to high SES males or females.

Next we consider the gender of the parents and ask whether the elite parent peer effect varies across the proportion of mothers compared to fathers with an elite degrees. According to Table A13 the peer effects are stronger when the parent with an elite degree is a father rather than a mother. For girls the peer effect of exposure to elite peers is again strongest when the father has an elite degree rather than the mother. This is also true for the high SES boys, whereas we see that for low SES boys the effect of exposure to peers whose mother has an elite degree is more of less equal to the effect of exposure to a peer whose father is elite.

## A2 Additional robustness checks

### A2.1 Other robustness checks

**Sensitivity of results to sample selection** We examine the extent to which our results are robust to changes in the sample composition including first born children and two-parent families; and schools with different admission mechanisms. First, the effect of exposure to elite educated peers may be different for first born children compared to the total sample, if for example children of higher birth order are more influenced by their older sibling than their school peers and their parents (Black et al., 2005b). Column (2) of Table A5 shows that indeed the peer effect is slightly higher for first borns, although the new estimates are not statistically different to the benchmark estimates.

**Measurement error in the elite peer variable** The incidence of marital breakup may be different across household socioeconomic status and it is possible that the rates of divorce or separation vary across the SES status of schools. This would cause a problem in our estimation as the treatment could have more measurement error in the low SES sample because it is based on all biological parents. Therefore the difference in coefficients between low and high SES may be driven by attenuation bias. We confirm that this is not a problem in Column (3) of Table A5 which restricts the sample to households who have not experienced divorce or separation by the year the student finishes middle school.

**Credit constraints** As argued earlier, the lack of tuition fees and wide availability of student grants and loans means that differential access to credit between low and high SES families is unlikely to be driving the SES gap in elite degree enrolment in the data. Nevertheless, it may be the case that for students attending high schools outside cities where elite degrees are offered, there are additional costs associated with moving to and finding accommodation in these cities. If low SES students do not have as many acquaintances or relatives in these cities as high SES students do, then this type of credit constraints may be one mechanism behind the SES gap in elite degree enrolment that the covariates included in the model do not control for.

To tease out the extent to which this is plausible, we re-estimate the model excluding students attending high school in Oslo. Oslo is the largest municipality in Norway, containing elite universities and a high exposure to elite educated families, and it is where this sort of mechanism is more likely to be at play. Column (4) of Table A5 show that the results are robust to this exclusion. These results also show that our benchmark results are not driven by students within Oslo naturally attending their local elite universities.

**Small schools** As mentioned earlier, our identification strategy may not be valid for areas with particularly small schools, where students may move together from a shared middle school to a shared high school. Column (5) of Table A5 suggests that our benchmark estimates are robust to dropping schools in the bottom decile of school size (where there are 31 or fewer students per cohort).

Counties across Norway differed in their admissions procedure for high school between a local catchment area and, more commonly competition based upon middle school GPA. Our benchmark analysis was repeated separately by the procedure for admissions to high schools but the results are almost identical in the two samples. For the full sample, the coefficient on treatment of the proportion of parents with an elite degree is 0.027 (standard error 0.004) and 0.026 (standard error 0.005) for areas with local catchment and school choice admissions, respectively.

**High school major** High school students in Norway on the academic pathway specialize in specific majors from the second year, including sciences, economics, mathematics, social sciences, languages and humanities. Dahl et al. (2023) highlight relatively higher returns for engineering, natural sciences and economics compared to social sciences and humanities. It could be that elite students select into high school majors with relatively high returns, which explains the transition to elite degree programmes. To test whether our results are driven by major choice, we drop from the sample high school students specialising in the majors most associated with the highest returns found by Dahl et al. (2023), excluding sciences, economics and mathematics. The results, in column (6) of Table A5 are very similar to our benchmark specification.

## A3 Simulating the consequence for intergenerational mobility from reassigning low (high) SES students into schools with high (low) exposure to elite peers

We imagine the consequences for intergenerational mobility from a policy which aims to balance exposure to elite peers across high and low SES students. The idea for the simulation is as follows.

Within each cohort, all schools were ranked by the proportion of parents with an elite degree. Starting with a school with the lowest exposure to elite peers (school a), a low SES student was randomly chosen from the set of low SES students to be reassigned into a school with the highest exposure to elite peers (school b). Simultaneously, a high SES student from school b was randomly selected from the set of high SES students to be reassigned into school a. The same procedure was repeated on the school with the second lowest exposure to elite peers (school c), where a low SES student, randomly picked from all low SES students within the school-cohort, was reassigned into a school with the second highest exposure to elite peers (school d) - and a high SES student was randomly chosen from the set of high SES students within school d from the same cohort, and moved into school c.

We varied the parameters of the simulation. Simulation 1 moved just one low and one high SES student, amongst schools from the bottom or top decile of the ranked exposure to elite peers, respectively. Simulation 2 chose the same set of schools with exposure to elite peers in the bottom or top decile but moved 5 low SES students and 5 high SES students from each school. Simulation 3 moved one low SES student from each school in the bottom half of the distribution of elite peer exposure and swapped with one high SES student from each school in the top half of the distribution (where again the low SES student in school a (c) is swapped with the high SES student in school b (d) etc.). Simulation 4 extended simulation 3 by moving 5 low SES students from each low exposure school and 5 high SES students from each high exposure school.

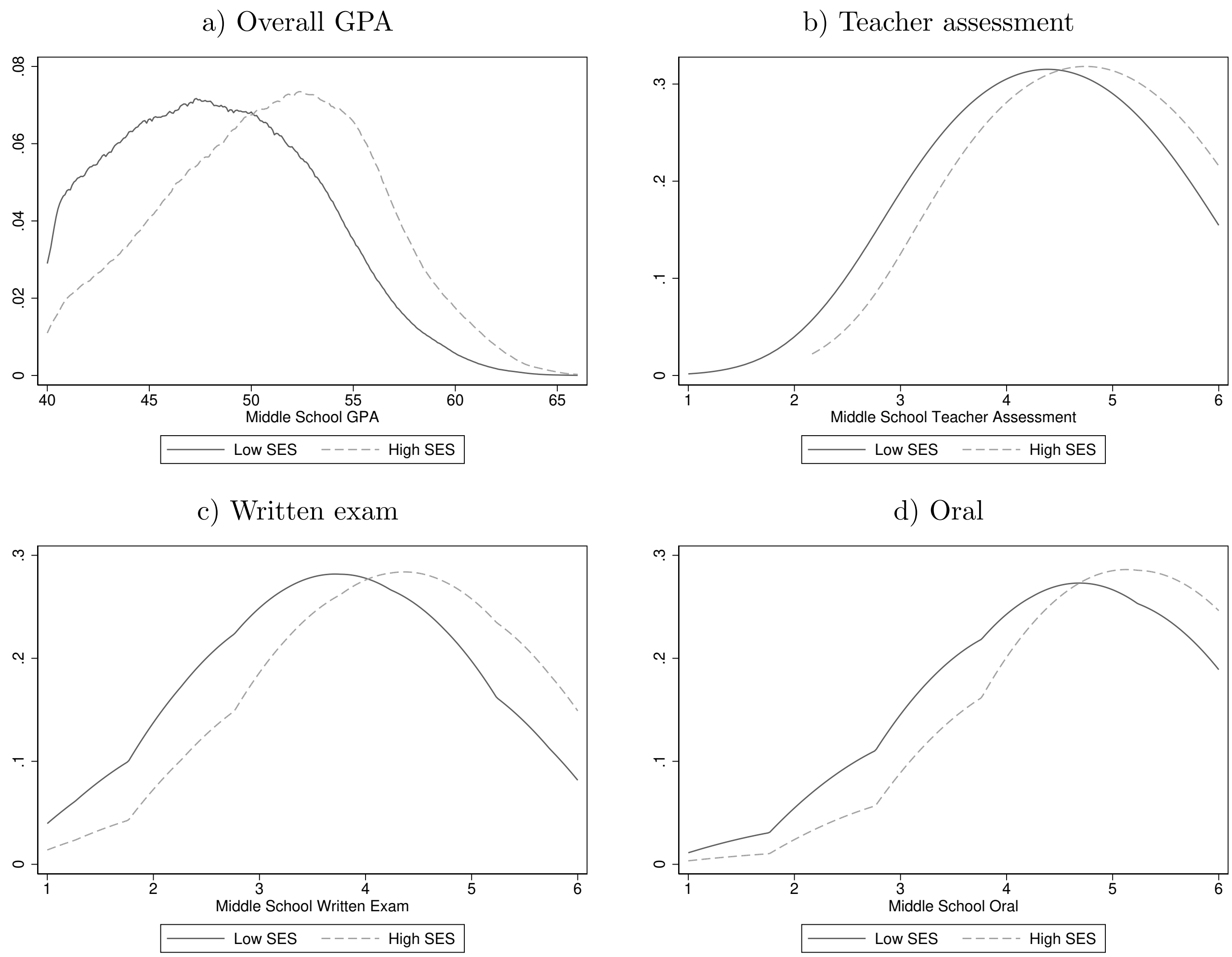
We repeat several simulations simply to show the sensitivity of our results to various reassignment strategies.

For each simulation, once the school re-allocations have taken place, we calculated the new mean exposure to parents with an elite degree within each school and cohort. Taking our estimates from Table 6 a new earnings percentile rank was calculated using the adjusted peer mean variable and assigning to any student who had been reassigned a school, the new school fixed effect. The parent percentile rank was then regressed on the new simulated earnings rank of the student to estimate the relationship between parent and student income under each of the four simulated reassignments.

This simulation does not aim to causally identify a change in intergenerational mobility, as this is not a general equilibrium model. Instead it is a useful exercise to understand whether intergenerational mobility increases or decreases from exposure to peers given the relatively higher effect of exposure to elite peers for the high SES students, compared to the low SES students.

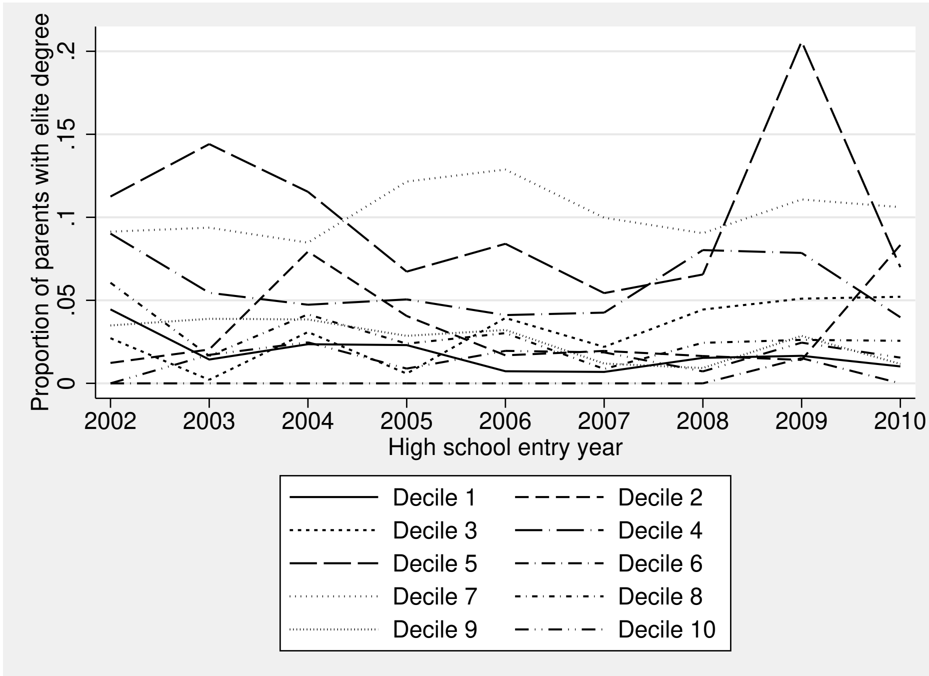
## A4 Appendix Figures

**Figure A1:** Distribution of middle school GPA taken at age 16 before entry to high school



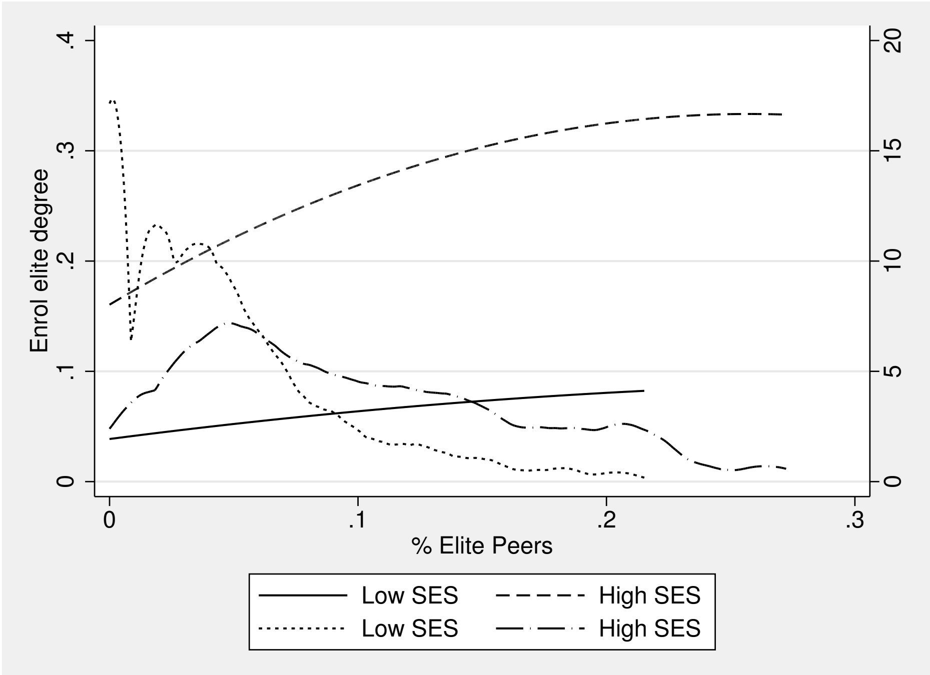

Notes: This graph plots the densities of middle school GPA scores in the low SES (solid line) and high SES samples (dashed line). The overall GPA in Panel A) is the sum of the teacher assessment (Panel B), written examinations (Panel C) and the oral examination (Panel D).

**Figure A2:** Time series of exposure to elite peers for 10 schools



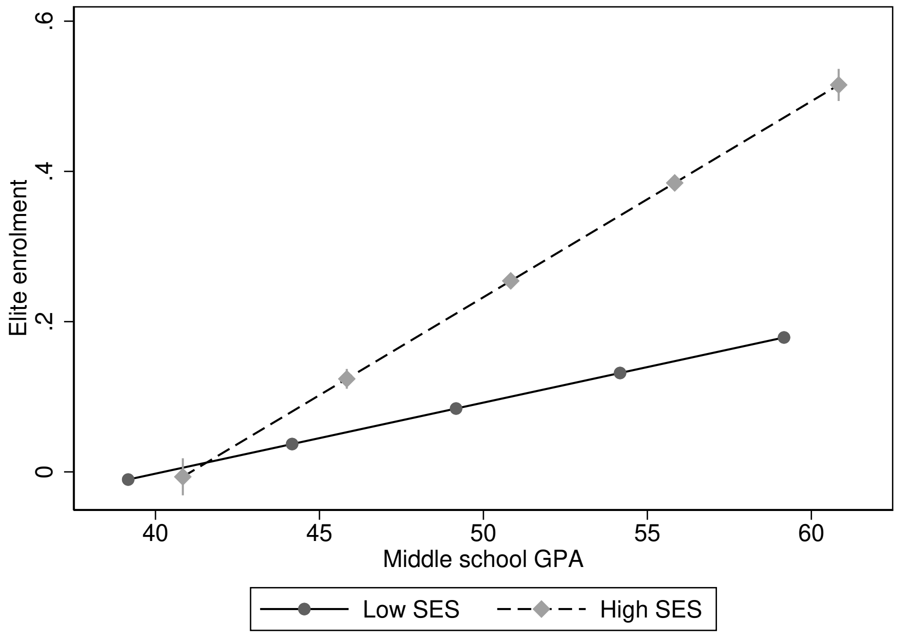

Notes: All schools were divided into deciles based on the average of their within cohort intake size across all years. One school was randomly chosen within each decile. The graph plots out the proportion of parents with an elite degree across the years for each of the ten randomly chosen schools.

**Figure A3:** Marginal effect of exposure to elite social networks implied from quadratic specification



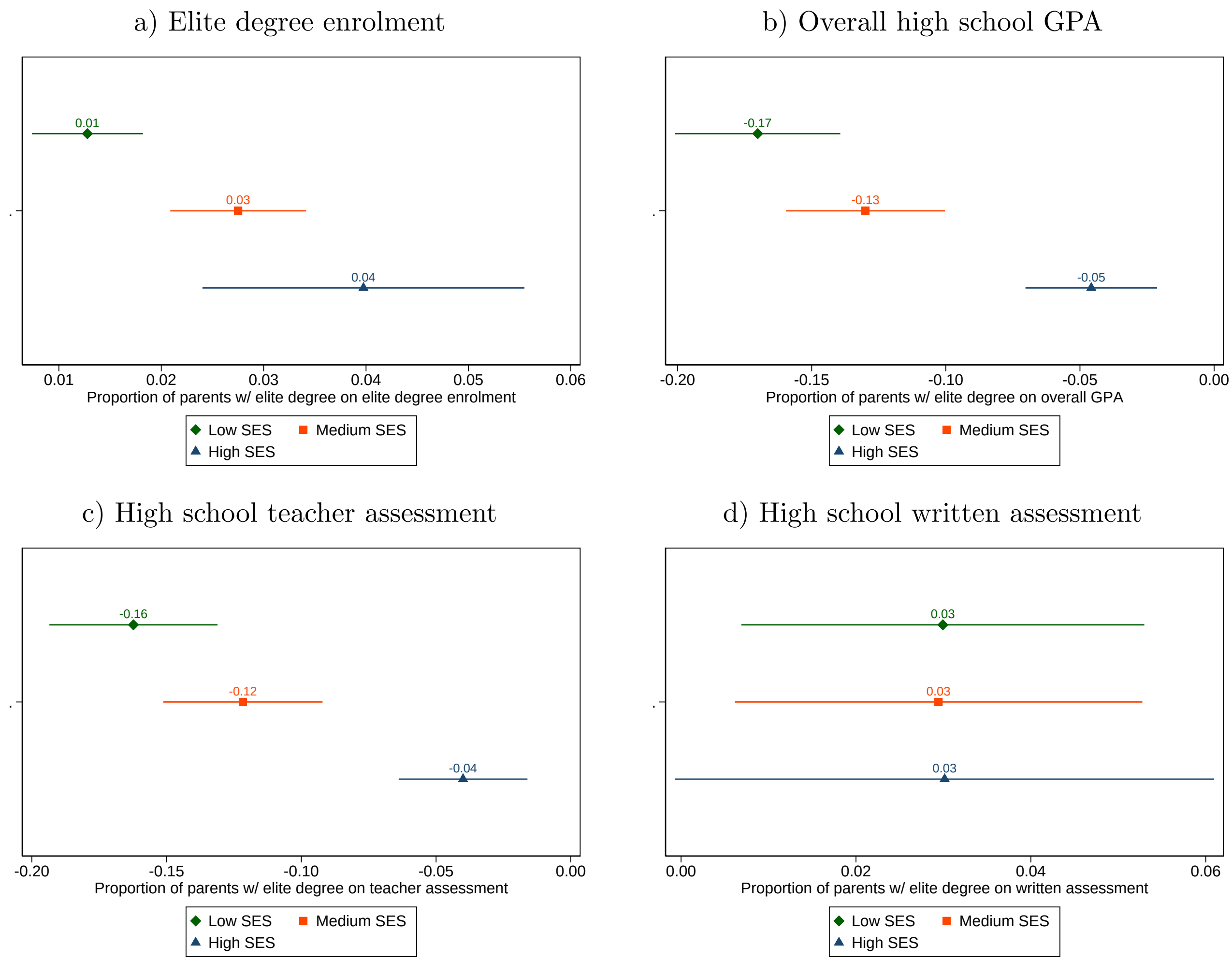

Notes: This graph plots the densities of $P_{-ics}$ in the low SES (dotted line) and high SES samples (dot-dashed line). It also plots the marginal effect of an increase in $P_{-ics}$ on the probability of enrolling in an elite degree as a function of $P_{-ics}$ as implied by estimates of $\beta_{11}$ and $\beta_{12}$ in equation (**??**). The marginal effect in the low SES (high SES) sample is plotted as a solid (dashed) line. The estimates of these coefficients are reported in Column (5) of Table 3.

**Figure A4:** Marginal effect of exposure to elite social networks: allowing for interaction between peer effect and middle school GPA



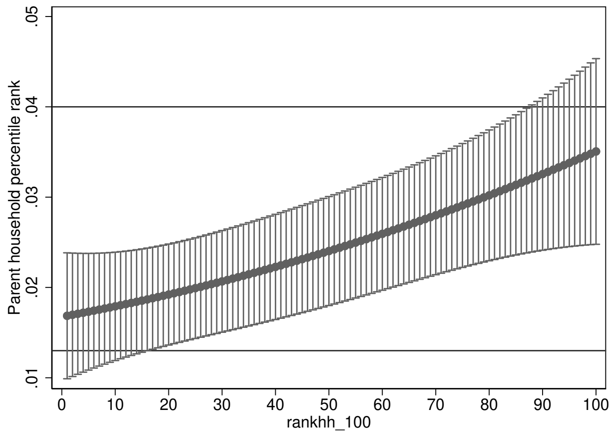

Notes: This graph plots the marginal effect of an increase in $P_{-ics}$ on the probability of enrolling in an elite degree as a function of middle school GPA. The predictions are based on the benchmark specification regression model augmented additionally with the interaction between GPA and $P_{-ics}$. The marginal effect in the low SES (high SES) sample is plotted as a solid (dashed) line.

**Figure A5:** Effect of exposure to elite peers on student outcomes by socioeconomic background



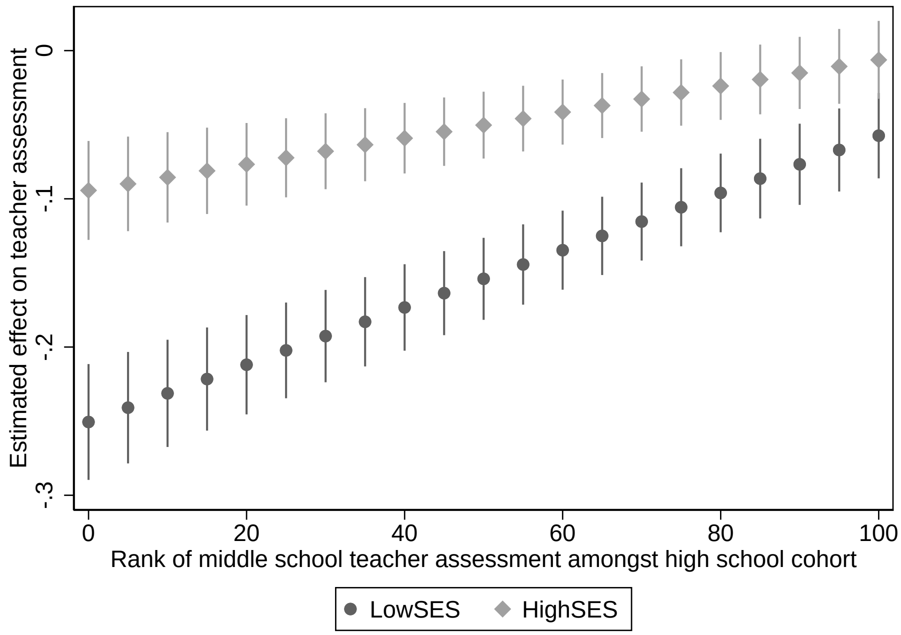

Notes: This graph plots the marginal effect of an increase in $P_{-ics}$ on student outcomes: the probability of enrolling in an elite degree; overall high school GPA; high school teacher assessment and high school written exams. The coefficients are estimated from regression Equation 1. See notes to Table 2 for details of the specification. The low SES sample is defined as the group of students who have at least one parent with no more than the compulsory level of education, but no parent with an elite education. The high SES sample is defined as the group of students who have at least one parent with a post-secondary education, but no parent with a compulsory level of education. The medium SES sample defines households with the education in between - where no parent left school at the compulsory age and no parent has an elite education.

**Figure A6:** Redefining parent SES by household income percentile rank



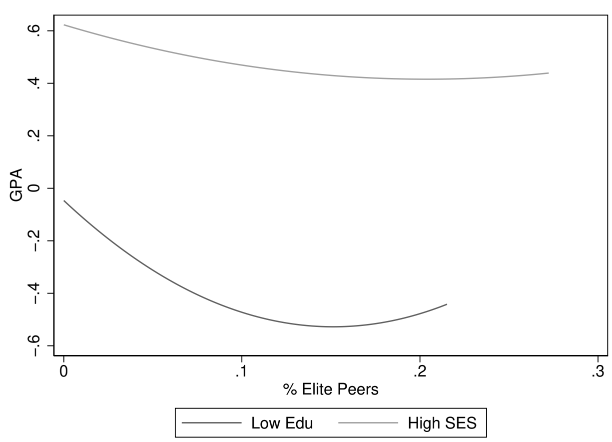

Notes: This graph plots the marginal effect (and 95% confidence intervals) of exposure to elite peers on elite degree enrolment, in the benchmark specification which is augmented by including an interaction between the peer variable $P_{-i}$ and a quadratic in the parent household income percentile rank. The horizontal lines intersect at the y-axis at points representing the benchmark estimates of the effect of exposure to elite peers for low SES defined students (the lower line) and high SES students (the upper line) defined by parents' education.

**Figure A7:** Marginal effect of exposure to elite social networks on high school teacher assessment across student middle school teacher assessment rank



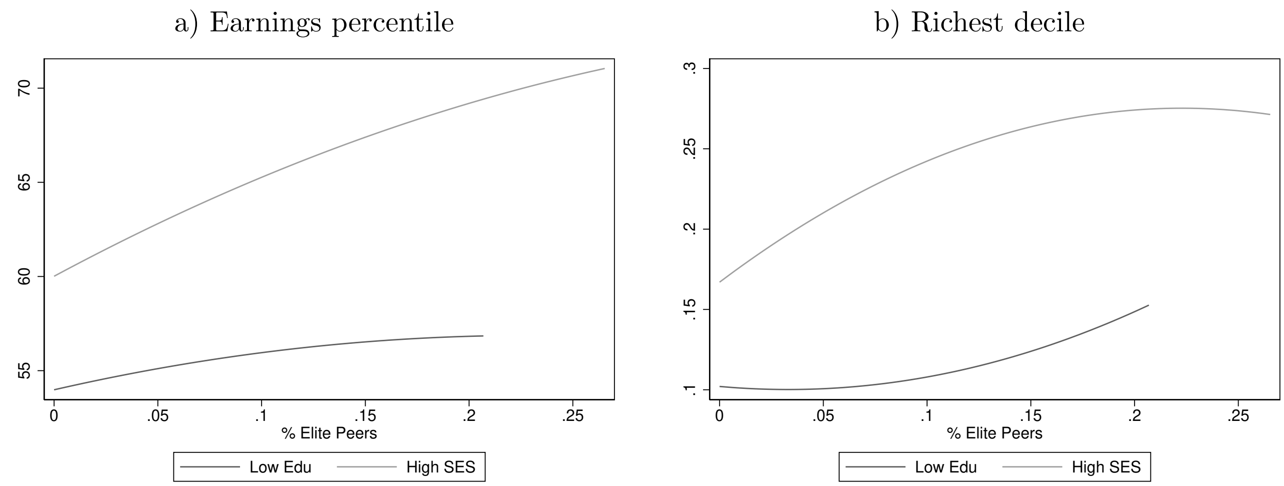

Notes: This graph plots the marginal effect of an increase in $P_{-ics}$ on the probability of enrolling in an elite degree as a function of the rank of the student's middle school teacher assessment amongst the high school cohort. Estimated on the benchmark specification including the rank of middle school GPA and an interaction between the rank and the proportion of parents from an elite educated background. The marginal effect in the low SES (high SES) sample is plotted as a dark grey circles (light grey diamonds).

**Figure A8:** Marginal effect of exposure to elite social networks implied from quadratic specification on high school GPA



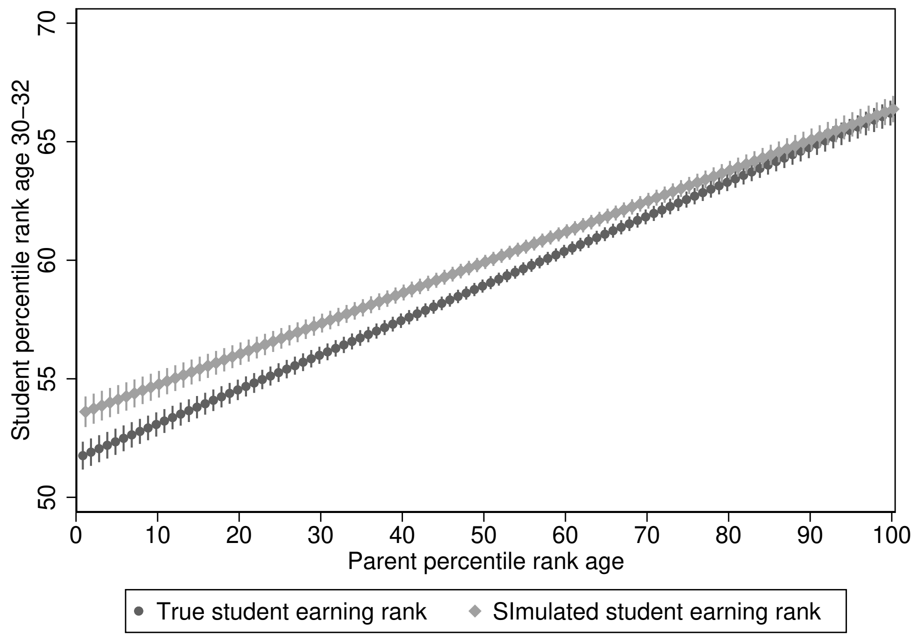

Notes: This graph plots the marginal effect of an increase in $P_{-ics}$ on the GPA as a function of $P_{-ics}$ as implied by estimates of $\beta_{11}$ and $\beta_{12}$ in equation (**??**) but with the dependent variable replaced with the overall high school GPA. The marginal effect in the low SES (high SES) sample is plotted as a dark (lighter) line. The estimates of these coefficients are reported in Column (6) of Table A7.

**Figure A9:** Marginal effect of exposure to elite social networks on student earnings age 30-32 implied from quadratic specification




Notes: This graph plots the marginal effect of an increase in $P_{-ics}$ on earnings as a function of $P_{-ics}$ as implied by estimates of $\beta_{11}$ and $\beta_{12}$ in equation (**??**) but with the dependent variable changed to the earnings percentile (Panel A) and an indicator for earning in the richest decile (Panel B). The marginal effect in the low SES (high SES) sample is plotted as a dark (lighter) line. The estimates of these coefficients are reported in Column (5) of Table A10.

**Figure A10:** The correlation between parent earnings rank percentile and student's (within-cohort) earnings rank percentile.




Notes: This graph plots the predicted relationship from a regression of parent earnings rank percentile (with parent earnings measured when the student is aged 15-19) on the student's earnings rank percentile (calculated within birth cohort) when aged 30-32. The bold symbols represent the relationship using the student's true earnings rank whilst the lighter symbols represent the relationship using the student's simulated earnings rank. The simulation, described in section A3 is represented in column (5) of Table A11 which swaps 5 low SES students from each school in the bottom half of the distribution of exposure to elite peers with 5 high SES students from each school in the top half of the distribution to exposure to elite peers. We predict the new earnings percentile rank for each student using estimates from Table 6 given the student's new exposure to elite peers and (for those reassigned students) new school fixed effect.

## A5 Appendix Tables

**Table A1:** Effect of exposure to elite families in high school on the probability of enrolling in an elite degree : Coefficients on control variables

| | (1) All | (2) Low SES | (3) High SES |
|---|---|---|---|
| Proportion of parents with elite degree (std) | 0.026*** | 0.013*** | 0.040*** |
| | (0.003) | (0.003) | (0.008) |
| Student is a female | -0.073*** | -0.053*** | -0.125*** |
| | (0.003) | (0.003) | (0.007) |
| Student is born in Norway | -0.011*** | -0.032*** | 0.013 |
| | (0.003) | (0.004) | (0.009) |
| Student's middle school GPA (std) | 0.132*** | 0.086*** | 0.255*** |
| | (0.004) | (0.003) | (0.007) |
| Proportion of student's own parent with an elite degree | 0.182*** | | 0.162*** |
| | (0.007) | | (0.021) |
| Student's parents are in top income decile | 0.027*** | 0.004 | 0.042*** |
| | (0.003) | (0.005) | (0.009) |
| *Mother's highest education level (ref = compulsory level)* | | | |
| High school | 0.015*** | 0.007 | 0.032*** |
| | (0.003) | (0.005) | (0.011) |
| University | 0.006** | | 0.006 |
| | (0.002) | | (0.010) |
| *Father's highest education level (ref = compulsory level)* | | | |
| High school | 0.018*** | 0.018*** | 0.008 |
| | (0.002) | (0.004) | (0.018) |
| University | 0.020*** | | 0.022** |
| | (0.002) | | (0.011) |
| Number of students | 177,219 | 58,328 | 20,018 |
| Number of schools | 556 | 524 | 459 |

**Table A2:** Oaxaca Binder decomposition of the SES gap in elite degree enrolment

| | **SES gap in characteristics** | | **SES gap in coefficients** | |
|---|---|---|---|---|
| | Gap | Contribution | Gap | Contribution |
| Fraction of elite peers | -0.015*** | 7.2% | -0.010*** | 4.8% |
| | (0.002) | | (0.003) | |
| Student's middle school GPA | -0.050*** | 24.2% | -0.140*** | 67.6% |
| | (0.001) | | (0.005) | |
| Fraction of own elite parent | -0.116*** | 56.0% | 0.022*** | -10.6% |
| | (0.011) | | (0.003) | |
| *Mother's highest education level (ref = compulsory level)* | | | | |
| High school | -0.001*** | -0.5% | -0.003** | 1.4% |
| | (0.000) | | (0.001) | |
| University | -0.013*** | 6.3% | 0.007** | -3.4% |
| | (0.005) | | (0.003) | |
| *Father's highest education level (ref = compulsory level)* | | | | |
| High school | 0.000*** | 0.0% | 0.001 | -0.5% |
| | (0.000) | | (0.001) | |
| University | -0.038*** | 18.4% | 0.020*** | -9.7% |
| | (0.008) | | (0.006) | |

Notes: This table reports a selected set of results from the Oaxaca decomposition of the gap in elite degree enrolment between the high SES and low SES groups of students. Specifically, we estimate the equation 1 in the sample pooling both low and high SES children, denoted by $g = L, H$ respectively. See notes to Table 2 for description of the regression and controls. For each covariate $X_{ig}$ included in the model, we construct two objects, reported in the first and second columns of the table respectively. The first, $\Delta(X)$, measures the gap in elite education enrolment between High and Low SES students explained by the gap in average characteristic X between the two groups. That is: $\Delta(X) = \beta^p_X(E^H(X_i) - E^L(X_i)$ where $\beta^p_X$ is the coefficient associated with variable $X$ in equation 1 estimated in the pooled sample and $E^g(X_i), g = H, L$ is the expected value of $X$ in each sample. The second, $\Omega(X)$, measures the gap in elite education enrolment between High and Low SES students explained by the gap in the effect of characteristic X between the two groups. That is: $\Omega(X) = (\beta^H_X - \beta^L_X)E^p(X_i)$ where $\beta^g_X$ is the coefficient associated with variable $X$ in equation 1 estimated in the sample of students $g = H, L.$ and $E^p(X_i)$ is the expected value of $X$ in the pooled sample.

**Table A3:** Placebo tests - Effect of elite peers on child birth outcomes

| | Point estimate | p-value | Number of students | Number of schools |
|---|---|---|---|---|
| Outcome variables: | | | | |
| Child birth weight | -3.303 | (0.584) | 169,864 | 554 |
| Low birth weight | 0.000 | (0.891) | 177,219 | 556 |
| Gestation | -0.027 | (0.584) | 157,669 | 552 |
| Height | -0.014 | (0.743) | 164,073 | 551 |
| Head circumference | 0.002 | (0.812) | 167,949 | 553 |
| Congenital malformation | -0.001 | (1.000) | 170,133 | 554 |
| Severe deformity | -0.002 | (0.5446) | 170,133 | 554 |

Notes: OLS estimates of the benchmark model (equation 1) on the full sample and where the dependent variables are predetermined characteristics of the student (indicated in the first column). Standard errors clustered at the school level and p-values adjusted using stepwise multiple hypothesis testing procedure that controls for family wise error rate. *** p<0.01, ** p<0.05, * p<0.1

**Table A4:** Defining "Elite" peers by occupational status and high income

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **A - Low SES students sample** | | | | |
| Proportion of parents w/elite degree | 0.013*** | | | 0.011*** |
| | (0.003) | | | (0.003) |
| Proportion of parents w/high prestige | | 0.004** | | -0.002 |
| | | (0.002) | | (0.003) |
| Proportion of parents w/elite income | | | 0.011*** | 0.008*** |
| | | | (0.002) | (0.002) |
| *Number of pupils* | *58,610* | *58,610* | *58,610* | *58,610* |
| *Number of schools* | *524* | *524* | *524* | *524* |
| **B - High SES students sample** | | | | |
| Proportion of parents w/elite degree | 0.040*** | | | 0.042*** |
| | (0.008) | | | (0.009) |
| Proportion of parents w/high prestige | | 0.011* | | -0.008 |
| | | (0.007) | | (0.007) |
| Proportion of parents w/elite income | | | 0.024*** | 0.005 |
| | | | (0.008) | (0.009) |
| *Number of pupils* | *20,018* | *20,018* | *20,018* | *20,018* |
| *Number of schools* | *459* | *459* | *459* | *459* |

*Notes*: Data source, Norwegian administrative data. Sample of students ending middle school and entering high school between 2002-2010. Dependent variable is indicator for studying for an elite (graduate) degree. Estimates of elite status defined by education (our benchmark specification - column 1); occupational prestige (column 2) and elite income (column 3). Column 4) includes all three measures of elite. High prestige is defined as working in a STEM profession, as a lawyer or a doctor. Elite income defined as household income in the top 5% of all high school parents in the sample (or the top 2-3% of the population). Results for low and high SES in Panels A) and B) respectively. Model controls for student Norwegian born, gender, middle school GPA, mother and father education and income in year before high school entry and fixed effects for school and year. Standard errors clustered at school level.

**Table A5:** Sensitivity analysis and interpretation

| | (1) Benchmark | (2) First born children | (3) Two-parent families | (4) Exclude OSLO | (5) Exclude small schools | (6) Drop high return high school majors |
|---|---|---|---|---|---|---|
| **A - Low SES students sample** | | | | | | |
| Proportion of parents with elite degree (std) | 0.013*** | 0.014*** | 0.013*** | 0.013*** | 0.013*** | 0.010*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| *Number of pupils* | *58,610* | *51,270* | *49,025* | *52,938* | *50,882* | *24,212* |
| *Number of schools* | *524* | *524* | *518* | *482* | *280* | *437* |
| **B - High SES students sample** | | | | | | |
| Proportion of parents with elite degree (std) | 0.040*** | 0.041*** | 0.034*** | 0.040*** | 0.040*** | 0.041*** |
| | (0.008) | (0.008) | (0.008) | (0.009) | (0.008) | (0.009) |
| *Number of pupils* | *20,018* | *15,439* | *17,435* | *16,444* | *19,153* | *10.131* |
| *Number of schools* | *459* | *449* | *450* | *418* | *279* | *364* |

Notes: OLS estimates of the coefficient on the variable measuring the proportion of elite educated parents in the student's youth cohort in different specifications in the low SES sample (Panel A) and in the high SES sample (Panel B). Column (1) refers to the benchmark specification from (equation 1) and also reported in Table 2. Column (2) refers to the benchmark specification estimated just for first born children. Column (3) drops the sample of divorced or separated households. Column (4) refers to the benchmark specification this time estimated on the subsample of schools outside of Oslo. Column (5) refers to the benchmark specification excluding schools in the bottom decile of the size distribution. Column (6) refers to the benchmark specification where we exclude high school students specialising in sciences, economics and mathematics. Standard errors clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

**Table A6:** Elite peer effect on GPA rank within the high school cohort

| | (1)<br>Full sample | (2)<br>Low SES | (3)<br>High SES |
|---|---|---|---|
| **A - Overall GPA** | | | |
| | -8.270*** | -10.215*** | -5.744*** |
| | (0.433) | (0.543) | (0.463) |
| *Number of observations* | *177,219* | *58,610* | *20,018* |
| **B - Components of GPA** | | | |
| Externally assessed written exam grades | -6.140*** | -7.152*** | -4.670*** |
| | (0.349) | (0.434) | (0.475) |
| *Number of observations* | *177,219* | *58,610* | *20,018* |
| Teacher-assessed internal grades | -8.272*** | -10.141*** | -5.757*** |
| | (0.437) | (0.542) | (0.474) |
| *Number of observations* | *177,219* | *58,610* | *20,018* |
| Semi-externally assessed oral exam grades | -4.233*** | -5.941*** | -2.666*** |
| | (0.347) | (0.460) | (0.530) |
| *Number of observations* | *149,488* | *49,414* | *17,189* |

Notes: OLS estimates of the effect of the proportion of parents with an elite degree in the student's school's cohort in the benchmark model controlling for average peer ability where the dependent variable is now a measure of academic performance. See notes to Table 2 for detailed list of controls. The measures of academic performance are the student's rank within their high school cohort on the overall high school GPA (row 1), average performance on externally assessed written exams across all three years of high school (row 2), average performance on teacher assessed grades across all three years of high school (row 3), and average performance on oral exams marked by an external examiner and the student's teachers across all three years of high school (row 4). All measures of performance are standardized to have mean 0 and standard deviation 1. Column (1) reports the coefficient on the proportion of parents with an elite degree estimated in the full sample, column (2) and column (3) report the same coefficient estimated in the low SES and high SES samples, respectively. Standard errors clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

**Table A7:** Validity of the empirical strategy: dependent variable is high school GPA

| | (1) Benchmark | (2) School-specific linear trends | (3) School-cohort fixed effects interacted | (4) 'Drop if more than random' | (5) Including teacher traits | (6) Quadratic specificaiton |
|---|---|---|---|---|---|---|
| **A - Low SES students sample** | | | | | | |
| Proportion of parents with elite degree (std) | -0.171*** | -0.178*** | | -0.339*** | -0.174*** | -0.215*** |
| | (0.016) | (0.016) | | (0.013) | (0.018) | (0.014) |
| Proportion of parents with elite degree squared | | | | | | 0.067*** |
| | | | | | | (0.009) |
| *Number of pupils* | *58,610* | *58,610* | | *28,181* | 37,270 | *58,610* |
| *Number of schools* | *524* | *524* | | *284* | *390* | *524* |
| **B - High SES students sample** | | | | | | |
| Proportion of parents with elite degree (std) | -0.046*** | -0.053*** | | -0.218*** | -0.028*** | -0.081*** |
| | (0.012) | (0.014) | | (0.035) | (0.013) | (0.019) |
| Proportion of parents with elite degree squared | | | | | | 0.016*** |
| | | | | | | (0.005) |
| *Number of pupils* | *20,018* | *20,018* | | *8,420* | 12,737 | *20,018* |
| **C - Low and High SES students sample** | | | | | | |
| Proportion of parents with elite degree (std) | | | -0.058*** | | | |
| | | | (0.006) | | | |
| Indicator for low SES | | | 0.052** | | | |
| | | | (0.024) | | | |
| Proportion of parents with elite degree * low | | | -0.146*** | | | |
| | | | (0.008) | | | |
| *Number of pupils* | | | *78,540* | | | |

Notes: OLS estimates effect of the proportion of elite educated parents in the student's youth cohort on high school GPA, in different specifications in the low SES sample (Panel A), in the high SES sample (Panel B) and pooled sample of low and high SES (Panel C). Column (1) the benchmark specification (equation 1) and Table 2. Column (2) is benchmark specification control also for school-specific linear trends. Column (3) includes fully interacted fixed effects for the school and cohort; estimating on the pooled sample of low and high SES students. Column (4) the benchmark specification estimated on the subsample of schools where variation in the elite peer variable evolves over time in a random way. Specifically, we drop the schools where the $R^2$ from a school-level regression of the proportion of elite educated peers on a quadratic in year is 1.05 times the $R^2$ from five regressions where cohorts are randomly re-ordered for each. See section 4. Column (5) the benchmark specification including additionally average traits of teachers within schools across cohorts: the proportion of females, the proportion of teachers from a professional or low skilled background and average age. The teacher background is defined by the occupation of their father. Column (6) refers to the benchmark specification augmented with a quadratic term in the elite peer variable. Regressions are weighted by school size. Standard errors clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

**Table A8:** Balance. Dependent variable is the Instrumental Variable.

| | (1) Low SES | (2) High SES |
|---|---|---|
| Proportion of parents with elite degree (std) | -0.009 | 0.010* |
| | (0.006) | (0.006) |
| Student is female | 0.001 | 0.006 |
| | (0.004) | (0.004) |
| Student is born in Norway | -0.004 | -0.001 |
| | (0.005) | (0.007) |
| Mother years of schooling | -0.001 | 0.001 |
| | (0.001) | (0.001) |
| Father years of schooling | -0.000 | -0.001 |
| | (0.001) | (0.001) |
| Middle school teacher assessments | -0.039 | -0.064 |
| | (0.063) | (0.101) |
| Middle school written assessments | -0.001 | -0.003 |
| | (0.006) | (0.009) |
| Middle school oral exams | -0.001 | -0.003 |
| | (0.005) | (0.008) |
| Middle school overall GPA | 0.015 | 0.061 |
| | (0.072) | (0.116) |
| Proportion of student's own parent with an elite degree | -0.013 | 0.012 |
| | (0.036) | (0.013) |
| Student's parents are in top income decile | 0.006 | -0.018** |
| | (0.006) | (0.007) |
| Number of pupils | 52,446 | 17,806 |
| Number of schools | 520 | 450 |

Notes: OLS estimates of a regression of the instrumental variable which is an indicator for a being assigned a maths examination through a lottery in years 2 or 3 of high school on the set of covariates reported and additionally school, cohort and programme fixed effects. The low SES sample in column (1) is defined as the group of students who have at least one parent with no more than the compulsory level of education, but no parent with an elite education. The high SES sample in column (2) is defined as the group of students who have at least one parent with a post-secondary education, but no parent with a compulsory level of education. Standard errors clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

**Table A9:** Mincer equation estimating correlation between elite degree programme and earnings

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Low SES | High SES | Low SES | High SES | Low SES | High SES |
| Dependent variable | Earnings percentile | | Richest decile | | Richest percentile | |
| Student ever enrolled in degree | 10.032*** | 9.120*** | 0.027*** | 0.048** | 0.008*** | 0.034*** |
| | (0.454) | (1.428) | (0.005) | (0.021) | (0.002) | (0.011) |
| Student enrolled in elite degree: | | | | | | |
| STEM | 25.699*** | 22.596*** | 0.243*** | 0.203*** | 0.040*** | 0.040*** |
| | (1.009) | (1.541) | (0.012) | (0.023) | (0.005) | (0.012) |
| Law | 24.143*** | 23.765*** | 0.146*** | 0.162*** | 0.030*** | 0.076*** |
| | (1.382) | (1.892) | (0.016) | (0.028) | (0.008) | (0.015) |
| Medicine | 39.014*** | 33.982*** | 0.642*** | 0.548*** | 0.218*** | 0.167*** |
| | (2.485) | (2.029) | (0.029) | (0.030) | (0.014) | (0.016) |
| *Number of students* | 20,454 | 6,765 | 20,454 | 6,765 | 20,454 | 6,765 |

Notes: Mincer-style regressions of earnings percentile age 30-32 (columns (1) and (2)); an indicator for earning in the top decile (columns (3) and (4)) and an indicator for earning in the top percentile (columns(5) and (6)) on indicators for a degree, an elite STEM degree, an elite law degree and an elite medicine degree. The omitted category is no degree. The low SES sample in columns (1), (3) and (5) is defined as the group of students who have at least one parent with no more than the compulsory level of education, but no parent with an elite education. The high SES sample in columns (2), (4) and (6) is defined as the group of students who have at least one parent with a post-secondary education, but no parent with a compulsory level of education. Sample of birth cohorts 1986-1988. Income is deflated to 2020. For the cohorts 1986; 1987 and 1988 income is measured ages 30-32; 30-31 and 30 respectively (see section 3). The regressions include a gender dummy and year of birth dummy variables as controls. Standard errors clustered at the school level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

**Table A10:** Validity of the empirical strategy: dependent variable is earnings in richest decile

| | (1) Benchmark | (2) Including family fixed effect | (3) School-specific linear trends | (4) 'Drop if more than random' | (5) Quadratic specifi-caiton |
|---|---|---|---|---|---|
| **A - Low SES students sample** | | | | | |
| Proportion of parents with elite degree (std) | 0.004 | -0.003 | 0.004 | 0.005 | 0.001 |
| | (0.004) | (0.036) | (0.005) | (0.007) | (0.005) |
| Proportion of parents with elite degree squared | | | | | 0.004 |
| | | | | | (0.003) |
| *Number of pupils* | *20454* | *20454* | *20454* | 9,710 | *20454* |
| *Number of schools* | *457* | *457* | *457* | 236 | *457* |
| | | | | | |
| **B - High SES students sample** | | | | | |
| Proportion of parents with elite degree (std) | 0.022*** | -0.045 | 0.036** | 0.040** | 0.039** |
| | (0.009) | (0.043) | (0.016) | (0.016) | (0.018) |
| Proportion of parents with elite degree squared | | | | | -0.005 |
| | | | | | (0.006) |
| *Number of pupils* | *6765* | *6765* | *6765* | 2,827 | *6765* |
| | *372* | *372* | *372* | 178 | *372* |

Notes: OLS estimates of the coefficient on the variable measuring the fraction of elite educated parents in the student's youth cohort in different specifications in the low SES sample (Panel A) and in the high SES sample (Panel B), on earning in the richest decile age 30-32. Column (1) refers to the benchmark specification (equation 1) and reported in Table 6. Column (2) refers to the benchmark specification where we also control for school-specific linear trends. Column (3) refers to the benchmark specification this time estimated on the subsample of schools where variation in the elite peer variable evolves over time in a random way. Specifically, we drop the schools where the $R^2$ from a school-level regression of the proportion of elite educated peers on a quadratic in year is 1.05 times the $R^2$ from five regressions where cohorts are randomly re-ordered for each. See section 4 for full details. Column (4) refers to the benchmark specification where we also control for a family fixed effect. Column (5) refers to the benchmark specification augmented with a quadratic term in the elite peer variable. Regressions are weighted by school size to take account of the parent peer variables group averages, taken from groups of different sizes. Standard errors clustered at the school level. Standard errors clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

**Table A11:** Simulating a reassignment of low (high) SES students into schools with a high (low) level of elite peers

| | (1) Benchmark | (2) Simulation 1 | (3) Simulation 2 | (4) Simulation 3 | (5) Simulation 4 |
|---|---|---|---|---|---|
| No. schools | | Top and bottom decile | | All schools | |
| No. children moved per school | | 1 | 5 | 1 | 5 |
| Parent percentile rank | 0.143*** | 0.139*** | 0.136*** | 0.139*** | 0.129*** |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| Constant | 53.090*** | 52.829*** | 53.032*** | 52.874*** | 53.474*** |
| | (0.332) | (0.332) | (0.332) | (0.332) | (0.334) |
| Observations | 30,849 | 30,849 | 30,849 | 30,849 | 30,849 |
| R-squared | 0.023 | 0.021 | 0.020 | 0.021 | 0.018 |

Notes: The table reports the coefficients from a regression of parent percentile rank in the earnings distribution (measured when the student was aged 15-19) on the student's (within birth cohort) percentile rank at age 30-32. See section A3 for full details of the reassignment exercise. Column (1) reports the benchmark regression. Columns (2) and (4) report the coefficients after a simulation which reassigns one low SES student (one high SES student) from each school with the lowest (highest) exposure to elite peers into the schools with the highest (lowest) exposure, (Simulations 1 and 3) respectively. Columns (3) and (5) instead swap 5 low SES students from the low exposure school with 5 high SES students in the high exposure school (Simulations 2 and 4). In columns (2) and (3) low SES students are moved out of schools in the bottom decile of the distribution of exposure to elite peers and into schools in the top decile of the distribution; whilst the high SES students move from schools in the top to the bottom decile. In columns (4) and (5), one or five low SES students are moved out of all schools in the bottom half of the distribution of exposure to elite peers whilst one or five high SES students are moved out of all schools in the top half of the distribution. Exactly which students are chosen to be reassigned is explained in section A3.

**Table A12:** Gender differences in effect of elite parent peers

| | (1) All | (2) Low SES | (3) High SES |
|---|---|---|---|
| **A - Sample of females** | | | |
| Proportion of parents with elite degree (std) | 0.018*** | 0.008*** | 0.032*** |
| | (0.003) | (0.002) | (0.008) |
| **B - Sample of males** | | | |
| Proportion of parents with elite degree (std) | 0.039*** | 0.025*** | 0.051*** |
| | (0.006) | (0.005) | (0.012) |
| Number of female pupils | 106,421 | 37,945 | 10,559 |
| Number of male pupils | 70,798 | 20,383 | 9,459 |

*Notes*: Data source, Norwegian administrative data. Sample of students ending middle school and entering high school between 2002-2010. OLS estimation where the dependent variable is an indicator for studying for an elite (graduate) degree. Model controls for student Norwegian born, gender, middle school GPA, mother and father education and income in year before high school entry and fixed effects for school and cohort. Panel A) estimated for the sample of female students and Panel B) for the sample of male students. Standard errors clustered at school level. Low SES (high SES) education household contains at least one parent with compulsory (elite) education.

**Table A13:** Gender differences in effect of elite parent peers

| | (1)<br>All | (2)<br>Low SES | (3)<br>High SES |
|---|---|---|---|
| **A - Sample of females** | | | |
| Proportion of mothers with elite degree (std) | 0.005* | 0.000 | 0.006 |
| | (0.003) | (0.003) | (0.007) |
| Proportion of fathers with elite degree (std) | 0.014*** | 0.007*** | 0.028*** |
| | (0.003) | (0.003) | (0.009) |
| **B - Sample of males** | | | |
| Proportion of mothers with elite degree (std) | 0.014*** | 0.014*** | 0.018** |
| | (0.004) | (0.005) | (0.009) |
| | 0.027*** | 0.013** | 0.035*** |
| Proportion of fathers with elite degree (std) | (0.005) | (0.005) | (0.013) |
| Number of female pupils | 106,421 | 37,945 | 10,559 |
| Number of male pupils | 70,798 | 20,383 | 9,459 |

*Notes*: Data source, Norwegian administrative data. Sample of students ending middle school and entering high school between 2002-2010. OLS estimation where the dependent variable is an indicator for studying for an elite (graduate) degree. The model allows a different coefficient between the peers' mother or father having an elite degree. Model controls for student Norwegian born, gender, middle school GPA, mother and father education and income in year before high school entry and fixed effects for school and cohort. Panel A) estimated for the sample of female students and Panel B) for the sample of male students. Standard errors clustered at school level. Low SES (high SES) education household contains at least one parent with compulsory (elite) education.

@cepeo_ucl

ucl.ac.uk/ioe/cepeo