

First generation elite: the role of school social networks: Supplemental appendix

Sarah Cattan

Institute for Fiscal Studies and IZA

Kjell G. Salvanes

Department of Economics, Norwegian School of Economics¹

Emma Tominey

Department of Economics, University of York, HCEO, FAIR and IZA

Online Appendix

A1 Robustness to alternative definitions 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. We now test the robustness of our findings to alternative definitions of SES based on parental education as well as on family income.

Figure A4 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.¹ Moving to an income-based definition of SES, **Figure A5** 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.

A2 **Borusyak and Hull (2023)** method to control for the expected treatment

Our identification assumption is that, conditional on school, cohort fixed effects and middle school GPA, the exposure to elite educated peers is random. The idea of **Borusyak and Hull (2023)** would be to simulate counterfactual peer measures by permutting students within the same school and middle school GPA bands across cohorts. This would lead to a school level estimate of the “expected” treatment. If our identification

¹FAIR, CEPR, CESifo, HCEO and IZA

¹The coefficients (standard error) are 0.024 (0.003) and 0.040 (0.008) for the low and high SES samples respectively.

assumptions are correct, regressing the elite enrolment only on the true exposure to elite peers, and this expected treatment will lead to a coefficient comparable to our benchmark.

In Section 5 we report estimates using this method which control also for the individual controls X . However to illustrate that we need to control just for the expected treatment (and not the controls X) we include [Table A3](#). Panel A) reports the coefficient from the benchmark model but includes only the essential controls (the continuous measure of MGPA, school and cohort fixed effects) as regressors. That is, the coefficient on P_{-ics} in panel A corresponds to coefficient β_1 in the following regression

$$Y_{ics} = \beta_1 P_{-ics} + \beta_2 MGPA_i + \lambda_s + \alpha_c + \epsilon_{ics}. \quad (1)$$

Panel B reports the coefficient on P_{-ics} and $\tilde{P}_s$, where the latter is labelled “expected treatment” in a regression of the outcome on these variables:

$$Y_{ics} = \delta_1 P_{-ics} + \delta_2 \tilde{P}_s + u_{ics} \quad (2)$$

Finally, in Table 2 the specification in column (2) applying the [Borusyak and Hull \(2023\)](#) method results in a smaller sample owing to some empty cells across the schools within middle school GPA bands. Panel C of [Table A3](#) repeats our benchmark specification on the smaller sample, to allow a full comparison and the estimates are almost unchanged.

A3 Additional validity and specification checks

A3.1 School-specific linear trends

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.² As shown in column (2) of [Table A4](#), the estimates of this specification are very similar to those from our benchmark specification (included in the first column of the table for easy comparison).

A3.2 Fully interacted school and cohort fixed effects

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

²That is, we estimate the following model: $Y_{ics} = \beta_1 P_{-ics} + X'_{ics} \beta_2 + \alpha_{s_1} + \alpha_{s_2} c \times D_{is} + \rho_c + \epsilon_{ics}$ where c is a cohort (linear) trend and D_s is an indicator for whether the student is in school s .

estimates the model 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. The results, presented in panel C, report 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.

A3.3 Drop if more than random

Our third test is in column (4) of [Table A4](#) is based on an idea proposed by [Hoxby \(2000\)](#) and referred to as ‘drop if more than random’ in this paper. A sensitivity analysis drops the schools whose variation in treatment does not appear to be random. These schools are selected as follows. 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 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.

A3.4 Control for changing teacher quality

A main concern with our identification strategy is that elite parents select into schools of improving quality. A fourth check aims to control for observable changes in school quality measured by the characteristics of teachers teaching in school s to cohort c - including 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). These estimates reported in column and (5) of [Table A4](#), are very similar to the benchmark estimates.

A3.5 Placebo tests

The estimates of coefficients β_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, middle school subject choice

and parents' income during middle school. We pick these outcomes because these student and parent 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. Each row in [Table A5](#) reports the coefficient on the elite peer variables in equation (1) where the dependent variable is a birth outcome; a middle school subject choice or mother or father income during middle school. As expected, the exposure to elite peers during high school is unrelated to outcomes measured before high school. An exception is the variable for fathers' income during middle school for which elite exposure during middle school is statistically significant at the 5% level. To ensure that this variable does not invalidate our estimates, [Table A6](#) runs the benchmark estimation controlling for fathers' income during middle school, showing that the point estimates are identical to our benchmark estimation.

A further placebo test adds measures of the elite peer composition in the period before (lead) and after (lag) the cohort entered high school. If our identification strategy is valid, adding additional measures of the proportion of students with elite educated parents in the cohorts before (leads) or cohorts after (lags) the students' cohort should have no direct effect on the decision to enrol in elite education. The placebo analysis is reported in [Table A7](#). [Figure A6](#) plots the estimates for the students' true exposure along with the coefficients and confidence intervals for two leads and lags in the peer exposure variable and confirms that these variables are not statistically significant predictors of elite degree enrolment.

A3.6 Sensitivity of results to sample selection - birth order

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., 2005](#)). Column (2) of [Table A8](#) shows that indeed the peer effect is slightly higher for first borns, although the new estimates are not statistically different to the benchmark estimates.

A3.7 Measurement error from marital breakup

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 A8](#) which restricts the sample to households who have not experienced divorce or separation by the year the student finishes middle school.

A3.8 Credit constraints - city of residence

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 A8](#) 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.

A3.9 Small schools

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 A8](#) 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).

A3.10 High school admissions mechanism

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.

A3.11 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 A8](#) are very similar to our benchmark specification.

A3.12 Measurement error

We consider whether, as [Angrist \(2014\)](#) points out, measurement error in parents' elite education may cause bias in our peer effect estimates and in particular inflate them ([Feld and Zoelitz, 2017](#)). To address if this is the case in our setting, we follow [Carrell et al. \(2018\)](#) and add measurement error to parents' elite education. Measurement error is added sequentially to the parents' elite education, at the rate of 1%, 5%, 10%, 25%, 50%, 75%, 90% and 100%. For each rate (e.g. 1%), a 1% sub-sample is chosen to be assigned error. Among the error sample, we randomly assign 2% of the sample to have both parents with an elite degree; 9% to have one parent with an elite degree and the remaining to have no parents with an elite degree, where these values represent the distribution in the benchmark sample. The new peer mean variables are calculated and the benchmark equation is re-estimated. [Figure A7](#) shows that, similarly to [Carrell et al. \(2018\)](#), as more measurement error is added to the peer mean variable, our estimates attenuate towards zero, suggesting that our estimates are not likely confounded by measurement error.

A3.13 Nonlinearities

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.³ First, panel d) of [Figure A8](#) tests for peer effects at the extensive margin, measuring elite peer exposure by an indicator taking the value of 1 if the student has any elite peers, and 0 otherwise. The results suggest that exposure to any elite peers raises elite enrolment by 0.8ppts and 5.1ppts for low and high SES students respectively.

If elite peer effects were non-linear in exposure to elite peers, the social gradient in elite peer effects could merely reflect that low SES students have lower average exposure to peers than high SES students. [Figure A9](#) plots the marginal effect of the proportion of elite families as implied by the estimates of a quadratic specification for each group, along with the densities of P_{-ics} . 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 β_{11} and β_{12} are then used to compute these marginal effects and are reported in column (3) of Table 2. There is little evidence of

³See [Feld and Zoelitz \(2017\)](#), [Lavy et al. \(2011\)](#), [Tincani \(2017\)](#) for evidence of non-linearities in peer effects.

non-linearity in the elite peer effect.⁴

To test whether the SES gradient in elite peer effects reflects a student ability gradient, we re-estimate the benchmark model allowing for the effect of elite peers on enrolment to vary with middle school GPA. [Figure A10](#) plots predicted elite degree enrolment and shows, for low ability students, the likelihood to enrol into an elite degree programme is very low for both low and high SES students. The SES gradient materialises and increases with student ability. Compared to high SES students of the same ability, the low SES students with medium to high ability face the greatest disadvantage in terms of the benefits of being exposed to elite peers.

A4 Heterogeneity analysis: Definitions of elite

Elite educated parents of peers may have high levels of income and consumption and possibly to hold prestigious occupations and these could be why high school students are influenced 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 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). 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. Within our sample 19% of families are present in the top 5% of the income distribution, and there is a social gradient such that 17% of low SES and 24% of high SES families are defined as wealthy peers.

The second measure of family eliteness 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 \(2018\)](#)). 1.7%, 1.4% and 2.3% of the total sample; low and high SES families respectively are defined as having occupational prestige.

Panels b) and c) of [Figure A8](#) plot the coefficients and confidence intervals from regressions which change the peer variable to be prestigious peers (those with an elite occupation) and wealthy peers (with an elite level of income) respectively. As expected, these two dimensions of elite peer matter for student enrolment.

⁴The coefficient on the quadratic term for high SES is negative and statistically significant, but 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.

A5 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.

The simulations take a low SES student from a school with the lowest exposure 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. 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.

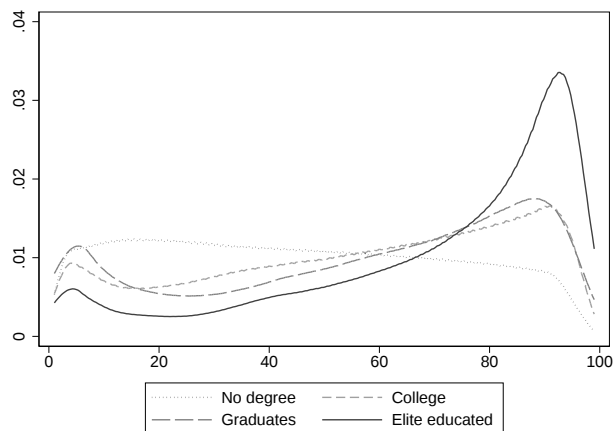
For each simulation, once the school re-allocations have taken place, we calculated the new mean exposure to parents with an elite degree within each school and cohort. Taking our estimates from Table 5 a new earnings percentile rank was calculated using the adjusted peer mean variable and assigning 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.

Table A16 shows that for all simulations, intergenerational mobility increases once the exposure to elite peers is re-balanced across low and high SES students, as indicated by 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 A15**. After the simulation, the intercept from the rank-rank regression 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.

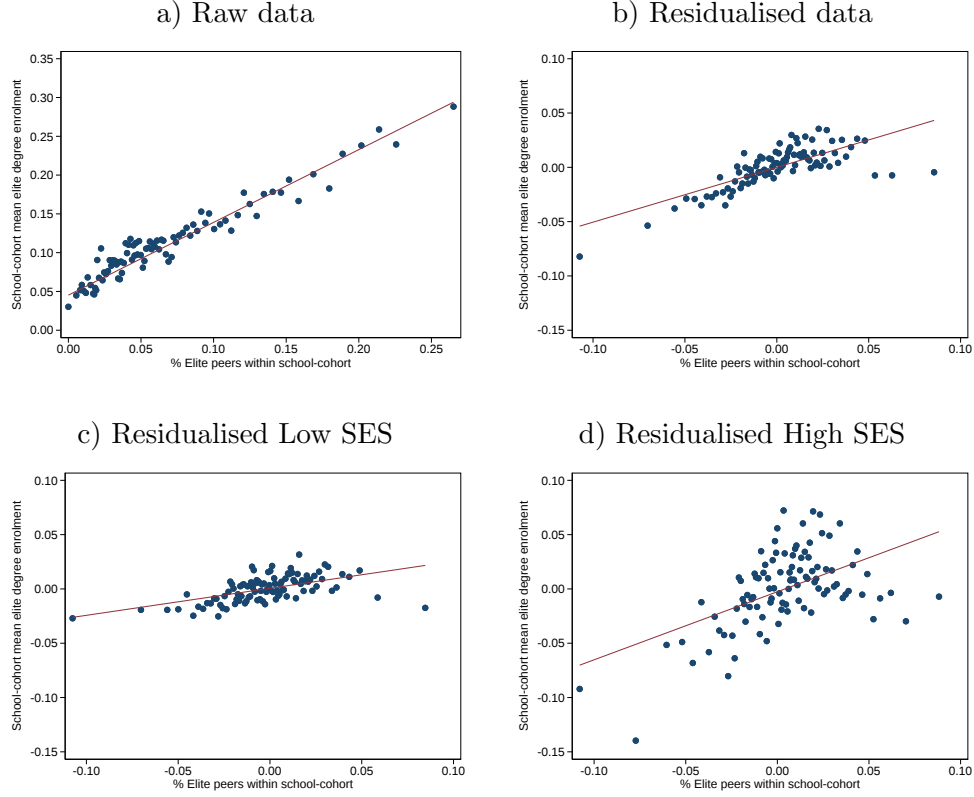
A6 Appendix Figures

Figure A1: Density of earnings percentiles by education level



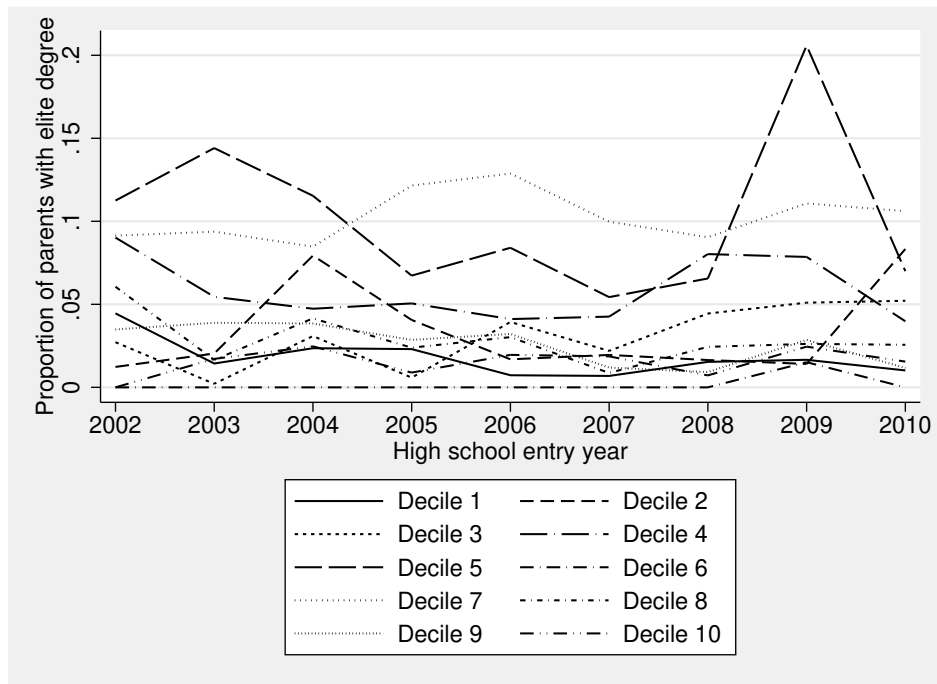
Notes: This graph plots the density of earnings percentiles across educational groups. Sample is the population of Norway aged 28-40 between 1993-2001. The percentile rank of earnings is calculated within each birth cohort.

Figure A2: Plotting school-cohort variation in elite peers against elite enrolment



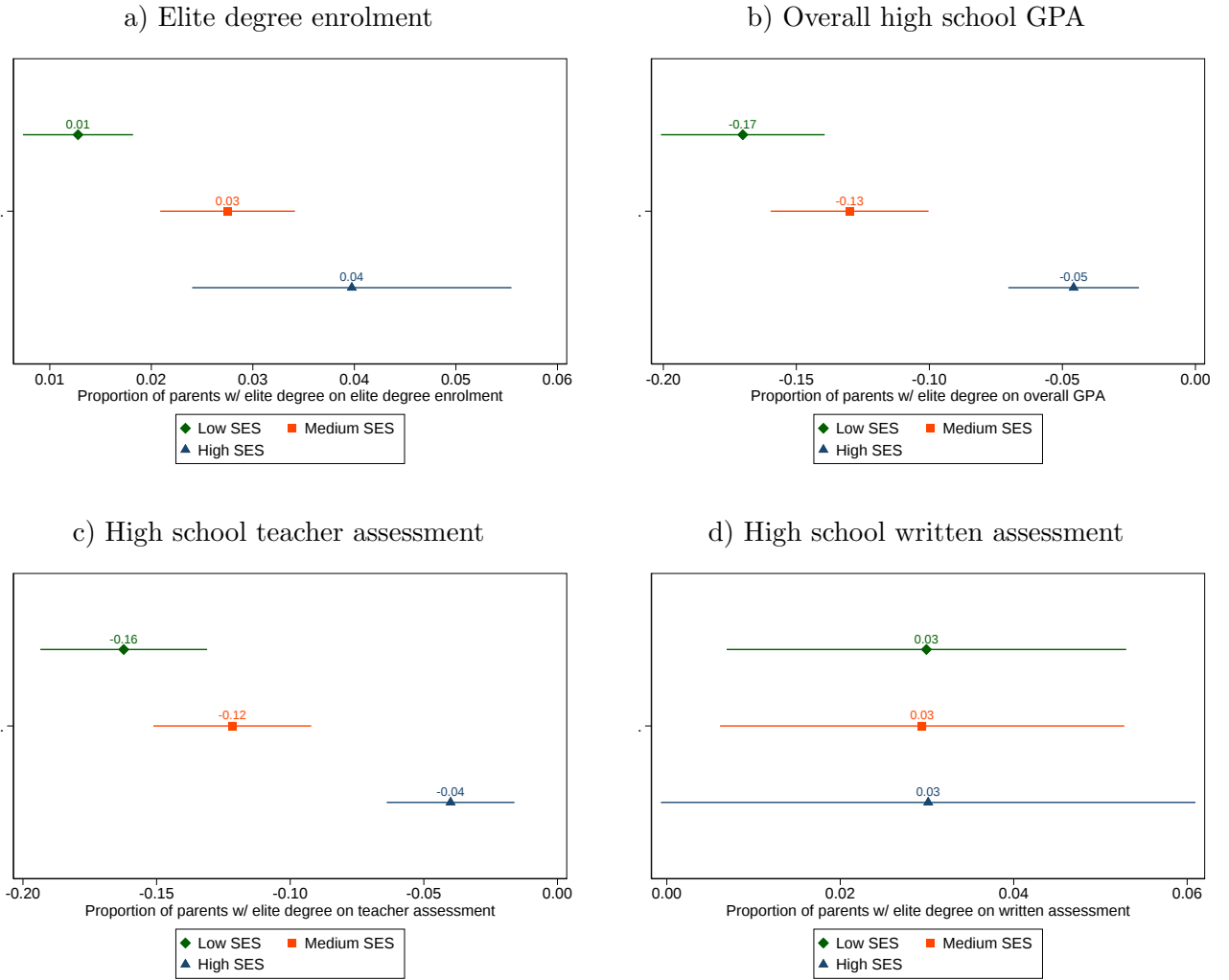
A residual (R_y) is predicted from a regression of elite degree enrolment on cohort and school fixed effects and middle school GPA. A residual (R_t) is predicted from a regression of exposure to elite peers on cohort and school fixed effects and middle school GPA. Bin scatter figures in panel a) plot the variance in the raw data of elite degree enrolment and elite peer exposure and panels b-d) plot R_y against R_t and the linear fit for the full, low and high SES samples.

Figure A3: Time series of exposure to elite peers for 10 schools



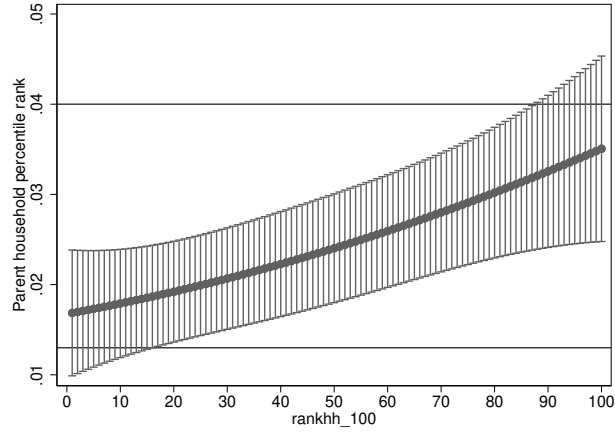
Notes: All schools were divided into deciles based on the average of their within cohort intake size across all years. One school was randomly chosen within each decile. The graph plots out the proportion of parents with an elite degree across the years for each of the ten randomly chosen schools.

Figure A4: Effect of exposure to elite peers on student outcomes by socioeconomic background



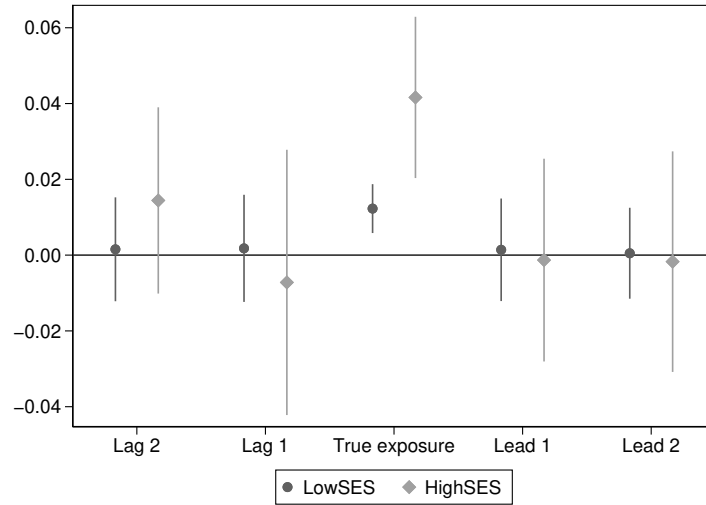
Notes: This graph plots the marginal effect of an increase in P_{-ics} on student outcomes: the probability of enrolling in an elite degree; overall high school GPA; high school teacher assessment and high school written exams. The coefficients are estimated from regression equation (1). See notes to 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 A5: Redefining parent SES by household income percentile rank



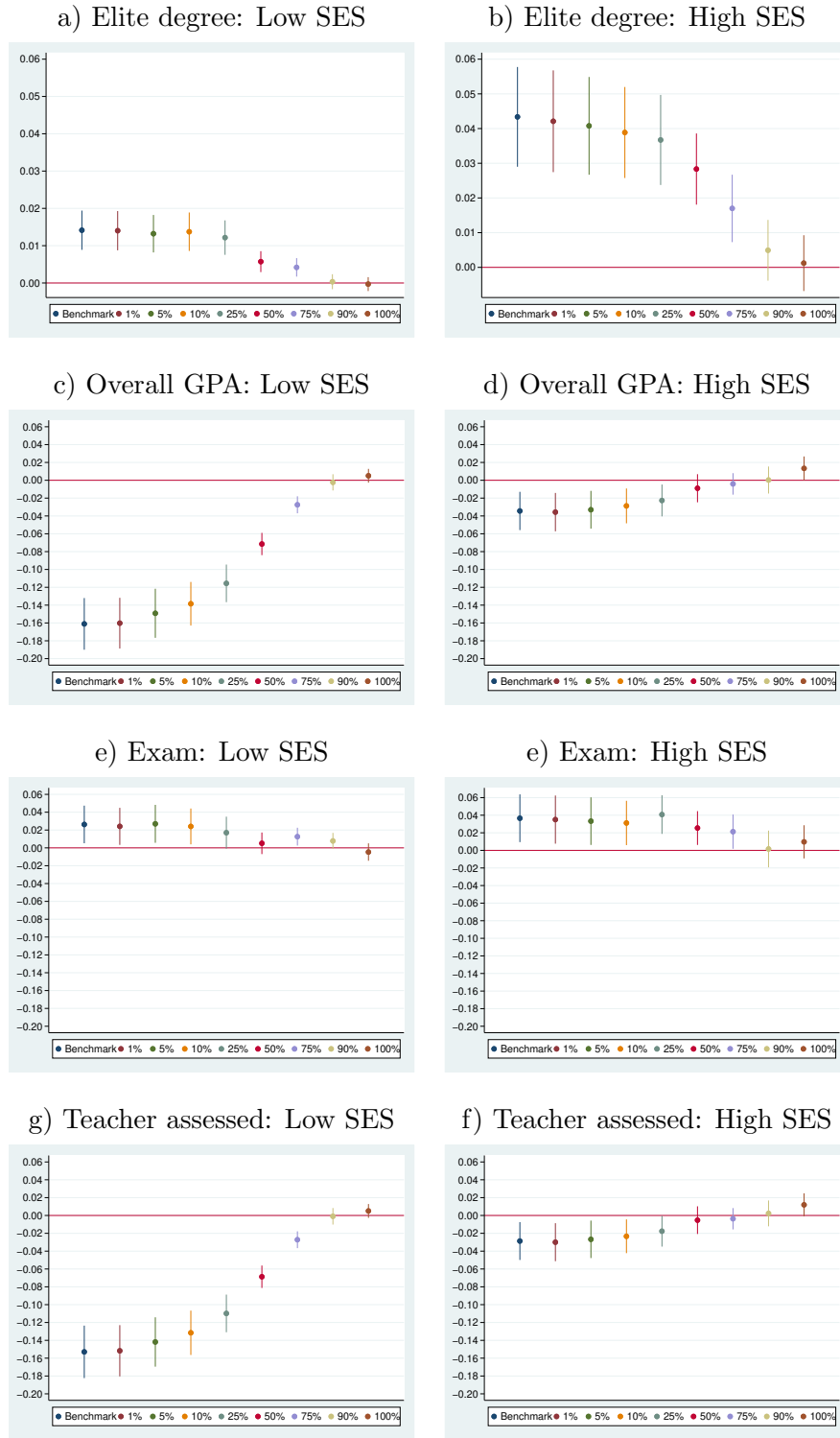
Notes: This graph plots the marginal effect (and 95% confidence intervals) of exposure to elite peers on elite degree enrolment, in the benchmark specification which is augmented by including an interaction between the peer variable P_i and a quadratic in the parent household income percentile rank. 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 A6: Placebo: adding lead and lags of peer exposure to the benchmark regression



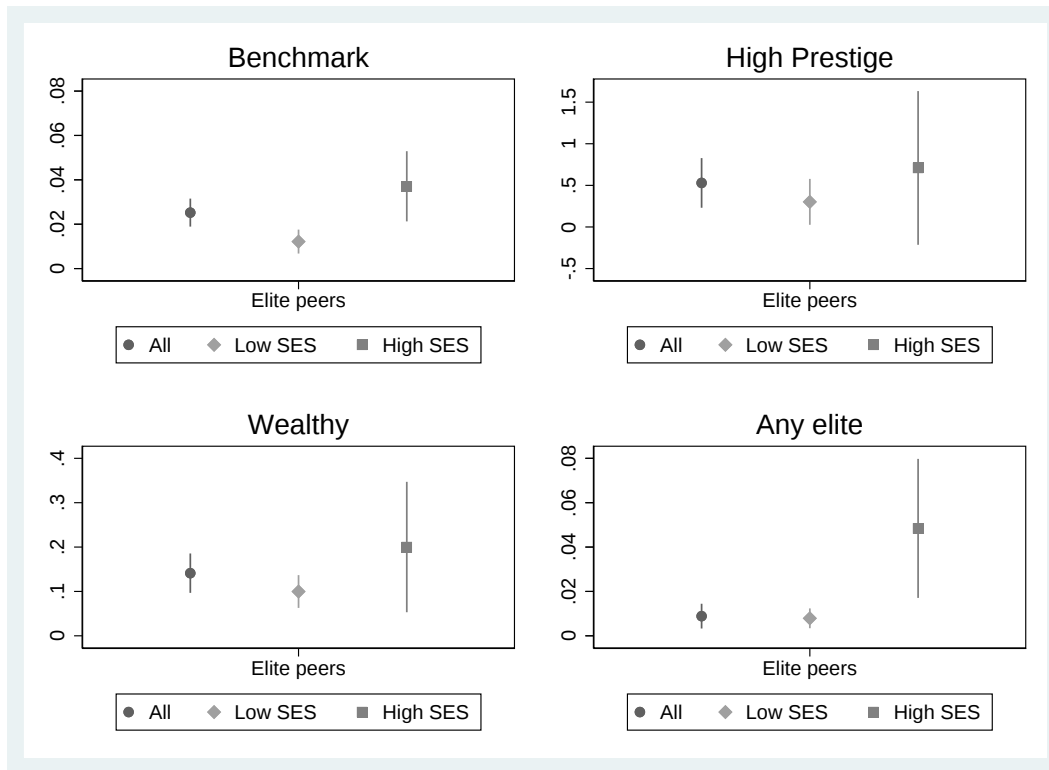
Notes: The figure plots the coefficients from columns (3) and (4) from [Table A7](#). These regressions augment the benchmark regression by including additionally two leads and two lags in the peer exposure variable. Specifically for each student, a lead (lag) is measured as the proportion of parents with an elite degree calculated across the students entering high school in the year after (before) the student.

Figure A7: Measurement error



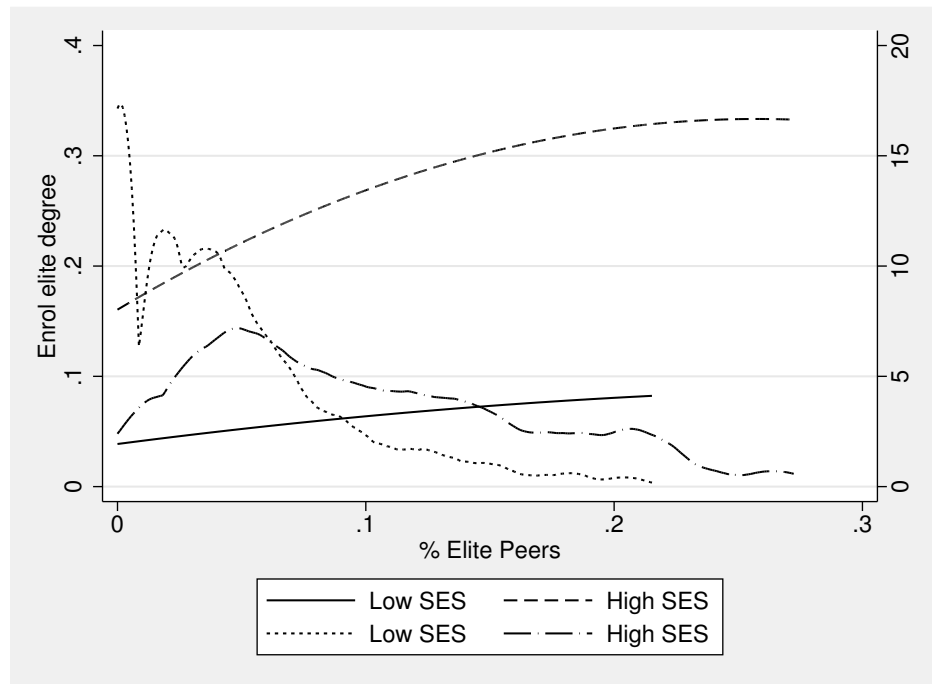
Notes: measurement error is sequentially added to the peer mean variable, at the rate of 1%, 5%, 10%, 25%, 50%, 75%, 90% and 100%. For each rate (e.g. 1%), a 1% sub-sample is chosen to be assigned error. Among the error sample, we randomly assign 2% of the sample to have both parents with an elite degree; 9% to have one parent with an elite degree and the remaining to have no parents with an elite degree, where these values represent the distribution in the benchmark sample. The new peer mean variables are calculated and the benchmark equation is re-estimated.

Figure A8: Estimated peer effect coefficients: types of Elite



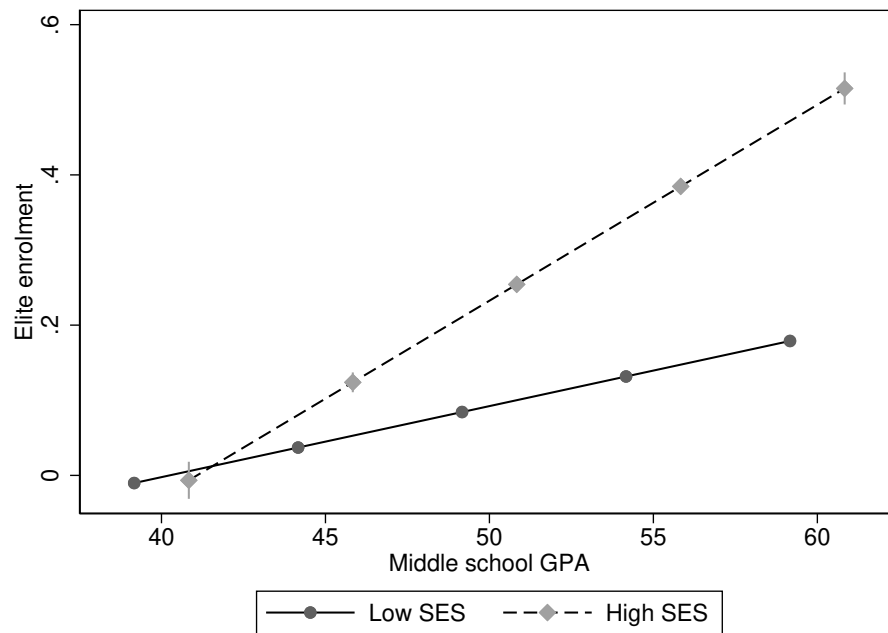
Notes: The figure plots the coefficients from analysis reporting the effect of the % of elite prestige peers in high school (panel b), the % of elite earning peers (panel c) and an indicator for being exposed to any elite educated peers (panel d). Each panel reports the coefficient on the elite peer variable across the total sample, low and high SES students.

Figure A9: Marginal effect of exposure to elite social networks implied from quadratic specification



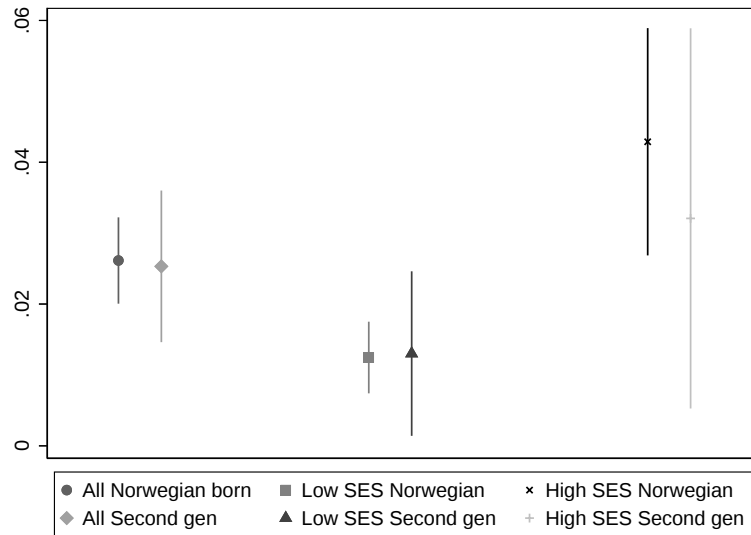
Notes: This graph plots the densities of P_{ics} in the low SES (dotted line) and high SES samples (dot-dashed line). It also plots the marginal effect of an increase in P_{ics} on the probability of enrolling in an elite degree as a function of P_{ics} as implied by estimates of β_{11} and β_{12} in the equation of footnote 28. 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 (6) of [Table A4](#).

Figure A10: Marginal effect of exposure to elite social networks: allowing for interaction between peer effect and middle school GPA



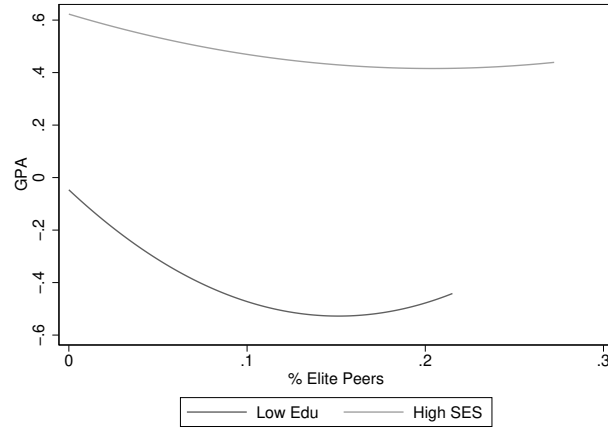
Notes: This graph plots the marginal effect of an increase in P_{-ics} on the probability of enrolling in an elite degree as a function of middle school GPA. The predictions are based on the benchmark specification regression model augmented additionally with the interaction between GPA and P_{-ics}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

Figure A11: Estimated peer effect coefficients heterogeneity analysis by parents' immigration status



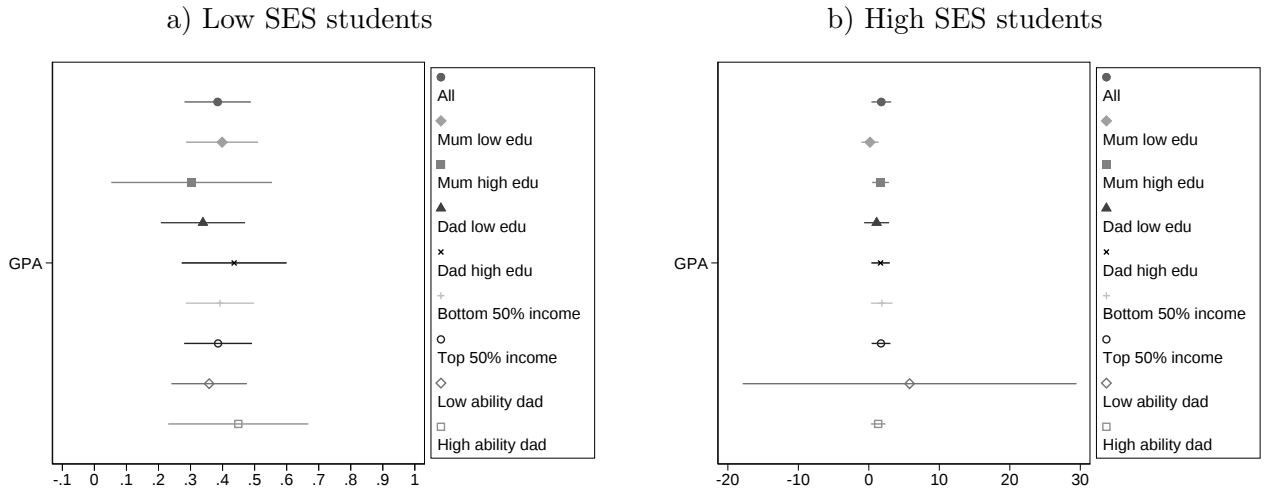
Notes: The figure plots the coefficient estimates from equation (1) allowing for heterogeneity in the effect of the proportion of elite educated peers in high school across Norwegian born students and Second-generation immigrant students.

Figure A12: Marginal effect of exposure to elite social networks implied from quadratic specification on high school GPA



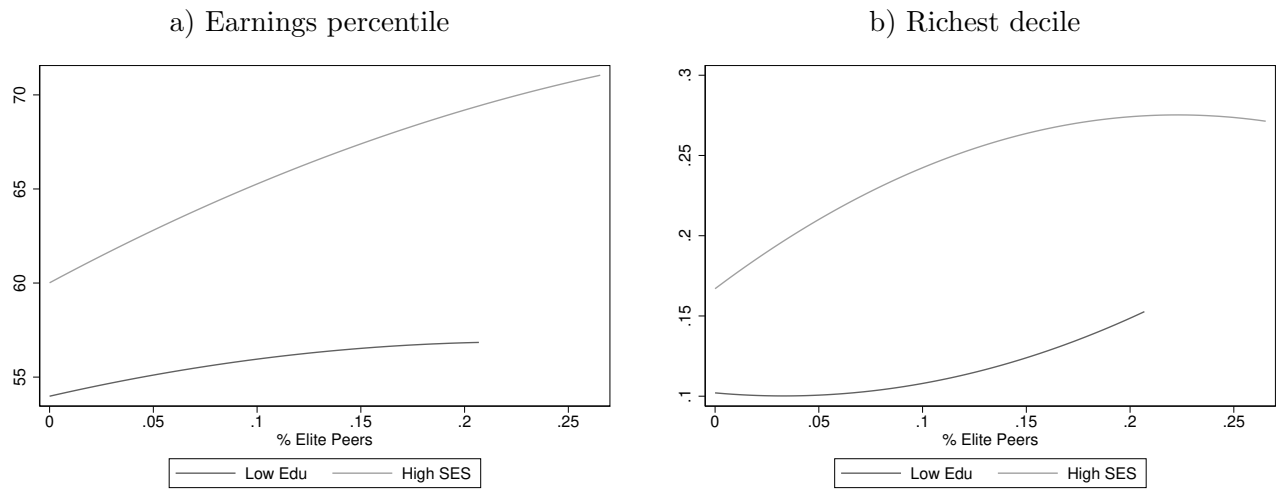
Notes: This graph plots the marginal effect of an increase in P_{ics} on the GPA as a function of P_{ics} as implied by estimates of β_{11} and β_{12} in the equation of footnote 28 but with the dependent variable replaced with the overall high school GPA. The marginal effect in the low SES (high SES) sample is plotted as a dark (lighter) line. The estimates of these coefficients are reported in Column (6) of [Table A10](#).

Figure A13: Parameter heterogeneity in effect of GPA on elite enrolment



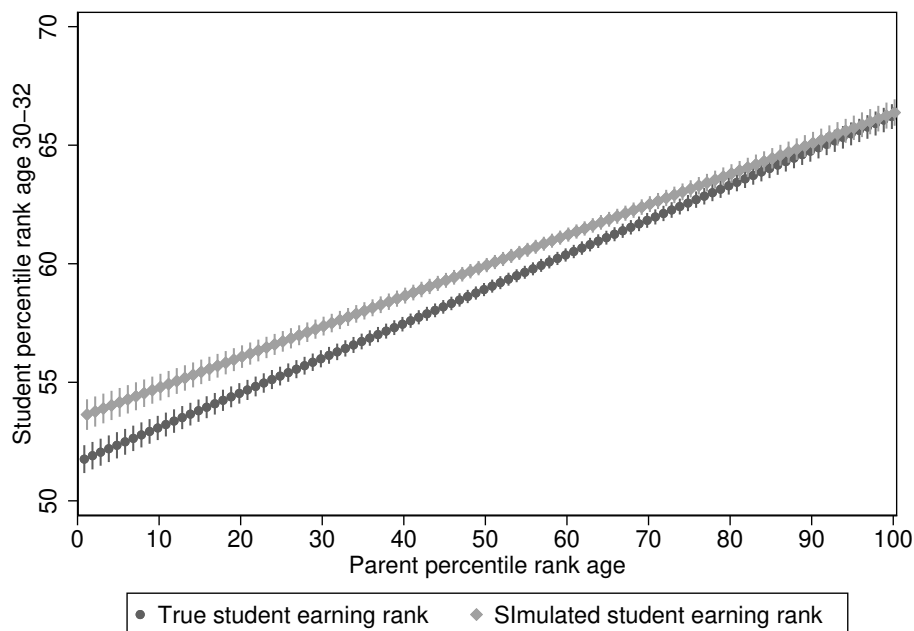
Notes: The figure plots the coefficients γ_2 from equation (5), estimated for the total sample and across sub-groups defined by the mother or father having a low (less than degree) and high (degree +) education level, earning in the bottom or top half of the income distribution or having a low or high ability father (defined as having a father whose IQ test in the bottom or top half of the distribution), across the low and high SES samples.

Figure A14: 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 β_{11} and β_{12} in the equation of footnote 28 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 A14](#).

Figure A15: 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 A5](#) is represented in column (5) of [Table A16](#) 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 5 given the student’s new exposure to elite peers and (for those reassigned students) new school fixed effect.

A7 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*** (0.003)	0.013*** (0.003)	0.040*** (0.008)
Student is a female	-0.073*** (0.003)	-0.053*** (0.003)	-0.125*** (0.007)
Student is born in Norway	-0.011*** (0.003)	-0.032*** (0.004)	0.013 (0.009)
Student's middle school GPA (std)	0.132*** (0.004)	0.086*** (0.003)	0.255*** (0.007)
Proportion of student's own parent with an elite degree	0.182*** (0.007)		0.162*** (0.021)
Student's parents are in top income decile	0.027*** (0.003)	0.004 (0.005)	0.042*** (0.009)
<i>Mother's highest education level (ref = compulsory level)</i>			
High school	0.015*** (0.003)	0.007 (0.005)	0.032*** (0.011)
University	0.006** (0.002)		0.006 (0.010)
<i>Father's highest education level (ref = compulsory level)</i>			
High school	0.018*** (0.002)	0.018*** (0.004)	0.008 (0.018)
University	0.020*** (0.002)		0.022** (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*** (0.002)	7.2%	-0.010*** (0.003)	4.8%
Student's middle school GPA	-0.050*** (0.001)	24.2%	-0.140*** (0.005)	67.6%
Fraction of own elite parent	-0.116*** (0.011)	56.0%	0.022*** (0.003)	-10.6%
<i>Mother's highest education level (ref = compulsory level)</i>				
High school	-0.001*** (0.000)	-0.5%	-0.003** (0.001)	1.4%
University	-0.013*** (0.005)	6.3%	0.007** (0.003)	-3.4%
<i>Father's highest education level (ref = compulsory level)</i>				
High school	0.000*** (0.000)	0.0%	0.001 (0.001)	-0.5%
University	-0.038*** (0.008)	18.4%	0.020*** (0.006)	-9.7%

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: $\Delta(X) = \beta_X^p(E^H(X_i) - E^L(X_i))$ where β_X^p 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: $\Omega(X) = (\beta_X^H - \beta_X^L)E^p(X_i)$ where β_X^g 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: Application of [Borusyak and Hull \(2023\)](#)

	(1)	(2)	(3)
	Total Sample	Low SES	High SES
A) Benchmark including essential controls			
% parents with elite education	0.035***	0.018***	0.048***
	(0.003)	(0.003)	(0.008)
N	177,219	58,328	20,018
B) Permute within schools and deciles of middle school GPA			
% parents with elite education	0.032***	0.019***	0.039***
	(0.001)	(0.002)	(0.003)
Expected treatment	0.137***	0.095***	0.083***
	(0.005)	(0.009)	(0.011)
N	165,910	53,329	19,516
C) Benchmark on the smaller permuted sample			
% parents with elite education	0.026***	0.011***	0.042***
	(0.003)	(0.003)	(0.008)
N	165,910	53,329	19,516

Notes: Panel A reports estimates of a regression of elite degree enrolment on peer elite exposure and essential controls of school, cohort fixed effects and middle school GPA as regressors. Panel B) reports the coefficient on the elite peer variable and “expected” treatment with no additional controls. Panel C) reproduces the benchmark estimates in column (1) of Table 2 but on a smaller sample comparable to Panel B. Regressions are weighted by school size. Standard errors clustered at the school level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A4: 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
A - All students					
Proportion of parents with elite degree (std)	0.026*** (0.003)	0.027*** (0.003)		0.022*** (0.004)	0.029*** (0.004)
<i>Number of pupils</i>	<i>177,219</i>	<i>177,219</i>		<i>83,465</i>	<i>111,309</i>
<i>Number of schools</i>	<i>556</i>	<i>556</i>		<i>313</i>	<i>400</i>
B - Low SES students sample					
Proportion of parents with elite degree (std)	0.013*** (0.003)	0.014*** (0.003)		0.010** (0.004)	0.015*** (0.004)
<i>Number of pupils</i>	<i>58,328</i>	<i>58,328</i>		<i>28,181</i>	<i>37,374</i>
<i>Number of schools</i>	<i>524</i>	<i>524</i>		<i>284</i>	<i>391</i>
C - High SES students sample					
Proportion of parents with elite degree (std)	0.040*** (0.008)	0.047*** (0.008)		0.038*** (0.013)	0.051*** (0.009)
<i>Number of pupils</i>	<i>20,018</i>	<i>20,018</i>		<i>8,420</i>	<i>12,679</i>
<i>Number of schools</i>	<i>459</i>	<i>459</i>		<i>240</i>	<i>347</i>
D - Low and High SES students sample					
Proportion of parents with elite degree (std)			0.050*** (0.004)		
Indicator for low SES			(0.014)		
Proportion of parents with elite degree * low			-0.031*** (0.003)		
<i>Number of pupils</i>			<i>78,540</i>		

Notes: The full sample, low SES, high SES and pooled low and high SES samples in panels A, B, C and D. Column (1) is benchmark specification (equation 1 and Table 2). Column (2) controls for school-specific linear trends. Column (3) includes fully interacted fixed effects for the school and cohort estimated on pooled low and high SES samples. Column (4) is the benchmark estimated on subsample of schools where variation in elite peers evolves over time in a random way. Specifically, we drop schools where the R^2 from a school-level regression of the % of elite 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 [subsection A3.3](#). Column (5) augments the benchmark with average traits of teachers within schools across cohorts: % of females, % from a professional or low skilled background and average age. Regressions are weighted by school size. Standard errors clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Table A5: Placebo tests - Effect of elite peers on child birth outcomes

	(1) Estimate (p-value)	(2) No. students	(3) No. schools
<i>Birth outcomes:</i>			
Birth weight	-3.011 (0.386)	169,864	554
Low birth weight	-0.000 (0.723)	177,219	556
Gestation	-0.011 (0.383)	157,669	552
Height	-0.010 (0.496)	164,073	551
Head circumference	0.005 (0.585)	167,949	553
Congenital malformation	-0.000 (0.999)	170,133	554
Severe deformity	-0.001 (0.340)	170,133	554
<i>Middle school subject choice:</i>			
Maths	0.003 (0.512)	177,219	556
English	-0.000 (0.954)	177,219	556
Norwegian	-0.001 (0.726)	177,219	556
Other	0.001 (0.243)	177,219	556
<i>Parent income during middle school:</i>			
Mother income	-0.006 (0.851)	175,862	555
Father income	0.014** (0.032)	173,666	556

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 A6: Benchmark regression controlling for fathers' middle school income

	(1)	(2)	(3)
	Total sample	Low SES	High SES
Parents w/elite degree	0.026*** (0.003)	0.013*** (0.003)	0.040*** (0.008)
Father's income	-0.001* (0.001)	0.000 (0.002)	0.000 (0.002)
Observations	<i>173,666</i>	<i>56,708</i>	<i>19,753</i>
Number of schools	556	524	457

Notes: OLS estimates of the benchmark model including additional variable of the father's income measured in middle school year. *** p<0.01, ** p<0.05, * p<0.1

Table A7: Placebo tests: lags and leads of exposure to elite peers

	(1)	(2)	(3)	(4)	(5)	(6)
	Total	Low SES	High SES	Total	Low SES	High SES
% Parents with elite degree	0.027*** (0.004)	0.014*** (0.003)	0.042*** (0.009)	0.027*** (0.004)	0.013*** (0.003)	0.042*** (0.011)
Elite exposure lag 1	-0.001 (0.004)	0.001 (0.005)	-0.012 (0.012)	-0.003 (0.005)	0.003 (0.007)	-0.007 (0.018)
Elite exposure lead 1	-0.001 (0.005)	-0.002 (0.005)	-0.005 (0.012)	0.004 (0.006)	0.002 (0.007)	-0.001 (0.014)
Elite exposure lag 2				0.009 (0.005)	0.003 (0.007)	0.014 (0.012)
Elite exposure lead 2				-0.001 (0.006)	0.002 (0.006)	-0.002 (0.015)
Observations	<i>137,286</i>	<i>44,783</i>	<i>15,647</i>	<i>90,365</i>	<i>29,493</i>	<i>10,420</i>
Number of schools		443	406	372	370	342

Notes: These regressions augment the benchmark regression by including additionally two leads and two lags in the peer exposure variable. Specifically for each student, a lead (lag) is measured as the proportion of parents with an elite degree calculated across the students entering high school in the year after (before) the student.

Table A8: Sensitivity analysis and interpretation

	(1) Benchmark	(2) First born children	(3) Two-parent families	(4) Exclude OSLO	(5) Exclude small schools	(6) Drop humanities general studies
A - All students						
Proportion of parents with elite degree (std)	0.026*** (0.003)	0.025*** (0.003)	0.026*** (0.003)	0.025*** (0.004)	0.026*** (0.003)	0.051*** (0.005)
<i>Number of pupils</i>	177,219	146,567	149,613	159,307	159,307	100,454
<i>Number of schools</i>	556	555	542	507	280	539
B - Low SES students sample						
Proportion of parents with elite degree (std)	0.013*** (0.003)	0.014*** (0.003)	0.014*** (0.003)	0.013*** (0.003)	0.013*** (0.003)	0.029*** (0.004)
<i>Number of pupils</i>	58,328	51,036	48,428	52,725	50,615	34,881
<i>Number of schools</i>	524	524	518	482	280	513
C - High SES students sample						
Proportion of parents with elite degree (std)	0.040*** (0.008)	0.041*** (0.008)	0.034*** (0.008)	0.040*** (0.009)	0.040*** (0.008)	0.065*** (0.011)
<i>Number of pupils</i>	20,018	15,439	17,415	16,444	19,153	10,236
<i>Number of schools</i>	459	449	450	418	279	433

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 full sample (Panel A), the low SES sample (Panel B) and in the high SES sample (Panel C). 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 A9: Elite peer effect on GPA rank within the high school cohort

	(1) Full sample	(2) Low SES	(3) High SES
A - Overall GPA			
	-8.147*** (0.431)	-10.071*** (0.553)	-5.604*** (0.462)
<i>Number of observations</i>	174,204	56,928	19,891
B - Components of GPA			
Externally assessed written exam grades	-6.026*** (0.338)	-7.036*** (0.435)	-4.469*** (0.480)
<i>Number of observations</i>	173,789	56,728	19,880
Teacher-assessed internal grades	-8.159*** (0.435)	-9.986*** (0.551)	-5.643*** (0.469)
<i>Number of observations</i>	174,198	56,932	19,888
Semi-externally assessed oral exam grades	-4.163*** (0.354)	-5.913*** (0.469)	-2.526*** (0.526)
<i>Number of observations</i>	145,648	47,390	17,022

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 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 rank on externally assessed written exams across all three years of high school (row 2), average rank on teacher assessed grades across all three years of high school (row 3), and average rank on oral exams marked by an external examiner and the student's teachers across all three years of high school (row 4). 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 A10: 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 specification
A - Low SES students sample						
Proportion of parents with elite degree (std)	-0.179*** (0.016)	-0.178*** (0.016)		-0.199*** (0.027)	-0.174*** (0.018)	-0.215*** (0.014)
Proportion of parents with elite degree squared						0.068*** (0.009)
<i>Number of pupils</i>	58,328	58,328		28,181	37,374	58,328
<i>Number of schools</i>	524	524		284	391	524
B - High SES students sample						
Proportion of parents with elite degree (std)	-0.046*** (0.012)	-0.053*** (0.014)		-0.080*** (0.017)	-0.031*** (0.013)	-0.081*** (0.019)
Proportion of parents with elite degree squared						0.016*** (0.005)
<i>Number of pupils</i>	20,018	20,018		8,420	12,679	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)			
<i>Number of pupils</i>			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) is 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 is 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 [subsection A3.3](#). 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 A11: Balance. Dependent variable is the Instrumental Variable.

	(1) Low SES	(2) High SES
Proportion of parents with elite degree	-0.005 (0.006)	0.010* (0.006)
Student is female	-0.005 (0.004)	0.006 (0.004)
Student is born in Norway	-0.006 (0.005)	-0.001 (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	-0.024 (0.062)	-0.064 (0.101)
Middle school written exams	0.001 (0.006)	-0.003 (0.009)
Middle school oral	0.000 (0.005)	-0.003 (0.008)
Middle school overall GPA	0.005 (0.072)	0.061 (0.116)
Proportion of student's own parents with an elite degree		0.012 (0.013)
Student's parents are in top decile	0.007 (0.006)	-0.018** (0.007)
Number of pupils	52,190	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 year 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 A12: Long-term implications for earnings

		(1)	(2)			(3)	(4)
		Low SES	High SES			Low SES	High SES
	A - Mincer regressions				B - Peer effect on earnings		
Top 1%	Student has a degree	0.007*** (0.002)	0.034*** (0.011)	Parents w/ elite degree	-0.003 (0.002)	0.010** (0.005)	
	Student has elite degree	0.052*** (0.004)	0.062*** (0.012)				
Top 10%	Student has a degree	0.029*** (0.005)	0.103*** (0.024)	Parents w/ elite degree	0.006 (0.004)	0.038*** (0.009)	
	Student has elite degree	0.237*** (0.009)	0.305*** (0.023)				
Top 25%	Student has a degree	0.081*** (0.007)	0.143*** (0.025)	Parents w/ elite degree	0.022** (0.007)	0.047*** (0.011)	
	Student has elite degree	0.427*** (0.013)	0.436*** (0.027)				
Top 50%	Student has a degree	0.207*** (0.008)	0.211*** (0.024)	Parents w/ elite degree	0.024** (0.007)	0.043*** (0.010)	
	Student has elite degree	0.438*** (0.015)	0.413*** (0.025)				
Observations		24,890	8,196		24,890	8,196	
No. schools		487	406		487	406	

Notes: Panel A runs 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. Panel B estimates the effect of exposure to elite peers during high school on earnings. The earnings measure as dependent variable changes across rows, from an indicator for earning in the top percentile, top decile, top quartile and top half of the income distribution. 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 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

Table A13: Mincer equation estimating correlation between elite degree programme and earnings

	(1) Low SES Earnings percentile	(2) High SES	(3) Low SES Richest decile	(4) High SES	(5) Low SES Richest percentile	(6) High SES
Student ever enrolled in degree	9.633*** (0.432)	8.581*** (1.305)	0.025*** (0.005)	0.069*** (0.020)	0.006*** (0.002)	0.032*** (0.010)
Student enrolled in elite degree: STEM	26.950*** (0.929)	22.884*** (1.424)	0.213*** (0.011)	0.205*** (0.021)	0.034*** (0.005)	0.019*** (0.011)
Law	26.010*** (1.337)	22.096*** (1.771)	0.148*** (0.016)	0.187*** (0.027)	0.038*** (0.007)	0.062*** (0.014)
Medicine	40.373*** (2.388)	35.268*** (1.884)	0.675*** (0.028)	0.606*** (0.028)	0.233*** (0.013)	0.154*** (0.015)
Observations	24,890	8,196	24,890	8,196	24,890	8,196

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 A14: Validity of the empirical strategy: dependent variable is earnings in richest decile

	(1) Benchmark	(2) Teacher traits	(3) School- specific linear trends	(4) 'Drop if more than random'	(5) Quadratic specifica- tion
A - Low SES students sample					
Proportion of parents with elite degree (std)	-0.003 (0.004)	0.090 (0.005)	0.007 (0.005)	0.009 (0.005)	0.001 (0.004)
Proportion of parents with elite degree squared					0.001 (0.003)
<i>Number of pupils</i>	<i>24,890</i>	<i>21,192</i>	<i>24,890</i>	<i>12,022</i>	<i>24,890</i>
<i>Number of schools</i>	<i>487</i>	<i>487</i>	<i>487</i>	<i>236</i>	<i>487</i>
B - High SES students sample					
Proportion of parents with elite degree (std)	0.012*** (0.008)	0.044 (0.011)	0.046** (0.011)	0.049** (0.019)	0.048** (0.014)
Proportion of parents with elite degree squared					-0.001 (0.006)
<i>Number of pupils</i>	<i>8,196</i> <i>406</i>	<i>7,141</i> <i>406</i>	<i>8,196</i> <i>406</i>	<i>3,520</i> <i>178</i>	<i>8,196</i> <i>406</i>

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 5. 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 A15: Mincer regression dropping dummy for having a degree

	(1)	(2)	(3)	(4)
	Low SES	High SES	Low SES	High SES
Elite degree	19.288** (0.699)	17.518*** (0.658)		
STEM			18.293*** (0.871)	15.322*** (0.779)
Law			17.164*** (1.304)	14.425*** (1.295)
Medicine			31.995*** (2.393)	28.008*** (1.499)
Observations	24,890	8,196	24,890	8,196

Notes: Mincer-style regressions of earnings percentile age 30-32 on indicators for an elite degree (columns 1-2) and an elite STEM degree, an elite law degree and an elite medicine degree (columns 2-3). The omitted category is no elite degree. The low SES sample in columns (1), (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), (4) 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. 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 A16: Simulating a reassignment of low (high) SES students into schools with a high (low) level of elite peers

	(1)	(2)	(3)	(4)	(5)
	Benchmark	Simulation 1	Simulation 2	Simulation 3	Simulation 4
No. Schools		Top and bottom decile		All schools	
No. children moved per school		1	5	1	5
Parent Percentile rank	0.146*** (0.004)	0.143*** (0.004)	0.143*** (0.004)	0.140*** (0.004)	0.139*** (0.004)
Constant	50.812*** (0.284)	50.764*** (0.284)	50.822*** (0.284)	50.860*** (0.284)	51.104*** (0.284)
Observations	44,394	44,394	44,394	44,394	44,394
R-squared	0.023	0.023	0.022	0.022	0.021

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 A5](#) 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 A5](#).

References

- Angrist, J. D. (2014). The perils of peer effects. *Labour Economics* 30, 98–108.
- Black, S., P. Devereux, and K. G. Salvanes (2005). 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.
- Borusyak, K. and P. Hull (2023). Nonrandom exposure to exogenous shocks. *Econometrica* 91(6), 2155–2185.
- Carrell, S. E., M. Hoekstra, and E. Kuka (2018). The long-run effects of disruptive peers. *American Economic Review* 108(11), 3377–3415.
- 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.
- Deming, D. J. and K. L. Noray (2018). Stem careers and the changing skill requirements of work. Technical report, National Bureau of Economic Research.
- 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.
- Hoxby, C. (2000). Peer effects in the classroom: Learning from gender and race variation. Technical report, National Bureau of Economic Research.
- 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.
- Tincani, M. (2017). Heterogeneous peer effects in the classroom. HCEO Working Paper 2017-006.