11-2019

Peer effects in education: When beliefs matter

Luca FACCHINELLO
Singapore Management University, lfacchinello@smu.edu.sg

Peer Effects in Education: When Beliefs Matter

Luca Facchinello*

Abstract

Recent literature explains the puzzling finding of zero or negative peer effects in academic achievement assuming that better peers negatively affect beliefs about ability (self-concept), motivation or peer interactions. This paper provides new evidence on such negative mechanisms, and on their impact on educational choices and attainment for students randomly assigned in compulsory school to classes with different cognitive ability. Using detailed longitudinal data on a nationally representative sample of Swedish compulsory school students, I find that students exposed to higher ability peers systematically underestimate their ability and are less likely to choose advanced subjects throughout compulsory school. While these students perform better, as measured by national test scores, they are assigned lower grades in subjects lacking national test scores, suggesting distortions in teachers' assessment of student performance. Negative effects persist after compulsory school: students exposed to better peers have lower well-being and GPA in high school. I find substantial heterogeneity in treatment effects. Students who interact with better peers and receive early grades suffer more severe grade distortions, but exhibit stronger positive performance spillovers, better sort into non-compulsory education, and attain more education with respect to students lacking early grades. Negative peer effects in self-concept and grades are concentrated among disadvantaged students, who also receive lower parental support when exposed to better peers. This paper shows that class composition can distort students' grades, self-concept and choices, and highlights the limits of assessing peer effects on test scores alone.

JEL codes: I21; I28; J24.

Keywords: Peer effects, Ability, Self-concept, Grades, Motivation, Education Choices, Educational Attainment, Long-run Effects.

---

*Affiliation: School of Economics, Singapore Management University. Address: 90 Stamford Road Singapore, 178903. Phone: +65 6828 0876. Email: lfacchinello@smu.edu.sg

Acknowledgements: Special thanks to Erik Lindqvist, who provided excellent feedback on the paper. I thank Tomoki Fujii, Seonghoon Kim, Jie Gong, Nicolas Jacquet, Juanna Joensen, Peter Nilsson, Björn Öckert, Sunha Myong, and Ulf Zölitz for their valuable comments. Lastly I thank seminar participants at SSE, SMU, SER, SOLE, and AASLE for their useful feedback. The usual disclaimers apply.

1 Introduction

Most of the peer effects literature in education focuses on how peers affect individual educational performance, and typically finds either positive, zero or non-linear effects.[1] A much smaller literature considers how exposure to better (or worse) peers affects long-run outcomes, similarly finding mixed results.[2] Zero and especially negative peer effects in educational performance or attainment are puzzling, and have been explained in different ways. Bui et al. (2014), Pop-Eleches et al. (2016), and Antecol et al. (2016) explain, to different extents, their results in the light of the frame of reference model, known in the psychology literature (Marsh [1987]) as the big-fish-in-a-little-pond effect. The basic idea is that students assigned to better classes form downward-biased beliefs about ability when they assess academic potential in relation to their peers. A different explanation is that underperforming with respect to peers leads to negative peer interactions or hurt school motivation and confidence. Pop-Eleches et al. (2016) finds evidence in favor of negative peer interactions, while negative effects on motivation and confidence are often used to explain the mixed results found in the selective schools and school integration literatures.[3] Similar mechanisms emerge explicitly in a small literature that finds that ranking higher in school positively affects academic choices and performance. Elsner and Ipshording (2017) provide evidence that rank affects perceived ability, while Murphy and Weinhardt (2018) find instead evidence in favor of higher rank improving confidence.[4]

This discussion raises two points. First, to better understand how students are affected by their peers we need clear evidence on intermediate channels like students' beliefs, motivation, and peer interactions. Second, if these negative channels are important, assessing peer effects on educational performance alone might be misleading. Exploiting rich administrative and survey data, I provide in this paper new evidence on such negative mechanisms, and on their impact on educational choices and attainment for Swedish compulsory school students

[1] Hanushek et al. (2003), Ammermueller and Pischke (2009), Imberman et al. (2012) and Pop-Eleches et al. (2016) find positive peer effects on student performance in primary or secondary school. Sund (2009), Lavy et al. (2012a), Burke and Sass (2013), Abdulkadiroglu et al. (2014), Bui et al. (2014) and Tincani (2017) find instead non-linear or zero effects. Antecol et al. (2016) finds negative effects in primary school. Booij et al. (2017) and Feld and Zölitz (2017) find non-linear peer effects in academic performance in college. See Sacerdote (2014) for a recent review.

[2] Cullen et al. (2006) finds no effects, Black et al. (2013) finds positive effects for specific categories and peer characteristics, while Carrell et al. (2018) find that disruptive peers negatively affect students' earnings.

[3] Angrist and Lang (2004), Cullen et al. (2006), Kling et al. (2007), Abdulkadiroglu et al. (2014).

[4] Notice that rank effects are estimated keeping peer performance fixed. Results from this literature are therefore not directly comparable to those in the peer effects literature.

randomly assigned to classes with different levels of cognitive ability.

Swedish education features progressive tracking: students choose whether to take general or advanced courses in math and English in late compulsory school (grades 7-9). More challenging advanced courses prepare for academic high school tracks, the only pathway to college. Swedish compulsory school students are assigned homework and test scores, rather than grades, during the academic year. Nationally norm-referenced grades, anchored to national tests in math, English and Swedish, are assigned at the end of each year in late compulsory school, and in earlier stages for some students. Given the scarce grading feedbacks and tracked education system, class comparisons might have played a big role in shaping students' beliefs about ability and education choices.

I use data from the ETF longitudinal study, which follows up until age 35 a nationally representative sample of students born 1967. Full classes are sampled in grade 6, when students take a battery of IQ tests. Students fill in detailed surveys in grades 6 and 10, irrespective of high school attendance. Detailed administrative records report their test scores, grades, high school choices, and educational attainment up to age 35.

To identify peer effects I exploit within-school variation in class ability, assuming that students are randomly assigned to classes and teachers. I provide three main arguments supporting the validity of this assumption. First, in the period I consider school administrators lacked early measures of academic performance when forming classes: Swedish pre-school was not part of the Swedish education system, did not involve formal schooling, and was generally attended only for a few hours a day. Similarly, administrators were unlikely to use early grades, phased out in the same period, to assign students to existing classes or to reassign students to existing classes due to class merges. Second, I show that, within school, peer ability is unrelated to student ability, individual background (including foreign status, SES, and detailed parental occupation), class size and teacher seniority. Finally, the level of within-school variation in class ability I observe is remarkably close to the level of variation generated by randomly reassigning students to classes in a Montecarlo-like simulation.

I find that students assigned to better peers systematically underestimate own ability and are less likely to choose advanced subjects during compulsory school. A one standard deviation increase in peer mean ability leads to a 3.6% standard deviation decrease in cognitive self-concept during grades 6-9, no effect on non-cognitive self-concept, and a 1.7 percentage points decrease in the probability of choosing advanced math in grades 7, 8 and 9. While students who interact with better peers perform better, as measured by national test scores unobserved by students, they are assigned lower grades in subjects lacking national test scores, suggesting distortions in teachers' assessment of student performance. A one standard deviation increase in peer mean ability leads to a 2.3% standard deviation increase in test scores, no effect on grades in nationally tested subjects, and a 2.6% standard deviation decrease in grades in untested subjects. Negative effects persist after compulsory school: students who attend high school and are assigned during compulsory school to better classes have 3% standard deviations lower school well-being and GPA. I do not find long-run peer effects on education choices and attainment for the average student.

I find substantial heterogeneity in peer effects by grade assignment and SES. Early grade assignment does not correct negative distortions in beliefs, and surprisingly worsens the grade distortions that students exposed to better peers suffer from. Students who receive (downward-biased) grades early on exhibit stronger positive performance spillovers, due to increased effort and parental support, when exposed to better peers. A one standard deviation increase in peer mean ability leads to a 8.8% standard deviation increase in English test scores, a 4.3% standard deviation increase in motivation, and respectively 5% and 2.1% higher likelihood to engage frequently in homework and receive help at home in grade 6 for early graded students. I find no effect for late graded students. Early graded students who are assigned to better classes improve their sorting to education: they are 2.2 percentage points less likely to go to high school, but 2.8 percentage points more likely to enroll in academic tracks and attain university education than their late graded counterparts. This leads early graded students to attain 0.12 additional years of non-compulsory education with respect to late graded students: the unconditional zero effect I find on educational attainment masks positive effects for early graded students, and negative effects for late graded students. This evidence is overall consistent with students positively reacting to distorted information about own ability. The negative effects I find in self-concept and grades are all concentrated among low SES students, defined as the children of working class parents. When exposed to better peers, low SES students are about 2.5 percentage points less likely to take advanced English and receive support at home during late compulsory school.

This paper contributes to several literatures. First, it contributes to the broad peer effects literature. The new evidence I provide on negative peer effects in self-concept, grading and choices can help to understand the mixed results found in the literature. While negative effects on self-concept have already been proposed as an explanation in papers finding zero or negative peer effects, my paper shows direct evidence on this channel. The grading channel has instead up to now been ignored in the literature, which mostly focuses on test scores. Second, the paper contributes to a small literature establishing how peers affect students in the long-run, and thus to the broader literature on the long-run effects of education interventions and inputs.[5] Third, the paper contributes to the tracking literature.[6] I show that interacting with better peers in a setup with progressive tracking have positive effects on performance, but negative effects on academic choices that carry over beyond compulsory school. Fourth, the paper contributes to the literature on grade feedbacks and belief updating (Zafar [2011], Stinebrickner and Stinebrickner [2012], Facchinello [2016]). I show that exposure to better peers leads to grading distortions when teachers lack objective measures of academic performance, and that students' choices strongly respond to updated beliefs.

Consistently with the core literature on peer effects in educational performance, reviewed by Sacerdote (2014), I find positive peer effects on academic performance. I do not find weaker positive spillovers in performance among low SES students, who have lower cognitive ability. This is at odds with results from the selective schools literature, which however considers very different institutional setups. Similarly to Cullen et al. (2006), I do not find evidence of long-run peer effects. I however find both positive and negative effects on educational attainment for specific categories of students, which is qualitatively consistent with the findings of Black et al. (2013). Most closely related to this paper are Bui et al. (2014) and Antecol et al. (2016), who find respectively zero and negative peer effects in performance and interpret their results in the light of the frame of reference model. My contribution is to explicitly show, rather than assume, the decrease in self-concept. My results are mostly consistent with Pop-Eleches et al. (2016), who find positive peer effects in performance, followed by negative effects on relative self-concept and peer interactions. My contribution with respect to their paper is showing that, on top of relative self-concept, better peers lead to decreases in absolute self concept. This mechanism has not yet received direct support in the economics literature. An important exception is found in the rank literature. Elsner and Ipshording (2017) find that ranking lower in high-school negatively affects beliefs about ability and GPA. Results in Murphy and Weinhardt (2018) do not support the self-concept mechanism, but appear instead to be consistent with a positive effect of rank on confidence. With the caveat mentioned before, my paper can contribute to this discussion. My results are not consistent with the confidence mechanism: I find that better peers do not affect relative assessment

[5] Long-run peer effects are studied by Cullen et al. (2006), Black et al. (2013), Elsner and Ipshording (2017) and Carrell et al. (2018). Chetty et al. (2011), Heckman et al. (2013), Fredriksson et al. (2013), Conti and Heckman (2016) provide evidence on long-run effects of education inputs and interventions.

[6] See for instance Meghir and Palme (2005), Hanushek and Wössmann (2006), Pekkarinen et al. (2009), Malamud and Pop-Eleches (2011), Betts (2011), Malamud and Pop-Eleches (2011), Guyon et al. (2012), Kerr et al. (2013).

of non-cognitive skills (mostly reflecting leadership) and do not decrease school motivation. My results directly support the perceived ability mechanism, and are more clear-cut than those in Elsner and Ipshording (2017), since in their setup lower self-concept might reflect lower GPA among low ranking students. Stinebrickner and Stinebrickner (2012) and Zafar (2011) find that students who get lower than expected grades are respectively more likely to drop out from college and switch to an easier major. I find a similar result in a completely different setup: this shows that findings at the college level carry over at the compulsory school level. My paper adds evidence on how these learning effects propagate over time. Finally, using similar data from the ETF project, Facchinello (2016) finds that short-run education choices of early graded students are more consistent with academic ability than those of late graded students. My paper finds evidence that being assigned to classes with better peers compounds the positive sorting effects of early grade assignment, and increases academic performance and effort.

The most important implication of my findings is that considering performance spillovers alone might provide an incomplete, and potentially misleading, picture of peer effects. Focusing on academic choices or educational attainment, that respond to both positive (performance spillovers) and negative (distortions in grades and beliefs) underlying mechanisms, can substantially improve evaluations of peer effects. The second implication is that the grading system should not be ignored when evaluating peer effects. The common practice of setting grades on a curve is particularly worrying: my results show that even in a setup where students are not graded on a curve class composition generates distorsions in grading, self-concept and choices. An important question is to what extent my findings generalize to other setups, and whether additional or different grading feedbacks can avoid these negative effects. My finding that providing additional early grades worsens distortions in grade assignment and does not prevent negative distortions in self-concept is in this sense not encouraging. Finally, this paper finds that negative effects of exposure to higher ability peers are concentrated among disadvantaged students, and adds thus another argument in favor of ability tracking.

The paper proceeds as follows. Section 2 describes the data and the education system, and presents descriptives. Section 3 discusses outcomes and potential mechanisms in the light of the literature. Section 4 discusses identification and treatment variation. Section 5 presents results, while Section 6 draws conclusions and discusses policy implications.

2 Setup

2.1 Data

I use data from the Swedish longitudinal research program Evaluation Through Follow-up (ETF), which since 1980 surveyed 10% nationally representative samples of classes, following up individual students at least until high school completion. The rich data include cognitive ability test scores, survey and administrative data.[7] I focus on the third wave of the study, a representative sample of classes attending grade 6 in 1980, roughly coinciding with the cohort born 1967. Several features make this wave suitable to answer the research question. First, students were assigned to classes independently of their ability, which allows identification of treatment effects. Second, surveys in this wave focus extensively on self-assessment, the main mechanism I explore in my analysis. Lastly, the data records education attainment until students are 35, allowing to investigate long-run effects.

During the spring term of grade 6, ETF administered standard IQ tests on verbal, mathematical, and spatial ability. Tests were taken at school by 90% of the sample. My measure of cognitive ability averages verbal and mathematical scores, and has high quality. First, it displays 0.78 correlation with the cognitive ability measure available for conscripted men at age 19, widely used in the literature and shown in Lindqvist and Vestman (2011) to be a good predictor for long-run labor market outcomes.[8] Second, since total test time for the IQ tests was short, my cognitive measure has less scope to be confounded by non-cognitive traits.[9] Third, ability is measured at age 13, a point in time when IQ should have stabilized, as discussed in Cunha and Heckman (2009). This reduces measurement error, but begs the question of whether the measure is exogenous to peer composition. This point is important, and is explicitly addressed in Section 4.1.

In grade 6 and 10 (when students are supposed to be in high school), respectively 92.5% and 77% of the sample answered a detailed student survey. The grade 10 survey both includes

[7] Classes were systematically sampled from municipalities (Swedish lowest administrative division) using stratified sampling. Strata were defined by population, share of left-wing voters, share working in the public sector, and share of immigrants. The three biggest municipalities in Sweden (Stockholm, Malmö, Gothenburg) were always part of the sample. See Emanuelsson (1979) for further details on the sampling scheme, and Härnqvist (2000) for a detailed description of the ETF project and data.

[8] To study the predictive power of the two ability measures, I regress both measures on college attainment among men. The ETF ability measure explains almost two thirds of the variation in long-term outcomes explained by the conscription measure (17% against 27%).

[9] Students had respectively 15, 27 and 22 seconds to answer each question in the test, assuming they didn't waste time. Borghans et al. (2008) show that reducing the time available for completing intelligence tests reduces differences in effort correlated with non-cognitive traits.

retrospective questions on grades 7 to 9, and questions on high school. The two surveys allow me to measure key outcomes in different grades, including self-concept, school motivation, expectations, and responses by parents, teachers and peers. To derive reliable measures for these outcomes, I construct scales and indexes using relevant items from the two surveys. All scales and indexes are weighted averages of standardized items, use factor/PCA loadings as weights, and are standardized for easiness of interpretation.[10]

The most important scales, related to self-concept, are measured in grade 6. The absolute cognitive self-concept scale is based on self-assessment items asking students how good they are in school, at keeping up in class, or at specific tasks (reading, spelling, sums, etc). I build similar scales for grades 7-9, and for grade 10 for students who enrolled into high school. The relative cognitive self-concept scale is instead based on items asking students to assess their skills relative to the class, including how well they perform academically, and how well they would perform in specific tasks (explaining a math problem to a classmate, making a project proposal, etc). The non-cognitive self-concept scale is based on items that ask similar questions for non-educational domains, like arranging a party, leading a play, seeing an injured person to the hospital, etc. Many of these items reflect leadership abilities. These scales allow me to understand to what extent students are aware of their relative ability, to precisely quantify big-fish-in-a-little-pond effects, and to look for spill-overs in the non-cognitive domain. Other scales are discussed in Appendix A.

ETF added administrative data from Statistics Sweden registers (including basic demographics, grades in the last years of compulsory and high school, final education at age 35), and other administrative data not available in the registers. This includes: special education status and teacher changes, which I use to test whether students are randomly assigned to classes and teachers; class section codes, that allow me to back out school configurations; choices in compulsory school, test scores and final grades, which I use as outcomes.

2.2 Education System

This section briefly describes the Swedish education system for the cohort born 1967, who attended compulsory school between 1974 and 1983.[11] The cohort I consider was not affected

[10] I use the term scale for measurement of a specific behavioral trait/construct, and index for a variable that simply aggregates information. Construction of scales and indexes follows standard psychometric practice, and is discussed in detail in Appendix A.

[11] The current education system is similar to the one I describe. The major differences are the full expansion of pre-school and the recent decentralization of education, described in Björklund et al. (2015).

by the big expansion following the Preschool Act of 1975, which compelled municipalities to provide free pre-schools for six-year-old children of working parents. According to Korpi (2007), less than 9% of children were in formal day care in 1970. Children were instead more likely to attend for a few hours a day play schools, a form of child care inspired by Fröbel's "children gardens", centered on child development rather than schooling.

Swedish education was centralized and national curricula set in detail school policies. Compulsory school started at age seven and lasted nine years. It was divided in three stages: early (grades 1-3), middle (grades 4-6), and late (grades 7-9) compulsory school. As reported by Holmlund and Böhlmark (2019), "schools were often organized around these stages into grade 1–6 schools, grade 7–9 schools ("middle schools"), or grade 1–9 schools spanning all years of compulsory education". Late compulsory school differed from earlier stages of education: the number of subjects increased, and students were no longer taught by a single teacher. Students had to choose whether to take math and English at general or advanced levels, and could switch courses later on. Advanced courses provided better preparation for academic high school tracks, creating a form of early tracking.

At the end of compulsory school, students could apply for academic or vocational high school tracks. Oversubscription was not rare, and resulted in competitive admission based on grade 9 GPA and/or advanced math attendance.[12] Vocational tracks lasted two years and provided specialized training; academic tracks lasted one to two years longer and prepared for college. Academic high school graduates were directly eligible for college, while vocational school graduates had to go through an additional year of education to gain eligibility. A centralized admission system competitively allocated about 80% of students to the available university-programs, on the basis of their GPA (or *SweSAT,* a college entry test similar to the American SAT) and ranking of choices.[13] College was tuition-free and a mix of grants and loans helped students pay for living expenses, making higher education accessible to all qualified students.

2.3 Descriptives

The ETF sample consists of roughly 9000 Swedish compulsory school students living in 29 out of 290 municipalities (10% of the target student cohort). I exclude all classes with less

[12] 23/10% of academic/vocational high school students in my sample fail to enroll in their preferred program.

[13] Other criteria worked in the opposite direction, favoring for instance students who had vocational education or came from the labor market. Nonetheless the system remained selective: Öckert (2002) reports that around 50% of the students were rejected admission to college in the period I study.

than 10 students (3% of the sample), generally special education classes. This sample, the *full sample*, consists of 8,792 students, 380 classes, and 258 schools. My preferred specification only uses within-school between-class variation in ability, hence only considers schools where more than one class was sampled. This sample, the *final sample,* consists of 4452 students, 204 classes, 88 schools. Table 1 shows a breakdown of the sample by number of classes sampled in each school. The *final sample* only uses the part of the sample corresponding to rows 2 to 5. While classes are mechanically larger in schools with more classes sampled (since these schools are bigger), cognitive ability does not appear to relate to the number of classes sampled, both in levels and variation. Appendix B.1 tests for differences between *full* and *final sample*, and concludes that the two samples are quite similar. Importantly, there are virtually no differences in the key behavioral variables I consider.

[Table 1 about here.]

In the following I provide descriptives on the *final sample*, and refer to tables in Section B.1.[14] Table B.1 provides descriptives on individual background. I find significant variation in ability, that display a coefficient of variation of 0.27. About 40% of students have working class parents, and are considered low SES according to the Swedish census definition. The period I consider predates the large intake of asylum seekers starting in 1980, so only 3% of students are born abroad and 6% have non-nordic parents. While only 10% of students attended day care (68% of them for less than two years), 86% attended play schools. Average class size in grade 6 is 24 students, there is little variation in class size, and teacher turnover up to grade 6 is low. These features are in line with the homogeneous nature of Swedish education.

Table B.2 focuses on education choices. When asked about what factors drove math and English choices in grade 6, and high school choices in grade 10, most students mention academic ability and preferences as driving factors. External influences like parents, teachers, and peers, summarized in the extrinsic choices scale, count far less in the choices of Swedish students. About 74% of students select advanced courses in grade 7. By grade 9, the percentage decreases, falling to 56% in math. Overall, 29% of students switch courses over grades 7-9, suggesting students are updating their beliefs about ability. Many students are assigned special education status in grade 6 (13%), but this is not the case in later grades.[15]

[14] Notice that in these tables I use modified scales where items are rescaled 0-100 rather than standardized. This is to help the reader appreciate levels and variation of each scale, which is not possible for scales defined on standardized items.

[15] Special education status is considered a choice since it requires parental agreement.

86% of students proceed into high school, and 42% enroll into an academic track.

Table B.3 shows that 79% of students regularly completes high school and 86% eventually receives a high school diploma. The difference is made up by diplomas achieved though Comvux, Swedish adult education. Finally, 14% of students complete 4-year university education.

Tables B.4 and B.5 provide descriptive statistics on student behavior and parent, teacher, and peer support. Table B.4 shows that in grade 6 most students assess their academic skills positively, but there is substantial variation. Average assessment of skills relative to the class displays lower variation, and students appear to over-rank their skills (absent sample selection effects, the mean should be 50). Students appear moderately motivated and engaged with school homework. About 73% of students expect to go on to high school, far below the actual enrollment rate, 85%. Parent and teacher support are quite high, and most students report positive peer interactions. Table B.5 shows that academic self-concept increases in grades 7-9, while motivation, parent and teacher support and peer interactions are stable.[16] Grades fairness is below 50, meaning that grades were perceived as too low. This reflects issues with grading discussed in Section 3.2. Levels of cognitive self-concept in grade 10 are similar to those in grades 7-9. Students' well-being, a composite index including school motivation, parent and teacher support, and peer interactions, is high and 92% of students expect to finish high school, consistent with the actual attainment rate of 89%.

[16]Notice however that, while items are similar, scales are not identical.

3 Outcomes and Mechanisms

The empirical model I estimate is the following:

$$Y_{ics} = \alpha + \beta \, \text{Ability}_{-i,c} + \epsilon_{ics} \tag{1}$$

where i indexes the student, c the classroom, and s the school. I am interested in β, the reduced form effect of peer ability on the set of outcomes Y_{ics}. The main outcomes in the analysis are educational choices and attainment. To understand the mechanisms that underlie these reduced form effects, I separately consider a set of intermediate outcomes, including academic performance, self-concept, motivation, and responses by peers, parents and teachers. I focus on education choices and attainment for several reasons. First, these outcomes are closely connected to labor market performance, hence are more relevant for policy makers interested in class formation or streaming policies. Second, better peers might positively affect academic performance but negatively impact student self-concept and motivation. Considering only some of these intermediate outcomes might provide a misleading picture of how peers affect students' outcomes. Third, as highlighted in Carrell et al. (2018), it is not clear that one's peers will necessarily affect outcomes years after those peers are gone. Peers effects on both academic performance or behavior might be short-lived.

3.1 Literature review

In the following I discuss empirical evidence on peer effects, and describe how each mechanism affects educational choices and attainment in the light of a model of human capital investment with uncertainty over ability.[17]

While there is close to no literature on early education choices, some studies credibly identify the link between early peer effects, college attainment and labor market outcomes. Cullen et al. (2006) show that winning lotteries for admission in Chicago Public High Schools leads to attending schools with substantially better peers, but does not improve standardized test scores, graduation, college attendance and attainment. Students who win lotteries to schools with substantially higher peer quality than their next-best option are 10.7 percentage points more likely to drop out by 12th grade than comparable lottery losers. Bifulco, Fletcher, and Ross (2011) find that students attending high school classes with a higher percentage

[17]These models, sometimes referred to as amended human capital models, are discussed in Aina et al. (2019). See Altonji (1993) and Altonji et al. (2012) for a recent exposition.

of peers with college-educated mothers are less likely to drop out, and more likely to attend college attendance. Bifulco et al. (2014) show that this effect diminishes over time, and find no evidence of effects in the labor market. Black, Devereux, and Salvanes (2013) show that higher peer father earnings in lower secondary school lead to reduced dropout and higher earnings for Norwegian men. Carrell et al. (2018) find that exposure to one additional disruptive student in elementary school reduces math and reading test scores in grades 9 and 10 by 2% standard deviations, and earning at age 24-28 by 3 percent. All in all, the pattern that emerges is similar to that found for short-run effects: peer effects might not be there, are generally positive when present, but can be negative for some students.

The literature on peer effects on academic performance is instead much larger, and has been the main focus of the peer effects literature. In a recent review, Sacerdote (2014) concludes that well-identified studies do not always find peer effects in student test scores and grades, that peer effects in performance are non-linear, and that social outcomes like crime or drinking behavior or career choices generally show larger effects than test scores do.[18] By reducing the cost of accumulating human capital, positive peer effects on academic performance lead students to weakly increase education investment and attainment.

Another strand of the peer effects literature highlights a mechanism operating in the opposite direction of positive performance spillovers. Students uncertain about own academic ability might assess own academic potential in relation to their peers, for instance their classmates.[19] When exposed to better peers, students' self-concept can thus fall, leading students to sub-optimally decrease investment in education. This is known in the psychology literature as the big-fish-in-a-little-pond effect (Marsh [1987]), and in the peer effects literature as the frame of reference model. The literature providing evidence on this mechanism is recent but growing. Bui et al. (2014) show that marginal Gifted and Talented students, exposed to higher achieving peers, do not improve their achievement in the short run. One of their explanations is that entering GT reduces a student's relative ranking within the class, generating negative impacts through an invidious comparison or frame of reference model of peer effects. This explanation is tested explicitly in Antecol et al. (2016), who find that

[18]Studies that find positive peer effects on student performance in primary or secondary school are Hanushek et al. (2003), Ammermueller and Pischke (2009), Imberman et al. (2012), and Pop-Eleches et al. (2016). Studies that find a non-linear or zero effect effect are Sund (2009), Lavy et al. (2012), Burke and Sass (2013), Abdulkadiroglu et al. (2014) and Tincani (2017). Antecol et al. (2016) finds negative effects in primary school. Booij et al. (2017) and Feld and Zölitz (2017) find non-linear peer effects in academic performance in college. See Sacerdote (2014) for a recent review.

[19]Recent literature shows that children have imperfect knowledge of their ability. See for instance Stinebrickner and Stinebrickner (2012, 2014), Zafar (2011), Bobba and Frisancho (2016), Facchinello (2016).

average classroom peer achievement adversely influences own student achievement in math and reading in primary schools in disadvantaged neighborhoods. All tests go in the direction predicted by the frame of reference model irrespective of test subject. Pop-Eleches et al. (2016) find that Romanian students who have access to higher achievement schools do more homework and perform better in a high stakes graduation test. However, children who just score above cutoff on average perceive themselves as weaker relative to their peers, and are more likely to have negative interactions with their peers.

The same mechanism is also a major candidate to explain the finding that, keeping peer ability (or performance) fixed, school rank negatively affects academic performance and education choices. Using AddHealth data, Elsner and Ipshording (2016) show that student's ordinal ability rank in U.S. high schools negatively affects the likelihood of smoking, drinking, having unprotected sex, and engaging in physical fights. Exploiting the same setup, Elsner and Ipshording (2017) find that one decile increase in within-cohort rank increases the probability a student believes to be more intelligent than the average by 2.3 percentage points, and the likelihood of attending college by 1.1 percentage points. Effects on high school completion and completion of a 4-year degree are positive but imprecisely estimated. The authors also find that students with higher rank put more effort in school and have higher GPA, form career expectations in line with actual outcomes 14 years later, and are less likely to suffer from mental distress. They find no evidence that ordinal rank is related to support from teachers, parents, and friends.

Using administrative UK data, Murphy and Weinhardt (2018) show that ordinal academic rank during primary school, conditional on relative achievement, has large effects on test scores, confidence and subject choice during secondary school. A one standard deviation increase in rank in a subject improves subsequent test scores in the same subject by 0.05 within student standard deviations. Moreover, being at the top of the class in a subject during primary school increases the probability an individual chooses that subject for O levels by almost 20 percent. The authors also find that higher primary school rank in a subject increases self-reported assessment (labelled confidence in the paper) in that subject during secondary school, and that, conditioning on prior attainment, parental investment is uncorrelated with prior rank. Since higher rank does not translate into lower school performance when rank information potentially leads to distortions (i.e., local rank is lower than national rank), the authors attribute increases in performance to higher confidence, a non-cognitive skill, rather than learning. Their test rests on the assumption that higher perceived ability in a subject leads to higher effort in that subject. Since the opposite might also be true, and all subjects are tested in the UK at age 14, the test they design might not be enough to tell the two mechanisms apart.

Finally, some mechanisms can theoretically affect educational choices and attainment in both directions. Peer ability might both decrease or increase student motivation, preferences for education, and thus educational investment. Students might feel less enthusiastic about school when their performance is systematically below their peers', and this might in turn negatively affect peer interactions. This mechanism is often used to explain the mixed results found in research on selective schools and school integration. In many cases marginal students attending selective or predominantly non-minority schools do not indeed appear to benefit from attending the selective schools (Angrist and Lang, 2004; Cullen et al., 2006; Kling et al., 2007; Abdulkadiroglu et al, 2014). Papers studying peer effects in tertiary education show evidence on a similar mechanism. For instance Booij et al. (2017) find that student achievement increases with mean peer GPA but decreases with its standard deviation. The authors show that in tracked groups low-ability students have more positive interactions with other students, and are more involved. Feld and Zölitz (2017) find consistent results in a similar setup.

Imitation or role model effects might instead increase student motivation, preferences for education, and educational investment. Evidence on this mechanism is scarce and non-conclusive. While Arcidiacono and Nicholson (2005) find that medical school students are not influenced by their peers' choices of specialty, De Giorgi et al. (2010) find that Bocconi undergraduate students are 4% more likely to choose economics if strongly connected peers choose economics as a major, in some cases leading to lower grades and expected earnings. Imitation effects might thus distort education choices, and lead to reductions in educational attainment.

Teachers and parents might counteract or amplify any of the effects described above. In higher ability classes, teachers might make lessons more challenging, or set apart more time for low ability students. Lavy et al. (2012b) finds that a higher proportion of low-ability students in the class results in a deterioration of teachers' pedagogical practices and in the relationship between teachers and students, increasing levels of violence and classroom disruption. Their findings are thus consistent with better peers positively affecting low-ability students. Whether parents are substitutes or complements in the education production function, and thus whether they counteract or amplify any of the effects listed above, is still an open question.

3.2 Peer effects mechanisms in the Swedish setup

As highlighted in Sacerdote (2014), "the size and nature of peer effects estimated are highly context specific". The relevance of the mechanisms described above depends indeed on specific features of the educational setup considered. Highly competitive setups that reward top performers and marginalize weak students are unlikely to foster motivation in students exposed to better peers. On the contrary, non-competitive setups with scarce grading feedbacks are more likely to foster distortions in beliefs about own ability. This last point has not yet been highlighted in the literature.

Swedish national school curricula, determining most school policies, emphasize cooperation and integration over competition. In such a setup it is less likely that better classes hurt student motivation and peer interactions. With regard to educational inputs, Swedish teachers are expected to follow teaching practices set in the national curricula. It it unlikely that weak students suffer from teachers raising standards in better classes.

The Swedish grading system deserves special attention. Compulsory school students were assigned homework and test *scores* during the school year. While norm-referenced formal grades were assigned in late compulsory school (grades 7-9), early grade assignment depended on municipalities. About 66% of students in the cohort I consider were assigned early grades at the end of grade 3, and 56% received grades at the end of grade 6.[20] Students took national standardized tests in Swedish, Math and English in specific grades. These scores helped teachers set grades for tested, and, to some extent, non-tested subjects. Notice that students did not observe their test scores. They only observed the final grades assigned at the end of the year.[21] Gustafsson and Erickson (2018) report that norm-referenced grading was later on abandoned for many reasons, including that "teachers did not understand how to use the system as intended, but instead applied the percentage distribution within classes without regard to the results on standard tests. Another objection was that the system invited competition rather than cooperation."

For the purpose of this study, it is important to highlight three points. First, the education

[20] These details are provided in Sjögren (2010). The rationale for the reform that abolished early grades was to make classes less competitive and more inclusive. Supporters of early grade abolition believed that early grades were demotivating low SES or weak students. Facchinello (2016) studies short and long-run effects of grading for students with different SES and ability, and finds no evidence for such demotivating effects.

[21] Teachers marked national tests and sent the scores to the government, who used a sample to determine the national grade distribution. Grades ranged from 1 to 5, corresponding to national score percentiles 0-7, 7-31, 31-69, 69-93, 93-100. Teachers had some freedom in deviating from test scores, to account for sickness during the test or effort put during the academic year.

system provided students limited information to assess their academic ability, with many students observing actual grades for the first time at age 14. Given the scarcity and type of feedback, class comparisons might have played a bigger role in the formation of beliefs about own ability.[22] If grades and peer comparisons are substitutes, early grade assignment could have made peer comparisons less relevant. Second, the imperfect grade setting process might have directly contributed to distorting beliefs about own ability. This might have been more true in subjects lacking national tests, where grades were more likely to reflect class composition. In this respect the role of early grading is ambiguous. Since students observed grades, and not scores, they received a mix of unbiased information for nationally tested subjects and potentially negative-biased information in non-tested subjects. Third, students who do not perform well in tests, or put less effort in school independently of own ability, might have been penalized more by the grading process. These hypotheses are tested empirically in Section 5.

3.3 Peer effects heterogeneity in the Swedish setup

Given the discussion above, the first dimension of heterogeneity I consider is early versus late grade assignment. Appendix B.2 shows that there are few differences between early and late-graded students. When interpreting heterogeneity by early grading assignment status, one can reasonably attribute those differences to the grading regime.

The second dimension of heterogeneity I consider is gender. In their systematic review about gender differences in preferences, Croson and Gneezy (2009) highlight that, while both men and women are often overconfident, men are more overconfident than women in uncertain situations. Men and women might thus differ in how their beliefs are framed when interacting with better peers. Appendix B.3 shows that, despite higher cognitive ability, women assess in grade 6 their academic skills in the same way men do, and surprisingly assess their cognitive and non-cognitive ability relative to classmates significantly below men. In late compulsory school, when grades are assigned, women exhibit higher self-concept than men do, but this reverses in high school, where women also display lower school well-being. I find extreme segregation in education choices: women shy away over time from advanced courses in math more than men do, are 14 percentage points more likely to choose advanced English, and, while they display overall higher educational attainment, only 9% of women attain STEM

[22] Consider for instance a student scoring 10 over 20 in a test. Assume her test performance is average in the whole student population. The same student might assess differently her test score, and academic ability, if she was assigned to a class with below- or above-average performance.

education, against 54% of men. These patterns are consistent with gender norms associating language and non-STEM education to women. Finally, women are 18% standard deviations more likely to report being graded fairly than men in grades 7-9, which might signal grading distortions or biases working against men in compulsory school.[23]

The third dimension of heterogeneity I consider is socio-economic status. As discussed above, low SES students might put less effort in school, and be more penalized by grading distortions than other students. Fredriksson et al. (2016) show that only high SES parents positively respond to the negative shocks induced by an increase in class size in Sweden. Parental responses to class composition might thus depend on SES. In Appendix B.4 I show that low SES students have significantly lower cognitive ability, receive lower parental support and early education inputs, and exhibit overall much lower academic performance than the rest of the students. In terms of education attainment, low SES students are 10 percentage points less likely to graduate from high school, 50% less likely to complete an academic track, and 58% less likely to attain university education than the rest of the students. Finally, low SES students are 11% standard deviation less likely to consider their grades in late compulsory school fair. Taken together with low teacher support and negative deviations between scores and grades, this might signal grading bias or grading distortions penalizing low SES students in compulsory school.

4 Identifying Variation

4.1 Identification

Manski (1993) identifies three main reasons why peers may display similar outcomes: peers have similar characteristics or are exposed to common shocks (correlated effects), peer's background affect individual outcomes (exogenous effects), peer's outcomes affect individual outcomes (endogenous effects). Only the last two constitute actual peer effects. The effect I aim to estimate is an exogenous effect: the effect of peer cognitive ability, a pre-determined characteristic, on individual outcomes. Identification of this effect requires exogenous, or conditionally exogenous variation in peer ability. If peer ability is randomly assigned at the cohort level, then β in equation 1 has a causal interpretation. This is likely not the case, since peer ability correlates with other factors that also affect education choices (own ability, SES, school quality, etc). My identification strategy relies instead on exogenous variation in

[23]Notice however that grade setting in Sweden also reflects effort put in school, higher among women.

peer ability stemming from random assignment of students to classes and teachers. My final specification is:

$$Y_{ics} = \alpha + \beta\,\text{Ability}_{-i,c} + \gamma\,Ability_i + \text{School}_s + \Delta\,X_{ics} + \epsilon_{ics} \tag{2}$$

where School_s are school fixed effects, and X_{ics} is a set of individual level controls.

I now discuss potential threats to identification. First of all, is it credible that students are randomly assigned to classes in the setup I consider? There are three reasons why this might be the case. First, in the period I consider Swedish pre-school was not yet formally part of the Swedish education system. It did not involve formal schooling, and most students attended it only for a few hours per day. There is thus little scope for information about student ability to be passed on to compulsory school administrators forming classes. Second, schools could not track students according to ability. If school administrators observed features related to educational achievement (gender, SES, etc), they had to balance those features between classes. This mechanism breaks the correlation between individual and peer characteristics, but generates lower variance in observed characteristics than under random assignment, a feature I leverage below. Third, formal grades could be assigned the earliest in grades 3 and 6, and were progressively abolished in the period I consider. It is unlikely that administrators would use grades to assign students to existing classes or grades, or to reassign students to existing classes due to class merges.

The second relevant issue is whether ability, which I observe when students are 13, is really exogenous. I am interested in the effect of peer ability on individual choices, intermediate outcomes and educational attainment, measured from grade 6 onwards, hence these outcomes cannot affect ability, stable since grade 3 (Cunha and Heckman [2007]). It is still possible that own ability in earlier grades was affected by peer ability. If this was the case, I should observe a (possibly positive) correlation between own and peer ability, which I will formally test. While there is not much literature on this matter, Black et al. (2013) show that a host of peer characteristics in middle school does not affect Norwegian men's IQ at age 19.

Can common shocks affect both class ability and my outcomes, confounding identification? While I am not aware of literature establishing a causal link between teacher quality and cognitive ability, good teachers could in theory raise both cognitive ability and educational achievement. I do not directly observe teacher characteristics, but can test whether peer ability correlates with class-level characteristics, including class size and teacher turnover, a proxy for teacher tenure.

The literature on early child interventions considers explicitly effects of educational programs on ability. Cunha et al. (2006) conclude that enriched and sustained interventions were able to raise IQ only for children exposed at the youngest ages (from 4 months to age 5 in the case of the Abecedarian program). School starts at age 7 in Sweden, so it doesn't seem likely that peers or teachers could affect students' cognitive skills.

Finally, ability measures typically contain measurement error. Since my ability measure includes both mathematical and verbal ability, it is more reliable from the psychometric point of view, but will still be obviously measured with error. $\text{Ability}_{-i,c}$ averages a variable ridden with classical measurement error: as long as peer ability is uncorrelated with own ability, there should be little scope for measurement error in peer ability and downward bias in β.

4.2 Evidence on random assignment

Angrist (2014) states that "research designs that manipulate peer characteristics in a manner unrelated to individual characteristics provide the most compelling evidence on the nature of social spillovers". Along this line, Tables C.1 to C.3 test whether peer ability is systematically related to predetermined factors that we have reason to think affect education choices and outcomes. Notice that I express peer ability in within-school standard deviations of class ability, which corresponds to the variation I use. Column 1 reports the coefficient β from the simple regression of each outcome on $\text{Ability}_{-i,c}$, corresponding to Equation 1. Column 2 instead reports the coefficient β from the following specification:

$$Y_{ics} = \alpha + \beta\,\text{Ability}_{-i,c} + \gamma\,\text{Ability}_{-i,s} + \text{School}_s + \epsilon_{ics} \tag{3}$$

which adds to Equation 2 a control for average school ability excluding i. [24] Column 1 confirms the selection problem mentioned above. Students with better peers are better off than other students. They have higher cognitive ability, lower likelihood of being disabled (Table C.1), are more likely to have attended play school, study in schools with lower teacher turnover and have better SES (Table C.2 and C.3). Column 2 of Tables C.1 to C.3 shows that within-school variation in peer ability is instead unrelated to pre-determined variables affecting educational achievement. In particular, I find no relationship between own and peer ability, which alleviates the concerns raised above. I also find no evidence that important school inputs like class size, teacher tenure and gender composition are manipulated by

[24] Guryan et al. (2009), and more recently Caeyers and Fafchamps (2016), show that tests of peer ability are mechanically negatively biased when omitting this control.

school administrators as a response to class ability composition. The only test failure is for the occupational dummy *professional,* positively associated with peer ability. This could either be a spurious finding, or suggest that parents with high social status are able to place their children in better classes. My results show that adding occupation dummies and other controls does not affect the peer ability coefficient, supporting the former explanation.

All in all balance tests appear to support the idea that students were randomly assigned to classes. But is the level of variation in ability that I observe consistent with random assignment? To answer this question I perform a Montecarlo-like simulation exercise. I randomly reassign students to classes within school 1000 times. At each iteration I recompute the within-school standard deviation of class ability and the regression coefficient for peer ability in the peer ability test with school fixed effects. Figure D.1 displays the empirical distribution of the within-school standard deviation of class ability under random assignment, and 5% rejection regions. The level of variation in ability I observe (1.54, dashed yellow line) is remarkably close to the distribution mean (1.38, black line). The associated p-value is 0.38. The level of variation in ability I observe is consistent with random assignment also conditioning on the number of classes sampled per school (see Table 1). This excludes the possibility of offsetting heterogeneous levels of variation due to sampling different numbers of classes per school. Figure D.2 shows the distribution of the peer ability test coefficient under random assignment. The value I observe in my sample (.018, dashed yellow line) is close to the distribution mean (0.000, black line), with an associated p-value of 0.39.[25]

4.3 Magnitude and relevance of variation in peer ability

Angrist (2014) points out that that a low level of naturally occurring variation in peer effects studies may result into weak instrument-type bias. This issue, especially relevant in studies exploiting between-cohort within-school variation in peer characteristics, does not appear to be a problem in my analysis. The variation in ability I exploit corresponds to around 25% of the cohort variation in student ability, and 64% of the between-class standard deviation in mean ability. Booij et al. (2017) highlight that the same issue can result in invalid extrapolation of peer effects. I exploit within-school variation in class ability, and always provide estimates referenced to that level of variation. This avoids extrapolating results outside the level of identifying variation.

[25]Since ability must be measured in levels in the simulation, the coefficient differs from the one in Table C.1, expressed instead in within-school standard deviations of class ability.

Another way to assess the relevance of class-level variation in ability is to relate it to self-concept, the most important mechanism I examine in my study. I first consider the difference between class and ability percentile rank. Figure D.3 shows that class percentile rank can differ substantially from national ranking. Students who form their beliefs on ability with reference to the class could end up with strongly biased beliefs.[26] Similarly, one might ask to what extent ability percentile rank would change if students were allocated to a different class in the same school. I perform a simple exercise: I recompute class rank for all students after they are randomly reassigned to a different class in the same school, and take the difference between actual and counterfactual rank. Figure D.4 shows that a significant fraction of students would see their ranking change when moved to a different class.[27]

One last point that deserves attention is to what extent students interact in the same class. I observe complete grade 6 classes, but do not observe directly the school grade configuration. Using administrative data, I estimate that 40% of the students in my sample are in schools offering grades 1-6, and would thus spend on average about 6 years together. The other 60% is likely attending schools offering all grades, and would thus spend on average 9 years together.[28] Altogether students should have been together at least 6 years, plenty of time to interact and learn from each other.

5 Results

5.1 Main results

Table 2 summarizes results for the main outcomes. Since some of the outcomes cannot be pinned down to one variable, I summarize them using PCA-based indexes. This alleviates concerns with multiple hypotheses testing. To get a better understanding of results, I separately examine effects on individual outcomes, presented in Appendix E.

The self-concept index for grades 6-9 includes all items used to define absolute and relative, cognitive and non-cognitive self-concept scales. The index for academic choices in grades 7-9 includes advanced elective choice in math and English, and special education status assignment. Finally, well-being indexes summarize overall school experience, and include

[26] Notice this measure is to some extent affected by class size: in smaller classes percentile rank will mechanically differ from national rank, due to the different support of the variable.

[27] Notice that in this case the problem of class size is strongly reduced, due to the fact that there is less variation in class size within school (the average within-school standard deviation in class size is 1.8).

[28] Other configurations are possible but not common. Students could also attend a school offering grades 4 to 9, or grades 6 to 9. See Holmlund and Böhlmark (2019) for details.

items related to school motivation, peer interaction, parent and teacher support.[29]

I report the coefficient β from equation 2 for each outcome, under different specifications. Specification (1) does not include controls apart from the pool correction suggested in Guryan et al. (2009). Specification (2) adds a control for own ability. Specification (3) is my preferred specification, and adds controls for gender, age, foreign born, foreign parents and parental occupation (10 dummies). The last column reports, as a reference, the mean for each outcome variable in the estimation sample. Finally, I cluster standard errors at the classroom level, where my treatment changes.

A one standard deviation increase in peer mean ability leads to a 4.6% standard deviation decrease in self-concept during grades 6-9, a 3% standard deviation decrease in the propensity to take academic choices, and no effect on school well-being in grade 6-9. In terms of academic performance, I find a marginally statistically significant positive spillover effect in academic performance, measured by test scores, no effect on grades for subjects where grading was anchored to national tests (Swedish, English, math), and a 2.6% standard deviation *decrease* in grades for the remaining subjects. Students who attend high school, and are assigned during compulsory school to better classes, report 3% standard deviation lower school well-being and GPA. Finally, peer ability does not seem to affect educational attainment. All effects are stable across specifications, supporting the validity of the identification strategy.

Let's now zoom on individual components. Tables E.1 to Table E.3 consider outcomes based on administrative data. Table E.1 shows that students do not switch course type for English, but are 1.7 percentage points less likely to choose advanced math throughout late compulsory school, and more likely to be assigned special education status, when they interact with higher ability peers. The latter could be due to resource allocation, or to teachers (and potentially parents) forming beliefs about ability using the class as a reference. Table E.2 shows that the marginally positive significant effect of peer ability in math, Swedish and English hides significant heterogeneity: English is the only subject where I detect positive and economically meaningful performance spillovers.[30] Results on grades in compulsory school are particularly interesting. I split subjects into two categories: those that involve at some point a national standardized test, and those that don't. Students with better peers are assigned lower grades throughout compulsory school only in subjects without national tests, and in subjects with national tests before the tests are carried out (there are no national tests

[29]Indexes are described in detail in Appendix A.2.

[30]A possible explanation is that it might be easier to learn from better peers in foreign languages, which involve group work and conversation, rather than in homework or essay-based subjects like Swedish and math.

in grade 7). This confirms what Gustafsson and Erickson (2018) report: teachers tended to set grades at the class rather than national level, penalizing in the process students assigned to better classes. The grading system might have thus reinforced any negative effect on self-concept due to peer interaction. Table E.3 confirms that there are no effects on high school choices and educational achievement.[31]

Tables E.4 and E.6 consider outcomes based on survey data. Table E.4 shows that students assigned to better classes assess their cognitive skills, relative to their class, 3.4% standard deviations lower. This finding implies that students are aware of the school performance of their peers, which should correlate with their cognitive ability. Finding an effect on this outcome is a necessary condition for any mechanism positing that students assess their ability with reference to their peers. Being in a class with higher cognitive ability has no effect on relative assessment of non-cognitive skills. This implies that students can distinguish different types of ability, and that self-concept is multi-dimensional. Students exposed to better peers assess their cognitive skills 3.6% standard deviation lower in absolute terms. Since school achievement, measured by scores in national tests, marginally increases for student interacting with better peers, I conclude that class comparisons are distorting students' beliefs about own ability. This is clear evidence of big-fish-in-a-little-pond effects (or equivalently the frame of reference model). Students who are assigned to better classes do school work at home more frequently, receive higher teacher support, and have slightly better peer interactions. In Swedish non-competitive setup, better peers seem to foster a positive school environment. There is instead no effect on school motivation and parental responses. Table E.5 considers effects in late compulsory school and high-school, when grades are released and students face an academically more challenging school environment. During grades 7-9, negative effects on self-concept persist, school motivation and home support remain unresponsive to peer ability, while the positive effect on peer interactions turns almost insignificant; students feel less supported by teachers, and retrospectively report unfairly receiving low grades. This last piece of information implies that by grade 10 students are aware of the compulsory school grade distortions. Table E.6 expands on choice protocols and expectations. In grade 6, students exposed to better peers are less likely to mention ability as a driver for elective choices in math, but not English, in line with actual choices, and are less likely to report they will attend high school. In grade 10, they are more likely to report

[31] Facchinello (2016) considers the same educational setup, and shows that high performance and academic choices in late compulsory school do not necessarily translate into choosing an academic track in high school. Controlling for ability, SES is a major determinant for enrollment in academic tracks.

ability, a constraint due to selective admissions, as a driver of high school choices. Their high school plans are not affected by peer ability, again consistently with observed choices.

[Table 2 about here.]

5.2 Heterogeneity

In line with the discussion in Section 3.3, Table 3 summarizes results within subsamples defined by early grade assignment, sex, and socioeconomics status. For each outcome, I report the coefficient β from equation 2 under my preferred specification, which controls for own ability, gender, age, foreign born, foreign parents and parental occupation. Table 4 presents the same results in fully interacted models, to formally test whether effects of peer ability differ by early grade assignment, sex, or socioeconomics status. Before inspecting results it is important to highlight two points. First, testing for differences within additional categories worsens multiple hypothesis testing concerns, so I will be conservative when interpreting results. Second, the sample is small, so I will only be able to detect large differences.

Columns 1 and 2 of Tables 3 and 4 reveal that students exposed to better peers do not choose electives or assess own ability differently depending on early grading feedbacks, while surprisingly most of the negative grade distortions in untested subjects are concentrated among early graded students. Additional grading information appears to worsen biases in teachers' grading: a one standard deviation increase in peer mean ability leads to 4.3% standard deviations decrease in grades in untested subjects for early graded students, and has no effect on late graded students. Furthermore, students exposed to better peers who receive early grades do not suffer from lower well-being in high-school, and achieve 0.12 years of additional non-compulsory education with respect to late graded students.

Tables E.7 and E.18 add details to the general picture. While differences are not significant, Tables E.13 and E.14 show that self-concept falls more for early than late graded students exposed to better peers. Positive peer effects in English performance, the only positive spillover I find, are much larger for early graded students (Tables E.9 and E.10), who also appear more motivated in grades 7-9 (Tables E.15 and E.16): a one standard deviation increase in peer mean ability leads to a 8.8% standard deviation increase in English test scores and a 4.3% standard deviation increase in motivation for early graded students. Effects for late graded students are insignificant. Different explanations are compatible with this result: teachers might more efficiently target or group together students when they observe early on student ability. Alternatively, weaker students exposed to better peers (and their families) might simply realize early on they need to put more effort, since they are assigned lower grades. Tables E.13 and E.14 show that students exposed to better peers are respectively 4.5% and 2.3% more likely to report they engage frequently in homework and receive more help at home in grade 6 when receiving early grades. Given the negative sign on teacher support, results are consistent with the second explanation. Finally, early graded students assigned to better classes are 2.2 percentage points less likely to go to high school than their late-graded counterparts, 2.8 percentage points more likely to enroll in academic tracks, and 2.7 percentage points more likely to attain university education than late graded students (Tables E.11 and E.12). A possible interpretation is that better peers compound the effects of early grade assignment: students react to (distorted) information about own ability sooner and better sort into non-compulsory education.[32] In conclusion, in the Swedish setup early grading does not correct the distortions in beliefs associated to interacting with better peers, worsens grade distortions, but reinforces positive performance spillovers and improves sorting in education, resulting in an improvement in long-term outcomes for early graded students.

Columns 1 and 2 of Tables 3 and 4 do not reveal big differences in peer effects between men and women. Looking into details, Tables E.7 and E.8 show that, when exposed to better peers, men shy away from advanced math more than women do. Tables E.9 and E.10 show that this results in an improvement in math grades for men that women do not experience. The same tables also show that the negative effect on grades in non-tested subjects is larger for men than for women. Tables E.13 and E.14 show that, while women's self-concept seem to fall more than men's, these differences are not statistically significant. Finally, Tables E.15 and E.16 show that the strong negative effect of peer ability on grades fairness is concentrated on men. Altogether, peer ability seems to affect men and women in similar ways, with a few differences possibly reflecting gender norms and the institutional setting.

Columns 1 and 2 of Tables 3 and 4 reveal stark differences in the effect of exposure to better peers between low and high SES students. The negative effects of exposure to better peers I found are concentrated among low SES students. When exposed to better peers, low SES students have 7.5% standard deviations lower self-concept, 5.3% standard deviations lower propensity to take academic choices, and 4.1% lower grades with respect to high SES students. While main effects are not statistically significant, I find that low SES students

[32] Facchinello (2016) evaluates the reform that abolished early grades, and finds a qualitatively similar result: short-run education choices of early graded students are more consistent with academic ability with respect to the choices of late graded students.

exposed to better peers have 5.3% standard deviations lower school well-being in grades 6-9 and 0.07 less years of education than their high SES counterparts.

Let's now consider individual outcomes. Tables E.7 and E.8 show that, on top of math, low SES students exposed to better peers are 2.6 percentage points less likely to choose advanced English. Tables E.13 and E.14 provide some detail on the negative difference found above in the broad measure of self-concept. In grade 6 low SES students' cognitive self-concept falls 7.8% standard deviations more than it is the case for the rest of the students. A similar result holds for non-cognitive self-concept, but differences are not statistically significant. It's interesting to notice that the zero effect on motivation found on the whole sample masks statistically significant positive effects for high SES students, and insignificant negative effects among low SES students exposed to better peers. Low SES students also do not share the benefit of additional teacher support I found for students assigned to better classes, entirely concentrated among high SES students. In grades 7-10 most of the differences revealed in grade 6 remain, but differences are no longer statistically significant. A relevant finding is that low SES students exposed to better peers are 2.4 percentage points less likely to receive parental support at home in late compulsory school. There is no effect for high SES students. I confirm Fredriksson et al. (2016) finding that parental responses to educational inputs depend on SES. Finally, Tables E.13 and E.14 show that the negative effects I found on ability's role in elective choices and on high school expectations is concentrated among low SES students. I conclude that negative effects on self-concept and grades due to exposure to better peers are disproportionally concentrated on academically and socially disadvantaged students, whose academic performance is however still positively affected by better peers.

[Table 3 about here.]

[Table 4 about here.]

6 Conclusions

6.1 Summary of results

This paper studies educational choices and attainment of Swedish students randomly assigned in compulsory school to classes with different levels of cognitive ability. Using rich administrative and survey data, I explore mechanisms beyond the traditionally studied peer effect in academic performance, focusing in particular on distortions in beliefs about ability and grades. I find that that students who interact with better peers systematically underestimate own ability and are less likely to choose advanced subjects in compulsory school. Students assigned to higher ability classes perform better in national standardized test scores, but are assigned lower grades in subjects lacking national tests. Even in a setup where national test scores anchor grade setting, teachers appear unable to distinguish absolute from relative performance, and assign grades that potentially contribute to distorting students' beliefs about ability. Negative effects persist after compulsory school: students assigned to better classes experience lower well-being and GPA in high school. In the long-run educational attainment is unaffected by peer ability for the average student.

I find substantial heterogeneity in the effect of exposure to better peers by grade assignment and SES, but no major difference by sex. Early grade assignment does not correct the distortions in beliefs and grading that students exposed to better peers suffer; if anything, it worsens them. Students positively react to distorted information about own ability: they exhibit stronger positive performance spillovers, possibly due to increased effort and parental support, and improve their sorting to education. The unconditional zero effect of peer ability on educational attainment I find masks positive effects for early graded students and negative effects for late graded students. The negative effects in self-concept and grading I find are disproportionally concentrated among low SES students, who also receive less parental support in late compulsory school, and display lower motivation, school well-being and educational attainment with respect to high SES students when exposed to better peers.

6.2 Contributions and comparison to the literature

This paper contributes to several literatures. First, it contributes to the broad peer effects literature. The new evidence I provide on negative peer effects in self-concept, grading and choices can help to understand the mixed results found in the literature. While negative effects on self-concept have already been proposed as an explanation in papers finding zero or negative peer effects, my paper shows direct evidence on this channel. The grading channel has instead up to now been ignored in the literature, which mostly focuses on test scores. Second, the paper contributes to a small literature establishing how peers affect students in the long-run, and thus to the broader literature on long-run effects of educational interventions.[33] Third, the paper contributes to the tracking literature.[34] I show that interacting with better peers in a setup with progressive tracking have positive effects on performance, but negative effects on academic choices that generate effects beyond compulsory school. Fourth, the paper directly contributes to the literature on grade feedbacks and belief updating (Zafar [2011], Stinebrickner and Stinebrickner [2012], Facchinello [2016]).

The core literature on peer effects in education, recently reviewed by Sacerdote (2014), finds either positive, zero or non-linear effects of peer performance on individual school achievement. Consistently with this literature, I find positive peer effects on academic performance in English, an interactive subject, but not in other subjects. While I do not consider effects within the ability distribution, I do not find weaker positive spillovers in performance among low SES students, who have lower cognitive ability and parental support. This is at odds with results from the selective schools literature (Angrist and Lang [2004], Cullen et al. [2006], Kling et al. [2007], Abdulkadiroglu et al. [2014]), which however considers the US, a more unequal and segregated setup. I instead find stronger spillovers among students who receive early grades. My findings are consistent with better classes compounding the information effects of early grade assignment, leading to increased effort and parental support for students who are assigned lower grades.

How do my results compare to those found in the literature on the long-run effects of peers? Similarly to Cullen et al. (2006), who considers the effect of marginally winning a lottery for admission to the Chicago Public High Schools, I do not find evidence of long-run effects for the average student. I however find both positive and negative effects on educational attainment for specific categories of students, which is qualitatively consistent with the findings of Black et al. (2013), who show systematic heterogeneity considering different peer effects. Altogether my results seems to confirm the general finding of the broader peer effect literature: peer effects might not be there, are generally positive when

[33] Long-run peer effects are studied by Cullen et al. (2006), Black et al. (2013), Elsner and Ipshording (2017) and Carrell et al. (2018). Chetty et al. (2011), Heckman et al. (2013), Fredriksson et al. (2013), Conti and Heckman (2016) provide evidence on long-run effects of education inputs and interventions.

[34] See for instance Meghir and Palme (2005), Hanushek and Wössmann (2006), Pekkarinen et al. (2009), Malamud and Pop-Eleches (2011), Betts (2011), Guyon et al. (2012), Malamud and Pop-Eleches (2011), Kerr et al. (2013).

present, and can be negative for specific categories of students.

Most closely related to this paper are Bui et al. (2014) and Antecol et al. (2016), who find respectively zero and negative peer effects in performance and interpret their results in the light of the frame of reference model. The negative effects on GPA and well-being I find in high school are in line with the findings of those papers. My contribution is to explicitly show, rather than assume, the decrease in self-concept. Pop-Eleches et al. (2016) find in the Romanian context positive peer effects in performance, followed by negative effects on relative self-concept and peer interactions. While I don't find negative effects on peer interactions, I also show that positive peer effects can coexist with negative effects in self-concept. My contribution with respect to their paper is showing that, on top of relative self-concept, better peers lead to decreases in absolute self concept, implying that better peers can lead to distortions in beliefs about ability. This effect, known in the psychology literature as the big-fish-in-a-little pond effect, has yet not received direct support in the economics literature.

An important exception is found in the literature that studies the effect of school rank, described in detail in Section 3.1. My results cannot be directly compared to those found in that literature, which isolates rank effects from the effect of average peer ability/performance. However they can shed light on an important point that remains open in that literature: do the positive effects of high rank in school stem from higher perceived ability (the main channel in Elsner and Ipshording [2017]) or confidence (as maintained in Murphy and Weinhardt [2018])? This point is quite important. The first scenario suggests distortions in beliefs that might be alleviated, for instance improving grading feedbacks. The second scenario implies a 0-sum game where top students win and the rest loses, which strengthens the case for ability tracking. My results clearly show that students' beliefs and grades are negatively biased when students are assigned to better classes, and support the perceived ability mechanism. My findings are instead not consistent with a confidence mechanism: better peers do not affect relative assessment of non-cognitive skills (mostly reflecting leadership) and do not decrease school motivation, save for positive exceptions for high SES and early graded students. My results thus provide support for the mechanism described in Elsner and Ipshording (2017), who find that ranking lower in high-school negatively affects the probability that a student believes to be more intelligent than average. Since the authors find a similar effect for GPA, their finding is not necessarily evidence of a big-fish-in-a-little pond effect, but could simply reflect the link between GPA and self-concept. My findings are in this sense more clear: while being assigned to a class with better peers has a positive effect on academic performance, measured by test scores, I find negative effects on students' beliefs about own ability, signaling systematic distortions in self-concept.

This paper also relates to the literature investigating how students react to grades and update beliefs about own ability. Stinebrickner and Stinebrickner (2012) and Zafar (2011) find that students who get lower than expected grades are respectively more likely to drop out from college and switch to an easier major. I find a similar result in a completely different setup: students assigned to compulsory school classes with better peers are systematically assigned lower grades, revise downward their assessment of cognitive ability, and switch as a consequence to easier courses in compulsory school. My paper thus shows that Stinebrickner and Stinebrickner (2012) and Zafar (2011)'s results hold at the compulsory school level, and adds evidence on how these learning effects propagate over time.

Another strand of the grading literature studies how providing national grades in compulsory (Sjögren [2010], Facchinello [2016]) or high school (Sofoklis and Megalokonomou [2015]) affects students' short and long-run outcomes. Using similar data from the ETF project, Facchinello (2016) finds that short-run education choices of early graded students are more consistent with academic ability than those of late graded students. This paper finds evidence that students assigned to classes with better peers compound the effects of early grade assignment: students learn about relative ability sooner, put more effort during compulsory school, and later on better sort into non-compulsory education.

6.3 Policy relevance and open questions

The most important implication of my findings is that evaluating effects of exposure to better peers considering performance spillovers alone might provide an incomplete, and potentially misleading, picture of peer effects. In the presence of negative effects in self-concept, academic choices or attainment might be more relevant targets for policy makers interested in designing tracking and streaming policies for two reasons. First, focusing on outcomes that respond to both positive performance spillovers and negative distortions in self-concept and grades can improve evaluations of peer effects. Second, while academic performance can be taken as a measure of human capital, the literature has shown that education choices at both the high school (Golsteyn and Stenberg [2017]) and college level (Kirkeboen et al. [2016], Altonji [2012]) have much larger implications for labor market outcomes.

The second implication is that the grading system should not be ignored when evaluating peer effects. In many education systems grades are set on a curve, which creates a mechanical negative correlation between peer ability and individual grades. In my setup this shifts grade signals and beliefs in the opposite direction of the positive gains detected in academic performance. Similarly, many education systems lack standardized tests, which again might foster distortions in beliefs and academic choices. An important question is to what extent my findings generalize to other setups, and whether additional and possibly different grading feedbacks can avoid these negative effects. My finding that providing additional grades early on worsens distortions in grade assignment and does not reduce negative effects on academic self-concept is in this sense not encouraging. Again, this might be related to the particular setting, and more research is needed.

Finally, special attention should be paid to disadvantaged students, for whom negative effects of exposure to higher ability peers seem to prevail. If my results generalize to other setups, this argument adds another point in favor of ability tracking.

References

Abdulkadiroglu, Atila, Joshua D. Angrist and Parag A. Pathak, "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," *Econometrica*, 82 (2014), 137-196.

Aina, Carmen, Eliana Baici, Giorgia Casalone, and Francesco Pastore, "Delayed graduation and university dropout: A review of theoretical approaches," GLO Discussion Paper Series 399, 2019.

Angrist, Joshua D., and Kevin Lang, "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program," *American Economic Review*, 94 (2004), 1613-1634.

Altonji, Joseph G., Erica Blom, and Costas Meghir, "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," *Annual Review of Economics*, 4 (2012), 185-223.

Ammermueller, Andreas, Järn-Steffen Pischke, "Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study," *Journal of Labor Economics*, 27 (2009), 315-348.

Arcidiacono, Peter, and Sean Nicholson, "Peer effects in medical school," *Journal of Public Economics*, 89 (2005), 327-350.

Betts, Julian R., "The Economics of Tracking in Education," in *Handbook of the Economics of Education,* Vol. 3, Eric A. Hanushek, Stephen Machin and Ludger Woessmann, eds. (Amsterdam: Elsevier, 2011).

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, "Under Pressure? The Effect of Peers on Outcomes of Young Adults," *Journal of Labor Economics*, 31 (2013), 119-153.

Bifulco, Robert, and Stephen L. Ross, "The Effect of Classmate Characteristics on Post-secondary Outcomes: Evidence from the Add Health," *American Economic Journal: Economic Policy*, American Economic Association, 3 (2011) 25-53.

Bifulco, Robert and Fletcher, Jason M., Jung Sun Oh, Stephen L. Ross, "Do high school peers have persistent effects on college attainment and other life outcomes?," *Labour Economics*, 29 (2014), 83-90.

Bobba, Matteo, and Verónica Frisancho, "Learning about Oneself: The Effects of Performance Feedback on School Choice," IDB Working Paper 7968, 2016.

Booij, Adam S., Edwin Leuven, and Hessel Oosterbeek, "Ability Peer Effects in University: Evidence from a Randomized Experiment," *Review of Economic Studies*, 84 (2017), 547-578.

Borghans, Lex, Huub Meijers, and Bas Ter Weel, "The Role Of Noncognitive Skills In Explaining Cognitive Test Scores," *Economic Inquiry*, 46 (2008), pages 2-12.

Bui, Sa A., Steven G. Craig, and Scott A. Imberman, "Is Gifted Education a Bright Idea? Assessing the Impact of Gifted and Talented Programs on Students," *American Economic Journal: Economic Policy*, 6 (2014), 30-62.

Burke, Mary A. , and Sass A. Tim , 2013. "Classroom Peer Effects and Student Achievement," *Journal of Labor Economics*, 31(1), 51 82.

Carrell, Scott E., Mark Hoekstra, and Elira Kuka, "The Long-Run Effects of Disruptive Peers," *American Economic Review*, 108 (2018), 3377-3415.

Caeyers, Bet, and Marcel Fafchamps, "Exclusion Bias in the Estimation of Peer Effects," *NBER Working Papers* 22565, 2016.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan, "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star," *The Quarterly Journal of Economics*, 126 (2011), 1593-1660.

Conti, Gabriella, James J. Heckman, and Rodrigo Pinto, "The Effects of Two Influential Early Childhood Interventions on Health and Healthy Behaviour," *Economic Journal*, 126 (2016), 28-65.

Cortina, Jose M., "What is coefficient alpha? An examination of theory and applications," *Journal of applied psychology*, 78 (1993), 98-104.

Cullen, Julie B., Brian A. Jacob, and Steven Levitt, "The Effect of School Choice on Participants: Evidence from Randomized Lotteries," *Econometrica*, , 74 (2006), 1191-1230.

Cunha, Flavio, James J. Heckman, and Lance Lochner, "Interpreting the Evidence on Life Cycle Skill Formation," in *Handbook of the Economics of Education*, Vol. 2, Eric A. Hanushek and Finish Welch, eds. (Amsterdam: Elsevier, 2006).

Cunha Flavio, and James J. Heckman, "The Technology of Skill Formation," *American Economic Review*, 97 (2007), 31-47.

Cunha, Flavio, and James J. Heckman, "The Economics and Psychology of Inequality and Human Development," *Journal of the European Economic Association*, 7 (2009), 320-364.

Croson Rachel, and Uri Gneezy, "Gender Differences in Preferences," *Journal of Economic Literature*, 47 (2009), 448-474.

De Giorgi Giacomo, Michele Pellizzari, and Silvia Redaelli, "Identification of Social Interactions through Partially Overlapping Peer Groups," *American Economic Journal: Applied Economics*, 2 (2010), 241-275.

Elsner, Benjamin, and Ingo E. Isphording, "A Big Fish in a Small Pond: Ability Rank and Human Capital Investment," *Journal of Labor Economics*, 35 (2017), 787-828.

Emanuelson, Ingemar, "Utvärdering genom uppföljning av elever - Ett nytt Individualstatistikproject," Stockholm School of Teacher Education Report 11, 1979.

Facchinello, Luca, "The Impact of Early Grading on Academic Choices: Mechanisms and Social Implications", Unicredit & Universities working paper series 73, 2016.

Feld Jan, and Ulf Zölitz, "Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects," *Journal of Labor Economics*, 35 (2017), 387-428.

Fredriksson, Peter, Björn Öckert, and Hessel Oosterbeek, "Parental Responses to Public Investments in Children: Evidence from a Maximum Class Size Rule," *Journal of Human Resources*, 51 (2016), 832-868.

Fruehwirth, Jane C., "Can Achievement Peer Effect Estimates Inform Policy? A View from Inside the Black Box," *The Review of Economics and Statistics*, 96 (2014), 514-523.

Furr, Mike, "Scale construction and psychometrics for social and personality psychology," 1st ed. (Newbury Park, CA: SAGE Publications, 2011).

Golsteyn, Bart H. H., and Anders Stenberg, "Earnings over the Life Course: General versus Vocational Education," *Journal of Human Capital*, 11 (2017), 167-212.

Guyon, Nina, Eric Maurin, and Sandra McNally, "The Effect of Tracking Students by Ability into Different Schools: A Natural Experiment," *Journal of Human Resources*, 47 (2012), 684-721.

Goulas, Sofoklis, and Rigissa Megalokonomou, "Knowing who you are: The Effect of Feedback Information on Short and Long Term Outcomes," University of Warwick Economic Research Papers 270019, 2015.

Guryan, Jonathan, Kory Kroft, and Matthew J. Notowidigdo, "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments," *American Economic Journal: Applied Economics*, 1 (2009), 34-68.

Gustafsson, Jan-Eric, and Gudrun Erickson, "Nationella prov i Sverige – tradition, utmaning, förändring," Acta Didactica Norge, 12 (2018), Art. 2.

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin, "Does peer ability affect student achievement?," *Journal of Applied Econometrics*, 18 (2003), 527-544.

Hanushek, Eric A., and Ludger Wössmann, "Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries," *Economic Journal*, 116 (2006), 63-76.

Härnqvist, Kjell, "A longitudinal program for studying education and career development," Göteborgs Universitet Report 01, 1998.

Heckman, James J., Rodrigo Pinto, and Peter Savelyev, "Understanding the Mechanisms through which an Influential Early Childhood Program Boosted Adult Outcomes," *American Economic Review*, 103 (2013), 2052-2086.

Hoekstra, Mark, Pierre Mouganie, and Yaojing Wang, "Peer Quality and the Academic Benefits to Attending Better Schools," *Journal of Labor Economics*, 36 (2018), 841-884.

Hoxby Caroline, "Peer effects in the classroom: Learning from gender and race variation," NBER Working Paper 7867, 2000.

Imberman, Scott, Adriana D. Kugler, and Bruce Sacerdote, "Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees," *American Economic Review*, 102 (2012), 2048-2082.

Jalava, Nina, Juanna S. Joensen, and Elin Pellas, "Grades and rank: Impacts of non-financial incentives on test performance," *Journal of Economic Behavior & Organization*, 115 (2015), 161-196.

Kaiser, Henry F., "An index of factor simplicity," *Psychometrika,* 39 (1974), 31-36.

Kerr, Sari P., Tuomas Pekkarinen, and Roope Uusitalo, "School Tracking and Development of Cognitive Skills," *Journal of Labor Economics*, 31 (2013), 577-602.

Kilpatrick, Jeremy, and Bengt Johansson, “Standardized mathematics testing in Sweden: The legacy of Frits Wigforss,” Nordic Studies in Mathematics Education, 2 (1993), 6-30.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad, “Field of Study, Earnings, and Self-Selection,” *The Quarterly Journal of Economics*, 131 (2016), 1057-1111.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, “Experimental Analysis of Neighborhood Effects,” *Econometrica*, 75 (2007), 83-119.

Korpi, Barbara Martin, “The politics of pre-school: Intentions and decisions underlying the emergence and growth of the Swedish pre-school”, Ministry of Education and Research, 2007.

Lavy, Victor, and Analia Schlosser, “Mechanisms and impacts of gender peer effects at school,” *American Economic Journal: Applied Economics*, 3 (2011), 1-33.

Lavy, Victor, Olmo Silva, and Felix Weinhardt, “The Good, the Bad and the Average: Evidence on Ability Peer Effects in Schools,” *Journal of Labor Economics*, 30 (2012a), 367-414.

Lavy, Victor, Daniele M. Paserman, and Analia Schlosser, “Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom,” *The Economic Journal,* 122 (2012b), 208-237.

Malamud, Ofer, and Cristian Pop-Eleches, “School tracking and access to higher education among disadvantaged groups,” *Journal of Public Economics*, 95 (2011), 1538-1549.

Manski, Charles, “Identification of Endogenous Social Effects: The Reflection Problem,” *The Review of Economic Studies*, 60 (1993), 531-542.

Marsh, Herbert W., “The big-fish-little-pond effect on academic self-concept,” *Journal of Educational Psychology*, 79 (1987), 280-295.

Meghir, Costas, and Palme Mårten, “Educational Reform, Ability, and Family Background,” *American Economic Review*, 95 (2005) 414-424.

Murphy, Richard, and Felix Weinhardt, “Top of the Class: The Importance of Ordinal Rank,” NBER Working Papers 24958, 2018.

Öckert, Björn, “Do university enrollment constraints affect education and earnings?,” IFAU Working Paper 16, 2002.

Pekkarinen, Tuomas, Roope Uusitalo, and Sari Kerr, “School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform,” *Journal of Public Economics*, 93 (2009), 965-973.

Sacerdote, Bruce, "Peer Effects in Education: How Might They Work, How Big are They and How much do we Know Thus Far?," in *Handbook of the Economics of Education,* Vol. 3, Eric A. Hanushek, Stephen Machin and Ludger Woessmann, eds. (Amsterdam: Elsevier, 2011).

Sacerdote, Bruce, "Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward?," *Annual Review of Economics*, 6 (2014), 253-272.

Sjögren, Anna, "Graded children – evidence of longrun consequences of school grades from a nationwide reform," IFAU Working Paper 7, 2010.

Stinebrickner, Todd, and Ralph Stinebrickner, "Learning about Academic Ability and the College Dropout Decision," *Journal of Labor Economics*, 30 (2012), 707-748.

Stinebrickner, Todd, and Ralph Stinebrickner, "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout". *Review of Economic Studies*, 81 (2014), 426-472.

Sund, Krister, "Estimating peer effects in Swedish high school using school, teacher, and student fixed effects," *Economics of Education Review*, 28 (2009),329-336.

Tincani, Michela, "Heterogeneous Peer Effects and Rank Concerns: Theory and Evidence," CESifo Working Paper 6331, 2017.

Zafar, Basit, "How Do College Students Form Expectations?," *Journal of Labor Economics*, 29 (2011), 301-348.

A Scales and Indexes

This section presents in detail the behavioral scales and aggregate indexes I use as outcomes in the analysis.[35] Following standard psychometric practice, I require scales to have high reliability, that is, a Cronbach alpha or Kaiser–Meyer–Olkin measure of sampling adequacy close to 0.7 or higher.[36] I use alpha and item-rest correlations to refine the choice of the items to include in each scale. I perform exploratory factor analysis, retaining one factor, and check that factor loadings are consistent with the construct I want to measure. All scales are weighted averages of standardized items, with factor loadings as weights.

In the following I present the items used in each scale, and the associated factor loadings. For each scale I report Kaiser–Meyer–Olkin measure of sampling adequacy (KMO) and Cronbach alpha for the standardized items. I also show the correlation between my scale and an alternative scale assigning equal weight to each standardized item. While the two scales have correlations close to unity, the factor-based scale is preferred since it yields higher precision in results estimates. Results using simple scales and indexes are similar to the results shown in the paper, and are available upon request.

A.1 Behavioral scales

Absolute cognitive self-concept (grade 6)

All scale items are taken from the grade 6 student survey. The scale is set to missing when respondents answer less than 4 items. KMO is 0.69 and alpha is 0.69. The correlation between factor-based scale and simple scale is 0.98.

1. I do well in school. Factor load: +0.56
2. I am good at sums. Factor load: +0.46
3. I am good at spelling. Factor load: +0.45
4. I would like to be better at doing sums. Factor load: -0.57
5. I would like to spell better. Factor load: -0.55

[35] I use the term scale for measurement of a specific behavioral trait. For variables that summarize different outcomes I use instead the word index. Consistently, I perform principal component analysis, rather than factor analysis, when dealing with indexes.

[36] As a reference, Kaiser (1994) described KMO values between 0.6 and 0.7 as mediocre, 0.70 to 0.79 middling, and above that meritorious. When using factor or PCA scores, alphas are interpreted as lower bounds for reliability (Furr [2011]). Scales with alphas around 0.7 are considered acceptable for scientific purposes. For further discussion on alpha see Cortina (1994).

6. I am bad at reading. Factor load: -0.40
7. It is hard to understand teacher explanations in class. Factor load: -0.28
8. I would like to do better in school. Factor load: -0.53

Relative cognitive self-concept (grade 6)

All scale items are taken from the grade 6 student survey, and have as a reference classmates. The scale is set to missing when respondents answer less than 3 items. KMO is 0.76 and alpha is 0.72. The correlation between factor-based scale and simple scale is almost 1.

1. On a 1-9 scale, compare yourself with the best and the poorest pupil in the class. Factor load: +0.62
2. On a 1-9 scale, how good would you be at explaining a math problem to an absent classmate? Factor load: +0.63
3. You have to take a lesson when the teacher is ill. On a 1-9 scale, how well can you cope with that? Factor load: +0.57
4. You have to make a proposal for a group project on, say, Africa. On a 1-9 scale, how would your peers rate it? Factor load: +0.60

Relative non-cognitive self-concept (grade 6)

All scale items are taken from the grade 6 student survey, and have as a reference group classmates. The scale is set to missing when respondents answer less than 2 items. KMO is 0.73 and alpha is 0.68. The correlation between factor-based scale and simple scale is 0.99.

1. You have to arrange a party for your class. On a 1-9 scale, how good would it be? Factor load: +0.60
2. You have to go to the hospital with a injured classmate. How well would you cope with that? Factor load: +0.44
3. You are elected into the student council. How well would you cope with that? Factor load: +0.60
4. You are asked to arrange and lead a play activity at school. How well would you cope with that? Factor load: +0.61

Extrinsic choices (grade 6)

All scale items are taken from the grade 6 student survey. The scale is set to missing when respondents answer less than 4 items. KMO is 0.65 and alpha is 0.82. The correlation between factor-based scale and simple scale is 0.99.

1. My math course choice depended to a large extent on parents. Factor load: +0.69
2. My math course choice depended to a large extent on peers. Factor load: +0.65
3. My math course choice depended to a large extent on teachers. Factor load: +0.70
4. My English course choice depended to a large extent on parents. Factor load: +0.69
5. My English course choice depended to a large extent on peers. Factor load: +0.66
6. My English course choice depended to a large extent on teachers. Factor load: +0.72

Motivation (grade 7-9)

All scale items are taken from the grade 10 student survey, and refer to grades 7-9. The scale is set to missing when respondents answer less than 3 items. KMO is 0.70 and alpha is 0.63. The correlation between factor-based scale and simple scale is 0.99.

1. I enjoyed most of my time in school. Factor load: +0.43
2. I did my absolute best even in boring subjects. Factor load: +0.58
3. I did my absolute best even in difficult tasks. Factor load: +0.60
4. I am happy with what I achieved in school. Factor load: +0.43
5. I gave up when the tasks were too hard. Factor load: -0.41
6. I had to learn a lot of unnecessary stuff. Factor load: -0.31

Absolute cognitive self-concept (grade 7-9)

All scale items are taken from the grade 10 student survey, and refer to grades 7-9. The scale is set to missing when respondents answer less than 3 items. KMO is 0.69 and alpha is 0.67. The correlation between factor-based scale and simple scale is 0.98.

1. I found it difficult to keep up with classwork. Factor load: -0.52
2. I often found it difficult to concentrate during lessons. Factor load: -0.41
3. The tasks I got from the teacher were too difficult. Factor load: -0.34

4. I had problems with reading. Factor load: -0.70
5. I had problems with writing. Factor load: -0.69
6. I had problems with speaking in front of a group/the class. Factor load: -0.33
7. I had problems with mathematics. Factor load: -0.31

Absolute cognitive self-concept (grade 10)

All scale items are taken from the grade 10 student survey, and refer to high school. The scale is set to missing when respondents answer less than 5 items. KMO is 0.60 and alpha is 0.69. The correlation between factor-based scale and simple scale is 0.98.

1. I have enough knowledge/skills in reading. Factor load: +0.42
2. I have enough knowledge in writing. Factor load: +0.50
3. I have enough knowledge/skills in speaking. Factor load: +0.42
4. I have enough knowledge/skills in mathematics. Factor load: +0.31
5. I have enough knowledge/skills in English. Factor load: +0.39
6. I have enough knowledge/skills in other subjects. Factor load: +0.34
7. I feel confident doing math. Factor load: +0.30
8. I feel confident reading. Factor load: +0.62
9. I feel confident writing. Factor load: +0.65
10. I feel confident speaking in a group. Factor load: +0.46

A.2 Aggregate indexes

Academic choices (grade 7-9)

All index items are taken from administrative records. The index is set to missing when respondents answer less than 5 items. KMO is 0.81 and alpha is 0.84. The correlation between PCA-based scale and simple scale is 0.98.

1. Advanced Math (grade 7). Component load: +0.37
2. Advanced Math (grade 8). Component load: +0.39
3. Advanced Math (grade 9). Component load: +0.36
4. Advanced English (grade 7). Component load: +0.40
5. Advanced English (grade 8). Component load: +0.41

6. Advanced English (grade 9). Component load: +0.40
7. Special education (grade 7). Component load: -0.20
8. Special education (grade 8). Component load: -0.18
9. Special education (grade 9). Component load: -0.14

Self-concept (grade 6-9)

Index items are either taken from the grade 6 students survey, or from the grade 10 survey (in this case they refer to grades 7-9). The index is set to missing when respondents answer less than 12 items. KMO is 0.86 and alpha is 0.85. The correlation between PCA-based scale and simple scale is 0.99.

1. I do well in school (grades 6). Component load: +0.26
2. I am good at sums (grades 6). Component load: +0.22
3. I am good at spelling (grades 6). Component load: +0.16
4. On a 1-9 scale, compare yourself with the best and the poorest pupil in the class (grades 6). Component load: +0.34
5. On a 1-9 scale, where would you be in the class if you really made an effort? (grades 6) Component load: +0.31
6. On a 1-9 scale, how good would you be at explaining a math problem to an absent classmate? (grades 6) Component load: +0.25
7. You have to take a lesson when the teacher is ill. On a 1-9 scale, how well can you cope with that? (grades 6) Component load: +0.25
8. You have to make a proposal for a group project on, say, Africa. On a 1-9 scale, how would your peers rate it? (grades 6) Component load: +0.16
9. You have to arrange a party for your class. On a 1-9 scale, how good would it be? (grades 6) Component load: +0.27
10. You have to go to the hospital with a injured classmate. How well would you cope with that? (grades 6) Component load: +0.22
11. You are elected into the student council. How well would you cope with that? (grades 6) Component load: +0.27
12. You are asked to arrange and lead a play activity at school. How well would you cope with that? (grades 6) Component load: +0.22
13. I often think that I would like to be better at doing sums (grades 6). Component load: -0.16

14. I often think that I would like to spell better (grades 6). Component load: -0.14
15. I often think that I am bad at reading (grades 6). Component load: -0.18
16. I often think it is hard to understand teacher explanations in class (grades 6). Component load: -0.13
17. I often think that I would like to do better in school (grades 6). Component load: -0.16
18. I found it difficult to keep up with classwork (grades 7-9). Component load: -0.16
19. I often found it difficult to concentrate during lessons (grades 7-9). Component load: -0.13
20. The tasks I got from the teacher were too difficult (grades 7-9). Component load: -0.9
21. I had problems with reading (grades 7-9). Component load: -0.19
22. I had problems with writing (grades 7-9). Component load: -0.18
23. I had problems with public speaking (grades 7-9). Component load: -0.14
24. I had problems with mathematics (grades 7-9). Component load: -0.13

School well-being (grade 6-9)

Scale items are either taken from the grade 6 students survey, or from the grade 10 survey (in this case they refer to grades 7-9). The index is set to missing when respondents answer less than 12 items. KMO is 0.80 alpha is 0.75. The correlation between PCA-based scale and simple scale is 0.97.

1. I always do my best, even when the tasks are boring (grade 6). Component load: 0.15
2. I got help at home with school work (grade 6). Component load: 0.04
3. My teacher cares about me (grade 6). Component load: 0.13
4. I like working together with other children in the class (grade 6). Component load: 0.06
5. Other children in the class like working together with me (grade 6). Component load: 0.12
6. I enjoyed most of my time in school (grades 7-9). Component load: 0.29
7. I did my absolute best even in boring subjects (grades 7-9). Component load: 0.22
8. I did my absolute best even in difficult tasks (grades 7-9). Component load: 0.23
9. I am happy with what I achieved in school (grades 7-9). Component load: 0.25
10. I was help with school work whenever needed (grades 7-9). Component load: 0.12
11. I often received help and support from teachers (grades 7-9). Component load: 0.19

12. I could ask teachers for help when needed (grades 7-9). Component load: 0.24
13. I got enough help from my teachers (grades 7-9). Component load: 0.27
14. I give up when the tasks were too hard (grade 6). Component load: -0.18
15. I often think about other things when I do math and writing in school (grade 6). Component load: -0.21
16. I have to learn a lot of unnecessary stuff (grades 6) Component load: -0.16
17. I often spend time on my own during breaks (grade 6). Component load: -0.06
18. I gave up when the tasks were too hard (grades 7-9). Component load: -0.25
19. I had to learn a lot of unnecessary stuff (grades 7-9). Component load: -0.19
20. The teacher should care more about me (grade 6). Component load: -0.08
21. I had problems relating to teachers (grades 7-9). Component load: -0.31
22. I had problems understanding teachers' explanations (grades 7-9). Component load: -0.30
23. I had problems getting help from the teacher (grades 7-9). Component load: -0.33

School well-being (grade 10)

All items are taken from the grade 10 survey, and refer to high school. The index is set to missing when respondents answer less than 4 items. KMO is 0.79 and alpha is 0.69. The correlation between PCA-based scale and simple scale is 0.99.

1. I feel great in high school. Component load: +0.41
2. I get enough help from my teacher. Component load: +0.36
3. I can ask the teacher for help when I don't understand. Component load: +0.36
4. I give up when the task is too hard. Component load: -0.35
5. I often feel lonely at school. Component load: -0.24
6. I often find it difficult to concentrate in class. Component load: -0.38
7. I often feel stressed. Component load: -0.33
8. I find it difficult to keep up in class. Component load: -0.39

B Descriptives

B.1 Full and final sample

The full sample consists of 8,792 students, 380 classes, and 258 schools. The final sample only includes schools with at least two classes, and consists of 4452 students, 210 classes and 88 schools. Since classes were systematically sampled from municipalities, more classes were sampled from large schools. As reported in Emanuelson (1979), all classes in municipalities with less than 17 classes were sampled. This implies that the *final sample* over-represents larger schools and small municipalities.

Table B.1 shows that students in the *final sample* have slightly lower verbal ability, but same general ability, are more likely to be low SES, male (preferences for boys appear to be higher in low SES families), attend play schools rather than day care, and attend larger classes. In terms of choices (see Table B.2), the only difference between the samples is that students in the *final sample* are less likely to choose advanced English courses. A possible explanation is that these courses might not have been provided in smaller municipalities. Finally, educational performance and behavior are statistically indistinguishable between the two samples (see Tables B.3 to B.5).

[Table B.1 about here.]

[Table B.2 about here.]

[Table B.3 about here.]

[Table B.4 about here.]

[Table B.5 about here.]

B.2 Differences by grade assignment

About half of the sample attended schools in municipalities where grades were assigned in grades 3 and 6, on top of grades 7-9. I label these students "early graded". Students in the remaining municipalities could still receive grades in grade 3 (40% did), but did not receive grades in grade 6.[37] I label these students "late graded". Grade postponement was in general favored by left-wing parties: early graded students thus attended schools in more conservative municipalities.

Early graded students have higher cognitive ability in grade 6, and attend larger classes (Table B.6). In terms of education choices (Table B.7), early graded students are more likely to report choosing electives in grade 7-9 according to ability, possibly due to the fact they observed early grades. Actual choices of early graded students slightly differ from those of late graded students: early graded students are less likely to choose advanced English courses in grades 7-9, but more likely to enroll in academic high school tracks. Interestingly, while early graded students perform about 13% standard deviations better in math and English national tests, their grades are virtually identical to those of late graded students. This is consistent with national grade distributions being applied at the school rather than national level, and implies systematic grade distortions. School attainment does not differ between early and late graded students, with one exception: early graded students are less likely to attain STEM education, consistent with higher enrollment in academic high school tracks (Table B.8).

In grade 6 (Table B.9) early graded students display higher cognitive and non-cognitive relative self-concept with respect to late graded students. I find a negative difference in terms of absolute cognitive self-concept, but the difference is not statistically significant. Early graded students might simply be more aware about their relative standing due to the early grades they observed. Finally, early graded students are more likely to frequently do schoolwork at home, appear less confident about attending high school, and receive less parental support. This picture is stable in grades 7 to 10 (Table B.10).

[Table B.6 about here.]

[Table B.7 about here.]

[Table B.8 about here.]

[37] My data records grade assignment only in grade 6.

[Table B.9 about here.]

[Table B.10 about here.]

B.3 Differences by sex

In grade 6 women display higher verbal and thus general ability than men, and are less likely to be assigned disability status (Table B.11). There are significant differences between men and women in both choice protocols and actual choices (Table B.12). With respect to men, women are 16% standard deviations more likely to mention external factors (parents, teachers, and peers) as drivers of elective choices, 3 percentage points less likely to mention academic ability in math, and 6 percentage points more likely to mention academic ability in English as drivers of elective choices than men. This pattern matches actual choices: women are 14 percentage points more more likely to choose advanced English courses through grades 7-9. Their math elective choices are initially indistinguishable from men's, but over time women shy away from advanced math courses more than men do. These patterns seem consistent with traditional gender norms associating language and math proficiency respectively to women and men. Men are up to three times more likely than women to receive special education status, which can depend both on differences in ability and gendered parental attitudes towards special education status. Finally, women's high school choices seem to depend more on ability (+10 percentage points) and less on preferences (-3 percentage points) than is the case for men. Women are as likely as men to enroll into high school, but 11 percentage points more likely to enroll in academic tracks.

Let's now consider academic achievement (Table B.13): women have 50% standard deviations higher grades in courses with national exams, 40% standard deviations higher grades in locally examined courses, while differences in overall national exams test scores are lower (11% standard deviations higher). The overall positive gap in national tests is mostly explained by Swedish, since there is no statistically significant difference in English performance, and women actually perform 10% standard deviations lower in national math tests. The positive gap decreases in high school, where women's GPA is 23% standard deviations higher than men's. In terms of final education, women are only 2 percentage points more likely than men to get a high school diploma, but are respectively 10 and 3 percentage points more likely to graduate from an academic track and get a university degree. An important point is that only 9% of female attain STEM education, against 54% of men, implying extreme segregation in education choices.[38]

Grade 6 behavior reveals some interesting differences (Table B.14). Despite higher cognitive ability and academic performance, women assess their academic skills in the same way men do. In relative terms, women rate themselves 10 percentage points lower than men do, and the same holds in the non-cognitive area (-13 percentage points).[39] This could be interpreted as women being less overconfident than men (Croson and Gneezy [2009]). Women appear both more motivated and conscientious than men (they do their homework more frequently), and are more likely to report they will attend high school. They have better interactions with peers (+3 percentage points) and receive slightly more support at home from parents (+3 percentage points).

In late compulsory school women assess their cognitive skills 11 percentage points higher than men do in absolute terms (Table B.15), possibly due to the higher grades they observe. Their school motivation remains higher than men (+22% standard deviations), but they no longer get more support at home from parents, and seem to have worse peer interactions than men do. The positive gaps reverse in high school: women have 18% standard deviations lower absolute self-assessment in the cognitive domain, 24% standard deviations lower overall student welfare, and are more pessimistic about finishing high school. A final point to stress is that men are 18% standard deviations less likely to report being graded fairly than women in grades 7-9. Taken together with the difference between test and grade performance, this might signal grading distortions or biases working against men in compulsory school. One must however notice that grade setting in Sweden also reflects effort put in school, higher among women.

[Table B.11 about here.]

[Table B.12 about here.]

[Table B.13 about here.]

[Table B.14 about here.]

[Table B.15 about here.]

[38]Note: STEM include vocational industrial courses.

[39]Notice that most items in the relative non-cognitive self-concept scale relate to leadership.

B.4 Differences by SES

Low SES students, defined according to the Swedish census definition as children of working class parents, make up 41% of the final sample. In grade 6 these students have about 40% standard deviations lower cognitive ability with respect to the rest of the students, and are 2.5 times more likely to be disabled (Table B.16). They have slightly higher age, implying some of them were retained in previous grades. In terms of educational investments, low SES students are 4 percentage points less likely to have attended day care, study in classes of similar size to that of other students, but are more likely to have changed teachers in grades 1-6. The fraction of low SES peers is higher for low SES students, implying some segregation at the school level.

Table B.17 shows that low SES students are 14% standard deviations more likely to mention external factors (parents, teachers, and peers) and about 7 percentage points less likely to mention academic ability as drivers of elective choices; they however are more likely to base high school choices on academic ability with respect to other students. In terms of actual choices, low SES students are around 2 times more likely to be assigned special education status than other students at all grades, are 22 percentage points less likely to choose advanced electives in grades 7-9, 9 percentage points less likely to enroll in high school, and 50% less likely than the rest of the students to attend academic high school (27 percentage points against 54 percentage points).

With regard to academic achievement in compulsory school (Table B.18), grades in subjects without national test, subjects with national tests, and national test scores are respectively 50%, 42% and 38% standard deviations lower for low SES students. High school grades are around 31% standard deviations lower, consistently with positive selection into high school. Final attainment of low SES students reflects initial choices: these students are 10 percentage points less likely to graduate from high school, 50% less likely to complete an academic track, and 58% less likely to attain university education than the rest of the students. Due to overrepresentation in vocational education, low SES students are slightly more likely to obtain a STEM diploma.

Table B.19 shows that in grade 6 low SES students assess their abilities 20 to 25% standard deviations below the rest of the students (especially in the cognitive domain), do slightly less frequently school work at home, are slightly more motivated at school than the rest of the students but far less likely to report they will attend high school (14 percentage points less, below the difference in actual choices). Finally, low SES students report slightly lower teacher (-2 percentage points) and parental (-4 percentage points) support. This picture is generally stable in grade 7-9 (Table B.20), when however school motivation and parental support become much lower than the rest of the students (-16% standard deviations and -12 percentage points). During high school, low SES students reduce the negative gap in self-concept (now -10% standard deviations), exhibit slightly higher school well-being, and are more likely than other students to report they will complete high school. A potential explanation is that low SES students disproportionally enroll into vocational tracks in high school, and enjoy the benefits of a more homogeneous peer group. Finally, low SES students are 11% standard deviations less likely to consider their grades in late compulsory school fair. Taken together with the low teacher support and grade pattern, this might signal grading bias or grading distortions penalizing low SES students in compulsory school.

[Table B.16 about here.]

[Table B.17 about here.]

[Table B.18 about here.]

[Table B.19 about here.]

[Table B.20 about here.]

C Balance Tests

[Table C.1 about here.]

[Table C.2 about here.]

[Table C.3 about here.]

D Simulations

[Figure D.1 about here.]

[Figure D.2 about here.]

[Figure D.3 about here.]

[Figure D.4 about here.]

E All Results

E.1 Whole sample

[Table E.1 about here.]

[Table E.2 about here.]

[Table E.3 about here.]

[Table E.4 about here.]

[Table E.5 about here.]

[Table E.6 about here.]

E.2 Heterogeneity

[Table E.7 about here.]

[Table E.8 about here.]

[Table E.9 about here.]

[Table E.10 about here.]

[Table E.11 about here.]

[Table E.12 about here.]

[Table E.13 about here.]

[Table E.14 about here.]

[Table E.15 about here.]

[Table E.16 about here.]

[Table E.17 about here.]

[Table E.18 about here.]

Tables and Pictures

Table 1
Descriptives by classes sampled per school

Sampled	Number of			Class	
classes	schools	students	size	ability (μ)	ability (sd)
1	170	3786	22.27	22.88	0.00
2	63	3001	23.82	22.51	1.55
3	17	1178	23.10	22.75	1.73
4	7	686	24.50	22.96	0.97
5	1	141	28.20	20.07	1.21

The table reports sample characteristics broken down by the number of classes sampled per school. Ability is the average of IQ-like test scores for inductive and verbal ability taken at the end of grade 6. Ability (standard deviations) is the within-school standard deviation of mean class ability, ability (μ).

Table 2
Summary of results

	(1)	(2)	(3)	Mean
Self-concept index (PC, grade 6-9)	-0.043***	-0.045***	-0.046***	-0.006
	(0.014)	(0.014)	(0.014)	(1.00)
Well-being index (PC, grade 6-9)	-0.004	-0.005	-0.006	0.007
	(0.015)	(0.016)	(0.016)	(0.95)
Academic choices index (PC, grade 7-9)	-0.030**	-0.033***	-0.034***	0.015
	(0.012)	(0.012)	(0.012)	(0.99)
National tests (std, grade 8-9)	0.025*	0.025*	0.023*	0.010
	(0.014)	(0.014)	(0.014)	(0.98)
GPA: nat. graded (std, grade 7-9)	-0.006	-0.009	-0.012	0.027
	(0.012)	(0.012)	(0.011)	(0.99)
GPA: loc. graded (std, grade 7-9)	-0.020*	-0.023**	-0.026**	0.021
	(0.011)	(0.011)	(0.011)	(0.99)
Well-being index (PC, grade 10)	-0.033*	-0.033*	-0.030*	-0.004
	(0.018)	(0.018)	(0.017)	(0.98)
GPA (std, high school)	-0.024*	-0.027**	-0.029**	-0.003
	(0.013)	(0.013)	(0.012)	(0.97)
Years of non-compulsory (-35)	0.005	0.001	-0.005	3.122
	(0.025)	(0.025)	(0.022)	(2.09)
N. of students (mean)	3945	3945	3945	3945
N. of classes	204	204	204	204
N. of schools	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability and school fixed effects. Specification (1) includes Guryan et al. (2009) pool correction, and corresponds to equation 2. Specification (2) replaces the pool correction with own ability. Specification (3) adds to specification (2) individual controls. Class ability is the average of verbal and inductive ability measures taken in grade 6. National tests is the average of Swedish, English and math test scores. All indexes are described in Appendix A.2. Standard errors are clustered at the class level.

Table 3
Summary of results: heterogeneity

	Grading		Sex		SES	
	Late	Early	Male	Female	High	Low
Self-concept index (PC, grade 6-9)	-0.043**	-0.046**	-0.049***	-0.055***	-0.016	-0.091***
	(0.017)	(0.022)	(0.018)	(0.019)	(0.016)	(0.025)
Well-being index (PC, grade 6-9)	-0.011	-0.002	-0.003	-0.021	0.013	-0.039
	(0.021)	(0.025)	(0.023)	(0.022)	(0.017)	(0.026)
Academic choices index (PC, grade 7-9)	-0.029*	-0.036*	-0.044**	-0.025**	-0.013	-0.066***
	(0.017)	(0.018)	(0.018)	(0.013)	(0.012)	(0.024)
National tests (std, grade 8-9)	0.014	0.035	0.017	0.041*	0.029*	0.027
	(0.017)	(0.022)	(0.016)	(0.022)	(0.015)	(0.021)
GPA: nat. graded (std, grade 7-9)	-0.004	-0.019	-0.019	-0.007	-0.009	-0.013
	(0.014)	(0.019)	(0.014)	(0.021)	(0.015)	(0.016)
GPA: loc. graded (std, grade 7-9)	-0.008	-0.043***	-0.038***	-0.017	-0.011	-0.052***
	(0.015)	(0.014)	(0.014)	(0.019)	(0.014)	(0.018)
Well-being index (PC, grade 10)	-0.058**	0.001	-0.030	-0.029	-0.021	-0.029
	(0.023)	(0.027)	(0.021)	(0.025)	(0.021)	(0.030)
GPA (std, high school)	-0.020	-0.034*	-0.026	-0.053**	-0.022	-0.031
	(0.016)	(0.019)	(0.019)	(0.020)	(0.015)	(0.021)
Years of non-compulsory (-35)	-0.063**	0.061**	0.010	-0.016	0.032	-0.041
	(0.031)	(0.030)	(0.035)	(0.040)	(0.032)	(0.031)
N. of students (mean)	1886	2059	1995	1950	2318	1476
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability, school fixed effects, own ability and individual controls, for different subsamples. Class ability is the average of verbal and inductive ability measures taken in grade 6. National tests is the average of Swedish, English and math test scores. All indexes are described in Appendix A.2. Standard errors are clustered at the class level.

Table 4
Summary of Results: heterogeneity tests

	Grading		Sex		SES	
	Late	Δ Early	Male	Δ Female	High	Δ Low
Self-concept index (PC, grade 6-9)	-0.043**	-0.003	-0.049***	-0.006	-0.016	-0.075***
	(0.017)	(0.028)	(0.018)	(0.024)	(0.016)	(0.026)
Well-being index (PC, grade 6-9)	-0.011	0.009	-0.003	-0.018	0.013	-0.053**
	(0.021)	(0.032)	(0.023)	(0.032)	(0.018)	(0.026)
Academic choices index (PC, grade 7-9)	-0.029*	-0.007	-0.044**	0.019	-0.013	-0.053**
	(0.017)	(0.025)	(0.018)	(0.020)	(0.012)	(0.025)
National tests (std, grade 8-9)	0.014	0.021	0.017	0.024	0.029*	-0.002
	(0.017)	(0.027)	(0.016)	(0.025)	(0.015)	(0.020)
GPA: nat. graded (std, grade 7-9)	-0.004	-0.015	-0.019	0.013	-0.009	-0.004
	(0.014)	(0.023)	(0.014)	(0.027)	(0.015)	(0.018)
GPA: loc. graded (std, grade 7-9)	-0.008	-0.035*	-0.038***	0.021	-0.011	-0.041**
	(0.015)	(0.021)	(0.014)	(0.025)	(0.014)	(0.020)
Well-being index (PC, grade 10)	-0.058**	0.059*	-0.030	0.000	-0.021	-0.008
	(0.023)	(0.035)	(0.021)	(0.031)	(0.021)	(0.033)
GPA (std, high school)	-0.020	-0.015	-0.026	-0.027	-0.022	-0.010
	(0.016)	(0.025)	(0.019)	(0.030)	(0.015)	(0.025)
Years of non-compulsory (-35)	-0.063**	0.124***	0.010	-0.026	0.032	-0.073*
	(0.031)	(0.043)	(0.035)	(0.058)	(0.032)	(0.042)
N. of students (mean)	3945	3945	3945	3945	3794	3794
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns 1, 3, 5 of the table report the coefficient for peer ability, expressed in within-school between-class standard deviations, for late graded, male, and high SES students. Columns 2, 4, 6 report the coefficient for the interaction between peer ability and early graded, female, and low SES status. All coefficients are from regressions of each outcome on peer ability, school fixed effects, own ability and individual controls, fully interacted by category. Class ability is the average of verbal and inductive ability measures taken in grade 6. National tests is the average of Swedish, English and math test scores. All indexes are described in Appendix A.2. Standard errors are clustered at the class level.

Table B.1
Student background by sample type

Sample:	Final	Full	Difference
Individual background			
General ability (0-40, grade 6)	22.52 (6.1)	22.65 (6.2)	−0.13 [0.23]
Verbal ability (0-40, grade 6)	22.79 (6.0)	22.95 (6.0)	−0.16 [0.09]
Inductive ability (0-40, grade 6)	22.27 (8.1)	22.36 (8.2)	−0.09 [0.51]
Disability	0.07 (0.3)	0.07 (0.3)	0.00 [0.22]
Female	0.48 (0.5)	0.49 (0.5)	−0.01 [0.05]
Age (grade 6)	12.93 (0.3)	12.93 (0.3)	−0.00 [0.86]
Born abroad	0.03 (0.2)	0.04 (0.2)	−0.00 [0.14]
Low SES	0.41 (0.5)	0.40 (0.5)	0.01 [0.08]
Parent not Nordic	0.06 (0.2)	0.06 (0.2)	−0.00 [0.19]
Day care	0.10 (0.3)	0.11 (0.3)	−0.02 [0.00]
Play school	0.86 (0.3)	0.83 (0.4)	0.03 [0.00]
Class background			
% female (-i, grade 6)	48.22 (9.9)	49.04 (9.9)	−0.82 [0.05]
% low SES (-i, grade 6)	41.42 (16.7)	40.10 (17.3)	1.32 [0.09]
Class size (grade 6)	24.46 (3.7)	24.04 (4.2)	0.42 [0.02]
Teacher changes (grade 1-6)	0.52 (0.8)	0.50 (0.8)	0.02 [0.64]

Standard errors are in parentheses and p-values are in brackets. The full sample consists of 8,792 observations. The final sample only includes schools with at least two classes, and consists of 4,452 observations. Ability measures are IQ-like scores measured at the end of grade 6, and general ability is the average of inductive and verbal ability. All the other information comes from administrative records.

Table B.2
Educational choices by sample type

Sample:	Final	Full	Difference
Choice protocols			
Extrinsic choices scale (0-100, grade 6)	41.76 (26.4)	41.15 (26.4)	0.62 [0.12]
Math choice: Ability (grade 6)	0.71 (0.3)	0.71 (0.3)	0.00 [0.66]
English choice: Ability (grade 6)	0.71 (0.3)	0.70 (0.3)	0.00 [0.38]
High school choice: Ability (grade 10)	0.61 (0.3)	0.60 (0.3)	0.00 [0.24]
High school choice: Preferences (grade 10)	0.80 (0.3)	0.80 (0.3)	0.00 [0.96]
Actual choices			
Advanced Math (grade 7)	0.74 (0.4)	0.74 (0.4)	−0.01 [0.27]
Advanced Math (grade 8)	0.65 (0.5)	0.66 (0.5)	−0.01 [0.17]
Advanced Math (grade 9)	0.56 (0.5)	0.56 (0.5)	0.00 [0.99]
Advanced English (grade 7)	0.75 (0.4)	0.77 (0.4)	−0.01 [0.04]
Advanced English (grade 8)	0.72 (0.4)	0.74 (0.4)	−0.01 [0.03]
Advanced English (grade 9)	0.68 (0.5)	0.69 (0.5)	−0.01 [0.13]
Special education (grade 6)	0.13 (0.3)	0.13 (0.3)	0.00 [0.85]
Special education (grade 7)	0.08 (0.3)	0.08 (0.3)	0.00 [0.84]
Special education (grade 8)	0.03 (0.2)	0.03 (0.2)	0.00 [0.44]
Special education (grade 9)	0.02 (0.1)	0.02 (0.1)	−0.00 [0.55]
High school enrollment (15-18)	0.86 (0.4)	0.85 (0.4)	0.00 [0.50]
Academic track enrollment (15-18)	0.42 (0.5)	0.43 (0.5)	−0.01 [0.36]

Standard errors are in parentheses and p-values are in brackets. The full sample consists of 8,792 observations. The final sample only includes schools with at least two classes, and consists of 4,452 observations. Choice protocols are derived from the grade 6 student survey, while actual choices come from administrative records.

Table B.3
School performance and attainment by sample type

Sample:	Final	Full	Difference
School performance			
English national test (0-100, grade 8)	68.41 (16.0)	68.40 (16.0)	0.01 [0.97]
Math national test (0-100, grade 9)	62.61 (15.5)	62.60 (15.9)	0.01 [0.98]
GPA: nat. graded (1-5, grade 7)	3.01 (0.8)	3.01 (0.8)	0.00 [0.97]
GPA: nat. graded (1-5, grade 8)	3.04 (0.8)	3.04 (0.8)	−0.00 [0.88]
GPA: nat. graded (1-5, grade 9)	3.16 (0.8)	3.17 (0.8)	−0.00 [0.75]
GPA: loc. graded (1-5, grade 7)	3.05 (0.6)	3.05 (0.6)	−0.00 [0.75]
GPA: loc. graded (1-5, grade 8)	3.09 (0.7)	3.09 (0.7)	−0.00 [0.59]
GPA: loc. graded (1-5, grade 9)	3.22 (0.7)	3.23 (0.8)	−0.01 [0.45]
GPA (1-5, high school)	3.23 (0.6)	3.24 (0.6)	−0.01 [0.25]
School attainment			
High school diploma (17-20)	0.79 (0.4)	0.78 (0.4)	0.00 [0.62]
Completes academic track (17-20)	0.38 (0.5)	0.38 (0.5)	−0.00 [0.66]
STEM education (-35)	0.33 (0.5)	0.32 (0.5)	0.01 [0.15]
High school diploma (-35)	0.86 (0.3)	0.86 (0.3)	0.00 [0.34]
University degree (-35)	0.14 (0.3)	0.15 (0.4)	−0.00 [0.31]
Years of non-compulsory (-35)	3.08 (2.1)	3.10 (2.1)	−0.01 [0.63]

Standard errors are in parentheses and p-values are in brackets. The full sample consists of 8,792 observations. The final sample only includes schools with at least two classes, and consists of 4,452 observations. All information comes from administrative records.

Table B.4
Grade 6 behavior by sample type

Sample:	Final	Full	Difference
Individual			
Abs. cogn. SC scale (0-100, grade 6)	59.82 (26.0)	60.01 (25.9)	−0.19 [0.58]
Rel. cogn. SC scale (0-100, grade 6)	57.41 (14.7)	57.55 (14.9)	−0.15 [0.48]
Rel. non cogn. SC scale (0-100, grade 6)	56.48 (13.8)	56.74 (13.9)	−0.26 [0.19]
I do my best at school (grade 6)	0.66 (0.5)	0.64 (0.5)	0.01 [0.02]
Frequent schoolwork at home (grade 6)	0.62 (0.3)	0.62 (0.3)	−0.00 [0.63]
Will attend high school (grade 6)	0.73 (0.4)	0.73 (0.4)	−0.00 [0.82]
School and family			
Helped at home (grade 6)	0.80 (0.4)	0.80 (0.4)	0.00 [0.90]
Helped by teachers (grade 6)	0.76 (0.4)	0.76 (0.4)	−0.01 [0.44]
Positive peer interactions (grade 6)	0.91 (0.3)	0.90 (0.3)	0.00 [0.31]

Standard errors are in parentheses and p-values are in brackets. The full sample consists of 8,792 observations. The final sample only includes schools with at least two classes, and consists of 4,452 observations. All information, apart from special education status, comes from the grade 6 student survey. Standardized variables are indexes built as the average of relevant survey questions, recodified on a 0-1 scale, and set to missing if less than half of the questions are answered. SC stands for self-concept, a measure of assessment of own abilities or skills. Details on each index are discussed in Appendix A.

Table B.5
Grade 7-10 behavior by sample type

Sample:	Final	Full	Difference
Individual			
Abs. cogn. SC scale (0-100, grade 7-9)	82.08 (16.1)	82.12 (16.2)	−0.04 [0.85]
Motivation scale (0-100, grade 7-9)	65.15 (25.5)	65.46 (25.4)	−0.31 [0.33]
Grades fairness (0-100, grade 7-9)	45.37 (17.4)	45.33 (17.7)	0.04 [0.82]
Abs. cogn. SC scale (0-100, grade 10)	80.48 (15.7)	80.57 (15.7)	−0.09 [0.64]
Well-being index (0-100, grade 10)	82.56 (18.4)	82.80 (18.6)	−0.24 [0.31]
Will complete high school (grade 10)	0.92 (0.3)	0.92 (0.3)	−0.00 [0.53]
School and family			
Helped at home (grade 7-9)	0.82 (0.4)	0.82 (0.4)	−0.00 [0.84]
Teacher support scale (0-100, grade 7-9)	76.78 (16.0)	76.88 (16.0)	−0.10 [0.62]
Positive peer interactions (grade 7-9)	0.90 (0.2)	0.90 (0.2)	0.00 [0.90]

Standard errors are in parentheses and p-values are in brackets. The full sample consists of 8,792 observations. The final sample only includes schools with at least two classes, and consists of 4,452 observations. All information, apart from special education status, comes from the grade 6 student survey. Standardized variables are indexes built as the average of relevant survey questions, recodified on a 0-1 scale, and set to missing if less than half of the questions are answered. SC stands for self-concept, a measure of assessment of own abilities or skills. Details on each index are discussed in Appendix A.

Table B.6
Differences in background by grades assignment

Grade assignment:	Early	Late	Difference
Individual background			
General ability (std, grade 6)	0.07 (1.0)	−0.12 (1.0)	0.19 [0.00]
Verbal ability (std, grade 6)	0.02 (1.0)	−0.08 (1.0)	0.10 [0.03]
Inductive ability (std, grade 6)	0.09 (1.0)	−0.12 (1.0)	0.21 [0.00]
Disability	0.07 (0.2)	0.08 (0.3)	−0.01 [0.33]
Female	0.48 (0.5)	0.48 (0.5)	0.00 [1.00]
Age (grade 6)	12.93 (0.3)	12.93 (0.3)	−0.00 [0.93]
Born abroad	0.03 (0.2)	0.04 (0.2)	−0.01 [0.15]
Low SES	0.42 (0.5)	0.41 (0.5)	0.01 [0.62]
Parent not Nordic	0.06 (0.2)	0.06 (0.2)	−0.00 [0.57]
Day care	0.10 (0.3)	0.09 (0.3)	0.01 [0.34]
Play school	0.85 (0.4)	0.87 (0.3)	−0.02 [0.16]
Class background			
% female (-i, grade 6)	48.23 (9.6)	48.22 (10.2)	0.01 [1.00]
% low SES (-i, grade 6)	41.97 (14.8)	40.83 (18.5)	1.13 [0.63]
Class size (grade 6)	24.85 (3.7)	24.04 (3.5)	0.82 [0.09]
Teacher changes (grade 1-6)	0.55 (0.8)	0.48 (0.9)	0.07 [0.53]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 4,452 observations. Ability measures are IQ-like scores measured at the end of grade 6, and general ability is the average of inductive and verbal ability. All the other information comes from administrative records.

Table B.7
Differences in educational choices by grades assignment

Grade assignment:	Early	Late	Difference
Choice protocols			
Extrinsic choices scale (grade 6)	0.00 (1.0)	0.04 (1.0)	−0.04 [0.38]
Math choice: Ability (grade 6)	0.72 (0.3)	0.70 (0.3)	0.02 [0.12]
English choice: Ability (grade 6)	0.72 (0.3)	0.70 (0.3)	0.02 [0.07]
High school choice: Ability (grade 10)	0.61 (0.3)	0.61 (0.3)	0.00 [0.85]
High school choice: Preferences (grade 10)	0.80 (0.3)	0.80 (0.3)	0.01 [0.58]
Actual choices			
Advanced Math (grade 7)	0.73 (0.4)	0.74 (0.4)	−0.01 [0.73]
Advanced Math (grade 8)	0.65 (0.5)	0.65 (0.5)	−0.00 [0.94]
Advanced Math (grade 9)	0.56 (0.5)	0.55 (0.5)	0.01 [0.56]
Advanced English (grade 7)	0.73 (0.4)	0.78 (0.4)	−0.04 [0.01]
Advanced English (grade 8)	0.70 (0.5)	0.75 (0.4)	−0.05 [0.00]
Advanced English (grade 9)	0.66 (0.5)	0.71 (0.5)	−0.05 [0.01]
Special education (grade 6)	0.13 (0.3)	0.14 (0.3)	−0.01 [0.76]
Special education (grade 7)	0.07 (0.3)	0.10 (0.3)	−0.02 [0.10]
Special education (grade 8)	0.03 (0.2)	0.03 (0.2)	−0.00 [0.54]
Special education (grade 9)	0.02 (0.1)	0.02 (0.1)	0.00 [0.53]
High school enrollment (15-18)	0.85 (0.4)	0.86 (0.4)	−0.00 [0.84]
Academic track enrollment (15-18)	0.44 (0.5)	0.40 (0.5)	0.04 [0.06]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 4,452 observations. Choice protocols are derived from the grade 6 student survey, while actual choices are taken from administrative records.

Table B.8
Differences in school performance and attainment by grades assignment

Grade assignment:	Early	Late	Difference
School performance			
English national test (std, grade 8)	0.06 (1.0)	−0.05 (1.0)	0.11 [0.04]
Math national test (std, grade 9)	0.06 (1.0)	−0.08 (1.0)	0.15 [0.02]
GPA: nat. graded (std, grade 7)	0.03 (1.0)	−0.03 (1.0)	0.07 [0.07]
GPA: nat. graded (std, grade 8)	0.02 (1.0)	−0.03 (1.0)	0.05 [0.22]
GPA: nat. graded (std, grade 9)	0.02 (1.0)	−0.03 (1.0)	0.05 [0.15]
GPA: loc. graded (std, grade 7)	0.02 (1.0)	−0.03 (0.9)	0.05 [0.18]
GPA: loc. graded (std, grade 8)	0.00 (1.0)	−0.02 (1.0)	0.02 [0.52]
GPA: loc. graded (std, grade 9)	−0.01 (1.0)	−0.01 (1.0)	0.00 [0.97]
GPA (std, high school)	−0.03 (1.0)	−0.01 (1.0)	−0.02 [0.64]
School attainment			
High school diploma (17-20)	0.78 (0.4)	0.79 (0.4)	−0.01 [0.67]
Completes academic track (17-20)	0.39 (0.5)	0.37 (0.5)	0.03 [0.16]
STEM education (-35)	0.31 (0.5)	0.34 (0.5)	−0.03 [0.06]
High school diploma (-35)	0.87 (0.3)	0.86 (0.3)	0.01 [0.51]
University degree (-35)	0.14 (0.4)	0.14 (0.3)	0.00 [0.71]
Years of non-compulsory (-35)	3.12 (2.1)	3.04 (2.1)	0.08 [0.32]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 4,452 observations. All information comes from administrative records.

Table B.9
Differences in grade 6 behavior by grades assignment

Grade assignment:	Early	Late	Difference
Individual			
Abs. cogn. SC scale (grade 6)	−0.03 (1.0)	0.02 (1.0)	−0.06 [0.15]
Rel. cogn. SC scale (grade 6)	0.04 (1.0)	−0.06 (1.0)	0.10 [0.01]
Rel. non cogn. SC scale (grade 6)	0.04 (1.0)	−0.08 (1.0)	0.11 [0.01]
I do my best at school (grade 6)	0.65 (0.5)	0.66 (0.5)	−0.00 [0.80]
Frequent schoolwork at home (grade 6)	0.67 (0.3)	0.57 (0.3)	0.09 [0.00]
Will attend high school (grade 6)	0.71 (0.5)	0.74 (0.4)	−0.03 [0.06]
School and family			
Helped at home (grade 6)	0.78 (0.4)	0.82 (0.4)	−0.04 [0.01]
Helped by teachers (grade 6)	0.76 (0.4)	0.75 (0.4)	0.01 [0.56]
Positive peer interactions (grade 6)	0.91 (0.3)	0.90 (0.3)	0.01 [0.33]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 4,452 observations. All information, apart from special education status, comes from the grade 6 student survey. Standardized variables are indexes built as the average of relevant survey questions, recodified on a 0-1 scale, and set to missing if less than half of the questions are answered. SC stands for self-concept, a measure of assessment of own abilities or skills. Details on each index are discussed in Appendix A.

Table B.10
Differences in grades 7-10 behavior by grades assignment

Grade assignment:	Early	Late	Difference
Individual			
Abs. cogn. SC scale (grade 7-9)	−0.00 (1.0)	−0.01 (1.0)	0.01 [0.85]
Motivation scale (grade 7-9)	−0.01 (1.0)	−0.01 (1.0)	−0.00 [0.93]
Grades fairness (std, grade 7-9)	−0.02 (1.0)	0.03 (1.0)	−0.05 [0.11]
Abs. cogn. SC scale (grade 10)	−0.01 (1.1)	−0.01 (1.0)	−0.00 [0.98]
Well-being index (grade 10)	−0.03 (1.0)	0.00 (1.0)	−0.03 [0.41]
Will complete high school (grade 10)	0.92 (0.3)	0.92 (0.3)	−0.00 [0.81]
School and family			
Helped at home (grade 7-9)	0.81 (0.4)	0.84 (0.4)	−0.03 [0.04]
Teacher support scale (grade 7-9)	−0.03 (1.0)	0.01 (1.0)	−0.04 [0.34]
Positive peer interactions (grade 7-9)	0.90 (0.2)	0.90 (0.2)	0.00 [0.60]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 4,452 observations. All information, apart from special education status, comes from the grade 6 student survey. Standardized variables are indexes built as the average of relevant survey questions, recodified on a 0-1 scale, and set to missing if less than half of the questions are answered. SC stands for self-concept, a measure of assessment of own abilities or skills. Details on each index are discussed in Appendix A.

Table B.11
Differences in background by sex

Sex:	Female	Male	Difference
Individual background			
General ability (std, grade 6)	0.02 (1.0)	−0.06 (1.0)	0.07 [0.05]
Verbal ability (std, grade 6)	0.05 (1.0)	−0.10 (1.0)	0.16 [0.00]
Inductive ability (std, grade 6)	−0.01 (0.9)	−0.01 (1.0)	0.00 [1.00]
Disability	0.05 (0.2)	0.10 (0.3)	−0.05 [0.00]
Age (grade 6)	12.92 (0.3)	12.93 (0.3)	−0.01 [0.34]
Born abroad	0.04 (0.2)	0.03 (0.2)	0.00 [0.42]
Low SES	0.40 (0.5)	0.42 (0.5)	−0.02 [0.13]
Parent not Nordic	0.05 (0.2)	0.06 (0.2)	−0.01 [0.10]
Day care	0.09 (0.3)	0.10 (0.3)	−0.00 [0.84]
Play school	0.87 (0.3)	0.85 (0.4)	0.01 [0.30]
Class background			
% female (-i, grade 6)	47.97 (9.6)	48.46 (10.2)	−0.49 [0.20]
% low SES (-i, grade 6)	41.49 (16.5)	41.36 (16.8)	0.13 [0.77]
Class size (grade 6)	24.39 (3.6)	24.52 (3.7)	−0.13 [0.21]
Teacher changes (grade 1-6)	0.52 (0.8)	0.52 (0.8)	0.01 [0.74]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 4,452 observations. Ability measures are IQ-like scores measured at the end of grade 6, and general ability is the average of inductive and verbal ability. All the other information comes from administrative records.

Table B.12
Differences in educational choices by sex

Sex:	Female	Male	Difference
Choice protocols			
Extrinsic choices scale (grade 6)	−0.06 (1.0)	0.10 (1.0)	−0.16 [0.00]
Math choice: Ability (grade 6)	0.69 (0.3)	0.73 (0.3)	−0.03 [0.00]
English choice: Ability (grade 6)	0.74 (0.3)	0.68 (0.3)	0.06 [0.00]
High school choice: Ability (grade 10)	0.65 (0.3)	0.55 (0.4)	0.10 [0.00]
High school choice: Preferences (grade 10)	0.78 (0.3)	0.82 (0.3)	−0.03 [0.00]
Actual choices			
Advanced Math (grade 7)	0.75 (0.4)	0.73 (0.4)	0.02 [0.28]
Advanced Math (grade 8)	0.64 (0.5)	0.66 (0.5)	−0.01 [0.37]
Advanced Math (grade 9)	0.54 (0.5)	0.57 (0.5)	−0.03 [0.06]
Advanced English (grade 7)	0.83 (0.4)	0.69 (0.5)	0.14 [0.00]
Advanced English (grade 8)	0.79 (0.4)	0.66 (0.5)	0.13 [0.00]
Advanced English (grade 9)	0.75 (0.4)	0.62 (0.5)	0.13 [0.00]
Special education (grade 6)	0.10 (0.3)	0.17 (0.4)	−0.07 [0.00]
Special education (grade 7)	0.05 (0.2)	0.12 (0.3)	−0.07 [0.00]
Special education (grade 8)	0.02 (0.1)	0.04 (0.2)	−0.02 [0.00]
Special education (grade 9)	0.01 (0.1)	0.03 (0.2)	−0.02 [0.00]
High school enrollment (15-18)	0.86 (0.3)	0.85 (0.4)	0.02 [0.11]
Academic track enrollment (15-18)	0.48 (0.5)	0.37 (0.5)	0.11 [0.00]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 4,452 observations. Choice protocols are derived from the grade 6 student survey, while actual choices are taken from administrative records.

Table B.13
Differences in school performance and attainment by sex

Sex:	Female	Male	Difference
School performance			
English national test (std, grade 8)	0.02 (1.0)	−0.02 (1.0)	0.04 [0.34]
Math national test (std, grade 9)	−0.05 (0.9)	0.05 (1.0)	−0.10 [0.01]
GPA: nat. graded (std, grade 7)	0.23 (1.0)	−0.22 (1.0)	0.45 [0.00]
GPA: nat. graded (std, grade 8)	0.26 (1.0)	−0.24 (1.0)	0.50 [0.00]
GPA: nat. graded (std, grade 9)	0.26 (0.9)	−0.25 (1.0)	0.52 [0.00]
GPA: loc. graded (std, grade 7)	0.18 (1.0)	−0.18 (1.0)	0.36 [0.00]
GPA: loc. graded (std, grade 8)	0.18 (1.0)	−0.18 (1.0)	0.36 [0.00]
GPA: loc. graded (std, grade 9)	0.23 (0.9)	−0.23 (1.0)	0.46 [0.00]
GPA (std, high school)	0.10 (1.0)	−0.13 (1.0)	0.23 [0.00]
School attainment			
High school diploma (17-20)	0.80 (0.4)	0.78 (0.4)	0.02 [0.09]
Completes academic track (17-20)	0.43 (0.5)	0.33 (0.5)	0.10 [0.00]
STEM education (-35)	0.09 (0.3)	0.54 (0.5)	−0.45 [0.00]
High school diploma (-35)	0.87 (0.3)	0.85 (0.4)	0.02 [0.08]
University degree (-35)	0.16 (0.4)	0.13 (0.3)	0.03 [0.00]
Years of non-compulsory (-35)	3.25 (2.1)	2.92 (2.1)	0.33 [0.00]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 4,452 observations. All information comes from administrative records.

Table B.14
Differences in grade 6 behavior by sex

Sex:	Female	Male	Difference
Individual			
Abs. cogn. SC scale (grade 6)	0.02 (1.0)	−0.03 (1.0)	0.04 [0.22]
Rel. cogn. SC scale (grade 6)	−0.05 (1.0)	0.03 (1.0)	−0.08 [0.02]
Rel. non cogn. SC scale (grade 6)	−0.08 (0.9)	0.04 (1.0)	−0.13 [0.00]
I do my best at school (grade 6)	0.71 (0.5)	0.61 (0.5)	0.09 [0.00]
Frequent schoolwork at home (grade 6)	0.66 (0.3)	0.58 (0.3)	0.08 [0.00]
Will attend high school (grade 6)	0.78 (0.4)	0.68 (0.5)	0.09 [0.00]
School and family			
Helped at home (grade 6)	0.82 (0.4)	0.78 (0.4)	0.03 [0.01]
Helped by teachers (grade 6)	0.77 (0.4)	0.75 (0.4)	0.02 [0.16]
Positive peer interactions (grade 6)	0.92 (0.3)	0.89 (0.3)	0.03 [0.01]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 4,452 observations. All information, apart from special education status, comes from the grade 6 student survey. Standardized variables are indexes built as the average of relevant survey questions, recodified on a 0-1 scale, and set to missing if less than half of the questions are answered. SC stands for self-concept, a measure of assessment of own abilities or skills. Details on each index are discussed in Appendix A.

Table B.15
Differences in grades 7-10 behavior by sex

Sex:	Female	Male	Difference
Individual			
Abs. cogn. SC scale (grade 7-9)	0.05 (0.9)	−0.06 (1.0)	0.11 [0.00]
Motivation scale (grade 7-9)	0.10 (1.0)	−0.13 (1.0)	0.22 [0.00]
Grades fairness (std, grade 7-9)	0.09 (0.9)	−0.09 (1.1)	0.18 [0.00]
Abs. cogn. SC scale (grade 10)	−0.04 (1.0)	0.03 (1.0)	−0.07 [0.05]
Well-being index (grade 10)	−0.14 (1.0)	0.13 (0.9)	−0.26 [0.00]
Will complete high school (grade 10)	0.91 (0.3)	0.94 (0.2)	−0.03 [0.00]
School and family			
Helped at home (grade 7-9)	0.83 (0.4)	0.81 (0.4)	0.02 [0.17]
Teacher support scale (grade 7-9)	−0.00 (1.0)	−0.02 (1.0)	0.02 [0.65]
Positive peer interactions (grade 7-9)	0.89 (0.2)	0.91 (0.2)	−0.02 [0.00]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 4,452 observations. All information, apart from special education status, comes from the grade 6 student survey. Standardized variables are indexes built as the average of relevant survey questions, recodified on a 0-1 scale, and set to missing if less than half of the questions are answered. SC stands for self-concept, a measure of assessment of own abilities or skills. Details on each index are discussed in Appendix A.

Table B.16
Differences in background by SES

SES:	Low	High	Difference
Individual background			
General ability (std, grade 6)	–0.27 (1.0)	0.18 (1.0)	–0.46 [0.00]
Verbal ability (std, grade 6)	–0.28 (1.0)	0.17 (1.0)	–0.45 [0.00]
Inductive ability (std, grade 6)	–0.21 (1.0)	0.15 (1.0)	–0.36 [0.00]
Disability	0.10 (0.3)	0.04 (0.2)	0.05 [0.00]
Female	0.47 (0.5)	0.49 (0.5)	–0.02 [0.13]
Age (grade 6)	12.94 (0.3)	12.91 (0.3)	0.03 [0.01]
Born abroad	0.02 (0.2)	0.03 (0.2)	–0.00 [0.59]
Parent not Nordic	0.06 (0.2)	0.05 (0.2)	0.00 [0.69]
Day care	0.07 (0.3)	0.11 (0.3)	–0.04 [0.00]
Play school	0.86 (0.3)	0.87 (0.3)	–0.01 [0.63]
Class background			
% female (-i, grade 6)	48.31 (9.8)	48.18 (10.0)	0.13 [0.78]
% low SES (-i, grade 6)	45.35 (15.8)	38.33 (16.5)	7.02 [0.00]
Class size (grade 6)	24.44 (3.6)	24.52 (3.7)	–0.08 [0.64]
Teacher changes (grade 1-6)	0.56 (0.9)	0.49 (0.8)	0.08 [0.06]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 5,006 observations. Ability measures are IQ-like scores measured at the end of grade 6, and general ability is the average of inductive and verbal ability. All the other information comes from administrative records.

Table B.17
Differences in educational choices by SES

SES:	Low	High	Difference
Choice protocols			
Extrinsic choices scale (grade 6)	0.09 (1.0)	−0.04 (1.0)	0.14 [0.00]
Math choice: Ability (grade 6)	0.67 (0.3)	0.74 (0.3)	−0.06 [0.00]
English choice: Ability (grade 6)	0.66 (0.3)	0.74 (0.3)	−0.08 [0.00]
High school choice: Ability (grade 10)	0.62 (0.3)	0.59 (0.4)	0.03 [0.02]
High school choice: Preferences (grade 10)	0.80 (0.3)	0.80 (0.3)	−0.01 [0.45]
Actual choices			
Advanced Math (grade 7)	0.63 (0.5)	0.82 (0.4)	−0.19 [0.00]
Advanced Math (grade 8)	0.53 (0.5)	0.75 (0.4)	−0.22 [0.00]
Advanced Math (grade 9)	0.43 (0.5)	0.65 (0.5)	−0.22 [0.00]
Advanced English (grade 7)	0.63 (0.5)	0.85 (0.4)	−0.22 [0.00]
Advanced English (grade 8)	0.59 (0.5)	0.82 (0.4)	−0.23 [0.00]
Advanced English (grade 9)	0.56 (0.5)	0.78 (0.4)	−0.22 [0.00]
Special education (grade 6)	0.18 (0.4)	0.10 (0.3)	0.08 [0.00]
Special education (grade 7)	0.11 (0.3)	0.06 (0.2)	0.05 [0.00]
Special education (grade 8)	0.04 (0.2)	0.02 (0.1)	0.02 [0.00]
Special education (grade 9)	0.03 (0.2)	0.01 (0.1)	0.01 [0.00]
High school enrollment (15-18)	0.81 (0.4)	0.90 (0.3)	−0.09 [0.00]
Academic track enrollment (15-18)	0.27 (0.4)	0.54 (0.5)	−0.27 [0.00]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 5,006 observations. Choice protocols are derived from the grade 6 student survey, while actual choices are taken from administrative records.

Table B.18
Differences in school performance and attainment by SES

SES:	Low	High	Difference
School performance			
English national test (std, grade 8)	−0.11 (1.0)	0.09 (1.0)	−0.21 [0.00]
Math national test (std, grade 9)	−0.14 (1.0)	0.12 (0.9)	−0.26 [0.00]
GPA: nat. graded (std, grade 7)	−0.24 (0.9)	0.20 (1.0)	−0.44 [0.00]
GPA: nat. graded (std, grade 8)	−0.23 (0.9)	0.19 (1.0)	−0.42 [0.00]
GPA: nat. graded (std, grade 9)	−0.26 (0.9)	0.20 (1.0)	−0.46 [0.00]
GPA: loc. graded (std, grade 7)	−0.28 (1.0)	0.23 (1.0)	−0.51 [0.00]
GPA: loc. graded (std, grade 8)	−0.28 (1.0)	0.22 (0.9)	−0.49 [0.00]
GPA: loc. graded (std, grade 9)	−0.28 (1.0)	0.21 (1.0)	−0.49 [0.00]
GPA (std, high school)	−0.20 (0.9)	0.11 (1.0)	−0.31 [0.00]
School attainment			
High school diploma (17-20)	0.73 (0.4)	0.84 (0.4)	−0.11 [0.00]
Completes academic track (17-20)	0.24 (0.4)	0.49 (0.5)	−0.25 [0.00]
STEM education (-35)	0.35 (0.5)	0.31 (0.5)	0.05 [0.00]
High school diploma (-35)	0.82 (0.4)	0.91 (0.3)	−0.10 [0.00]
University degree (-35)	0.08 (0.3)	0.19 (0.4)	−0.11 [0.00]
Years of non-compulsory (-35)	2.52 (1.9)	3.57 (2.1)	−1.05 [0.00]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 5,006 observations. All information comes from administrative records.

Table B.19
Differences in grade 6 behavior by SES

SES:	Low	High	Difference
Individual			
Abs. cogn. SC scale (grade 6)	−0.15 (1.0)	0.11 (1.0)	−0.25 [0.00]
Rel. cogn. SC scale (grade 6)	−0.16 (1.0)	0.12 (0.9)	−0.28 [0.00]
Rel. non cogn. SC scale (grade 6)	−0.13 (1.0)	0.07 (1.0)	−0.20 [0.00]
I do my best at school (grade 6)	0.68 (0.5)	0.65 (0.5)	0.03 [0.06]
Frequent schoolwork at home (grade 6)	0.60 (0.3)	0.63 (0.3)	−0.03 [0.02]
Will attend high school (grade 6)	0.65 (0.5)	0.79 (0.4)	−0.14 [0.00]
School and family			
Helped at home (grade 6)	0.78 (0.4)	0.82 (0.4)	−0.04 [0.00]
Helped by teachers (grade 6)	0.75 (0.4)	0.77 (0.4)	−0.02 [0.11]
Positive peer interactions (grade 6)	0.91 (0.3)	0.91 (0.3)	−0.00 [0.99]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 5,006 observations. All information, apart from special education status, comes from the grade 6 student survey. Standardized variables are indexes built as the average of relevant survey questions, recodified on a 0-1 scale, and set to missing if less than half of the questions are answered. SC stands for self-concept, a measure of assessment of own abilities or skills. Details on each index are discussed in Appendix A.

Table B.20
Differences in grades 7-10 behavior by SES

SES:	Low	High	Difference
Individual			
Abs. cogn. SC scale (grade 7-9)	−0.15 (1.0)	0.11 (0.9)	−0.26 [0.00]
Motivation scale (grade 7-9)	−0.10 (1.1)	0.06 (1.0)	−0.16 [0.00]
Grades fairness (std, grade 7-9)	−0.06 (1.0)	0.05 (0.9)	−0.11 [0.00]
Abs. cogn. SC scale (grade 10)	−0.06 (1.0)	0.04 (1.0)	−0.10 [0.00]
Well-being index (grade 10)	0.03 (1.0)	−0.03 (1.0)	0.06 [0.09]
Will complete high school (grade 10)	0.93 (0.3)	0.91 (0.3)	0.02 [0.05]
School and family			
Helped at home (grade 7-9)	0.75 (0.4)	0.87 (0.3)	−0.12 [0.00]
Teacher support scale (grade 7-9)	−0.08 (1.0)	0.06 (1.0)	−0.14 [0.00]
Positive peer interactions (grade 7-9)	0.90 (0.2)	0.91 (0.2)	−0.01 [0.09]

Standard errors are in parentheses and p-values are in brackets. Statistics are based on the final sample, consisting of 5,006 observations. All information, apart from special education status, comes from the grade 6 student survey. Standardized variables are indexes built as the average of relevant survey questions, recodified on a 0-1 scale, and set to missing if less than half of the questions are answered. SC stands for self-concept, a measure of assessment of own abilities or skills. Details on each index are discussed in Appendix A.

Table C.1
Does peer ability predict student ability?

Test	(1)	(2)	Mean
General ability (0-40, grade 6)	1.038***	0.028	22.648
	(0.036)	(0.043)	(6.195)
Verbal ability (0-40, grade 6)	0.816***	0.059	22.953
	(0.044)	(0.064)	(5.965)
Inductive ability (0-40, grade 6)	1.245***	-0.005	22.360
	(0.052)	(0.083)	(8.162)
Has ability measure	-0.011*	-0.008	0.935
	(0.006)	(0.006)	(0.246)
Special education (grade 6)	-0.005	0.008	0.134
	(0.004)	(0.005)	(0.340)
Disability	-0.008***	-0.005	0.067
	(0.003)	(0.009)	(0.250)
School FE	No	Yes	
N. of students (mean)	8163	4676	8163
N. of classes	366	204	366
N. of schools	250	88	250

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient from regressions of each tested variable on class ability (excluding student i, expressed in within-school between-class standard deviations). Column (1) corresponds to equation 3. Column (2) adds school FE and standardized school ability (excluding student i), and corresponds to equation 4. Standard errors, reported in parentheses, are clustered at the class level.

Table C.2

Does peer ability predict individual and class characteristics?

Test	(1)	(2)	Mean
Female	0.002	0.003	0.491
	(0.003)	(0.005)	(0.500)
Age (grade 6)	-0.006*	-0.002	12.928
	(0.003)	(0.004)	(0.332)
Born abroad	-0.001	0.001	0.036
	(0.002)	(0.003)	(0.187)
Parent not Nordic	-0.004*	0.004	0.061
	(0.002)	(0.004)	(0.239)
Low SES	-0.044***	-0.006	0.399
	(0.005)	(0.007)	(0.490)
Day care	-0.000	0.000	0.113
	(0.003)	(0.005)	(0.317)
Play school	0.013**	0.007	0.826
	(0.005)	(0.006)	(0.379)
Class size (grade 6)	0.048	0.013	23.987
	(0.141)	(0.154)	(4.227)
Teacher changes (grade 1-6)	-0.086***	-0.067	0.484
	(0.026)	(0.044)	(0.794)
School FE	No	Yes	
N. of students (mean)	7848	4506	7848
N. of classes	366	204	366
N. of schools	250	88	250

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient from regressions of each tested variable on class ability (excluding student i, expressed in within-school between-class standard deviations). Column (1) corresponds to equation 3. Column (2) adds school FE and standardized school ability (excluding student i), and corresponds to equation 4. Standard errors, reported in parentheses, are clustered at the class level.

Table C.3

Does peer ability predict parental occupation?

Test	(1)	(2)	Mean
Not employed/unknown	-0.008***	-0.004	0.055
	(0.002)	(0.003)	(0.228)
Farmer	0.002	0.003	0.030
	(0.002)	(0.003)	(0.170)
Unskilled employee (goods)	-0.010***	0.004	0.086
	(0.002)	(0.003)	(0.280)
Unskilled employee (services)	-0.012***	-0.002	0.096
	(0.002)	(0.004)	(0.294)
Skilled employee (goods)	-0.015***	-0.004	0.172
	(0.003)	(0.005)	(0.377)
Skilled employee (services)	-0.001	-0.002	0.023
	(0.001)	(0.002)	(0.151)
Lower non-manual (low)	-0.002	-0.001	0.044
	(0.002)	(0.004)	(0.206)
Lower non-manual (high)	-0.004*	-0.003	0.082
	(0.002)	(0.003)	(0.275)
Intermediate non-manual	0.013***	-0.003	0.200
	(0.004)	(0.005)	(0.400)
Entrepreneur	0.006**	-0.000	0.095
	(0.003)	(0.004)	(0.294)
Professional	0.031***	0.012***	0.116
	(0.004)	(0.004)	(0.320)
School FE	No	Yes	
N. of students (mean)	8451	4863	8451
N. of classes	366	204	366
N. of schools	250	88	250

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient from regressions of each tested variable on class ability (excluding student i, expressed in within-school between-class standard deviations). Column (1) corresponds to equation 3. Column (2) adds school FE and standardized school ability (excluding student i), and corresponds to equation 4. Standard errors, reported in parentheses, are clustered at the class level.

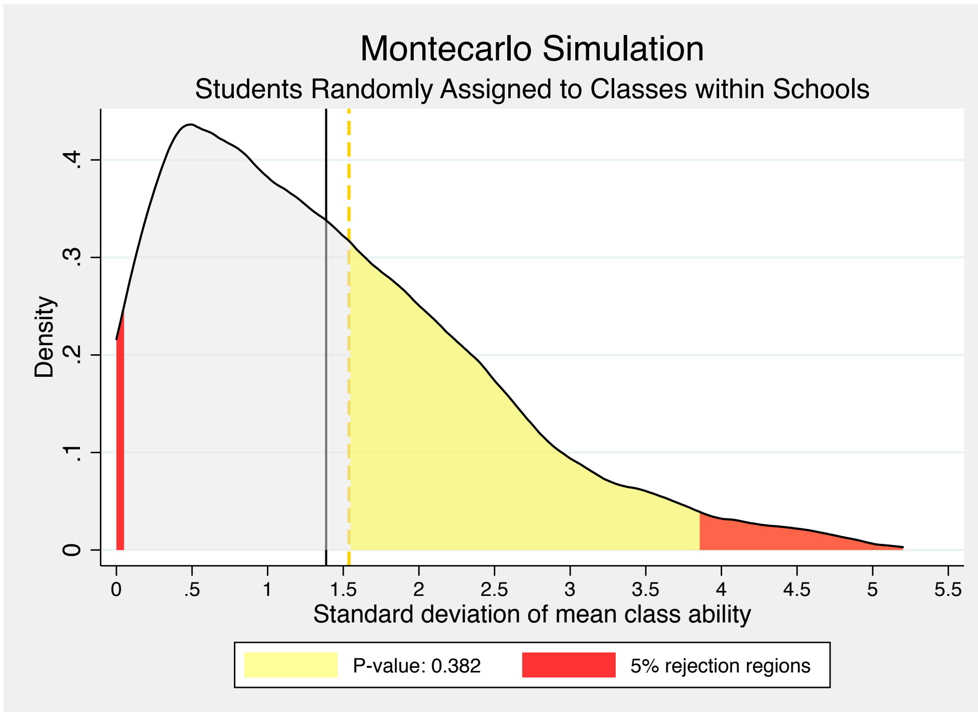

Figure D.1

Standard deviation of mean class ability under random student assignment

Note: The Figure shows the result of a Montecarlo simulation that randomly assigns within each school students to classes with the same class size as in the sample 1000 times. The plot shows the distribution of the standard deviation of mean class ability under random assignment, its mean (solid black line), and 5% two-tailed rejection region. The dashed yellow line marks the value observed in the sample, and the yellow region is the associated p-value.

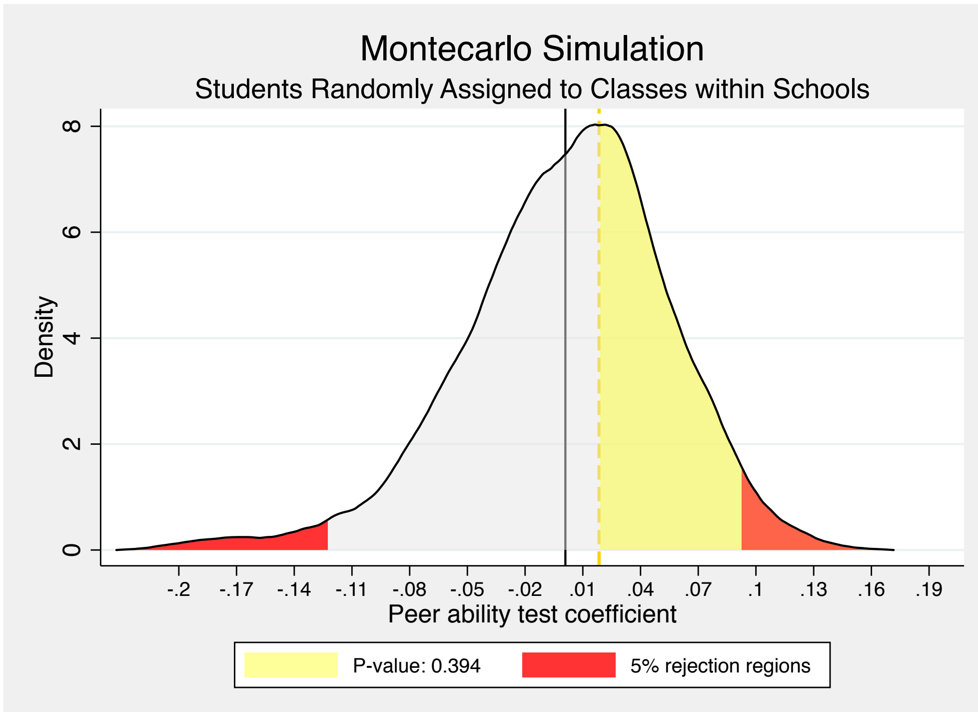

Figure D.2
Peer ability test coefficient under random student assignment

Note: The Figure shows the result of a Montecarlo simulation that randomly assigns within each school students to classes with the same class size as in the sample 1000 times. The plot shows the distribution of the peer ability test coefficient under random assignment, its mean (solid black line), and 5% two-tailed rejection region. The dashed yellow line marks the value observed in the sample, and the yellow region is the associated p-value.

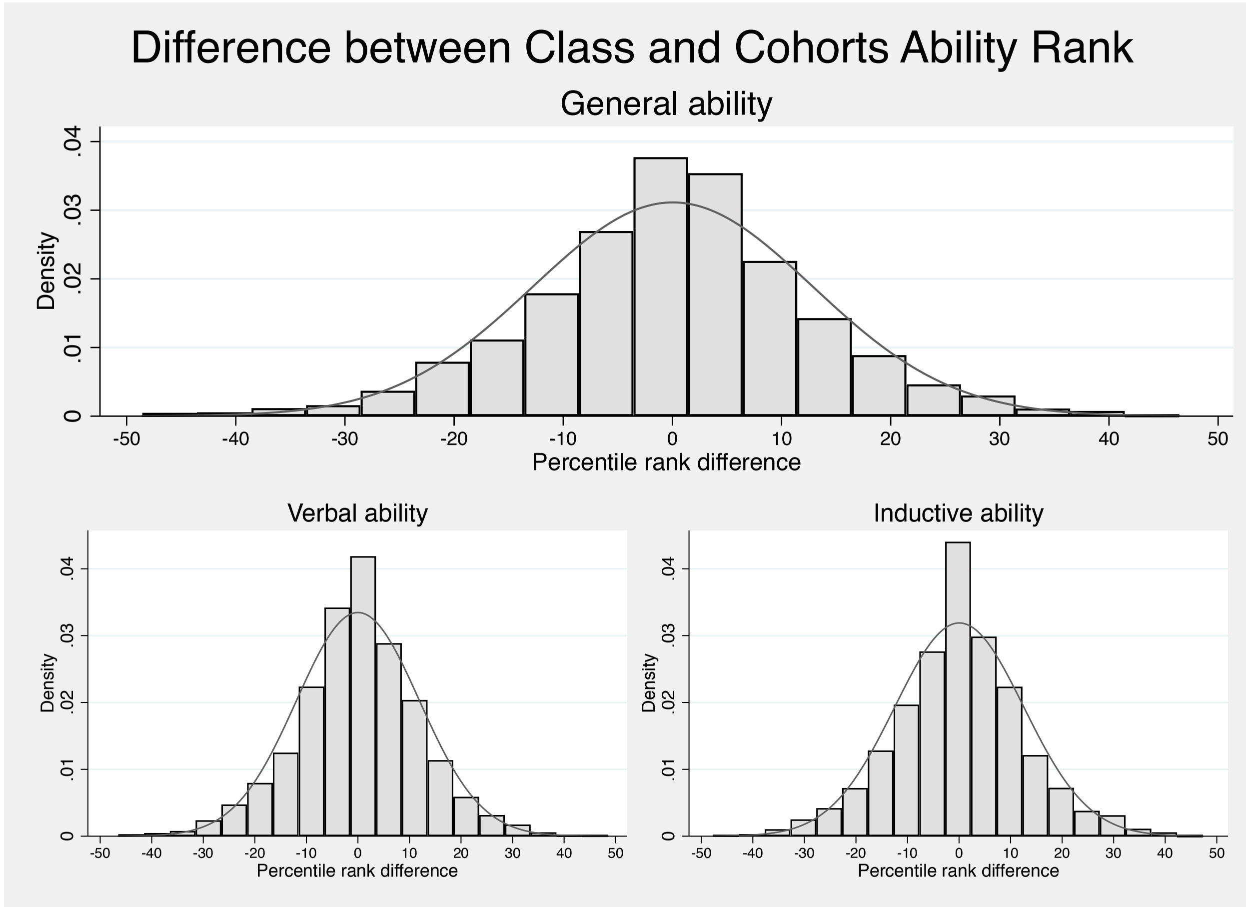

Figure D.3

Comparison of class and national ability percentile ranks

Note: The Figure plots the difference between student percentile ability rank in the class and in the whole cohort, for different ability measures. Students who rank high in their class might rank much lower nationally, and vice versa. Students who form beliefs about own ability with reference to the class might have biased beliefs.

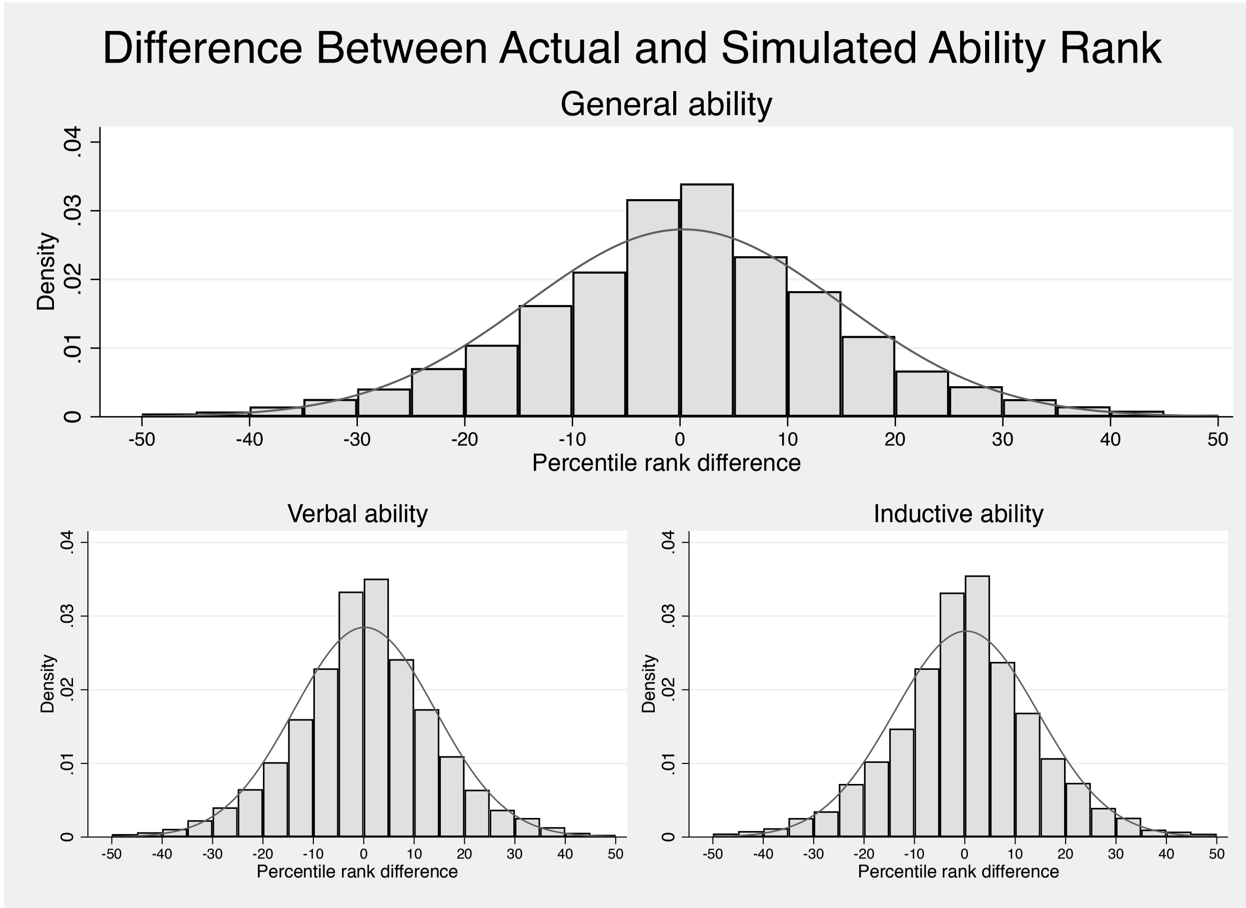

Figure D.4
Comparison of actual and counterfactual class ability rank

Note: The Figure plots the difference between actual class ability rank, and the ability rank the same students would have if they were moved to a random class in the same school. Students who rank high in their class might rank much lower in a different class, and vice versa. Students who form beliefs about own ability with reference to the class might have biased beliefs.

Table E.1
Effects on compulsory school choices

	(1)	(2)	(3)	Mean
Advanced Math (grade 7)	-0.015**	-0.016**	-0.017**	0.755
	(0.007)	(0.007)	(0.007)	(0.43)
Advanced Math (grade 8)	-0.016**	-0.017***	-0.017***	0.666
	(0.006)	(0.006)	(0.006)	(0.47)
Advanced Math (grade 9)	-0.013*	-0.014*	-0.015**	0.565
	(0.007)	(0.007)	(0.007)	(0.50)
Advanced English (grade 7)	-0.007	-0.008	-0.008	0.768
	(0.007)	(0.006)	(0.006)	(0.42)
Advanced English (grade 8)	-0.003	-0.004	-0.004	0.737
	(0.006)	(0.006)	(0.006)	(0.44)
Advanced English (grade 9)	-0.007	-0.008	-0.008	0.691
	(0.006)	(0.006)	(0.006)	(0.46)
Special education (grade 7)	0.006**	0.007**	0.007**	0.077
	(0.003)	(0.003)	(0.003)	(0.27)
Special education (grade 8)	0.004**	0.004**	0.004**	0.028
	(0.002)	(0.002)	(0.002)	(0.17)
Special education (grade 9)	0.004**	0.004**	0.004**	0.017
	(0.002)	(0.002)	(0.002)	(0.13)
N. of students (mean)	4310	4310	4310	4310
N. of classes	204	204	204	204
N. of schools	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability and school fixed effects. Specification (1) includes Guryan et al. (2009) pool correction, and corresponds to equation 2. Specification (2) replaces the pool correction with own ability. Specification (3) adds to specification (2) individual controls. Class ability is the average of verbal and inductive ability measures taken in grade 6. Standard errors are clustered at the class level.

Table E.2
Effects on school performance and grades

	(1)	(2)	(3)	Mean
English national test (std, grade 8)	0.042**	0.040**	0.040**	0.007
	(0.018)	(0.018)	(0.018)	(1.00)
Swedish national test (std, grade 8)	-0.019	-0.015	-0.017	0.007
	(0.017)	(0.018)	(0.019)	(0.99)
Math national test (std, grade 9)	0.014	0.012	0.010	0.022
	(0.019)	(0.019)	(0.018)	(0.97)
GPA: nat. graded (std, grade 7)	-0.024*	-0.027*	-0.029**	0.029
	(0.014)	(0.014)	(0.014)	(0.99)
GPA: nat. graded (std, grade 8)	0.002	-0.001	-0.004	0.024
	(0.012)	(0.012)	(0.011)	(0.98)
GPA: nat. graded (std, grade 9)	0.001	-0.002	-0.004	0.022
	(0.013)	(0.012)	(0.012)	(0.99)
GPA: loc. graded (std, grade 7)	-0.023**	-0.026**	-0.029**	0.024
	(0.012)	(0.012)	(0.012)	(0.99)
GPA: loc. graded (std, grade 8)	-0.025*	-0.027**	-0.030**	0.021
	(0.013)	(0.013)	(0.012)	(0.99)
GPA: loc. graded (std, grade 9)	-0.014	-0.017	-0.019*	0.016
	(0.011)	(0.011)	(0.010)	(0.98)
GPA (std, grade 9)	-0.010	-0.012	-0.016	0.018
	(0.011)	(0.011)	(0.010)	(0.98)
GPA (std, high school)	-0.024*	-0.027**	-0.029**	-0.003
	(0.013)	(0.013)	(0.012)	(0.97)
N. of students (mean)	3834	3834	3834	3834
N. of classes	204	204	204	204
N. of schools	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability and school fixed effects. Specification (1) includes Guryan et al. (2009) pool correction, and corresponds to equation 2. Specification (2) replaces the pool correction with own ability. Specification (3) adds to specification (2) individual controls. Class ability is the average of verbal and inductive ability measures taken in grade 6. Standard errors are clustered at the class level.

Table E.3
Effects on non-compulsory education

	(1)	(2)	(3)	Mean
High school enrollment (15-18)	0.003	0.002	0.002	0.865
	(0.005)	(0.005)	(0.005)	(0.34)
Academic track enrollment (15-18)	-0.000	-0.001	-0.003	0.432
	(0.007)	(0.006)	(0.006)	(0.50)
High school diploma (17-20)	0.007	0.007	0.007	0.797
	(0.006)	(0.006)	(0.006)	(0.40)
Completes academic track (17-20)	0.002	0.001	-0.000	0.392
	(0.007)	(0.007)	(0.006)	(0.49)
University degree (-35)	0.002	0.002	0.000	0.144
	(0.004)	(0.004)	(0.004)	(0.35)
Years of non-compulsory (-35)	0.005	0.001	-0.005	3.122
	(0.025)	(0.025)	(0.022)	(2.09)
N. of students (mean)	4302	4302	4302	4302
N. of classes	204	204	204	204
N. of schools	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability and school fixed effects. Specification (1) includes Guryan et al. (2009) pool correction, and corresponds to equation 2. Specification (2) replaces the pool correction with own ability. Specification (3) adds to specification (2) individual controls. Class ability is the average of verbal and inductive ability measures taken in grade 6. Standard errors are clustered at the class level.

Table E.4

Effects on self-concept, motivation and school Inputs (Grade 6)

	(1)	(2)	(3)	Mean
Abs. cogn. SC scale (PF, grade 6)	-0.035**	-0.036**	-0.036**	-0.004
	(0.015)	(0.015)	(0.015)	(1.00)
Rel. cogn. SC scale (PF, grade 6)	-0.030**	-0.032**	-0.034***	-0.006
	(0.013)	(0.013)	(0.013)	(0.99)
Rel. non cogn. SC scale (PF, grade 6)	-0.006	-0.007	-0.007	-0.016
	(0.016)	(0.016)	(0.016)	(0.99)
I do my best at school (grade 6)	0.005	0.005	0.005	0.658
	(0.007)	(0.006)	(0.007)	(0.47)
Frequent schoolwork at home (grade 6)	0.027***	0.027***	0.027***	0.619
	(0.007)	(0.007)	(0.007)	(0.34)
Positive peer interactions (grade 6)	0.013***	0.013***	0.014***	0.906
	(0.004)	(0.004)	(0.005)	(0.29)
Helped at home (grade 6)	0.009	0.009	0.010	0.804
	(0.006)	(0.006)	(0.006)	(0.40)
Helped by teachers (grade 6)	0.022**	0.021**	0.021**	0.756
	(0.009)	(0.009)	(0.010)	(0.43)
N. of students (mean)	4258	4258	4258	4258
N. of classes	204	204	204	204
N. of schools	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability and school fixed effects. Specification (1) includes Guryan et al. (2009) pool correction, and corresponds to equation 2. Specification (2) replaces the pool correction with own ability. Specification (3) adds to specification (2) individual controls. Class ability is the average of verbal and inductive ability measures taken in grade 6. All scales are described in Appendix A.1. Standard errors are clustered at the class level.

Table E.5
Effects on self-Concept, motivation and school Inputs (grades 7-10)

	(1)	(2)	(3)	Mean
Abs. cogn. SC scale (PF, grade 7-9)	-0.027*	-0.029**	-0.031**	0.013
	(0.014)	(0.014)	(0.013)	(0.98)
Motivation scale (PF, grade 7-9)	0.022*	0.021*	0.019	0.006
	(0.012)	(0.012)	(0.012)	(1.00)
Grades fairness (std, grade 7-9)	-0.056***	-0.057***	-0.059***	-0.000
	(0.015)	(0.015)	(0.015)	(0.98)
Positive peer interactions (grade 7-9)	-0.005*	-0.005*	-0.005*	0.900
	(0.003)	(0.003)	(0.003)	(0.17)
Helped at home (grade 7-9)	-0.005	-0.005	-0.007	0.830
	(0.006)	(0.006)	(0.006)	(0.38)
Teacher support scale (PF, grade 7-9)	-0.032*	-0.033*	-0.034**	0.003
	(0.017)	(0.017)	(0.017)	(0.99)
Abs. cogn. SC scale (PF, grade 10)	-0.009	-0.010	-0.012	-0.005
	(0.015)	(0.015)	(0.015)	(1.01)
Well-being index (PC, grade 10)	-0.033*	-0.033*	-0.030*	-0.004
	(0.018)	(0.018)	(0.017)	(0.98)
N. of students (mean)	3352	3352	3352	3352
N. of classes	204	204	204	204
N. of schools	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability and school fixed effects. Specification (1) includes Guryan et al. (2009) pool correction, and corresponds to equation 2. Specification (2) replaces the pool correction with own ability. Specification (3) adds to specification (2) individual controls. Class ability is the average of verbal and inductive ability measures taken in grade 6. National tests is the average of Swedish, English and math test scores. Scales and indexes are described in Appendix A. Standard errors are clustered at the class level.

Table E.6
Effects on choice protocols and expectations

	(1)	(2)	(3)	Mean
Extrinsic choices scale (PF, grade 6)	-0.015	-0.013	-0.013	0.014
	(0.016)	(0.016)	(0.016)	(1.00)
Math choice: Ability (grade 6)	-0.009**	-0.009**	-0.010**	0.710
	(0.004)	(0.004)	(0.004)	(0.27)
English choice: Ability (grade 6)	-0.003	-0.004	-0.004	0.708
	(0.004)	(0.004)	(0.004)	(0.29)
Will attend high school (grade 6)	-0.012**	-0.013**	-0.013**	0.731
	(0.006)	(0.006)	(0.006)	(0.44)
High school choice: Ability (grade 10)	0.012**	0.012**	0.011**	0.604
	(0.005)	(0.005)	(0.005)	(0.35)
High school choice: Preferences (grade 10)	0.000	-0.000	-0.000	0.797
	(0.006)	(0.006)	(0.006)	(0.29)
Will complete high school (grade 10)	-0.000	-0.000	0.001	0.919
	(0.004)	(0.004)	(0.004)	(0.27)
N. of students (mean)	3729	3729	3729	3729
N. of classes	204	204	204	204
N. of schools	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability and school fixed effects. Specification (1) includes Guryan et al. (2009) pool correction, and corresponds to equation 2. Specification (2) replaces the pool correction with own ability. Specification (3) adds to specification (2) individual controls. Class ability is the average of verbal and inductive ability measures taken in grade 6. National tests is the average of Swedish, English and math test scores. Scales are described in Appendix A.1. Standard errors are clustered at the class level.

Table E.7
Effects on compulsory school choices

	Grading		Sex		SES	
	Late	Early	Male	Female	High	Low
Advanced Math (grade 7)	-0.014	-0.019**	-0.021**	-0.013	-0.016**	-0.020
	(0.011)	(0.008)	(0.009)	(0.008)	(0.006)	(0.012)
Advanced Math (grade 8)	-0.021**	-0.013	-0.024***	-0.012	-0.012*	-0.028**
	(0.008)	(0.010)	(0.008)	(0.008)	(0.006)	(0.011)
Advanced Math (grade 9)	-0.021**	-0.007	-0.029***	-0.002	-0.009	-0.026**
	(0.009)	(0.012)	(0.009)	(0.010)	(0.008)	(0.012)
Advanced English (grade 7)	-0.005	-0.011	-0.011	-0.005	-0.000	-0.017
	(0.010)	(0.009)	(0.009)	(0.007)	(0.007)	(0.011)
Advanced English (grade 8)	0.006	-0.014	0.003	-0.012	0.010	-0.027**
	(0.008)	(0.009)	(0.008)	(0.008)	(0.006)	(0.011)
Advanced English (grade 9)	-0.009	-0.007	-0.008	-0.009	0.003	-0.026**
	(0.008)	(0.009)	(0.008)	(0.009)	(0.007)	(0.011)
Special education (grade 7)	0.011**	0.001	0.012**	-0.000	0.004	0.008
	(0.004)	(0.004)	(0.005)	(0.004)	(0.003)	(0.006)
Special education (grade 8)	0.003	0.006***	0.004	0.005**	0.003	0.007*
	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
Special education (grade 9)	0.003	0.005**	0.005*	0.003	0.005***	0.003
	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)
N. of students (mean)	2061	2249	2213	2096	2492	1640
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability, school fixed effects, own ability and individual controls, for different subsamples. Class ability is the average of verbal and inductive ability measures taken in grade 6. Standard errors are clustered at the class level.

Table E.8
Effects on compulsory school choices

	Grading		Sex		SES	
	Late	Δ Early	Male	Δ Female	High	Δ Low
Advanced Math (grade 7)	-0.014	-0.005	-0.021**	0.009	-0.016**	-0.004
	(0.011)	(0.013)	(0.009)	(0.010)	(0.006)	(0.012)
Advanced Math (grade 8)	-0.021**	0.008	-0.024***	0.012	-0.012*	-0.015
	(0.008)	(0.012)	(0.008)	(0.011)	(0.006)	(0.012)
Advanced Math (grade 9)	-0.021**	0.014	-0.029***	0.027**	-0.009	-0.017
	(0.009)	(0.014)	(0.009)	(0.012)	(0.008)	(0.012)
Advanced English (grade 7)	-0.005	-0.006	-0.011	0.005	-0.000	-0.017
	(0.010)	(0.013)	(0.009)	(0.010)	(0.007)	(0.010)
Advanced English (grade 8)	0.006	-0.020*	0.003	-0.015	0.010	-0.037***
	(0.008)	(0.012)	(0.008)	(0.011)	(0.007)	(0.011)
Advanced English (grade 9)	-0.009	0.001	-0.008	-0.001	0.003	-0.029**
	(0.008)	(0.012)	(0.008)	(0.012)	(0.007)	(0.011)
Special education (grade 7)	0.011**	-0.010	0.012**	-0.012	0.004	0.004
	(0.004)	(0.006)	(0.005)	(0.007)	(0.003)	(0.007)
Special education (grade 8)	0.003	0.003	0.004	0.001	0.003	0.005
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.006)
Special education (grade 9)	0.003	0.002	0.005*	-0.002	0.005***	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.004)
N. of students (mean)	4310	4310	4310	4310	4132	4132
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns 1, 3, 5 of the table report the coefficient for peer ability, expressed in within-school between-class standard deviations, for late graded, male, and high SES students. Columns 2, 4, 6 report the coefficient for the interaction between peer ability and early graded, female, and low SES status. All coefficients are from regressions of each outcome on peer ability, school fixed effects, own ability and individual controls, fully interacted by category. Class ability is the average of verbal and inductive ability measures taken in grade 6. Standard errors are clustered at the class level.

Table E.9
Effects on school performance and grades

	Grading		Sex		SES	
	Late	Early	Male	Female	High	Low
English national test (std, grade 8)	0.001	0.088***	0.015	0.059**	0.041*	0.046*
	(0.022)	(0.026)	(0.025)	(0.027)	(0.022)	(0.025)
Swedish national test (std, grade 8)	-0.032*	0.002	-0.028*	-0.000	-0.000	-0.032
	(0.019)	(0.036)	(0.017)	(0.029)	(0.020)	(0.030)
Math national test (std, grade 9)	0.004	0.017	0.037*	-0.007	0.013	0.009
	(0.019)	(0.031)	(0.021)	(0.024)	(0.021)	(0.027)
GPA: nat. graded (std, grade 7)	-0.010	-0.047**	-0.040**	-0.022	-0.012	-0.045**
	(0.017)	(0.021)	(0.016)	(0.023)	(0.017)	(0.018)
GPA: nat. graded (std, grade 8)	-0.003	-0.003	-0.010	-0.001	-0.003	-0.006
	(0.015)	(0.017)	(0.014)	(0.020)	(0.015)	(0.017)
GPA: nat. graded (std, grade 9)	-0.003	-0.006	-0.008	0.000	-0.011	0.010
	(0.015)	(0.018)	(0.014)	(0.022)	(0.014)	(0.017)
GPA: loc. graded (std, grade 7)	-0.005	-0.051***	-0.036**	-0.029	-0.008	-0.057***
	(0.017)	(0.015)	(0.016)	(0.019)	(0.016)	(0.018)
GPA: loc. graded (std, grade 8)	-0.015	-0.044***	-0.041***	-0.025	-0.018	-0.055***
	(0.018)	(0.016)	(0.015)	(0.021)	(0.014)	(0.019)
GPA: loc. graded (std, grade 9)	-0.006	-0.032**	-0.040***	-0.001	-0.010	-0.043**
	(0.016)	(0.014)	(0.014)	(0.018)	(0.013)	(0.018)
GPA (std, grade 9)	-0.003	-0.029**	-0.034**	0.001	-0.009	-0.033*
	(0.015)	(0.014)	(0.013)	(0.018)	(0.013)	(0.018)
GPA (std, high school)	-0.020	-0.034*	-0.026	-0.053**	-0.022	-0.031
	(0.016)	(0.019)	(0.019)	(0.020)	(0.015)	(0.021)
N. of students (mean)	1834	2000	1971	1863	2239	1444
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability, school fixed effects, own ability and individual controls, for different subsamples. Class ability is the average of verbal and inductive ability measures taken in grade 6. Standard errors are clustered at the class level.

Table E.10
Effects on school performance and grades

	Grading		Sex		SES	
	Late	Δ Early	Male	Δ Female	High	Δ Low
English national test (std, grade 8)	0.001	0.087**	0.015	0.044	0.041*	0.004
	(0.022)	(0.034)	(0.025)	(0.039)	(0.022)	(0.031)
Swedish national test (std, grade 8)	-0.032*	0.034	-0.028*	0.027	-0.000	-0.032
	(0.019)	(0.041)	(0.017)	(0.025)	(0.020)	(0.023)
Math national test (std, grade 9)	0.004	0.013	0.037*	-0.044*	0.013	-0.005
	(0.019)	(0.037)	(0.021)	(0.025)	(0.021)	(0.026)
GPA: nat. graded (std, grade 7)	-0.010	-0.037	-0.040**	0.017	-0.012	-0.034*
	(0.017)	(0.027)	(0.016)	(0.026)	(0.017)	(0.019)
GPA: nat. graded (std, grade 8)	-0.003	0.000	-0.010	0.010	-0.003	-0.003
	(0.015)	(0.023)	(0.014)	(0.026)	(0.015)	(0.022)
GPA: nat. graded (std, grade 9)	-0.003	-0.003	-0.008	0.008	-0.011	0.021
	(0.015)	(0.024)	(0.015)	(0.028)	(0.014)	(0.019)
GPA: loc. graded (std, grade 7)	-0.005	-0.045**	-0.036**	0.008	-0.008	-0.049**
	(0.017)	(0.022)	(0.016)	(0.027)	(0.016)	(0.023)
GPA: loc. graded (std, grade 8)	-0.015	-0.029	-0.041***	0.016	-0.018	-0.037*
	(0.018)	(0.024)	(0.015)	(0.027)	(0.014)	(0.021)
GPA: loc. graded (std, grade 9)	-0.006	-0.026	-0.040***	0.040*	-0.010	-0.033
	(0.016)	(0.021)	(0.014)	(0.023)	(0.013)	(0.021)
GPA (std, grade 9)	-0.003	-0.025	-0.034**	0.035	-0.009	-0.024
	(0.015)	(0.021)	(0.013)	(0.023)	(0.013)	(0.020)
GPA (std, high school)	-0.020	-0.015	-0.026	-0.027	-0.022	-0.010
	(0.016)	(0.025)	(0.019)	(0.030)	(0.015)	(0.025)
N. of students (mean)	3834	3834	3834	3834	3684	3684
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns 1, 3, 5 of the table report the coefficient for peer ability, expressed in within-school between-class standard deviations, for late graded, male, and high SES students. Columns 2, 4, 6 report the coefficient for the interaction between peer ability and early graded, female, and low SES status. All coefficients are from regressions of each outcome on peer ability, school fixed effects, own ability and individual controls, fully interacted by category. Class ability is the average of verbal and inductive ability measures taken in grade 6. Standard errors are clustered at the class level.

Table E.11
Effects on non-compulsory education

	Grading		Sex		SES	
	Late	Early	Male	Female	High	Low
High school enrollment (15-18)	0.013**	-0.009	0.010	-0.004	0.004	-0.003
	(0.006)	(0.006)	(0.008)	(0.007)	(0.006)	(0.009)
Academic track enrollment (15-18)	-0.016**	0.012	-0.008	0.004	-0.001	-0.009
	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.009)
High school diploma (17-20)	0.020**	-0.007	0.014*	0.002	0.009	0.001
	(0.009)	(0.006)	(0.008)	(0.009)	(0.007)	(0.010)
Completes academic track (17-20)	-0.013	0.014*	-0.010	0.015	0.002	-0.006
	(0.008)	(0.008)	(0.007)	(0.011)	(0.008)	(0.010)
University degree (-35)	-0.013**	0.014**	-0.003	0.004	0.005	-0.008
	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.005)
Years of non-compulsory (-35)	-0.063**	0.061**	0.010	-0.016	0.032	-0.041
	(0.031)	(0.030)	(0.035)	(0.040)	(0.032)	(0.031)
N. of students (mean)	2057	2245	2212	2090	2484	1641
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability, school fixed effects, own ability and individual controls, for different subsamples. Class ability is the average of verbal and inductive ability measures taken in grade 6. Standard errors are clustered at the class level.

Table E.12
Effects on non-compulsory education

	Grading		Sex		SES	
	Late	Δ Early	Male	Δ Female	High	Δ Low
High school enrollment (15-18)	0.013**	-0.022**	0.010	-0.014	0.004	-0.007
	(0.006)	(0.009)	(0.008)	(0.011)	(0.006)	(0.011)
Academic track enrollment (15-18)	-0.016**	0.028***	-0.008	0.013	-0.001	-0.008
	(0.007)	(0.010)	(0.008)	(0.012)	(0.008)	(0.012)
High school diploma (17-20)	0.020**	-0.027**	0.014*	-0.012	0.009	-0.008
	(0.009)	(0.011)	(0.008)	(0.011)	(0.007)	(0.012)
Completes academic track (17-20)	-0.013	0.026**	-0.010	0.025*	0.002	-0.008
	(0.008)	(0.011)	(0.007)	(0.014)	(0.008)	(0.012)
University degree (-35)	-0.013**	0.027***	-0.003	0.007	0.005	-0.012
	(0.006)	(0.008)	(0.006)	(0.009)	(0.006)	(0.008)
Years of non-compulsory (-35)	-0.063**	0.124***	0.010	-0.026	0.032	-0.073*
	(0.031)	(0.043)	(0.035)	(0.058)	(0.032)	(0.042)
N. of students (mean)	4302	4302	4302	4302	4125	4125
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns 1, 3, 5 of the table report the coefficient for peer ability, expressed in within-school between-class standard deviations, for late graded, male, and high SES students. Columns 2, 4, 6 report the coefficient for the interaction between peer ability and early graded, female, and low SES status. All coefficients are from regressions of each outcome on peer ability, school fixed effects, own ability and individual controls, fully interacted by category. Class ability is the average of verbal and inductive ability measures taken in grade 6. Standard errors are clustered at the class level.

Table E.13
Effects on self-concept, motivation and school inputs (Grade 6)

	Grading		Sex		SES	
	Late	Early	Male	Female	High	Low
Abs. cogn. SC scale (PF, grade 6)	-0.024	-0.047*	-0.028	-0.058***	-0.006	-0.084***
	(0.018)	(0.024)	(0.019)	(0.019)	(0.017)	(0.027)
Rel. cogn. SC scale (PF, grade 6)	-0.029	-0.037*	-0.026	-0.050***	-0.013	-0.053**
	(0.018)	(0.019)	(0.017)	(0.018)	(0.016)	(0.025)
Rel. non cogn. SC scale (PF, grade 6)	-0.008	-0.004	-0.027	0.009	0.014	-0.045*
	(0.019)	(0.025)	(0.019)	(0.022)	(0.019)	(0.027)
I do my best at school (grade 6)	-0.004	0.014	0.003	0.009	0.018**	-0.010
	(0.008)	(0.010)	(0.008)	(0.010)	(0.008)	(0.011)
Frequent schoolwork at home (grade 6)	0.004	0.050***	0.024***	0.027***	0.026***	0.030***
	(0.009)	(0.010)	(0.008)	(0.009)	(0.008)	(0.011)
Positive peer interactions (grade 6)	0.011*	0.017**	0.015**	0.013**	0.008	0.021***
	(0.006)	(0.007)	(0.007)	(0.006)	(0.006)	(0.007)
Helped at home (grade 6)	-0.002	0.021**	0.018**	0.001	0.005	0.018*
	(0.008)	(0.010)	(0.007)	(0.010)	(0.007)	(0.011)
Helped by teachers (grade 6)	0.033**	0.010	0.025**	0.016	0.028**	0.005
	(0.013)	(0.013)	(0.011)	(0.011)	(0.011)	(0.013)
N. of students (mean)	2038	2220	2187	2071	2457	1618
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability, school fixed effects, own ability and individual controls, for different subsamples. Class ability is the average of verbal and inductive ability measures taken in grade 6. All scales are described in Appendix A.1. Standard errors are clustered at the class level.

Table E.14
Effects on self-concept, motivation and school inputs (grade 6)

	Grading		Sex		SES	
	Late	Δ Early	Male	Δ Female	High	Δ Low
Abs. cogn. SC scale (PF, grade 6)	-0.024	-0.022	-0.028	-0.030	-0.006	-0.078***
	(0.018)	(0.030)	(0.019)	(0.025)	(0.017)	(0.030)
Rel. cogn. SC scale (PF, grade 6)	-0.029	-0.008	-0.026	-0.024	-0.013	-0.041
	(0.018)	(0.026)	(0.017)	(0.024)	(0.016)	(0.028)
Rel. non cogn. SC scale (PF, grade 6)	-0.008	0.004	-0.027	0.036	0.014	-0.059*
	(0.019)	(0.031)	(0.019)	(0.025)	(0.019)	(0.031)
I do my best at school (grade 6)	-0.004	0.019	0.003	0.006	0.018**	-0.029**
	(0.008)	(0.013)	(0.008)	(0.012)	(0.008)	(0.014)
Frequent schoolwork at home (grade 6)	0.004	0.045***	0.024***	0.003	0.026***	0.004
	(0.009)	(0.013)	(0.008)	(0.009)	(0.008)	(0.011)
Positive peer interactions (grade 6)	0.011*	0.006	0.015**	-0.002	0.008	0.013*
	(0.006)	(0.009)	(0.007)	(0.009)	(0.006)	(0.007)
Helped at home (grade 6)	-0.002	0.023*	0.018**	-0.017	0.005	0.014
	(0.008)	(0.012)	(0.007)	(0.012)	(0.007)	(0.011)
Helped by teachers (grade 6)	0.033**	-0.023	0.025**	-0.009	0.028**	-0.023*
	(0.013)	(0.018)	(0.011)	(0.011)	(0.011)	(0.014)
N. of students (mean)	4258	4258	4258	4258	4075	4075
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns 1, 3, 5 of the table report the coefficient for peer ability, expressed in within-school between-class standard deviations, for late graded, male, and high SES students. Columns 2, 4, 6 report the coefficient for the interaction between peer ability and early graded, female, and low SES status. All coefficients are from regressions of each outcome on peer ability, school fixed effects, own ability and individual controls, fully interacted by category. Class ability is the average of verbal and inductive ability measures taken in grade 6. All scales are described in Appendix A.1. Standard errors are clustered at the class level.

Table E.15
Effects on self-concept, motivation and school inputs (grades 7-10)

	Grading		Sex		SES	
	Late	Early	Male	Female	High	Low
Abs. cogn. SC scale (PF, grade 7-9)	-0.042**	-0.017	-0.022	-0.047***	-0.021	-0.038
	(0.019)	(0.017)	(0.018)	(0.017)	(0.017)	(0.024)
Motivation scale (PF, grade 7-9)	-0.005	0.043**	0.023	0.007	0.032*	0.005
	(0.016)	(0.019)	(0.022)	(0.022)	(0.017)	(0.025)
Grades fairness (std, grade 7-9)	-0.056***	-0.061***	-0.099***	-0.019	-0.045**	-0.064*
	(0.021)	(0.021)	(0.024)	(0.020)	(0.022)	(0.034)
Positive peer interactions (grade 7-9)	-0.007	-0.003	-0.007	-0.006	-0.003	-0.009
	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.006)
Helped at home (grade 7-9)	-0.009	-0.005	-0.010	-0.007	-0.001	-0.024**
	(0.008)	(0.009)	(0.007)	(0.010)	(0.007)	(0.011)
Teacher support scale (PF, grade 7-9)	-0.023	-0.045*	-0.049**	-0.033	-0.020	-0.064**
	(0.023)	(0.026)	(0.022)	(0.025)	(0.019)	(0.028)
Abs. cogn. SC scale (PF, grade 10)	-0.032	0.011	-0.011	-0.010	0.001	-0.041
	(0.019)	(0.021)	(0.019)	(0.021)	(0.021)	(0.028)
Well-being index (PC, grade 10)	-0.058**	0.001	-0.030	-0.029	-0.021	-0.029
	(0.023)	(0.027)	(0.021)	(0.025)	(0.021)	(0.030)
N. of students (mean)	1579	1773	1584	1768	2028	1212
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability, school fixed effects, own ability and individual controls, for different subsamples. Class ability is the average of verbal and inductive ability measures taken in grade 6. National tests is the average of Swedish, English and math test scores. Scales and indexes are described in Appendix A. Standard errors are clustered at the class level.

Table E.16
Effects on self-concept, motivation and school inputs (grades 7-10)

	Grading		Sex		SES	
	Late	Δ Early	Male	Δ Female	High	Δ Low
Abs. cogn. SC scale (PF, grade 7-9)	-0.042**	0.026	-0.022	-0.025	-0.021	-0.017
	(0.019)	(0.026)	(0.018)	(0.023)	(0.017)	(0.030)
Motivation scale (PF, grade 7-9)	-0.005	0.048*	0.023	-0.016	0.032*	-0.027
	(0.016)	(0.025)	(0.022)	(0.035)	(0.017)	(0.030)
Grades fairness (std, grade 7-9)	-0.056***	-0.004	-0.099***	0.079**	-0.045**	-0.019
	(0.021)	(0.029)	(0.024)	(0.032)	(0.022)	(0.043)
Positive peer interactions (grade 7-9)	-0.007	0.004	-0.007	0.000	-0.003	-0.007
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)	(0.007)
Helped at home (grade 7-9)	-0.009	0.004	-0.010	0.003	-0.001	-0.023*
	(0.008)	(0.012)	(0.007)	(0.013)	(0.007)	(0.013)
Teacher support scale (PF, grade 7-9)	-0.023	-0.022	-0.049**	0.016	-0.020	-0.043
	(0.023)	(0.034)	(0.022)	(0.032)	(0.019)	(0.026)
Abs. cogn. SC scale (PF, grade 10)	-0.032	0.043	-0.011	0.002	0.001	-0.041
	(0.019)	(0.029)	(0.019)	(0.026)	(0.021)	(0.035)
Well-being index (PC, grade 10)	-0.058**	0.059*	-0.030	0.000	-0.021	-0.008
	(0.023)	(0.035)	(0.021)	(0.031)	(0.021)	(0.033)
N. of students (mean)	3352	3352	3352	3352	3240	3240
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns 1, 3, 5 of the table report the coefficient for peer ability, expressed in within-school between-class standard deviations, for late graded, male, and high SES students. Columns 2, 4, 6 report the coefficient for the interaction between peer ability and early graded, female, and low SES status. All coefficients are from regressions of each outcome on peer ability, school fixed effects, own ability and individual controls, fully interacted by category. Class ability is the average of verbal and inductive ability measures taken in grade 6. National tests is the average of Swedish, English and math test scores. Scales and indexes are described in Appendix A. Standard errors are clustered at the class level.

Table E.17
Effects on choice protocols and expectations

	Grading		Sex		SES	
	Late	Early	Male	Female	High	Low
Extrinsic choices scale (PF, grade 6)	-0.042*	0.018	-0.022	-0.007	-0.013	-0.003
	(0.022)	(0.023)	(0.024)	(0.021)	(0.020)	(0.027)
Math choice: Ability (grade 6)	-0.013**	-0.006	-0.006	-0.012**	-0.005	-0.018***
	(0.005)	(0.006)	(0.004)	(0.006)	(0.005)	(0.006)
English choice: Ability (grade 6)	-0.006	-0.000	-0.004	-0.001	0.003	-0.007
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.007)
Will attend high school (grade 6)	-0.013	-0.014*	-0.006	-0.016*	-0.003	-0.031***
	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)	(0.010)
High school choice: Ability (grade 10)	0.008	0.014*	0.002	0.017**	0.008	0.004
	(0.008)	(0.007)	(0.009)	(0.007)	(0.007)	(0.010)
High school choice: Preferences (grade 10)	0.006	-0.005	0.002	0.004	0.003	-0.005
	(0.006)	(0.009)	(0.007)	(0.008)	(0.006)	(0.010)
Will complete high school (grade 10)	-0.007	0.006	0.001	0.003	0.006	-0.007
	(0.006)	(0.005)	(0.006)	(0.006)	(0.007)	(0.006)
N. of students (mean)	1777	1952	1851	1878	2204	1380
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the coefficient for peer ability, expressed in within-school between-class standard deviations, in the regression of each outcome on peer ability, school fixed effects, own ability and individual controls, for different subsamples. Class ability is the average of verbal and inductive ability measures taken in grade 6. National tests is the average of Swedish, English and math test scores. Scales are described in Appendix A.1. Standard errors are clustered at the class level.

Table E.18
Effects on choice protocols and expectations

	Grading		Sex		SES	
	Late	Δ Early	Male	Δ Female	High	Δ Low
Extrinsic choices scale (PF, grade 6)	-0.042*	0.060*	-0.022	0.015	-0.013	0.009
	(0.022)	(0.032)	(0.024)	(0.032)	(0.020)	(0.030)
Math choice: Ability (grade 6)	-0.013**	0.007	-0.006	-0.006	-0.005	-0.012*
	(0.005)	(0.008)	(0.004)	(0.006)	(0.005)	(0.007)
English choice: Ability (grade 6)	-0.006	0.006	-0.004	0.002	0.003	-0.010
	(0.006)	(0.008)	(0.006)	(0.008)	(0.005)	(0.008)
Will attend high school (grade 6)	-0.013	-0.001	-0.006	-0.011	-0.003	-0.028**
	(0.008)	(0.011)	(0.008)	(0.013)	(0.008)	(0.013)
High school choice: Ability (grade 10)	0.008	0.005	0.002	0.015	0.008	-0.004
	(0.008)	(0.011)	(0.009)	(0.012)	(0.007)	(0.011)
High school choice: Preferences (grade 10)	0.006	-0.011	0.002	0.001	0.003	-0.008
	(0.006)	(0.011)	(0.007)	(0.010)	(0.006)	(0.010)
Will complete high school (grade 10)	-0.007	0.013	0.001	0.002	0.006	-0.012
	(0.006)	(0.008)	(0.006)	(0.009)	(0.007)	(0.009)
N. of students (mean)	3729	3729	3729	3729	3584	3584
N. of classes	204	204	204	204	204	204
N. of schools	88	88	88	88	88	88

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns 1, 3, 5 of the table report the coefficient for peer ability, expressed in within-school between-class standard deviations, for late graded, male, and high SES students. Columns 2, 4, 6 report the coefficient for the interaction between peer ability and early graded, female, and low SES status. All coefficients are from regressions of each outcome on peer ability, school fixed effects, own ability and individual controls, fully interacted by category. Class ability is the average of verbal and inductive ability measures taken in grade 6. National tests is the average of Swedish, English and math test scores. Scales are described in Appendix A.1. Standard errors are clustered at the class level.