The Impact of Early Grading on Academic Choices: Mechanisms and Social Implications

- JOB MARKET PAPER -

Luca Facchinello*

Abstract

Does early grading affect educational choices? To answer this question, I exploit a curriculum reform which postponed grade assignment in Swedish compulsory schools. The staggered implementation of the reform allows me to identify short- and long-term effects of early grading, for students with different academic ability and socioeconomic status (SES). When graded early on, high-ability students (especially if high-SES) exhibit higher grades in compulsory school, and are more likely to choose academic courses. Low-ability students react in the opposite way, with particularly negative reactions among low-SES students. High school attainment increases for high-ability low-SES students; college attainment decreases for low-ability low-SES students. None of these effects carry over to the labor market. This suggests that early grades improve the match between early education choices and academic ability, and reduce over-investment in education. I show that the short-term effects are consistent with predictions from a learning model in which children are uncertain about academic ability, have different priors depending on SES, and use grading information to re-optimize educational choices. I find no evidence of demotivating effects for low-ability students, an alternative mechanism through which grades might affect education choices, and the main motivation behind the grading reform.

JEL codes: I21; I28; J13; J24
Keywords: Grades, Ability, Uncertainty, Learning, Sequential Choice, School Choice, Social Background, Educational Attainment, Dropout, Difference in Differences

---

*Department of Economics, Stockholm School of Economics. Email: luca.facchinello@hhs.se

Acknowledgements: a special thanks to my two advisors, Erik Lindqvist and Juanna Joensen, who provided excellent guidance and sound advice throughout the paper. I thank John Bound, Charlie Brown, Kerem Coşar, Susan Dynarski, Tore Ellingsen, Jeffrey Smith and Kevin Stange for their valuable comments. Lastly I thank seminar participants at Stockholm School of Economics, University of Michigan (CIERS and Economics Lunch Seminar), SITE, Brucchi-Luchino and Collegio Carlo Alberto for their useful feedback. I gratefully acknowledge financial support from the Swedish Foundation for Humanities and Social Sciences (Riksbankens Jubileumsfond) grant P12-0968 enabling the data collection for this project. The usual disclaimers apply.

Contents

1 Introduction

While education is traditionally seen in economics as a form of investment with known costs and returns (Becker, 1994; Ben-Porath, 1967), recent models of education choice (e.g., Altonji, 1993) have highlighted the role of uncertainty in educational investment: the expected return of any education choice depends ex-ante on the probability of graduation, and thus on academic ability. When students are uncertain about ability, information, in the form of school grades, might affect their choices.

The role of grades on education choice has been studied almost exclusively at the college level (Stinebrickner & Stinebrickner, 2012; Zafar, 2011). Little is known about how grades affect students at early stages of education, when children have less information on their academic ability, and are still unconstrained by previous choices.

In this paper I investigate how assigning grades in early compulsory school affects educational choices and attainment of Swedish students. To investigate mechanisms I compare the empirical results to the predictions of a sequential choice learning model calibrated to the data.

The institutional setup and the data are particularly suitable to answer the research question. In Sweden, students used to receive the first formal grades in school year 3, at age 10. Grades were based on students' rankings in national standardized tests, and thus provided different information from the test scores students received during the year. In 1969 a reform allowed municipalities to postpone grade assignment to school years 6 or 7. In 1982, a second reform compelled all municipalities to postpone grade assignment to school year 8.[1] The reforms, gradually implemented over time in different municipalities, provide a source of exogenous variation in grade assignment.

I use detailed survey and register data on cohorts born in 1967 and 1972. The 1967 cohort comprises treated students, who were living in municipalities where grading started in middle compulsory school (school year 6), and control students, who lived in municipalities where grades were assigned starting from late compulsory school (school year 7). Students born in 1972 started receiving grades in late compulsory school (in school year 8) in both treatment and control municipalities. If the education choices of students in treatment and control municipalities trend in the same way over time, it is possible to disentangle the effect of early grade assignment from pre-existing differences between the two sets of municipalities. I provide evidence that trends in educational attainment are the same in treatment and control municipalities, for cohorts who did not receive early

[1] In 2012 the reform was reversed, and grades in school year 6 were reinstated. Currently grading in school year 4 is being discussed.

grades. I also show that pre-treatment differences in determinants of education appear in general to persist over time.

To guide the empirical analysis, I set up a model of early education choice that captures the most important features of the institutional setup.[2] In the model, ability determines optimal effort and education choices during compulsory school. For low-ability students it is optimal to exert low effort and enroll into vocational high school. For high-ability students it is instead optimal to exert higher effort and enroll into academic education. Children are uncertain about their cognitive ability. Their priors reflect aggregate ability distributions: high-SES children are on average endowed with higher ability than low-SES children.[3] Grades reveal information about true ability, and allow students to re-optimize educational choices. As in the institutional setup, grades can be assigned starting from middle or late compulsory school, while they are never assigned in early compulsory school.

The calibrated model shows that early grade assignment results in better sorting of students into education, that is, in choices closer to first best. However, students with the same ability react differently to the ability signals, due to the different priors about ability. Low (high) SES students who receive low (high) ability signals confirm their priors, and thus react strongly to the information. Students who receive signals inconsistent with their priors form imprecise posteriors, thus their responses are weaker. The model solution implies different reactions to early grading for very low ability students, low ability students, and high ability students. When they receive early grades, students with very low ability increase effort in compulsory school, are more likely to choose vocational high school, and thus less likely to drop out of high school. Low-ability students on average reduce effort in compulsory school, and are more likely to choose vocational education paths. These responses appear to be stronger for low-SES students, who are more sensitive to low ability signals. When graded early on, high-ability students increase effort in late compulsory school if they are low-SES, and decrease it if they are high-SES. All high-ability students are more likely to choose academic high school, but only low-SES students increase college attainment as a result of early grading: some high-ability high-SES students fail to access college due to early reductions in effort.

The model guides the empirical analysis: I present the effects of early grades for students with different SES and academic ability. SES is proxied by parental education. Ability is measured using cognitive ability tests administered to both cohorts in school year 6, before grade assignment.

To investigate empirically the effects of early grading on short-term effort, I focus on

[2] The model builds on the theoretical framework outlined in Altonji et al (2012)
[3] Régner (2002) discusses biases about ability for low SES students in the psychology literature.

two outcomes: grades and academic course choices in late compulsory school. Higher grades require higher effort; academic courses are more demanding. Results are broadly consistent with model's predictions: when graded early on, low-ability students, especially if low-SES, receive lower grades and are less likely to choose academic courses in late compulsory school. High-ability students exhibit instead higher grades in late compulsory school, but do not revise course choices. The pattern found in the model is thus reproduced by the data, with the difference that high-ability high-SES students are putting more effort, instead of reducing it.[4]

I consider thereafter effects of early grades on high school choices and attainment. I find an increase in high school enrollment for all students.[5] Contrary to model predictions, I do not observe changes in high school track choice. Why is this the case? I propose as an explanation that preferences for education might attenuate the effects of early grades. My data shows that, controlling for ability, high-SES students' academic high school enrollment rates are 20 percentage points higher than those of low-SES students. At the same time grade differences in late compulsory school between high- and low-SES students are at most one fourth of a grade: SES appears thus to strongly influence high school choices in Sweden, independently of ability. I find effects on educational attainment only for low-SES students. Early grading leads to a 3 percentage points decrease in college attainment for low-ability low-SES students, and a 6 percentage points increase in high school attainment for high-ability low-SES students (mostly due to a reduction in dropout).

Do the effects found on education carry over to the labor market? I find that early grades do not affect income at ages 33-40, but they increase upward income mobility for low-ability low-SES students. These students showed the strongest reductions in school grades and educational attainment. This suggests that early grades improved the match between early education choices and academic ability, and reduced over-investment in education. Methodologically I confirm the importance of evaluating education policy in the long-run: limiting the analysis to short-term or intermediate education outcomes would have led to different conclusions.

The idea that early grades could motivate/demotivate children in compulsory school was the main motivation behind the grading reform. I empirically investigate this alternative mechanism through which grades could affect education choice. I test for effects of early grades on student motivation and attitudes toward school. These outcomes are

[4] The result can be easily reconciled with the model assuming that different college majors require different ability levels.

[5] This is due to the increase in effort during compulsory school for high-ability students. While low-ability students reduced effort, the weakest students could have increased effort when graded early on.

measured from survey responses, in the year in which grades where assigned, and in late compulsory school. I find no evidence of early grading discouraging or motivating students, which is consistent with grades simply revealing information to the students.

I conclude that early grades allow students to better sort into education, and thus lead to an increase in efficiency. At the same time the grading policy increases inequality in educational attainment, and reduces effort in compulsory school among low-ability students. The final judgement on the policy depends on the objectives of the policy-maker.

Early grading has relevant effects in the Swedish education system, in which students are explicitly sorted into academic tracks that provide access to college (a *tracked* education system). To what extent do my results generalize to different setups? As knowledge production is cumulative, early education choices constrain late choices for all students (e.g., college preparation affects college enrollment). Assigning grades early on might thus affect students' education choices and attainment also in non-tracked (*comprehensive*) education systems.[6]

Results are consistent with the learning mechanism outlined by Stinebrickner & Stinebrickner (2012) and Zafar (2011), who find that college students react to grading information. Students who get lower (higher) than expected grades are more (less) likely to drop out/switch to an easier major. A limitation of this literature is that, due to the college setup, it is not possible to tell whether students are learning from grades about academic ability or previous preparation. In my setup grades were assigned when children were 13, so there is less concern that students are learning about previous preparation rather than ability. Moreover I show that students' reactions to grade assignment are consistent with a model in which students learn about ability.

My paper is also related to the grading standards literature, which stresses the role of ability in students' responses to grades. Becker & Rosen (1992) and Betts (1998) show theoretically that higher grading standards encourage high ability students to put more effort, while students below standard might be discouraged. Betts & Grogger (2003) empirically confirm the heterogeneous effects of increasing grading standards at the high school level, while Figlio & Lucas (2004) find that higher standards lead to positive results on test scores, with effects that depend on the ability of the student relative to the class. In my setup untreated students do not observe grades, but only test scores. Absent grades, low-SES students are likely to have lower grading standards than high-SES students (for

[6] Early grade assignment has a bigger impact in tracked education systems because students face early choices, and benefit more of timely information about ability. This point has not received much attention in the tracking literature (e.g., Brunello & Checchi, 2007).

instance because the difficulty of the tests follow class ability), so that introducing grades should lead to positive effects for high-SES students and negative effects for low-SES students. My results do not confirm this, as I find different reactions to grades within SES.

The grading reform I consider has been previously studied in economics by Sjögren (2010) and in the educational psychology literature by Alli Klapp (2014, 2015).[7] Sjögren's paper uses administrative data to study long-run effects (final education and income) of the overall grading reform. She finds evidence of a positive effect of early grading on educational attainment for girls, and a negative effect for high-SES students. Differences in educational attainment are found also before and after the reform took place, which casts some doubts on the robustness of the results. My paper focuses on the mechanisms through which grades affect education choice, and is motivated by a learning model. Results appear to be more robust, as tests for parallel trends in educational attainment do not fail. This is likely due to the different cohorts used: Sjögren needs to assume parallel trends over two decades, while I only need to assume parallel trends within a 5-year period.

The paper proceeds as follows. In Section 2 I describe the data, the education system, and the grading reforms. In Section 3 I set up the sequential choice learning model that guides the empirical analysis, and illustrate the solution of the model. Section 4 discusses the model's results. In Section 5 I turn to the empirical analysis, and discuss identification, inference and robustness. Section 6 discusses empirical results, while Section 7 relates them to the literature. Section 8 draws conclusions.

2 Institutional Setup

2.1 Data

I use survey data matched to administrative data. The surveys are part of *Evaluation Through Follow-up* (ETF), a longitudinal project which surveys every 5 years representative samples of Swedish students enrolled in compulsory school. I use waves 3 and 4 of the study, corresponding to cohorts born approximately in 1967 and 1972.[8] The 1967 cohort was followed from 1980, when students were in school year 6 (most students were 13 at the time). The 1972 cohort was followed from 1982, when students were in school year 3 (most students were 10 at the time).

[7] Klapp's papers are descriptive regression-control studies

[8] In the following I will refer to the two samples as 1967 and 1972 cohort.

Each sample consists of roughly 9000 Swedish compulsory school students (10% of the targeted population) living in 29 (out of 290) municipalities, the lowest administrative division in Sweden. Whole classes were systematically sampled from municipalities, and the same municipalities were extracted in both waves.[9] The final sample is thus a repeated cross-section, which allows me to implement a difference in differences identification strategy.

The survey data contains relevant information for the analysis. First, sampled students took standard intelligence tests in verbal, logical and spatial ability in school year 6, before end-of-the-year grades were assigned. The tests are exactly the same for both cohorts, which grants comparability of the intelligence measures over time. At the time of the tests students were 13, a point in which IQ should have already stabilized (Cunha & Heckman, 2009). I can thus investigate the effects of early grading using proper measures of ability, rather than previous performance measures. Second, grades and course choices in compulsory school are recored from school registers. This allows me to inspect the effect of early grading right after grades were assigned. Third, children filled in detailed surveys in school years 6 and 10 (the first year of high school). They were asked questions about own ability, courses and high school track choices, well being and motivation in school. I use children responses about stress, anxiety, and motivation as outcomes to understand whether early grades had motivating/demotivating effects on the children, a main concern in the policy debate. Finally, parents were surveyed when children received their first survey. They were asked questions about school choices and priorities. This evidence helps to understand whether and to what extent choices of parents living in early grading municipalities differ from those of parents living in municipalities where early grades were abolished.

I match to the sample high quality register data from Statistics Sweden. For both cohorts I observe parental education, income and demographics. These variables allow me to test for compositional change in the sample, and allow to increase precision in the main specification. The registers record educational attainment, income, and income mobility at ages 33-40 for both cohorts. This allows me to evaluate how the short- and medium-run effects of early grading transmit to the labor market.

[9]Municipalities are drawn using stratified sampling. Strata are defined by population, fraction of left-wing voters, fraction working in the public sector and fraction of immigrants. The three biggest municipalities in Sweden (Stockholm, Malmö, Gothenburg) are always part of the sample. Further details on the sampling scheme can be found in Emanuelsson (1979).

2.2 The Education System

Table 1 summarizes the Swedish educational setup for the two cohorts in my sample. Compulsory school (*Grundskola*) started at age 7 and lasted 9 years. It was formally divided in three stages, that could also entail physically changing schools: early compulsory school (grades 1-3), middle compulsory school (grades 4-6), and late compulsory school (grades 7-9). Standardized end-of-the-year grades were released at the end of each education cycle, and in every year during late compulsory school. Early grades were over time abolished. The next section provides details about the grades and the grading reforms.

Table 1: Structure of Swedish education

<table>
<tr><th></th><th colspan="2">Compulsory school</th><th colspan="2">Non Compulsory school</th></tr>
<tr><th></th><th>Early and Middle</th><th>Late</th><th>High School</th><th>College</th></tr>
<tr><td>Age</td><td>7-12</td><td>13-15</td><td>16-19</td><td rowspan="5">Selection:
- HS track
- GPA or SweSAT

Funding:
- Free tuition and grants
- Loans for living expenses</td></tr>
<tr><td>School Year</td><td>1-6</td><td>7-9</td><td>10-12</td></tr>
<tr><td>Grades</td><td>(3), (6)</td><td>(7), 8, 9</td><td>10-12</td></tr>
<tr><td>Choices</td><td>-</td><td>General or advanced courses</td><td>Vocational or academic track</td></tr>
<tr><td>Selection</td><td>-</td><td>-</td><td>GPA and course choices</td></tr>
</table>

The education system was tracked. In the spring of school year 6 children had to choose whether to take math and English at the advanced or general level in the next school year. Academic electives provided better preparation for academic tracks in high school. Students were allowed to switch course type over time. At the end of compulsory school, students could enroll in either academic or vocational high school tracks. Vocational tracks lasted two years, provided professional training, and did not allow direct access to college. Academic high school lasted three or four years, prepared for college, and was selective.[10]

After academic high school graduation (or taking one more year of high school after vocational school) students became eligible to apply to college. A student quota, set by the government, limited access to college. Slots were competitively assigned to the stu-

[10] A high grade 9 GPA and advanced math electives in compulsory school could be used as admission requirement.

dents with highest GPA or *SweSAT* (a college entry test similar to the American SAT).[11] College was tuition-free, and a mix of grants and income-contingent loans allowed admitted students to pay for living expenses. Higher education was thus both meritocratic and competitive. Appendix B.2 presents detailed evidence on education choices and attainment for sampled students.

2.3 Grades and the Reform

Standardized grades in math, English and Swedish were assigned at the end of specific school years during compulsory school. Grades were norm-referenced at the national level: they represented student performance with reference to the whole student cohort.[12] Given that only homework and test scores were assigned during the school year, grades provided students with additional information about school performance. In particular grades provided a first idea of their chances of admission to college, which was restricted by a quota system.

The school year in which grades were first assigned was over time postponed from school year 3, when students were 10, to school year 8, when they were 15.[13] Up to 1968 grades were assigned in school years 3, 6, 7, 8 and 9. In 1969 a curriculum reform (*Curriculum Lgr* 69) allowed municipal school boards to abolish "early" grades, that is, grades in school years 3 and 6. As a substitute for the abolished grades the reform introduced parent-teacher conferences, non-compulsory biannual meetings in which teachers evaluated pupil improvement over the year. Sjögren (2010) reports that supporters of early grade abolition were concerned about early grades harming low SES or poorly performing students. The idea behind the reform was that of making the class environment less competitive and more inclusive.

Since 1969 more and more municipalities took the chance to abolish grades in the early school years, but the issue was contentious. Left parties (Social Democrats and Communists) in general favored early grades abolition, while right-wing parties (Center party and Moderate Party) leaned towards keeping the early grades (this is confirmed in Figure B.2 on page 81). In the end the government, led by a socialist majority, chose

[11] Öckert (2002) reports that around 50% of the students were rejected admission to college in the period I study, confirming the selective nature of Swedish higher education.

[12] Tests were corrected by the teachers. The government used the scores to determine the national grade distribution. When assigning final grades, teachers could deviate from test scores, if they thought the student test performance did not reflect proficiency.

[13] In 2012 grades were reintroduced in school year 6, and the government is considering assigning grades also in school year 4.

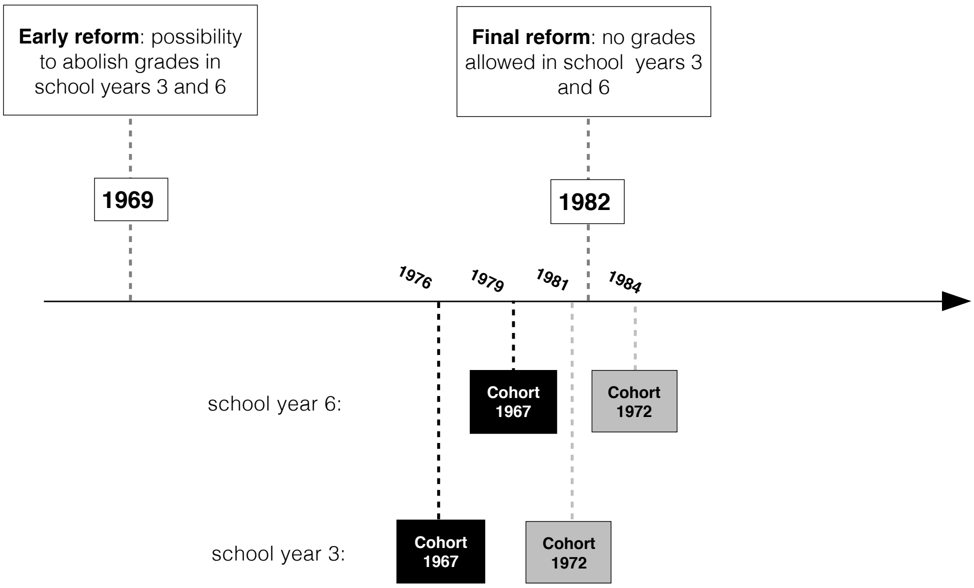


Figure 1: Grading reform timeline and sampled cohorts

to abolish "early grading" in all municipalities: starting from 1982 (*Curriculum Lgr* 80) end-of-the-year grades were released only starting from school year 8, when children were 15.

Figure 1 shows in a timeline how the reforms affected the two cohorts in the sample. Half of the municipalities in the 1967 cohort sample were assigning grades in school year 6, while the rest had abolished them.[14] Grade assignment in school year 3 is not recorded in my data for this cohort, as the children were followed from school year 6. Using information provided in Sjögren (2010), I can assume that the municipalities assigning grades in year 6 could have been also assigning grades in year 3. However municipalities not assigning grades in year 6 should have also abolished grades in school year 3. No municipality in the 1972 cohort sample was assigning grades in school year 6. While the final reform was effective the year after the children born in this cohort were in school year 3, my data reports that no grades were assigned in school year 3. Finally end-of-the-year grades were assigned for all cohorts and municipalities in school years 8 and 9.[15]

In the following I emphasize the role of grades in school year 6, rather than school year 3. First, treatment status in my analysis is based on grade assignment in school year 6. Second, grades at the age of 13 are arguably more relevant than grades at age 10, the end of early compulsory school. At that stage, grades might be more informative of

[14] Figure B.3 shows in a map which sampled municipalities were assigning early grades.

[15] Differently from earlier school years, they were assigned two times per year, at the end of each semester. Details are taken from Skolverket.

effort, or preferences for education, rather than academic ability. Finally, after school year 6 students had to choose whether to take math and English at the general or advanced level. Grades in school year 6 should thus be more relevant for education choices.

Table 2: Grade assignment

	1967 Cohort	**1972 Cohort**
Early Grading Municipalities (Treatment)	(Year 3) **Year 6** Year 7 Year 8 Year 9	Year 8 Year 9
Late Grading Municipalities (Control)	Year 7 Year 8 Year 9	Year 8 Year 9

Table 2 summarizes the grading structure. I label "treatment municipalities" those municipalities that were assigning grades in school year 6 before the final reform, "control municipalities" those not assigning grades in school year 6 before the final reform. The treatment is thus receiving early grades in school year 6 (and potentially 3), which holds for students born in 1967 who lived in early grading municipalities.

3 Model

The model presented in this section investigates how early grading affects students' education choices and attainment when grades convey information about ability. The qualitative predictions of the model are compared to empirical results in Section 6.

3.1 Structure of the Model

The model focuses on the link between early education choices, educational attainment, and lifetime income. I model explicitly early phases of education, and treat non-compulsory education and the labor market as realizations. The structure of the model is illustrated in Figure 2. Compulsory education is divided, as in my setup, into three periods: early compulsory school (t_1), middle compulsory school (t_2), and late compulsory school (t_3). In

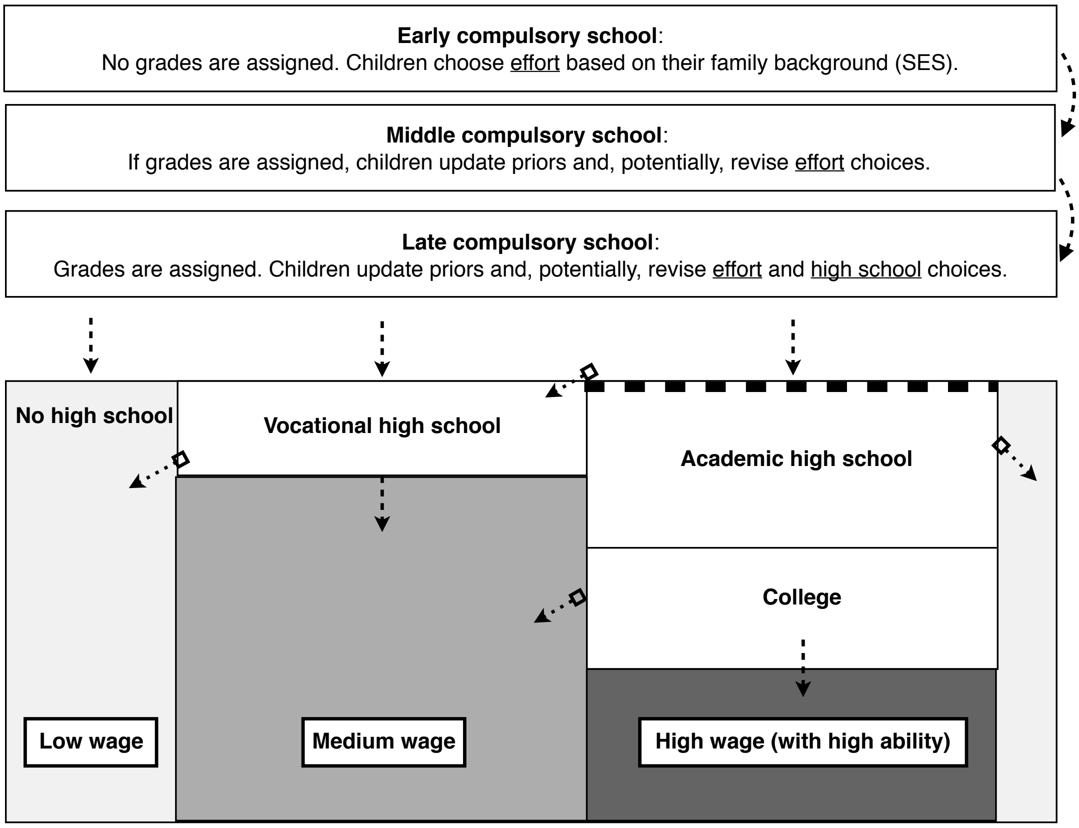


Figure 2: Structure of the model

each period student i chooses how much effort to exert: $e_{it} \in \{1, 2, 3\}$. Effort choices and academic ability (a_i) determine the stock of knowledge (k_{it}) the student accumulates:[16]

$$k_{it} = \omega_t(\alpha a_i + \beta e_{it}) + \delta k_{it-1}. \tag{1}$$

After the end of compulsory education students have three choices. They can go to work (E_1) and earn low wages (w_1). They can enroll into vocational high school (E_2) and study for two years, or enroll into academic high school (E_3) and study for 3 years. Both high school tracks grant medium wage ($w_2 = w_3$) upon graduation. Academic high school is the only option that gives access to college (E_4), which lasts 4 years and grants upon graduation wages that increase with academic ability: $w_4 = f(a_i)$.[17]

Completing higher levels of education and accessing academic high school requires

[16]The three stages of education have different lengths in my setup. Weights ω_t adjust the length of each stage to mimic the actual setup.

[17]As this is a stylized model, returns to education do not reflect the substantial wage heterogeneity documented in the literature (Arcidiacono, 2004; Hussey et al, 2011).

higher knowledge (and thus higher ability and effort) at the end of period 3. Notice that knowledge is here not productive per se, as wages fully depend on attained education and ability. This is consistent with a signaling model where employers are uncertain about workers' ability, but observe attained education (Spence, 1974). Higher education is attained in equilibrium only by high-ability workers, who fetch higher wages in the market. The knowledge thresholds at time 3 are the following:

$$\bar{k}^{E_2} < \underline{k}^{E_3} < \bar{k}^{E_3} < \bar{k}^{E_4}, \tag{2}$$

where $\underline{k}^{E_j}$ and $\bar{k}^{E_j}$ are respectively the entry and attainment requirements for education level E_j. Failure to meet the thresholds results in dropout (assumed at the midpoint of each education level), and thus in foregone earnings. Given that high school grants the same wage independently of track, it is optimal to enter academic high school only under the expectation to be able to complete college.[18] Academic ability indeed determines optimal education and effort choices. Low-ability students ($a_i \in \{1, 2, 3\}$) optimally choose vocational school, and put during compulsory school levels of effort inversely proportional to their ability: to reach the same education level, a weaker student needs to exert higher effort in school. The optimal education choice of high-ability students ($a_i \in \{4, 5\}$) is academic high school, and thus college. To attain college education they need to exert higher effort in compulsory school.

Students are uncertain about academic ability: $\tilde{a}_{it} \sim f_t(a_i)$. They have priors reflecting the ability distribution by SES: $f_1(a_i) = f(a_i|SES)$. In particular low-SES students have on average lower ability than high-SES students. Grades are unbiased ability signals, and allow students to update their priors about academic ability. They can be assigned in middle compulsory school, and are always assigned in late compulsory school, before students choose high school track:

$$g_{i2} = d(a_i + \epsilon_2) \text{ with } \epsilon_2 \sim N\left(0, \sigma_2^2(\epsilon)\right), \tag{3}$$

$$g_{i3} = d(a_i + \epsilon_3) \text{ with } \epsilon_3 \sim N\left(0, \sigma_3^2(\epsilon)\right), \tag{4}$$

where d is a function that maps the normal values into the discrete ability scale. Grades assigned in late compulsory school are more precise than grades assigned in middle compulsory school: $\sigma_2^2(\epsilon) > \sigma_3^2(\epsilon)$. This reflects the fact that more grades are assigned in the last period of compulsory school. Table 3 makes explicit the information structure in the three periods. If early grades are not assigned in period 2, students' beliefs remain

[18] Notice also that I do not model entry to college, and simply consider people staying out as college dropouts.

Table 3: Information structure

	Early grades	Late grades
$f_1(a_i)$	$f(a_i\|SES)$	$f(a_i\|SES)$
$f_2(a_i)$	$f(a_i\|g_{i2}, SES)$	$f(a_i\|SES)$
$f_3(a_i)$	$f(a_i\|g_{i3}, g_{i2}, SES)$	$f(a_i\|g_{i3}, SES)$

unchanged: $f_2(a_i) = f(a_i|SES)$. Otherwise they are updated: $f_2(a_i) = f(a_i|g_{i2}, SES)$. In period 3 grades are always assigned, so that $f_3(a_i) = f(a_i|g_{i3}, SES)$ if no grades are assigned in middle compulsory school, and $f_3(a_i) = f(a_i|g_{i3}, g_{i2}, SES)$ with early grades. Finally, when students update their priors about ability in period τ, they revise their beliefs about accumulated or future knowledge at any time t:

$$\tilde{k}_{it,\tau} = \sum_{j=1}^{5} P_\tau(a_i = j) \times \left[\omega_t(\alpha j + \beta e_{it}) + \delta \widetilde{k}_{it-1,\tau}\right]. \tag{5}$$

3.2 Optimal Choice

Given their information about ability in period τ, students consider optimal effort and education choices in any subsequent period $t \geq \tau$. They choose the education level that gives the highest utility, and the associated optimal effort level $e^{E\star}_{it,\tau}$:

$$e^{\star}_{it,\tau} = \arg\max_{e^{E_s\star}_{it,\tau}} \left\{ V_i^{E_1}(e^{E_1\star}_{it}, V^{E_2}_{i,\tau}(e^{E_2\star}_{it,\tau}), V^{E_3}_{i,\tau}(e^{E_3\star}_{it,\tau}) \right\}. \tag{6}$$

The value of compulsory school, V_i^1, does not depend on ability, and is thus not indexed by time. It is maximized when effort is set to the lowest level, so it is a constant: $V_i^{1\star} = k$. As vocational and academic high school have access and attainment requirements, values $V^{E_2}_{i,\tau}$ and $V^{E_3}_{i,\tau}$ depend on students' beliefs about ability. They are indexed by the time index τ, as $\tilde{k}_{i3,\tau}$ changes when new information is revealed:

$$\begin{aligned} V^{E_2}_{i,\tau} = {} & \textstyle\sum_{t=\tau}^{3} C(e^{E_2\star}_{it,\tau}) + P(\tilde{k}_{i3,\tau} \geq \bar{k}^{E_2}) \times U((L-2) \times w_2) \\ & + P(\tilde{k}_{i3,\tau} < \bar{k}^{E_2}) U((L-1) \times w_1)) \end{aligned} \tag{7}$$

$$\begin{aligned} V^{E_3}_{i,\tau} = {} & \textstyle\sum_{t=\tau}^{3} C(e^{E_3\star}_{it}) + P(\tilde{k}_{i3,\tau} \geq \bar{k}^{E_4}) \times U((L-7) \times w_3(a_i)) + \\ & P(\bar{k}^{E_3} \leq \tilde{k}_{i3,\tau} < \bar{k}^{E_4}) \times U((L-5) \times w_2) + P(\underline{k}^{E_3} \leq \tilde{k}_{i3,\tau} < \bar{k}^{E_3}) \times U((L-2) \times w_1) \\ & P(\tilde{k}_{i3,\tau} < \underline{k}^{E_3}) \times \left[P(\tilde{k}_{i3,\tau} \geq \bar{k}^{E_2}) \times U((L-2) \times w_2 + P(\tilde{k}_{i3,\tau} < \bar{k}^{E_2}) \times U((L-1) \times w_1)\right]. \end{aligned} \tag{8}$$

C is a convex cost function, U is a concave utility function, and L is the number of working years.

The effect of grades

When students are assigned grades they update their priors in the direction of their true ability level. Figures A.5 to A.9 in Appendix A.2 show priors and posterior distributions of ability after grades are assigned. Updating can have two effects: an *"income"* and a *"substitution"* effect. When the student realizes she has higher (lower) ability than expected, she revises the level of knowledge accumulated upward (downward). Provided the optimal education choice has not changed, the student will need to put weakly less (more) effort to reach the level of non-compulsory education she was targeting, an *"income effect"*:

$$\left.\frac{\partial e_{it}^{E_s\star}}{\partial \tilde{a}_{it}}\right|_{E_t^\star = E_{t-1}^\star} = \frac{\partial \tilde{k}_{i3}^t}{\partial \tilde{a}_{it}} \times \frac{\partial e_{it}^{E_s\star}}{\partial \tilde{k}_{i3}^t} \leq 0. \tag{9}$$

If after observing the signal expected ability is high (low) enough to alter optimal educational choice, the student will instead revise effort choices upward (downward), a *"substitution effect"*:[19]

$$\left.\frac{\partial e_{it}^{E_s\star}}{\partial \tilde{a}_{it}}\right|_{E_t^\star \neq E_{t-1}^\star} = \frac{\partial \tilde{k}_{i3}^t}{\partial \tilde{a}_{it}} \times \frac{\partial e_{it}^{E_s\star}}{\partial \tilde{k}_{i3}^t} \geq 0. \tag{10}$$

4 Model's Results

Before discussing the model's predictions, it is important to be clear about the purpose of the model. First, the model is meant to qualitatively assess the effect of early grades in the specific setup I consider. I calibrate to the data the key parameters of the model, ability distributions and education payoffs. I set thresholds for educational attainment such that higher levels of education require both higher ability and effort. Parameters with no direct counterpart in the data (knowledge production function, precision of grade signals, and value function parameters) are fixed to specific values.[20] Appendix A.1 contains further details on calibration, and provides evidence on model assumptions. Second, I do not estimate the model. While this might be an interesting direction for future research, my

[19] Higher education levels always require higher knowledge.

[20] Results remain qualitatively the same when slightly changing the parameters. Extreme parameterizations lead to different predictions, but are also inconsistent with the data observed.

aim here is to generate qualitative predictions of the effect of early grading in a learning model, rather than fitting the data.

I solve numerically the model under three different information setups: late grade assignment, early grade assignment and, as a benchmark, full information.[21] In Table 4 I show as a reference optimal effort and education choices by ability level under full information. The *"income effect"* is clear for both low- and high-ability students: for

Table 4: Optimal choices under full information

	a_i	e_{i1}	e_{i2}	e_{i3}	E	$V_{i,1}^E$
Low-ability	1	Medium	Medium	Medium	Vocational	106.53
	2	Medium	Medium	Low	Vocational	112.63
	3	Low	High	Low	Vocational	115.84
High-ability	4	High	High	Medium	Academic	126.49
	5	Medium	High	Medium	Academic	155.48

higher levels of ability it is optimal to put less effort. The *"substitution effect"* appears when ability changes from 3 to 4: students need to put higher effort early on in order to be able to attain college education.

4.1 Effort in Compulsory School

Figure 3 shows optimal effort choices in t_1, before grades are assigned. The fact that additional information will arrive in t_2 might change effort choices before grades are released. This is not the case in the model. Under uncertainty about ability, it is always optimal for both low- and high-SES students to keep effort at a medium level. This is due to three reasons. First, uncertainty favors higher effort early on: putting low effort in the beginning might actually prevent the student from entering academic high school, and thus college. Second, even if the student learns that she is high-ability in time, she would then need to compensate for previous low effort levels: effort cost is convex, so this behavior would not be optimal. Finally, knowledge production is cumulative, so it is better to exert higher effort early on, when effort is more productive.

In middle compulsory school students can be assigned grades. Figure A.10 in Appendix A.3 compares posterior distributions of ability for low and high SES students who get the same grades in t_2. While all students update priors in the right directions, updates differ by SES. Low (high) SES students who receive low (high) grades confirm their priors, and thus their posterior distributions have higher densities on low (high) ability levels.

[21]Appendix A.2 presents the simulation and solution methods.

Students who receive grades different from their priors form instead posterior distributions with higher weight on intermediate values of ability. These posteriors are thus also less precise.

Figure 4 shows the effect of early grading on effort choices in t_2, by aggregate (low or high) ability and SES. Results for each ability level, reported in Appendix A.3, are useful to better interpret the aggregate picture, so in the following I refer to both pictures. Early grading changes optimal behavior in middle compulsory school only for high-SES students: students who observe signals consistent with high-ability put higher effort (see Figure A.11). As shown in table 4, this is consistent with optimal education choice: for high-ability students it is optimal to exert high effort in middle compulsory school, and then reduce effort in late compulsory school. Low-SES students do not react differently at this stage, independently of ability. Their priors are set lower, and hence posteriors about ability are less sensitive to the high grades they observe.

In t_3 all students are graded. Figure 5 shows that high-SES students with high-ability strongly react to the additional grades, and put lower effort. Together with the reaction in middle school, this can be overall interpreted as a negative "*income effect*". High-ability low-SES students react to early grades in the opposite way: they increase effort. Against their priors, these students realize they are high-ability. They thus switch education and effort choices (*"a substitution effect"*). Low-ability students reduce effort when graded early on. Figure A.12 shows that the strongest reductions are found among low-SES students. The aggregate effect for low-ability students masks a positive *"income effect"* among lowest ability students. Figure A.12 shows that, when graded early on, these students put more effort to reach the same education level they targeted (an *"income effect"*). This effect is strongest among low-SES students.

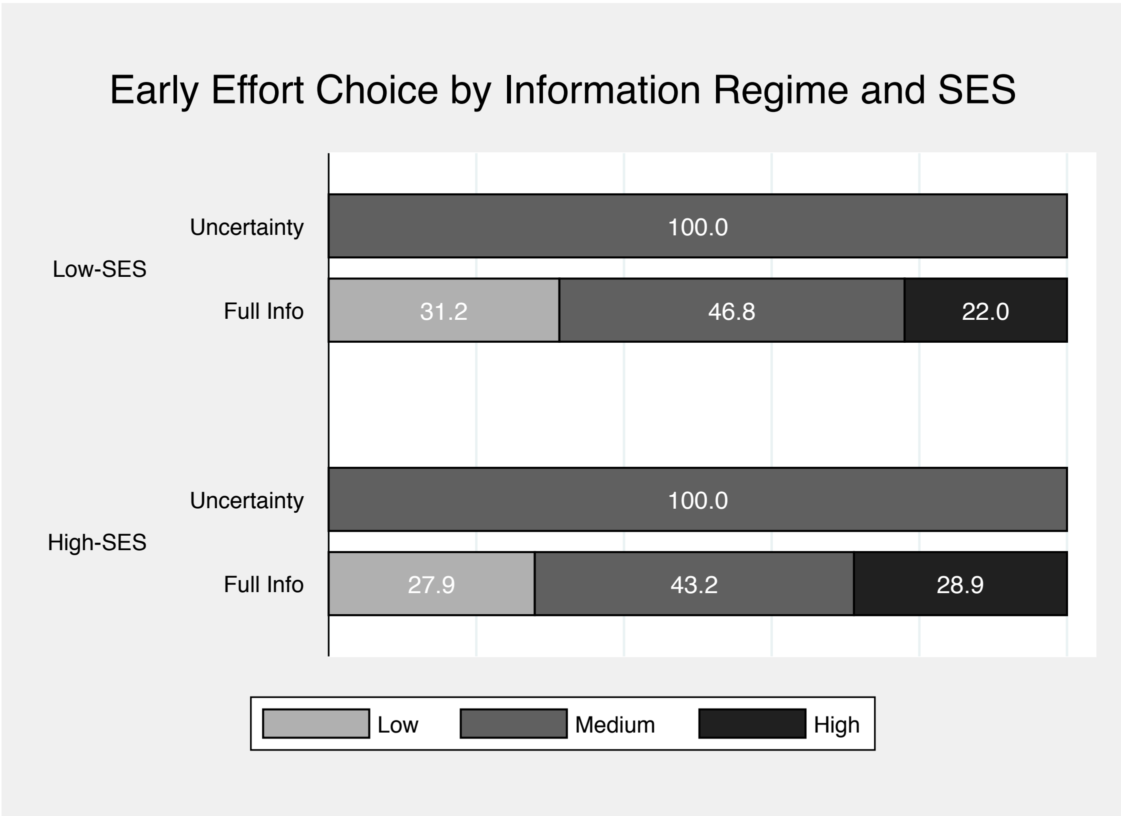


Figure 3: Early effort choice by SES and information regime

Note: The Figure plots effort distributions in early compulsory school. Since in this period no grades are assigned, choices are the same for both low and high ability students, and can only differ by SES. Assigning early or late grades does not change effort choices in t_1.

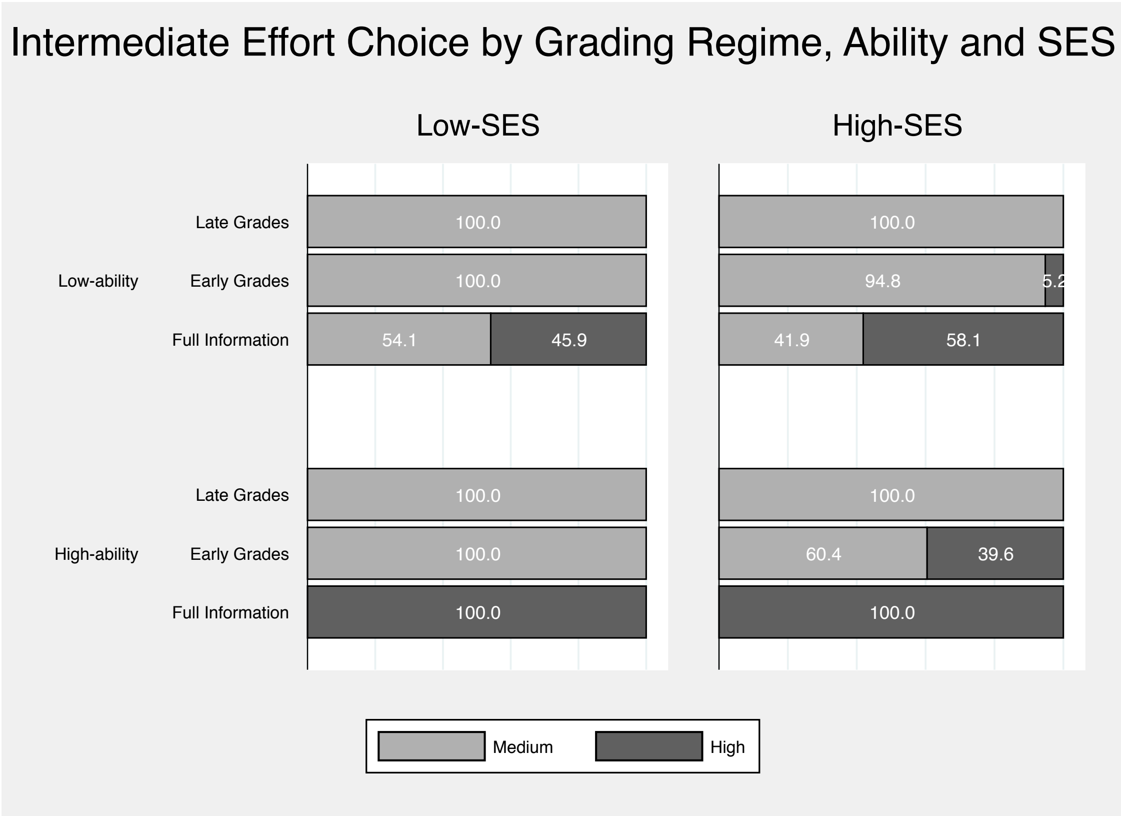


Figure 4: Intermediate effort choice by aggregate ability and SES for different grading regimes

Note: The Figure plots effort distributions in middle compulsory school. Results are presented for high-ability students, for whom it is optimal to follow an academic education path, and low-ability students, whose optimal choice is vocational high school. SES affects students's priors about ability, and thus optimal choices.

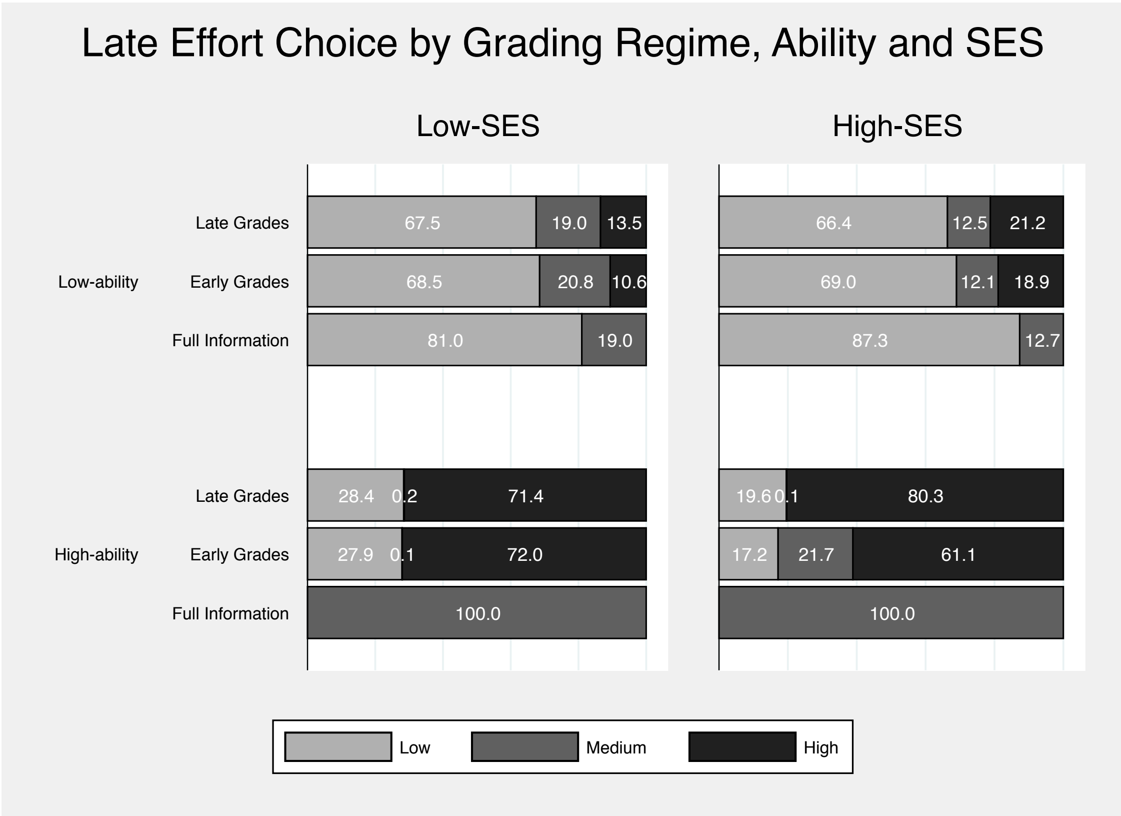


Figure 5: Late effort choice by aggregate ability and SES for different grading regimes

Note: The Figure plots effort distributions in late compulsory school. Results are presented for high-ability students, for whom it is optimal to follow an academic education path, and low-ability students, whose optimal choice is vocational high school. SES affects students's priors about ability, and thus optimal choices.

4.2 Education

Figures 6 and A.13 show high school choices in the three information regimes. When graded early on, low-ability students are less likely to choose academic education paths. The effect is stronger for low-SES students. All high-ability students are instead more likely to choose academic high school with early grades. Among students with high (but not top) ability, the reaction is stronger for high-SES students.

Figures 7 and A.14 show final education distributions for the different grading setups. The effects of early grades mirror those observed for education choice. The main difference is that some high (but not top) ability students with high-SES fail to attain college, and only complete academic high school. These students observed signals consistent with top ability early on, lowered effort, and thus failed to graduate from college (see Figure A.16). No such effect is found for low-SES students, who are actually less likely to dropout of both high school (see Figure A.15) and college (see Figure A.16).

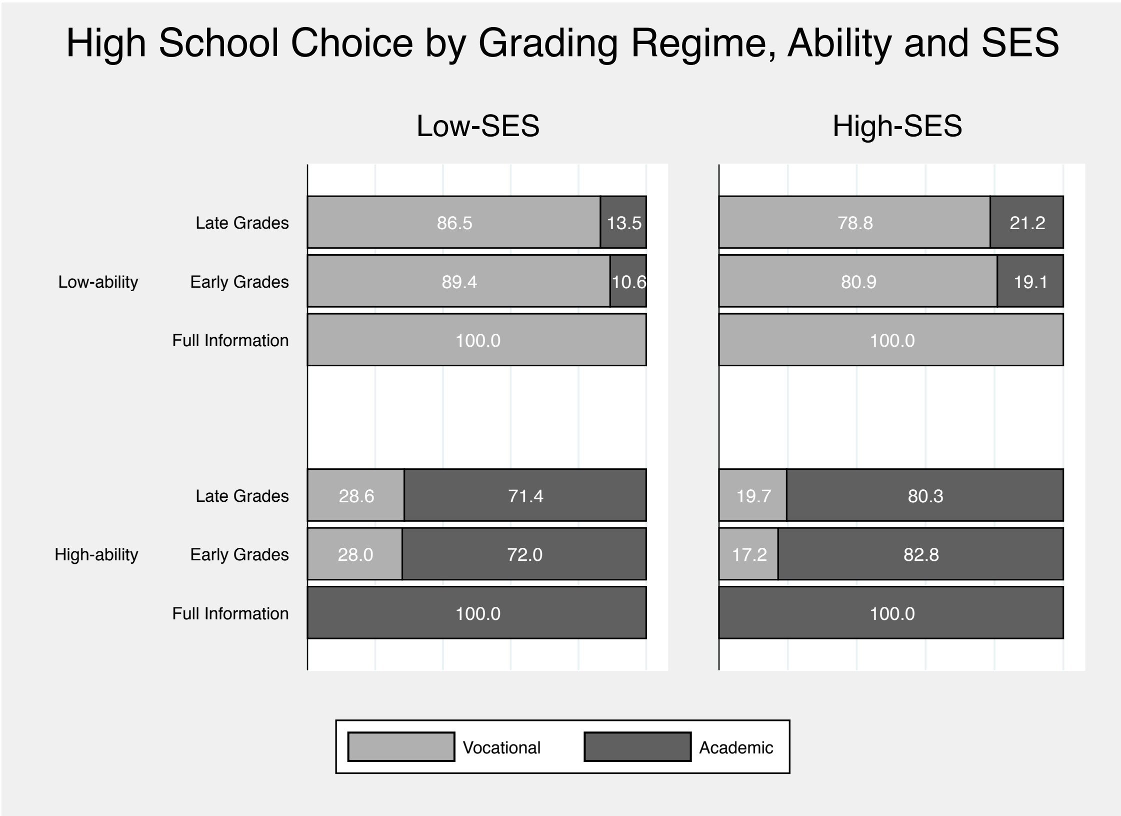


Figure 6: High school choice by aggregate ability and SES for different grading regimes

Note: The Figure plots high school choice distributions. Results are presented for high-ability students, for whom it is optimal to follow an academic education path, and low-ability students, whose optimal choice is vocational high school. SES affects students's priors about ability, and thus optimal choices.

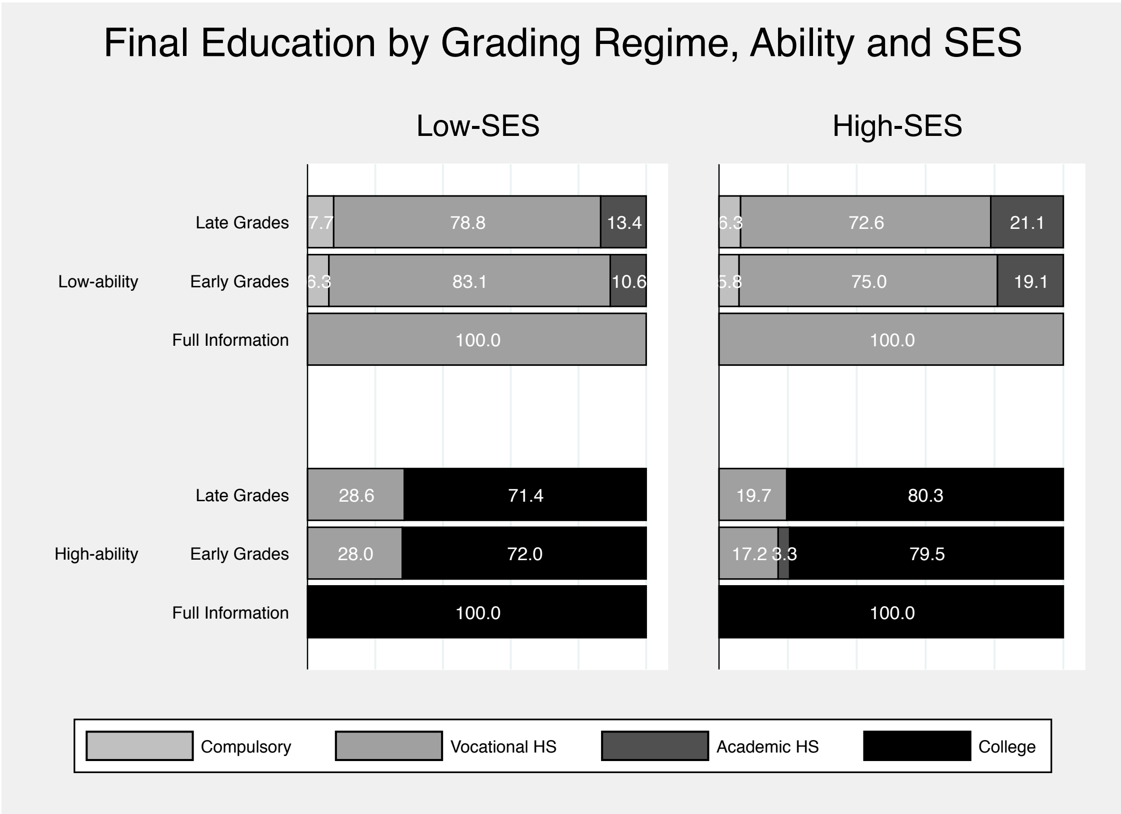


Figure 7: Final education by aggregate ability and SES for different grading regimes

Note: The Figure plots final education distributions. Results are presented for high-ability students, for whom it is optimal to follow an academic education path, and low-ability students, whose optimal choice is vocational high school. SES affects students's priors about ability, and thus optimal choices.

4.3 Summary of Results

In Table 5 I summarize the effects of early grading (the treatment) on education choices, educational attainment, and income. The effects are reported for each ability and SES group, and are compared to the baseline scenario (late grading, in brackets).

In general early grades lead to an overall reduction in effort.[22] Only high-ability low-SES students - for whom positive *"substitution effects"* prevail - increase effort when graded early on. While the mean reduction in effort is the same for all low-ability students, effects are qualitatively different by SES. As seen before, there are weaker negative *"income effects"* for high-SES students, and stronger positive and negative *"income effects"* for low-SES students. The biggest negative *"income effect"* on effort is found for high-ability high-SES students, who are the most sensitive to high grade signals.

Staying out of high school is never optimal in the model, even when students realize they have lower ability than expected. The rational choice for these low-ability students is to enroll into vocational school, and later dropout if they fall short of the required preparation. This does not change with early grades, which instead have nontrivial effects on high school track choices: all low ability-students are less likely to enroll into academic tracks, and the opposite is true for high-ability students. Positive reactions are strongest for high-SES students, while negative reactions are more pronounced for low-SES students.

Even if they reduced effort after observing early grades, all low-ability students benefit of the early information, due to the different choices taken at the end of compulsory school: dropout rates decrease, in particular for low-SES students. This translates one to one into an increase in high school attainment. Finally, college attainment increases for high-ability low-SES students, and slightly decreases for high-ability high-SES students.

In the long-run the effects of early grades on education translate into small increases in lifetime income for all students, with the exception of high-ability high-SES students. Effects on income are pretty small, and close to 0. Most of the gains in utility are due to the early reductions in effort, so that early grading improves on average the welfare of all students.

All in all the simulations show that assigning grades earlier leads to choices and education outcomes more consistent with academic ability, with responses differing by SES. Lowest ability students are more likely to increase effort when graded early on, especially if low-SES. Low to medium ability students reduce effort in compulsory school, in particular if low-SES, but are more likely to choose vocational tracks, which they are able to

[22] I take a weighted average of middle and late effort choices in order to provide a more complete picture on the effects of early grades on effort choice.

complete. High-ability low-SES students increase effort in compulsory school, are more likely to choose academic paths, and to attain college. For high-ability high-SES students, *"income effects"* tend to prevail: these students put less effort when they observe high grades, which leads some of them to fail to graduate from college.

Table 5: Summary of the effects of early grade assignment

Outcome:		**Low-Ability**		**High-Ability**	
	All Sample	**Low-SES**	**High-SES**	**Low-SES**	**High-SES**
Effort in late compulsory school	-0.03 [1.92]	-0.03 [1.55]	-0.03 [1.62]	0.01 [2.36]	-0.07 [2.51]
HS Enrollment	0.00 [1.00]	0.00 [1.00]	0.00 [1.00]	0.00 [1.00]	0.00 [1.00]
Academic track HS Enrollment	-0.01 [0.40]	-0.03 [0.13]	-0.02 [0.21]	0.01 [0.71]	0.02 [0.80]
HS Dropout	-0.01 [0.04]	-0.01 [0.08]	-0.00 [0.06]	0.00 [0.00]	0.00 [0.00]
Attains HS	0.01 [0.96]	0.01 [0.92]	0.00 [0.94]	0.00 [1.00]	0.00 [1.00]
Attains College	-0.00 [0.30]	0.00 [0.00]	0.00 [0.00]	0.01 [0.71]	-0.01 [0.80]
Income (0-1 scale)	0.00 [0.75]	0.00 [0.67]	0.00 [0.68]	0.00 [0.85]	-0.00 [0.88]
Utility	0.51 [117.01]	0.62 [107.83]	0.39 [107.20]	0.23 [128.91]	0.67 [133.17]

Values in brackets represent outcomes when only late grades are assigned. Effort is defined on a 1-3 scale (1 is low effort). Income is a measure of lifetime income, and assumes everybody starts working right after finishing their education or dropping out.

5 Empirics

In this section I discuss identification of the effect of early grading on education choices. I then briefly discuss inference in my setup, and lastly provide evidence on identifying assumptions.

5.1 Identification

The decision to assign early grades was taken by municipal school boards, and, as previously discussed, correlates with the political color of the municipality. Treatment assignment is thus likely not random with respect to education outcomes. A simple comparison of outcomes between grading and non-grading municipalities would pick up systematic differences between the two sets of municipalities, and thus bias OLS.

In Appendix B.3 I test for differences in pre-treatment variables between graded and non-graded municipalities in the 1967 cohort. Table B.7 shows that in graded municipalities children are less likely to be foreign born, score better in the ability tests, and are less likely to switch classes over compulsory school. In terms of school level variables (changes of teachers, class size, kindergarten) there are no big differences, in line with the homogeneous nature of Swedish education. Parents in grading municipalities (Tables B.8 to B.12) are less likely to divorce and more likely to be married. They are slightly poorer, less educated, and more likely to be employed in low-skill jobs or agriculture. When asked about how they chose math and English courses, and the priorities of Swedish education, parents give very similar answers. The only differences, the weight they put on the role of parents in school choice and critical thinking in school, do not seem to imply a different preference for children educational attainment. Altogether it appears that there are some small differences in determinants of education choice between the two sets of municipalities. The differences in parental education seem to reflect a different structure of the economy, rather than different preferences for education.

A simple cross-sectional comparison of outcomes for treated and untreated municipalities would likely lead to a negative bias, due to the pre-existing differences between treated and control units. Given that I observe treatment and control group before and after the final reform, when early grades were abolished, I can "control" for any persistent difference between the two sets of municipalities. If outcomes trend in the same way in the two municipalities (*parallel trends assumption*), it is possible to isolate the effect of early grades. This situation is pictured in Figure 8: while the two sets of municipalities exhibit differences in outcomes unrelated to grade assignment, these differences are stable over time. Observed outcomes for the 1967 treated cohort can be compared to the

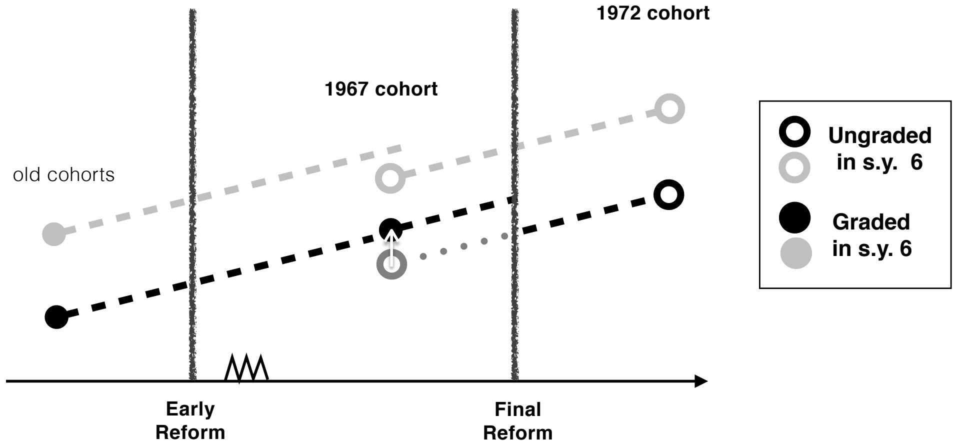


Figure 8: Difference in differences identification strategy

counterfactual outcomes that would have been observed for the same set of municipalities absent the treatment (early grades). This counterfactual is given by the trend observed for the ungraded municipalities, assumed to be the same for treated municipalities. The effect of early grading is represented in the picture by the white arrow. The empirical specification that implements the difference in differences identification strategy is the following:

$$Y_{imc} = \alpha + \beta_{as}\text{Graded}_m \times 1967_c + 1967_c + \text{Munic}_m + \Delta X_{imc} + \epsilon_{imc} \tag{11}$$

$$a \in \{\text{Low ability, High ability}\};\ s \in \{\text{Low SES, High SES}\},$$

where i indexes the individual, m the municipality, and c the cohort. Munic_m is a vector of fixed effects that captures persistent cross-sectional differences between municipalities. 1967_c is a dummy that controls for the trend in outcomes. The variable $\text{Graded}_m \times 1967_c$ picks up any differences in outcomes between grading and non-grading municipalities, that are not persistent, or the same, over time. Under the parallel trends assumption β_{as} represents the causal effect of early grading. Consistently with the model, the effect is allowed to differ by ability and SES, indexed respectively by a and s in equation 11. SES and ability are measured respectively using parental education and ability tests administered in school year 6.[23] Notice that any determinant of the outcome that changes over time in a different way between the two sets of municipalities will also enter β_{as}, and thus bias the coefficient. Observable compositional change can be controlled for in

[23] Appendix B.1 provides further details on ability and SES measures, and on the way I discretize them to match the model.

the regression by adding X_{imc}, a vector of time varying pre-treatment controls. These covariates also increase precision of the estimates.

5.2 Inference

Sample size is large (around 18000 observation), but the treatment, grade assignment, varies at the municipal level. There are 29 municipalities in my sample, and half of them are treated before the reform. I conservatively cluster standard errors at the municipal level, rather than at the municipal-cohort level, which would result in twice as many clusters.[24] While the standard solution is to use cluster robust standard errors (Arellano, 1987; White, 1984), the number of clusters must be high for these standard errors to be unbiased. Cameron et al (2008) show that cluster-robust standard errors are downward biased in samples with few balanced (equally sized) clusters. They instead propose to use Cluster Bootstrap-t methods with null hypothesis imposed, and find that these methods yield the right p-values even with relatively few clusters (as few as 20). In a recent working paper MacKinnon & Webb (2014) confirm the good performance of the Cluster Bootstrap-t in the realistic case in which clusters are unbalanced. The Cluster Bootstrap-t is shown to perform well when treatment has enough variance.

My sample consists of 29 municipalities, both small and big. Treatment is given by the interaction between belonging to the cohort born 1967 and studying in an early grading municipality, which holds for about a quarter of the sample. There are thus enough clusters and treatment variation to believe that the Cluster Bootstrap-t should guarantee unbiased standard errors in my analysis. So in all my specifications I bootstrap standard errors using the method suggested by Cameron et al (2008). I also use sample weights to recover nationally representative estimates.

5.3 Testing for Identifying Assumptions

Difference in differences identifies the causal effect of assigning early grades under a specific set of assumptions. The most important one, as discussed before, is the parallel trends assumption: outcomes should trend similarly in both early grading and late grading municipalities. The assumption is more credible when the treated and untreated populations are not so different, especially in terms of "characteristics that are thought to be associated with the dynamics of the outcome variable" (Abadie 2005). This was shown to be the case above. In Appendix C.1 I use administrative data from Statistics Sweden to test whether

[24] This is suggested in Bertrand, Duflo, and Mullainathan (2004) for the case of panels. My final dataset is instead a cluster-panel, so there should be less correlation between clusters over time.

education and its determinants evolve in the same way in the two sets of municipalities: all tests pass. In particular trends in education for cohorts who went through compulsory school when all municipalities had abolished early grades (cohorts born 1969 onwards) appear to be parallel. The evidence thus supports the main assumption underlying the identification strategy.

A testable assumption of the identification strategy is that differences between treatment and control group in determinants of the outcome should be stable over time (e.g., there should be no compositional change). In the same way, response rates should be the same between treated and controls units over time (e.g., there should be no differential attrition).[25] In Appendix C.2 I test for differential attrition and compositional change in the sample. First, there is no differential response to the student surveys and, importantly, I find no differential attrition in availability of SES and ability data. Second, it appears that the cross-sectional differences between grading non-grading municipalities are broadly stable over time. I find compositional change in specific parental occupations and education levels. Therefore in my final specification I also control for occupational dummies and parental education.[26]

A further assumption in the difference in differences setup is that the treated population should not change as a reaction to treatment assignment. In my setup this means that the students born 1967 should not enroll into different schools to get/avoid early grades. As catchment areas determined the compulsory school the student attended, parents had to relocate to a different municipality if they desired a different grading policy for their children. Alternatively they could send their children to a private school. The first scenario seems highly unlikely, while private schools were not common in that period.

Finally it is important for identification that treatment and control group do not undergo different shocks over time. The presence of concurrent education reforms would be a problem in my setup if they were implemented at the municipal level. During the period I consider, schooling was quite centralized, with *national curricula* determining most of school policies. There is thus little scope for additional policies being differentially implemented in the two sets of municipalities. On top of that, the two cohorts I use in my analysis received their education in a relatively stable educational system: Sweden had already implemented the reforms of the 60s for the 9-year inclusive compulsory school, while the market-oriented school reforms of the 90s did not affect these cohorts.[27]

[25] Both compositional change and differential attrition can lead to biased difference in differences coefficients (Blundell & Costa Dias, 2009).

[26] Results are robust to excluding income and parental controls. I include those variables to increase precision of the estimates.

[27] The reforms are described respectively by Meghir & Palme (2005) and Björklund et al (2005).

6 Empirical Results

The outcomes in the empirical analysis match those of the model. This allows me to understand whether empirical findings are consistent with students learning about their academic ability from grades. I thus investigate the effect of early grades on short-term effort choices, high school choices and attainment, and, finally, educational attainment and income. I also consider an alternative mechanism through which grades might affect education choices: grades might motivate/demotivate students, and thus affect their welfare.

I present difference in differences estimates from specification 11, which I re-parametrize to directly get coefficients for each ability − SES cell. In all specifications I control for ability (verbal and inductive ability normalized to the cohort-treatment level), basic demographics (gender, birth year, foreign status, special education), SES (income, parental occupation dummies, and education) and school-level variables (class size and teacher changes). For every outcome I report the point estimate, the p-value in parentheses, and, as a reference, the sample mean in brackets.[28]

There are two caveats when interpreting results. First, estimates are not very precise, so I can not detect very small effects. Second, I test many hypotheses, which in principle creates problems of false null rejection. Notice that the two problems go in opposite directions, and that the multiple hypothesis testing problem is less severe than it seems: most of the outcomes are strongly correlated, or can be considered different proxies for the same underlying variable (e.g., grades and course choices proxy for effort choice). Keeping this in mind, when I interpret results I focus on the overall picture rather than on single coefficients.

6.1 Effort in Compulsory School

In Tables 6 and 7 I investigate effects of early grades on school effort. The first Table reports effects on math and English course choices, which can be interpreted both as effort choices (academic courses are more challenging), and as early school choices reflecting future track selection (advanced courses are good preparation for academic high school). The second Table reports effects on grades in late compulsory school, which are straightforward measures of school effort.[29]

Low-ability students, especially those with low-SES, reacted to early grade assignment by switching to non-academic English, which can be interpreted as a reduction in effort

[28]The wild cluster bootstrap with null imposed does not yield standard errors.

[29]This is especially true of Swedish, a subject that does not involve any additional choice.

(columns 2 and 3 in Table 6). The switches appear in grade 8, the first time in which the students could respond to grades released at the end of school year 6, and persist in school year 9.[30] Switches in course choice are observed for English, but not for math. One possible explanation is that parents and children already had feedback in math due to the correction of exercises. At this proficiency level parents could probably test children's math skills more than their English proficiency.[31] High-ability students did not revise course choices when graded early on.

Low-ability low-SES students exhibit worse math performance when graded early on (see column 2 and 3 of Table 7). High-ability high-SES students show instead higher English and Swedish grades when they receive the early grades. One can clearly see from the standardized Swedish test, which has more variation due to the different scale, that all low-ability students performed worse after being assigned early grades, while high-ability students performed better. Negative effects are stronger for low-SES students, positive effects are instead more pronounced for the high-SES students. In the aggregate no effect is found, as both positive and negative effects are summed up. This confirms the importance of looking at heterogeneous effects. The pattern found in the model is thus reproduced by the data: low (high) ability students are putting less (more) effort, and effects are stronger for low (high) SES students. However, high-ability high-SES students are putting more effort, rather than reducing it, as in the model. This implies that *"substitution"*, rather than *"income effects"*, are prevailing. This can be easily rationalized within the model, assuming that different majors require different ability levels. Then it is easy to see that these students would react to high grades by further increasing effort.

[30] The courses were chosen at the end of school year 6 for year 7, before final grades were released.

[31] The parents of the treated students were born in the 40s: at that time English proficiency was less widespread among parents than it is now the case in Sweden.

Table 6: Effects on course choices (school years 7-9):
Summary of difference in differences estimates

Outcome:		**Low-ability**		**High-ability**	
	All Sample	**Low-SES**	**High-SES**	**Low-SES**	**High-SES**
Advanced Math	0.00	-0.03	0.01	0.04	0.04
(school year 7)	(0.94)	(0.70)	(0.84)	(0.50)	(0.47)
	[0.73]	[0.54]	[0.72]	[0.90]	[0.95]
Advanced Math	-0.01	-0.00	-0.04	0.02	-0.00
(school year 8)	(0.80)	(1.00)	(0.26)	(0.40)	(0.94)
	[0.66]	[0.43]	[0.64]	[0.87]	[0.96]
Advanced Math	0.02	0.01	-0.00	0.03	0.05
(school year 9)	(0.64)	(0.74)	(0.97)	(0.58)	(0.16)
	[0.57]	[0.32]	[0.53]	[0.76]	[0.90]
Advanced English	0.00	-0.03	-0.01	0.05	0.05
(school year 7)	(0.90)	(0.60)	(0.88)	(0.15)	(0.21)
	[0.75]	[0.57]	[0.76]	[0.91]	[0.97]
Advanced English	-0.05**	-0.06***	-0.07	-0.01	-0.01
(school year 8)	(0.02)	(0.01)	(0.19)	(0.48)	(0.63)
	[0.73]	[0.53]	[0.73]	[0.91]	[0.97]
Advanced English	-0.06*	-0.07**	-0.08	-0.05	-0.01
(school year 9)	(0.06)	(0.02)	(0.12)	(0.14)	(0.58)
	[0.68]	[0.46]	[0.65]	[0.87]	[0.95]

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

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets. All specifications control for basic demographics, relative ability measures (standardized at the treatment-cohort level) and parental background.

Table 7: Effects on grades (school years 8 and 9):
Summary of difference in differences estimates

Outcome:		**Low-ability**		**High-ability**	
	All Sample	**Low-SES**	**High-SES**	**Low-SES**	**High-SES**
Math Grade	-0.02	-0.08*	-0.04	0.05	0.07
(school year 8)	(0.75)	(0.07)	(0.64)	(0.14)	(0.32)
	[3.04]	[2.70]	[2.87]	[3.35]	[3.59]
Math Grade	-0.11	-0.15**	-0.13	-0.04	-0.06
(school year 9)	(0.12)	(0.02)	(0.20)	(0.68)	(0.42)
	[3.20]	[2.86]	[3.03]	[3.54]	[3.73]
English Grade	0.06	0.00	0.10	0.02	0.17***
(school year 8)	(0.25)	(0.98)	(0.11)	(0.79)	(0.00)
	[3.05]	[2.70]	[2.86]	[3.37]	[3.61]
English Grade	0.05	-0.03	0.12	0.05	0.16***
(school year 9)	(0.21)	(0.57)	(0.11)	(0.48)	(0.00)
	[3.18]	[2.82]	[3.05]	[3.46]	[3.73]
Swedish Grade	0.03	-0.04	-0.03	0.12	0.17***
(school year 8)	(0.47)	(0.38)	(0.58)	(0.13)	(0.00)
	[3.06]	[2.64]	[2.91]	[3.43]	[3.68]
Swedish Grade	0.06	-0.02	0.03	0.14**	0.18***
(school year 9)	(0.17)	(0.76)	(0.64)	(0.01)	(0.01)
	[3.17]	[2.69]	[3.03]	[3.54]	[3.86]
Swedish Test	0.18	-3.64***	-1.54**	4.49***	6.14***
(school year 9)	(0.84)	(0.00)	(0.01)	(0.00)	(0.00)
	[15.84]	[13.43]	[13.25]	[19.96]	[18.51]

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

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets. Math and English pool together grades for advanced and general courses. All specifications control for basic demographics, relative ability measures (standardized at the treatment-cohort level) and parental background.

6.2 Education and Income

In Table 8 I report effects of early grades on high school choices, educational attainment, and income. Contrary to what the model predicts, early grades do not lead to different high school track choices. I observe instead an increase in enrollment for all students. While this can be surprising (on average low-ability students reduced effort in compulsory school), it is possible that lowest ability students increased effort early on, and thus decided to enroll into high school. This *"income effect"* was discussed in the model in Section 4.

When looking at educational attainment, I find an increase in high school attainment at age 17-20 for high-ability low-SES students, mostly explained by a reduction in high school dropout. In the long-run this effect becomes smaller and close to insignificant. In Sweden adult education programs (*Komvux*) allow people to complete further education: in the counterfactual scenario of late grading students might still have been able to finish their high school education. Moreover, I find that low-ability low-SES students are less likely to attain college. These effects are qualitatively consistent with model's predictions: a reduction in dropout due to higher effort in compulsory school, and less low-ability students ending up with an academic education.

Why do the short-run effects of early grades do not pass on to high school track choice, and why is educational attainment not affected for high SES-students? I propose as an explanation that preferences for education might attenuate the effects of early grades. In Appendix B.2 I show that, controlling for ability, academic high school enrollment rates of high-SES students are 20 percentage points higher than those of low-SES students. At the same time grade differences in late compulsory school between high- and low-SES students are at most $\frac{1}{4}$th of a grade. SES appears thus to strongly influence high school choices in Sweden, independently of ability.

While it is important to assess how early grades affect education outcomes to understand mechanisms, a full evaluation of the policy requires looking at long-run outcomes. Early grade assignment does not significantly affect income at ages 33-40, a good proxy of lifetime income in the Sweden labor market (Börklund, 1993). This is consistent with the theoretical model, which also generated very small effects on lifetime income. Early grading leads instead to an increase in upward income mobility among low-ability low-SES students, who displayed the strongest downward revisions in education choices.[32] I conclude that, from the perspective of the labor market, early grades simply allowed students to better sort by ability into education. For low-ability low-SES students this implies a reduction of over-investment in education, and potentially an in increase in total earnings.

[32] I consider upward mobile a student if she is 15 percentile ranks above the parents' income percentile rank.

Table 8: Effects on high school choices, educational attainment and income: Summary of difference in differences estimates

Outcome:		**Low-ability**		**High-ability**	
	All Sample	**Low-SES**	**High-SES**	**Low-SES**	**High-SES**
HS Enrollment	0.04**	0.03*	0.06**	0.03*	0.03**
(age 15-18)	(0.02)	(0.08)	(0.03)	(0.06)	(0.03)
	[0.89]	[0.85]	[0.92]	[0.93]	[0.97]
Academic HS Track	0.02	0.01	0.01	0.02	0.04
(age 15-18)	(0.55)	(0.82)	(0.85)	(0.68)	(0.13)
	[0.47]	[0.20]	[0.44]	[0.59]	[0.81]
HS Dropout	0.00	0.02	-0.01	-0.05**	0.02
(age 17-20)	(0.98)	(0.46)	(0.53)	(0.02)	(0.34)
	[0.13]	[0.18]	[0.12]	[0.10]	[0.08]
Attains HS	0.00	-0.02	0.02	0.06***	-0.01
(age 17-20)	(0.79)	(0.49)	(0.35)	(0.01)	(0.82)
	[0.79]	[0.72]	[0.83]	[0.86]	[0.91]
Attains HS	0.00	-0.02	0.03	0.02	-0.00
(age 33-40)	(0.94)	(0.19)	(0.10)	(0.14)	(0.87)
	[0.92]	[0.88]	[0.95]	[0.96]	[0.98]
College or more	-0.02	-0.03*	0.02	-0.04	0.00
(age 33-40)	(0.28)	(0.06)	(0.59)	(0.23)	(0.94)
	[0.43]	[0.22]	[0.42]	[0.52]	[0.75]
Gross income	3.28	11.49	-3.64	-6.76	0.78
(age 33-40)	(0.61)	(0.21)	(0.75)	(0.62)	(0.95)
	[259.11]	[223.88]	[256.31]	[269.89]	[330.13]
↑ *Income mobility*	0.04**	0.08***	0.02	0.01	0.02
(age 33-40)	(0.02)	(0.00)	(0.54)	(0.60)	(0.60)
	[0.34]	[0.38]	[0.27]	[0.45]	[0.28]

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

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets. HS Enrollment is measured at ages 16-18, HS attainment at age 40. Income is measured at ages 33-40. ↑ Income mobility is 1 when student income is 15 ranks above family income rank. All specifications control for basic demographics, relative ability measures (standardized at the treatment-cohort level) and parental background.

6.3 Student Welfare

Part of the policy debate in Sweden, and in other countries that considered early grades abolition, revolved around the concern that grades might demotivate (motivate) students who put low (high) effort independently of their ability, and create a competitive environment where weak students fare worse. If this was the case students would derive disutility (utility) from low (high) grades, and their preferences for education would be affected by their performance.

To validate this alternative mechanism I investigate the effects of early grading on self-reported child welfare. Outcomes are taken from the student surveys. The first survey was administered in school year 6, before final grades were assigned. It should pick up potential effects due to the more competitive/challenging environment. The second survey was assigned in school year 10, and asked many retrospective questions about how children were feeling in late compulsory school, when I observe most of the effects of early grades. Tables 9 and 10 show that, all in all, early grades did not significantly affect student welfare.[33] The only statistically significant effects are found for low-ability low-SES students, who are less likely to report that they do well in school before getting the grades, and also are less likely to report that they enjoyed late compulsory school (school years 7-9). While the first finding is not negative per se, since it shows that these students were more conscious of their school performance, the second one might be more concerning for policy-makers. However similar outcomes pertaining to school welfare show a 0 effect also for these students, so I am more inclined to consider the finding a spurious effect.

[33] As explained before I cannot detect small effects, but I can state that there appears to be no major effect.

Table 9: Effects on behavior in school year 6:
Summary of difference in differences estimates

Outcome:		**Low-ability**		**High-ability**	
	All Sample	**Low-SES**	**High-SES**	**Low-SES**	**High-SES**
I do well in school	-0.04	-0.07**	-0.08	-0.01	0.03
	(0.15)	(0.02)	(0.13)	(0.71)	(0.23)
	[0.73]	[0.61]	[0.69]	[0.84]	[0.89]
Parents think I do well in school	-0.02	-0.02	-0.06	-0.03	0.03
	(0.27)	(0.17)	(0.11)	(0.19)	(0.11)
	[0.89]	[0.84]	[0.86]	[0.94]	[0.96]
I do my best, even if boring	-0.02	-0.03	-0.03	-0.01	-0.01
	(0.30)	(0.29)	(0.16)	(0.77)	(0.88)
	[0.71]	[0.73]	[0.71]	[0.71]	[0.67]
I want to improve in school	-0.00	-0.01	-0.02	0.04	-0.03
	(0.91)	(0.77)	(0.53)	(0.36)	(0.51)
	[0.59]	[0.71]	[0.66]	[0.47]	[0.44]
I dislike answering questions	-0.02	-0.03	-0.01	0.00	-0.04*
	(0.34)	(0.21)	(0.79)	(0.87)	(0.07)
	[0.16]	[0.21]	[0.18]	[0.14]	[0.10]
I learn useless stuff at school	0.01	0.03	0.03	-0.00	-0.01
	(0.69)	(0.33)	(0.37)	(0.97)	(0.80)
	[0.38]	[0.39]	[0.38]	[0.40]	[0.34]
I get disappointed if I get bad scores	0.01	-0.02	0.02	0.04	0.03
	(0.65)	(0.58)	(0.33)	(0.34)	(0.26)
	[0.68]	[0.63]	[0.70]	[0.69]	[0.75]

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

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets. All specifications control for basic demographics, relative ability measures (standardized at the treatment-cohort level) and parental background.

Table 10: Effects on behavior in late compulsory school:
Summary of difference in differences estimates

Outcome:		**Low-ability**		**High-ability**	
	All Sample	**Low-SES**	**High-SES**	**Low-SES**	**High-SES**
I enjoyed grades 7-9	-0.01	-0.03*	-0.01	-0.01	0.01
	(0.41)	(0.07)	(0.65)	(0.69)	(0.67)
	[0.72]	[0.71]	[0.71]	[0.75]	[0.74]
I was worried in grades 7-9	0.01	-0.01	0.05	0.01	-0.00
	(0.79)	(0.72)	(0.65)	(0.75)	(0.98)
	[0.12]	[0.13]	[0.10]	[0.11]	[0.12]
I am happy with grades 7-9	0.00	-0.01	0.00	0.02	0.00
	(0.81)	(0.89)	(0.86)	(0.65)	(0.86)
	[0.75]	[0.68]	[0.71]	[0.82]	[0.85]
I got help at home in grades 7-9	-0.01	-0.01	-0.02	-0.01	0.02
	(0.66)	(0.52)	(0.30)	(0.62)	(0.33)
	[0.71]	[0.67]	[0.75]	[0.68]	[0.76]
I did my best even if boring	0.01	-0.00	0.05	0.02	-0.00
	(0.62)	(0.93)	(0.14)	(0.28)	(1.00)
	[0.47]	[0.50]	[0.47]	[0.47]	[0.45]
I did my best even if hard	-0.00	-0.01	0.04	-0.02	0.02
	(1.00)	(0.44)	(0.17)	(0.26)	(0.55)
	[0.71]	[0.71]	[0.70]	[0.71]	[0.74]
I learned useless stuff at school	0.03	0.04	0.03	0.05	0.01
	(0.19)	(0.27)	(0.19)	(0.17)	(0.84)
	[0.53]	[0.57]	[0.52]	[0.54]	[0.47]
I was stressed at school	0.01	0.01	0.00	0.00	0.02
	(0.65)	(0.76)	(0.86)	(0.93)	(0.39)
	[0.20]	[0.20]	[0.22]	[0.19]	[0.19]

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

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets. All specifications control for basic demographics, relative ability measures (standardized at the treatment-cohort level) and parental background.

7 Discussion

In the following I assess how my results compare to previous literature on the effects of grading information. Stinebrickner & Stinebrickner (2012) study dropout behavior in Berea college, an institution with free tuition and subsidized boarding catering to disadvantaged students. They find that dropout is strongly explained by students revising downward their priors on academic perforamnce. Similarly Zafar (2011) finds that Northwestern undergraduates revise downward their beliefs, and switch to easier majors, when they observe grades lower than predicted. In Zafar's paper the deviation between expected and realized academic performance is taken as an "information metric" that identifies new information about students' "own unobserved academic ability." In fact this information might reflect, as explicitly recognized in Stinebrickner's paper, college preparation rather than academic ability. In both studies it is not possible to determine whether the updates are on academic ability or the stock of knowledge accumulated. While both problems signal the need for better selection into college, from the policy perspective they have quite different implications.[34]

The results of my paper are in line with the learning mechanism outlined by Stinebrickner & Stinebrickner (2012) and Zafar (2011) at the college level. In particular, the responses I find along the ability distribution are consistent with students revising their priors about ability. In my setup grades were assigned when children were 13 years old, so there is less concern that the update is on previous preparation, rather than ability. On top of that I show that students with the same SES (which could proxy for early effort), but different ability levels, react differently to grade assignment. This is consistent with students learning about ability rather than previous preparation. My paper shows both theoretically and empirically that the reaction to grades differs by SES. In the above-mentioned two papers there is no variation in SES, as sampled students are either low-SES (Berea college) or high-SES (Northwestern undergraduate students).

My paper is also related to the grading standards literature, which stresses the role of ability in students' responses to grades. Becker & Rosen (1992) and Betts (1998) show theoretically that higher grading standards encourage high ability students to put more effort, while students below standard might be discouraged. Betts & Grogger (2003) empirically confirm the heterogeneous effects of increasing grading standard at the high school level, while Figlio & Lucas (2004) find that higher standards lead to positive results on test scores, with effects that depend on the ability of the student relative to the

[34] Failure in evaluating own ability calls for a revision of grading information. Failure in college preparation requires to revise curricula in earlier education tiers.

class. In my setup untreated students do not observe grades, but only test scores. Absent grades, low-SES students are likely to have lower grading standards than high-SES students (for instance because the difficulty of the tests follow class ability). Introducing grades should thus lead to positive effects for high-SES students and negative effects for low-SES students. My results do not confirm this, and are rather consistent with students learning about their ability from grades.

The grading reform I analyze has been previously studied by Sjögren (2010), who uses administrative data to study long-run effects (final education and income) of the overall reform using difference in differences. She finds evidence of a positive effect of early grading on educational attainment for girls, and a negative effect for high-SES students. Differences in educational attainment are found also before and after the reform took place, which casts some doubts on the robustness of the results. My paper focuses on the mechanisms through which grades affect education choice: the theoretical model I develop shows that average treatment effects could mask substantial heterogeneity in the response to grades, which suggests conditioning the analysis at least by ability. I do not find any negative effect of early grades for high-SES students, while I do not study effects by gender, as this would exacerbate the multiple hypothesis testing problem. Results appear to be more robust, as trends in educational attainment for treated and control municipalities appear to be parallel in the refutability tests. This is likely due to the different cohorts used: Sjögren's sample comprises twenty cohorts, so she needs to assume parallel trends over two decades. My sample uses cohorts who studied just before and after the final reform. I only need to assume that trends between the two sets of municipalities are the same within a 5-year window, which I show to be the case.

8 Conclusion

In this paper I investigate the effect of early grades on students' education choices and attainment. I exploit the staggered implementation of a curriculum reform, which postponed grade assignment in Swedish compulsory schools, to estimate both short- and long-run effects of early grading. To investigate mechanisms I compare empirical results to the predictions of a sequential choice learning model based on the setup.

In the model children are uncertain about academic ability, and their priors differ by socioeconomic status (SES). Grades are ability signals that allow children to re-optimize educational choices. The calibrated model shows that early grading results into choices closer to first best for all students: low-ability students reduce effort in compulsory school and are more likely to choose vocational high school. High-ability students increase effort in school and are more likely to choose academic education paths. Stronger responses are found for students who observe information consistent with their priors, so that effects differ by SES.

The empirical results of my analysis are in line with the theoretical predictions for effort choices in late compulsory school. When graded early on, low-ability low-SES students are more likely to get lower grades and switch to easier courses in compulsory school than high-SES students with similar levels of ability. High-ability students, especially if high-SES, are more likely to get higher grades in late compulsory school when graded early on. Contrary to what the model predicts, early grades do not affect high school track choices and educational attainment for high-SES students. I find that high school attainment increases by 6 pp for high-ability low-SES students, while college attainment decreases by 3 pp for low-ability low-SES students. What explains the differences between model predictions and empirical findings at the high school level? The data suggests that SES strongly influences high school choices in Sweden, independently of ability. This might attenuate the effects of early grades. None of the effects found on education carry over to the labor market. In particular I find no effects on lifetime income, measured at ages 33-40. This suggests that early grading information simply improved the match between early education choices and ability, and reduced over-investment in education. Finally I find no evidence of demotivating effects for low-SES students, one of the main concerns that motivated the grading reforms.

The key economic implication of my results is that students are uncertain about their ability in early stages of education, when I show that grades affect their choices. This contrasts with the workhorse models of education choice (Becker, 1994; Ben-Porath, 1967), that assume complete information and thus no ex-ante uncertainty in the returns to non-compulsory education. From the policy point of view, I establish that early grading leads to a better match between education and ability, but increases inequality in educational attainment and reduces effort in compulsory school for low-ability students. Whether early grading is a desirable policy depends thus on the objective function of the policy-maker.

With regard to future research directions, it is possible to expand the scope of the analysis by looking at further sources of heterogeneity, which relate to different mechanisms. First, as whole classes are sampled in my data, I can look at differential effects of grading depending on relative ability. If students judge ability against their immediate peers, average ability students in high-ability (low-ability) classes might react more positively (negatively) to early grades. Second, average ability students might get more information out of grades, as it is less likely that they get top or bottom scores in homework and tests.[35] Third, there could be different responses to grading information along the gender dimension, as boys are found to generally be more overconfident than girls in ability (Bertrand, 2011). Before exploring these additional sources of heterogeneity it is however important to correct standard errors for multiple hypothesis testing. This requires some additional work, as standard errors are already bootstrapped to deal with the small number of clusters.

Lastly, it is worthy to investigate theoretically alternative mechanisms through which grades affect education choices, including the ones outlined above. For instance what happens when ability signals reflect knowledge rather than ability? This can be the case if parents and students are not able to distinguish the ability component of the grade from previous effort choices. I argue this is a plausible mechanism when parents do not observe children's effort, but do observe the final grades.

[35] Relatedly, Stange (2012) reports that the students for whom college grades, and the option to dropout in college, have the highest value, are moderate ability students, who have the strongest uncertainty about finishing college.

Bibliography

[1] Abadie Alberto, 2005. "Semiparametric Difference-in-Differences Estimators," *Review of Economic Studies*, Oxford University Press, vol. 72(1), pages 1-19.

[2] Altonji Joseph G, 1993. "The Demand for and Return to Education When Education Outcomes Are Uncertain," *Journal of Labor Economics*, University of Chicago Press, vol. 11(1), pages 48-83, January.

[3] Altonji Joseph G. & Blom Erica & Meghir Costas, 2012. "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," *NBER Working Papers* 17985, National Bureau of Economic Research, Inc.

[4] Angrist Joshua D. & Pischke Jörn-Steffen. "Mostly harmless econometrics: An empiricist's companion," Princeton university press, 2008.

[5] Arcidiacono Peter, 2004. "Ability Sorting and the Returns to College Major," *Journal of Econometrics*, Elsevier, vol. 121(1-2), pages 343-375.

[6] Arellano Manuel, 1987 "Computing Robust Standard Errors for Within-Group Estimators," *Oxford Bulletin of Economics and Statistics* 49, 431–434.

[7] Azmat Ghazala & Iriberri Nagore, 2010. "The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment using High School Students," *Working Papers* 444, Barcelona Graduate School of Economics.

[8] Bandiera Oriana & Larcinese Valentino & Rasul Imran, 2008. "Blissful ignorance? Evidence from a natural experiment on the effect of individual feedback on performance," mimeo.

[9] Basit Zafar, 2011. "How Do College Students Form Expectations?," *Journal of Labor Economics*, University of Chicago Press, vol. 29(2), pages 301 - 348.

[10] Becker Gary S., 1994. "Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education," 3rd Edition, *NBER Books*, National Bureau of Economic Research, Inc, number beck94-1, August.

[11] Ben-Porath Yoram, 1967. "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, University of Chicago Press, vol. 75, pages 352.

[12] Betts Julian R, 1998. "The Impact of Educational Standards on the Level and Distribution of Earnings," *American Economic Review*, American Economic Association, vol. 88(1), pages 266-75, March.

[13] Betts Julian R. & Grogger Jeff, 2003. "The impact of grading standards on student achievement, educational attainment, and entry-level earnings," *Economics of Education Review*, Elsevier, vol. 22(4), pages 343-352, August.

[14] Becker W. & Rosen S., 1990. "The Learning Effect Of Assessment And Evaluation In High School," University of Chicago - Economics Research Center 90-7, Chicago - Economics Research Center.

[15] Bertrand Marianne, 2011. "New Perspectives on Gender," *Handbook of Labor Economics*, Elsevier.

[16] Bertrand Marianne & Duflo Esther & Mullainathan Sendhil, 2004. "How Much Should We Trust Differences-in-Differences Estimates?," *The Quarterly Journal of Economics*, MIT Press, vol. 119(1), pages 249-275, February.

[17] Blundell Richard & Costa Dias Monica, 2009. "Alternative Approaches to Evaluation in Empirical Microeconomics," *Journal of Human Resources*, University of Wisconsin Press, vol. 44(3).

[18] Börklund Anders, 1993. "A Comparison between Actual Distributions of Annual and Lifetime Income: Sweden 1951-89," *Review of Income and Wealth*, International Association for Research in Income and Wealth, vol. 39(4), pages 377-86, December.

[19] Björklund Anders & Clark Melissa A. & Edin Per-Anders & Fredriksson Peter & Krueger Alan B., 2005. "The market comes to education in Sweden: an evaluation of Sweden's surprising school reforms," Russell Sage Foundation Publications.

[20] Borghans Lex & Lee Duckworth Angela & Heckman James J. & Ter Weel Bas, 2008. "The Economics and Psychology of Personality Traits," *Journal of Human Resources*, University of Wisconsin Press, vol. 43(4).

[21] Borghans Lex & Meijers Huub & Weel Bas Ter, 2008. "The Role Of Noncognitive Skills In Explaining Cognitive Test Scores," *Economic Inquiry*, Western Economic Association International, vol. 46(1), pages 2-12, 01.

[22] Brunello Giorgio & Checchi Daniele, 2007. "Does school tracking affect equality of opportunity? New international evidence," *Economic Policy*, CEPR;CES;MSH, vol. 22, pages 781-861, October.

[23] Colin Cameron & Jonah B. Gelbach & Douglas L. Miller, 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, MIT Press, vol. 90(3), pages 414-427, August.

[24] Cunha Flavio & Heckman James, 2007. "The Technology of Skill Formation," *American Economic Review*, American Economic Association, vol. 97(2), pages 31-47, May.

[25] Cunha Flavio & Heckman James, 2009. "The Economics and Psychology of Inequality and Human Development," *Journal of the European Economic Association*, MIT Press, vol. 7(2-3), pages 320-364, 04-05.

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

[27] Figlio David N. & Lucas Maurice E., 2004. "Do high grading standards affect student performance?," *Journal of Public Economics*, Elsevier, vol. 88(9-10), pages 1815-1834, August.

[28] Härnqvist Kjell, 1994a. "Den sociala selektionen till gymnasieskola och högskola," Rapport 1994:10, Göteborgs universitet, Institutionen för pedagogik.

[29] Härnqvist Kjell, 1998. "A longitudinal program for studying education and career development," Report 1998: 01, Göteborgs universitet, Institutionen för pedagogik (1998).

[30] Hoxby Caroline, 2011. "Program Report," NBER Economics of Education Program, Issue 1, NBER Reporter.

[31] Jalava Nina & Joensen Schrøter Juanna & Pellas Elin, 2015. "Grades and rank: Impacts of non-financial incentives on test performance," *Journal of Economic Behavior & Organization*, Elsevier, vol. 115(C), pages 161-196.

[32] Hussey J. Andrew & Swinton H. Omari, 2011. "Estimating the Ex Ante Expected Returns to College," *American Economic Review*, American Economic Association, vol. 101(3), pages 598-602, May.

[33] Klapp Alli & Cliffordson Christina & Gustafsson Jan-Eric. "The effect of being graded on later achievement: Evidence from 13-year olds in Swedish compulsory school," *Educational Psychology* ahead-of-print (2014): 1-19.

[34] Klapp Alli. "Does grading affect educational attainment? A longitudinal study," *Assessment in Education: Principles, Policy & Practice* ahead-of-print (2015): 1-22.

[35] Kolenikov Stanislav & Angeles Gustavo, 2005. "On reuse of clusters in repeated studies," presented at the American Statistical Association, Minneapolis, MN.

[36] Lechner Michael, 2011. "The Estimation of Causal Effects by Difference-in-Difference Methods," Foundations and Trends in Econometrics, now publishers, vol. 4(3), pages 165-224, November.

[37] Lindqvist, Erik & Roine Vestman, "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment," *American Economic Journal: Applied Economics* 3.1 (2011): 101-28.

[38] MacKinnon James & G. & Matthew D. Webb, 2014. "Wild Bootstrap Inference for Wildly Different Cluster Sizes," *Working Papers* 1314, Queen's University, Department of Economics.

[39] Meghir Costas & Palme Mårten, 2005. "Educational Reform, Ability, and Family Background," *American Economic Review*, American Economic Association, vol. 95(1), pages 414-424, March.

[40] Mora Ricardo & Reggio Iliana, 2012. "Treatment effect identification using alternative parallel assumptions," *Economics Working Papers* we1233, Universidad Carlos III, Departamento de Economía.

[41] Öckert, Björn, 2002. "Do university enrollment constraints affect education and earnings?," *Working Paper Series* 2002:16, IFAU - Institute for Evaluation of Labour Market and Education Policy.

[42] Öckert, Björn, 2010. "What's the value of an acceptance letter? Using admissions data to estimate the return to college," *Economics of Education Review*, Elsevier, vol. 29(4), pages 504-516, August.

[43] OECD (2010), "How many students drop out of tertiary education?," in Highlights from Education at a Glance 2010, OECD Publishing.

[44] Régner Isabelle & Pascal Huguet & Jean-Marc Monteil. "Effects of socioeconomic status (SES) information on cognitive ability inferences: When low-SES students make use of a self-threatening stereotype," *Social Psychology of Education* 5.3 (2002): 253-269.

[45] Sjögren Anna, 2010. "Graded children – evidence of longrun consequences of school grades from a nationwide reform," *Working Paper Series* 2010:7, IFAU - Institute for Evaluation of Labour Market and Education Policy.

[46] Spence Michael, 1973. "Job market signaling," *The Quarterly Journal of Economics*, 355-374.

[47] Todd Stinebrickner & Ralph Stinebrickner, 2012. "Learning about Academic Ability and the College Dropout Decision," *Journal of Labor Economics*, University of Chicago Press, vol. 30(4), pages 707 - 748.

[48] Stange Kevin, 2012. "An Empirical Investigation of the Option Value of College Enrollment," *American Economic Journal: Applied Economics*, American Economic Association, vol. 4(1), pages 49-84, January.

[49] Svensson Allan, 1995. "Att välja eller välja bort naturvetenskap och teknik," NOT-projektet, 3.

[50] Todd Petra & Wolpin Kenneth, 2003. "On the specification and estimation of the production function for cognitive achievement," *The Economic Journal*, 113 (485), F3–F33.

[51] White H., 1984. "Asymptotic Theory for Econometricians," San Diego: Academic Press.

A Numerical Model

A.1 Evidence on assumptions and calibration

The model makes precise assumptions about choice protocols, the distribution of ability in the population, school selection, and payoffs to education. In this section I provide evidence supporting model assumptions, and discuss calibration.

The basic assumption underlying the model is that students are forward looking in the education choices. Table A.1 reports summary statistics on the items that surveyed students considered important when choosing high school. Apart from preferences for the chosen program, the items that rank highest are study plans, ability and grades. This shows that students were forward-looking in their choices, and considered feasibility of the chosen track important in their choices.

Table A.1: Survey evidence on HS choice, 1967 cohort

	Mean	**Obs**
Chose HS after interest	0.80	6195
Chose HS after study plans	0.64	6099
Chose HS after ability	0.60	6093
Chose HS after grades	0.49	6117
Chose HS after parents	0.23	6099
Chose HS after peers	0.07	6098

Data from grade 10 survey. All variables represent agreement with the statement and are coded from 0 to 1 (1 represents full agreement).

In the model high SES students are assumed to have higher levels of ability than low SES students. Figure A.1 confirms this empirically.[36] While low-SES students have normal ability distributions, high-SES students display right-skewed distributions. In Table A.2 I calibrate the data to the discrete distribution in column 1. The resulting distributions by SES are then used to simulate ability distributions in the model.

[36]This is consistent with the evidence on early differences in ability through the socioeconomic gradient shown by Cunha & Heckman (2009).

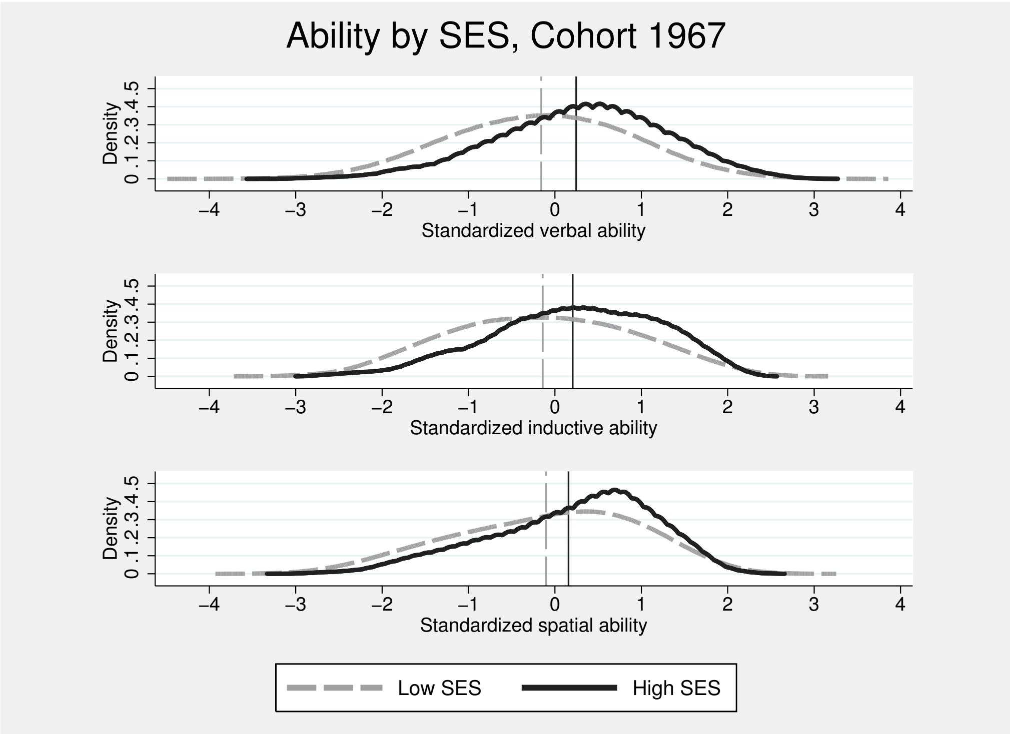


Figure A.1: Differences in standardized ability by SES

Note: The SES division is based on parental education. Ability measures are taken from tests administered in school year 6, and are standardized at the treatment-cohort level.

Table A.2: Distribution of discretized ability by SES, 1967 cohort

	All	Low SES	High SES
Lowest ability	0.10	0.13	0.06
Low ability	0.20	0.24	0.14
Medium ability	0.30	0.31	0.28
High ability	0.25	0.22	0.29
Highest ability	0.15	0.10	0.23

The population ability distribution is constrained to the bins in column 1. The distributions by SES are generated using the same cut-points.

Figure A.2 shows ability levels by completed education. Students who attained high school have higher ability than those who dropped out of high school school, or never enrolled. Students who have a college education have much higher ability levels, which is consistent with the assumptions I make in the model, and is not surprising given that college is highly selective.

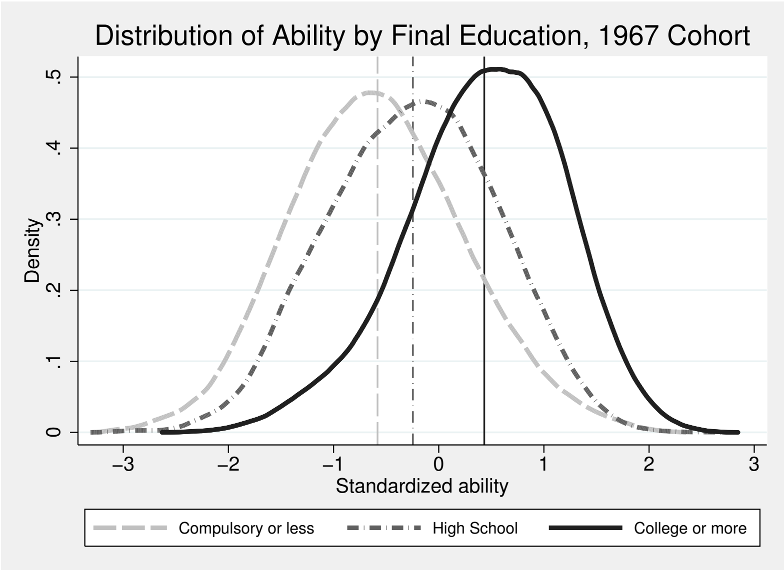


Figure A.2: Standardized ability by final attained education.

Note: Ability measures are taken from tests administered in school year 6, and are standardized at the treatment-cohort level.

Table A.3 summarizes income premia for each education choice. While these are not causal estimates, they might be representative of the information that young students use when assessing their education goals. High school graduates exhibit higher incomes than students with compulsory education. As assumed in the model, the income of students with academic high school are not substantially different from those of students with vocational high school. The wages of college graduates[37] are instead quite higher. In Figure A.3 I plot the wages of the students in the sample by discretized ability. There is little variation in wages by ability for students with compulsory school or high school. However there seems to be complementarity between income and ability for college graduates. In the model I thus allow the wage premium for college to depend on ability, and use the estimates as payoffs.

[37] Here including also 2-year short college

Table A.3: Income by Final Education,
1967 Cohort

Completed education	Gross Income	Premium
Compulsory school	184.40	0.00
Vocational HS	221.21	0.20
Academic HS	226.22	0.23
College	290.67	0.58

Before-tax income measured at ages 33-40, in thousands kronor

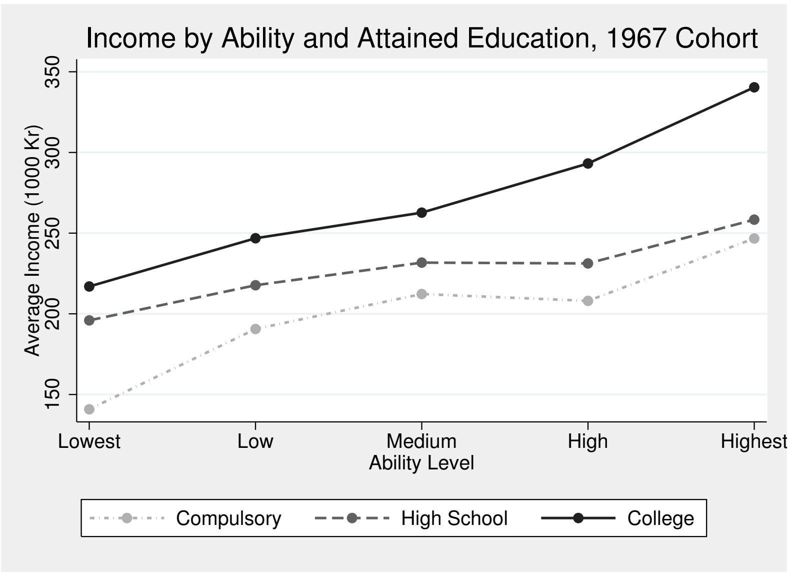


Figure A.3: Average gross income by ability and attained education, 1967 cohort

Note: Before-tax income measured at ages 33-40, in thousands kronor. Ability is discretized to a 1-5 scale, as in the model.

I calibrate the knowledge production function, $k_{it} = \omega_t(\alpha a_i + \beta e_{it}) + \delta k_{it-1}$, using the following parameters:

Table A.4: Production Function Parameters

weights	coefficients
$\omega_1 = 4/9$	$\alpha = 1$
$\omega_2 = 1/9$	$\beta = 2.5$
$\omega_2 = 5/9$	$\gamma = 1.1$

Table A.5 reports minimum effort and ability levels required to access and attain each education level. The knowledge thresholds are found substituting the values for each education level into equation 12:

$$k_3 = \alpha a \times (\omega_1\gamma^2 + \omega_2\gamma + \omega_3) + \beta(\omega_1\gamma^2 \times e_1 + \omega_2\gamma \times e_2 + \omega_3 \times e_3). \tag{12}$$

Table A.5: Minimum ability and effort for Educational Attainment (Knowledge Thresholds)

	$\bar{k}^{E_2}$	$\underline{k}^{E_3}$	$\bar{k}^{E_3}$	$\bar{k}^{E_4}$
a	2	3	3	4
e_1	Medium	Medium	Medium	Medium
e_2	Medium	Medium	Medium	Medium
e_3	Low	Low	Medium	High

The value of vocational school, $V_{i,t}^{E_2}$, was shown in Section 3.1 to be:

$$\begin{aligned} V_{i,\tau}^{E_2} = \textstyle\sum_{t=\tau}^{3} -\omega_\tau \times \underline{\gamma}_E (e_{it,\tau}^{E_2\star})^{\bar{\gamma}_E} + P(\tilde{k}_{i3,\tau} \geq \bar{k}^{E_2}) \times U((L-2) \times w_2) \\ +P(\tilde{k}_{i3,\tau} < \bar{k}^{E_2}) U((L-1) \times w_1)). \end{aligned} \tag{13}$$

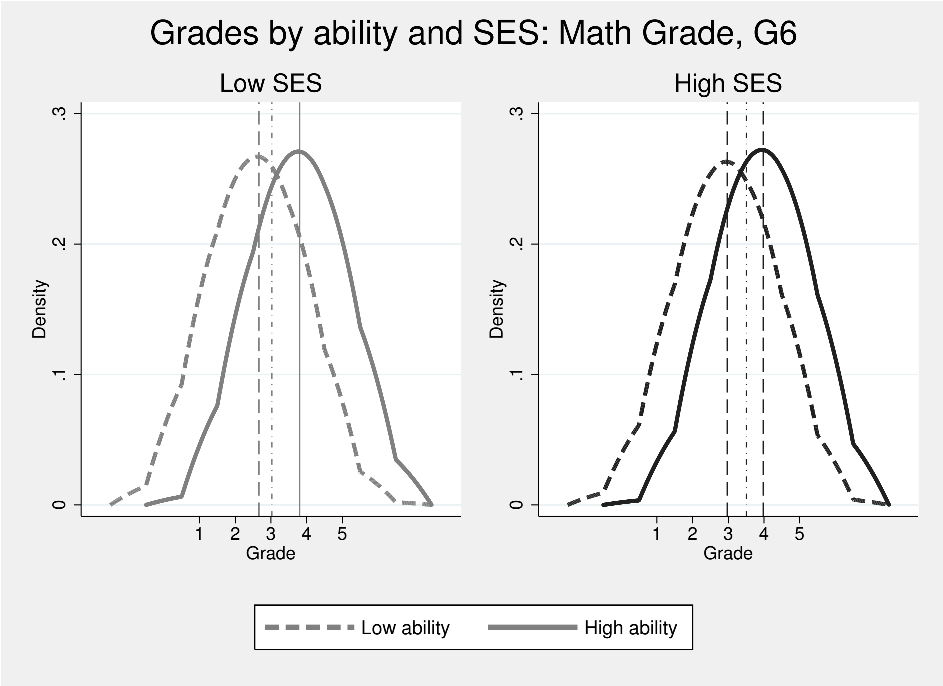
The parameters I use for effort disutility are: $\underline{\gamma}_E = 6$ and $\bar{\gamma}^E = 1.6$. The parameters for income utility, $U(I) = \underline{\gamma}_I (I_i^{E_2\star})^{\bar{\gamma}_I}$, are: $\underline{\gamma}_I = 2.8$ and $\bar{\gamma}^I = 0.9$. Effort costs are thus convex, and income utility is concave.

In the model high-ability students are more likely to observe higher grades than low-ability students do when graded early on. I confirm this in Figure A.4, where I plot grades at the end of school year 6 (the treatment) for treated students born 1967. The vertical black line represents the average grade for each SES: it could be considered the prior grade the student is expected to get, before information about ability is revealed. For low-SES students the average grade is closer to the mean for low-ability students. The opposite is true for high-SES students. This reflects the different composition in ability within SES.

Table A.6: Factors Affecting HS admission, Cohort 1967

	Admitted in HS at first Choice
Adv Math (s.y. 9)	0.05***
	(0.01)
Adv English (s.y. 9)	0.00
	(0.02)
Math Grade (s.y. 9)	0.04***
	(0.01)
English Grade (s.y. 9)	0.02***
	(0.01)
Swedish Grade (s.y. 9)	0.03***
	(0.01)
Mean	0.83
R^2	0.05
Observations	7884

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The data does not record which type of school the student was applying for. Standard errors clustered at the class level.



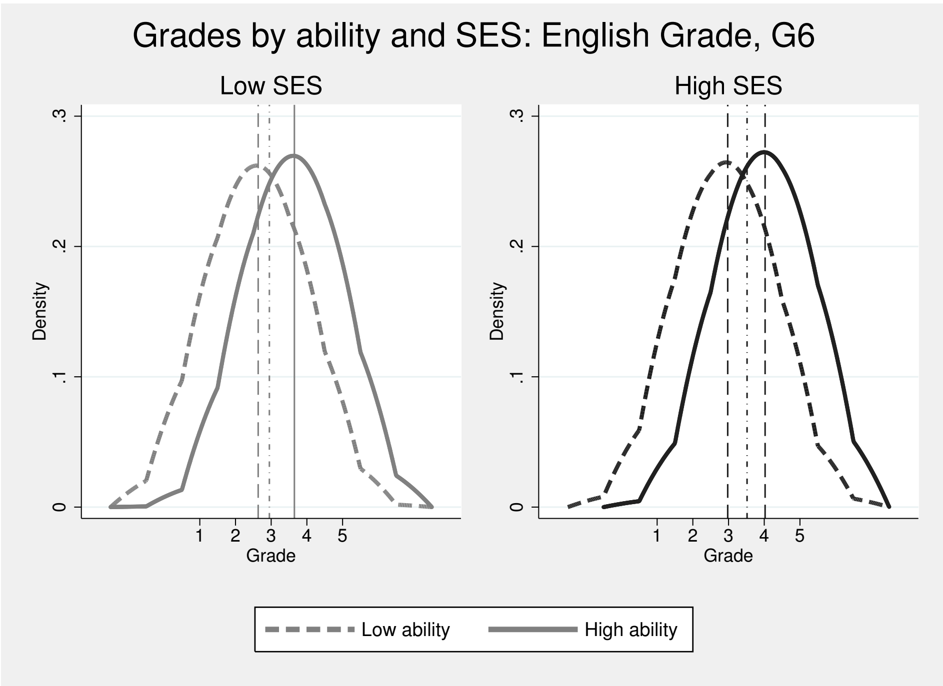


Figure A.4: Grades by Ability and SES in school year 6

Note: Distributions over-smoothed for illustrational clarity. Dash-dot vertical lines represent averages for each SES cell. The other vertical lines represent averages for each ability−SES cell.

Finally In the model I assume that grades are unbiased. In Table A.7 I try to assess this empirically. While there is a strong relationship between the standardized test and the final grade (the coefficient is close to 1), it appears that SES has an independent positive effect on final grades, controlling for ability. This could be due to a positive bias towards wealthier students, but could also be related to the fact that high-SES students put more effort into schooling. Given that final grades corrected for discrepancies between yearly and test performance, it is still possible that they are unbiased. Notice that the magnitude of this higher bound effect is actually small: one child over ten/twenty gets a higher grade if categorized as high-SES with respect to a low-SES child.

Table A.7: Testing for biases in final grade assignment

	Adv Math grade (year 9)	Adv English grade (year 8)
Regressor of interest:		
High-SES	0.09***	0.04***
	(0.02)	(0.01)
Controls:		
Normalized test score	0.69***	0.63***
	(0.01)	(0.01)
Normalized ability	0.06***	0.11***
	(0.02)	(0.01)
1967 cohort	–0.03	0.01
	(0.02)	(0.02)
R^2	0.60	0.60
Observations	6535	8867

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Normalized test score and Ability refer to math tests in column 1, and English tests in column 2. The SES division is based on parental education. Standard errors are clustered at the class level.

A.2 Solution Method

Simulation of ability signals

I extract true ability and unbiased ability signals (See Figure A.5) from a multivariate normal with covariance matrix:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma_1^2(\epsilon) & 0 \\ 0 & 0 & \sigma_2^2(\epsilon) \end{bmatrix}$$

I discretize the normal draws using the SES-specific distributions shown in Table A.2. I assume the following:

- The grade signals are a sum of true ability and noise:
  - $g_{i2} = a_i + \epsilon_2$ with $\epsilon_2 \sim N\left(0, \sigma_2^2(\epsilon)\right)$
  - $g_{i3} = a_i + \epsilon_3$ with $\epsilon_3 \sim N\left(0, \sigma_3^2(\epsilon)\right)$
  - $cov(a_i, \epsilon_2) = cov(a_i, \epsilon_3) = cov(\epsilon_2, \epsilon_3) = 0$
- Late grades are more precise than grades assigned in school year 6:
  - $corr(a_i, g_{i2}) = 0.7$
  - $corr(a_i, g_{i3}) = 0.8$

I need to find $\sigma_1^2(\epsilon)$ and $\sigma_2^2(\epsilon)$ such that $k_2 = corr(a_i, g_{i2}) = 0.7$ and $k_3 = corr(a_i, g_{i3}) = 0.8$: $k_t = corr(a_i, g_{it}) = corr(a_i, a_i + \epsilon_t) = \dfrac{1+0}{\sigma(a_i) + \sigma(a_i + \epsilon_t)} = \dfrac{1}{\sigma(a_i) \times \sigma(a_i + \epsilon_j)} = \dfrac{1}{\sigma(a_i + \epsilon_j)}$ and $\sigma^2(a_i + \epsilon_t) = 1 + \sigma_t^2(\epsilon)$. So $k_t = \dfrac{1}{\sqrt{1 + \sigma_t^2(\epsilon)}}$, thus $\sigma_t^2(\epsilon) = \dfrac{1}{k_t^2} - 1$. Because $\sigma^2(a_i + \epsilon_t) = \dfrac{1}{k_t^2}$ it follows that $corr(g_{i2}, g_{i3}) = \dfrac{cov(a_i + \epsilon_2, a_i + \epsilon_3)}{\sigma(a_i + \epsilon_2) \times \sigma(a_i + \epsilon_3)} = \dfrac{1}{\sqrt{\dfrac{1}{k_2^2}} \times \sqrt{\dfrac{1}{k_3^2}}} = k_2 \times k_3$.

I simulate the joint ability and grade distributions 1000 times to get three sets of posterior distributions:

- $f(a_i | g_{i2}, SES)$, plotted in Figure A.6
- $f(a_i | g_{i3}, SES)$, plotted in Figure A.7
- $f(a_i | g_{i3}, g_{i2}, SES)$, plotted in Figures A.8 and A.9

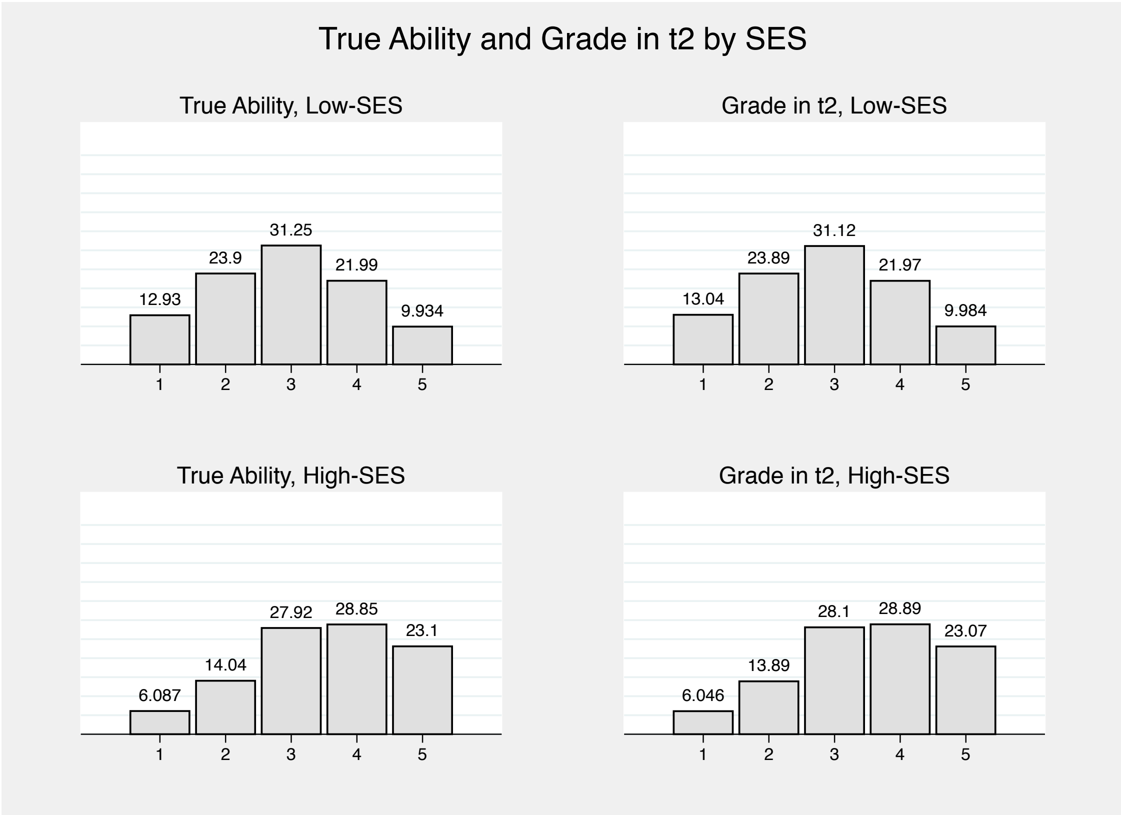


Figure A.5: Ability and grade signals in t_2

Note: The Figure plots simulated distributions of ability and grades in middle compulsory school, for low- and high-SES students.

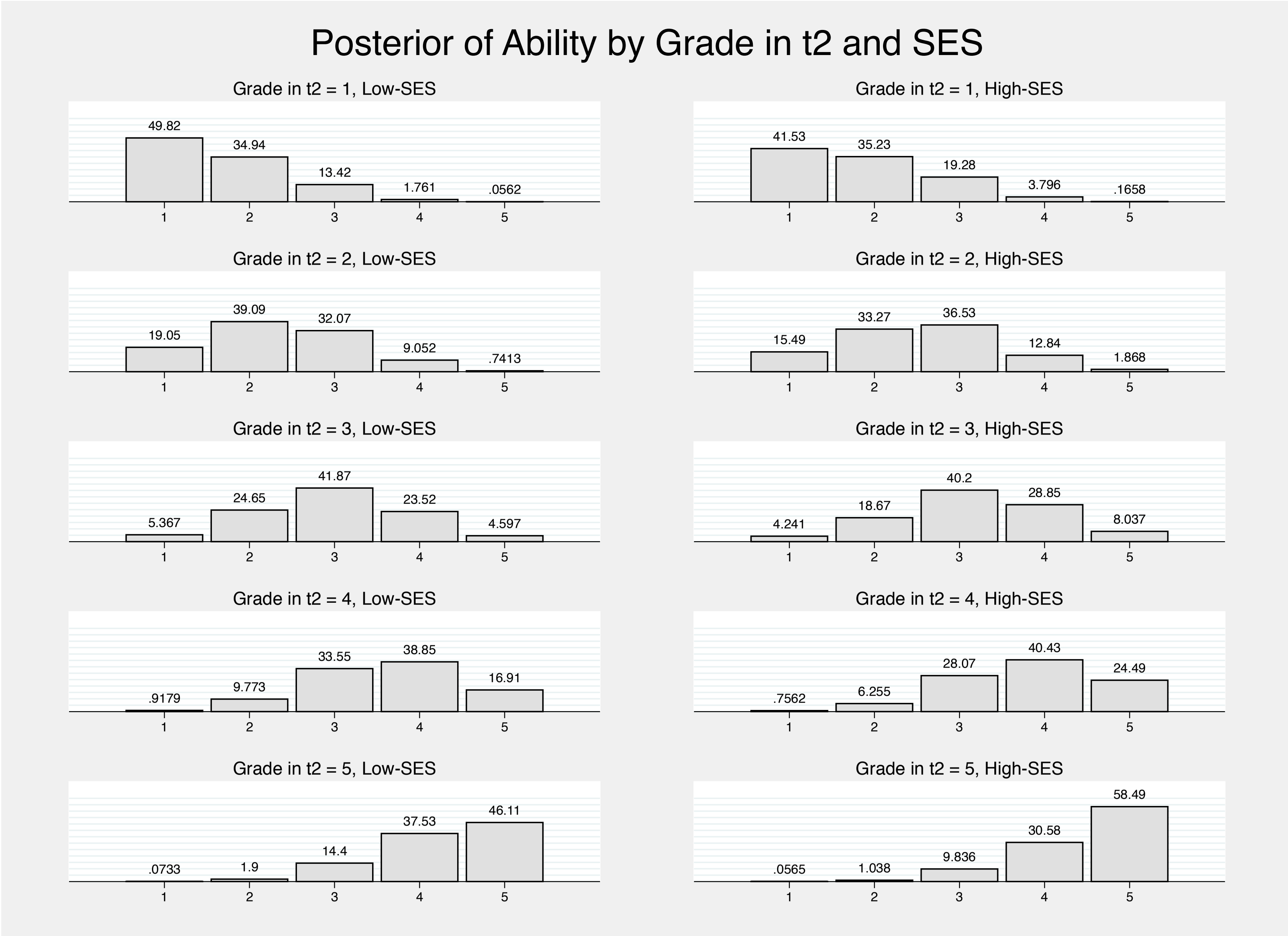


Figure A.6: Posterior distributions by grade signal in t_2

Note: The Figure plots the updated distributions of ability after students observe specific grades in middle compulsory school. Updates differ by SES due to the different priors about ability.

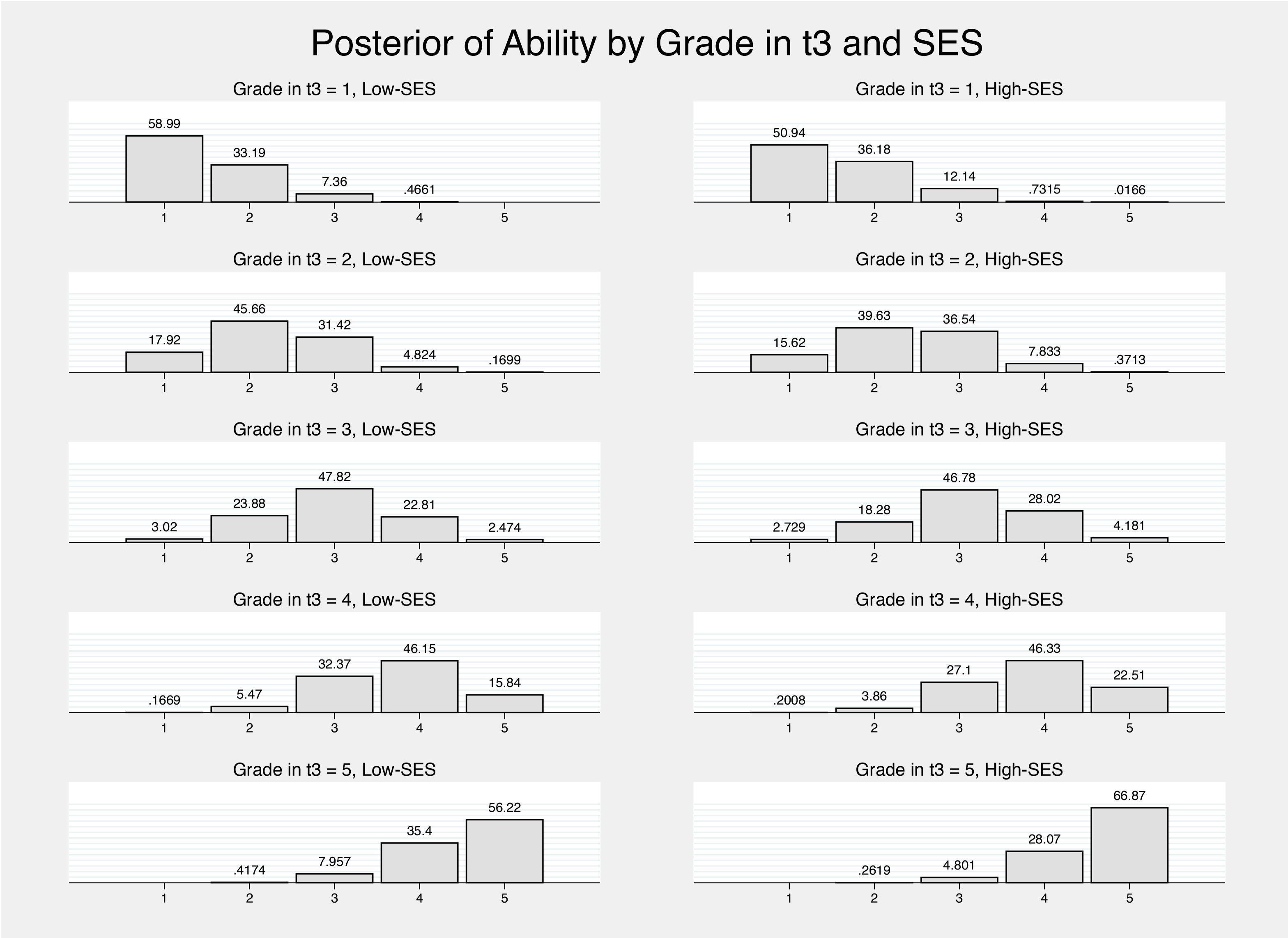


Figure A.7: Posterior distributions by grade signal in t_3

Note: The Figure plots the updated distributions of ability after students observe specific grades in late compulsory school, when only late grades are assigned. Updates differ by SES due to the different priors about ability.

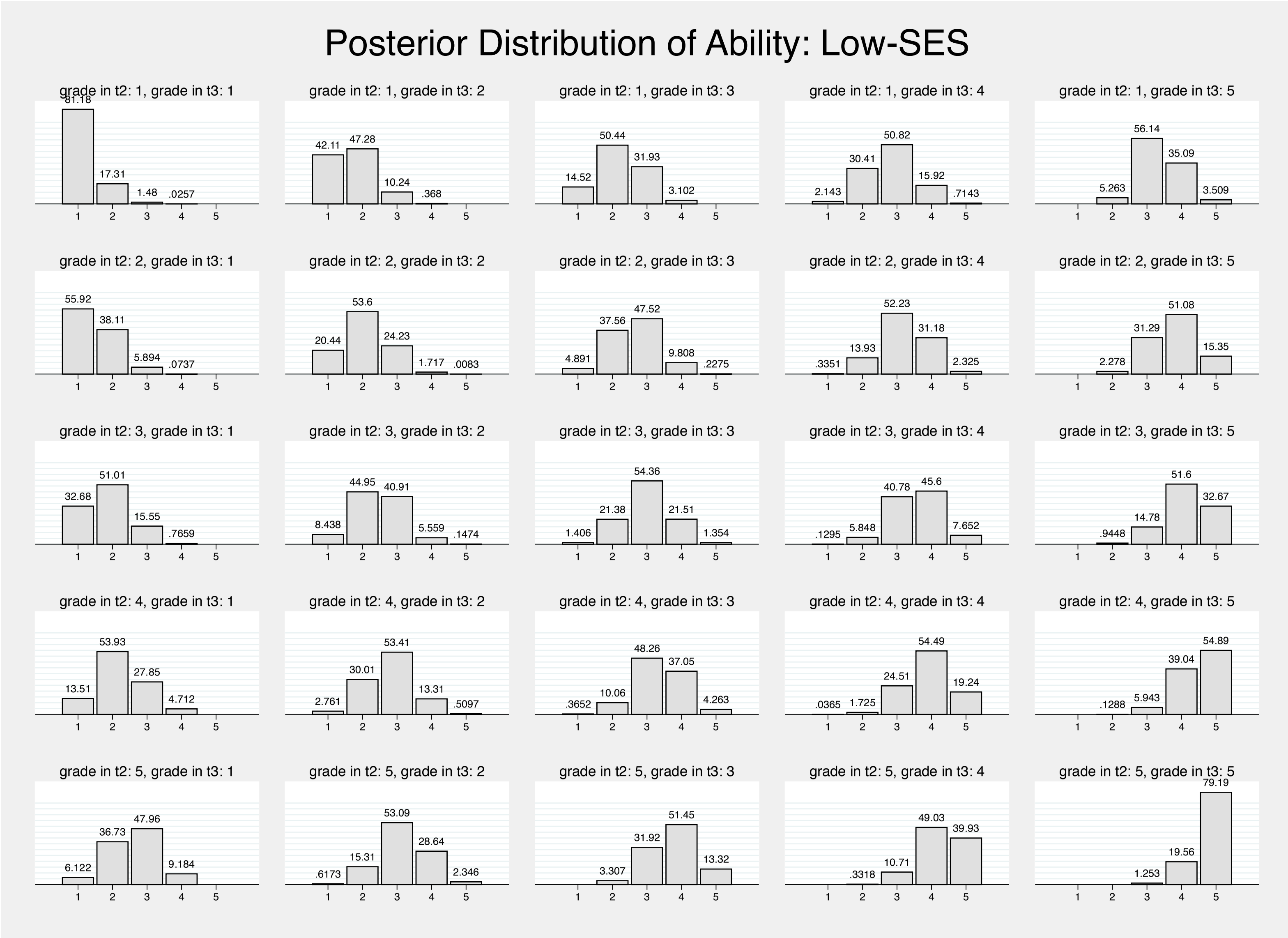


Figure A.8: Posterior distributions after grades in t_2 and t_3 are assigned: Low-SES students

Note: The Figure plots the updated distributions of ability after low-SES students observe specific grades in middle and late compulsory school. Updates differ by SES due to the different priors about ability.

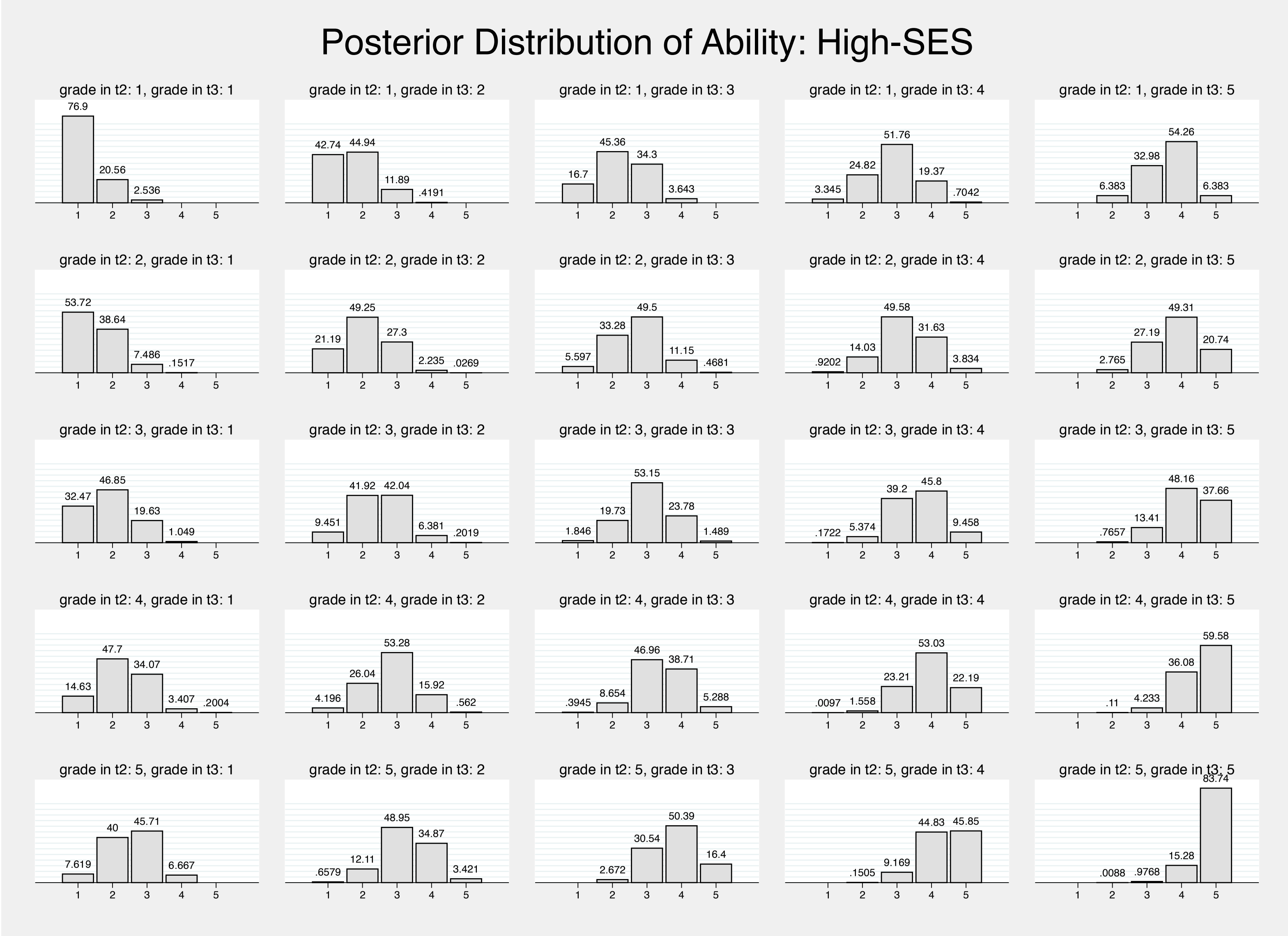


Figure A.9: Posterior distributions after grades in t_2 and t_3 are assigned: High-SES students

Note: The Figure plots the updated distributions of ability after high-SES students observe specific grades in middle and late compulsory school. Updates differ by SES due to the different priors about ability.

Solution strategy

I solve by backward induction the optimization problem in three different information scenarios:

1. In the case of full information about ability (first best). The solution is found for 5 ability levels. SES has no role in individual choice, but aggregate outcomes will differ due to the different distribution of ability by SES.

2. When only late grades are released. The solution is found for 2 (SES) x 5 (g_{i3}) = 10 cases.

3. When early grades are released. The solution is found for 2 (SES) x 5 (g_{i2}) x 5 (g_{i3}) = 50 cases

Solution when early grades are assigned

At the end of t_3 choose optimally $E^\star$, given any $[SES, e_{i1}, g_{i2}, e_{i2}, g_{i3}, e_{i3}]$ vector. There are (2 x 3 x 5 x 3 x 5 x 3) x 3 = 4050 cases. 1350 solutions are optimal, given $f(a_i|g_{i3}, g_{i2}, SES)$. In the same stage choose optimally e_{i3}, given any $[SES, e_{i1}, g_{i2}, e_{i2}, g_{i3}]$ vector. There are (2 x 3 x 5 x 3 x 5) x 3 = 1350 cases. 450 solutions are optimal given $f(a_i|g_{i3}, g_{i2}, SES)$.

In t_2 choose optimally e_{i2}, given any realized $[SES, e_{i1}, g_{i2}]$ vector. Use $f(a_i|g_{i2}, SES)$ to assign the proper weight to each of the 5 potential grades that can be assigned in t_3. $Emax_2$ thus summarizes 450 cases into 450/5=90 cases, before g_{i2} is assigned. There are indeed (2 x 3 x 5) x 3 = 90 cases. 30 solutions are optimal, given $f(a_i|g_{i2}, SES)$.

In t_1 choose optimally e_1, given $[SES]$. Use $f_1(a_i)$ to assign the proper weight to each of the 5 potential grades that might be assigned in t_2. Thus $Emax_1$ summarizes 30 cases into 30/5=6 cases, before grades are assigned. There are indeed (2) x 3 = 6 cases. 2 solutions are optimal, given $f_1(a_i)$, one for each SES. So in the end I find 2 x 25 $[e_1^\star, e_2^\star, e_3^\star, E^\star]$ contingent plans, for each SES and ability signal realized.

Solution when late grades are assigned

At the end of t_3 choose optimally $E^\star$, given any $[SES, e_{i1}, e_{i2}, g_{i3}, e_{i3}]$ vector. There are (2 x 3 x 3 x 5 x 3) x 3 = 810 cases. 270 cases are optimal, given $f(a_i|g_{i3}, SES)$. In t_3

choose optimally e_{i3}, given any $[SES, e_{i1}, e_{i2}, g_{i2}]$ vector. There are (2 x 3 x 3 x 5) x 3 = 270 cases. 90 solutions are optimal given $f(a_i|g_{i3}, SES)$.

In t_1 and t_2 choose optimally $[e_{i1,}, e_{i2}]$, given SES. Use $f_1(a_i)$ to assign the proper weight to each of the 5 potential grades (mirroring ability type) that can be assigned in t_3. Thus $Emax_1$ summarizes 90 cases into 90/5=18 cases, before grades are assigned. There are indeed (2 x 3) x 3 = 18 cases. 2 solutions are optimal given $f_1(a_i)$, one for each SES. So in the end I find 2 x 5 $[e_1^\star, e_2^\star, e_3^\star, E^\star]$ contingent plans, for each SES and ability signal realized.

Realizations

I append the datasets created in the simulation phase, and take a random sample. I merge the final dataset to first and second best solutions. The merge is on $[SES, g_{i2}, gi_3]$ for the solution with early grades, $[SES, g_{i3}]$ for the solution with late grades, and $[a_i]$ for the first best. I use true ability, the knowledge production function, and education thresholds, to determine final outcomes. This gives me a distribution of realized outcomes for each SES and ability level. At this point I can assess how the information structure affects final outcomes.

A.3 Additional Simulation Results

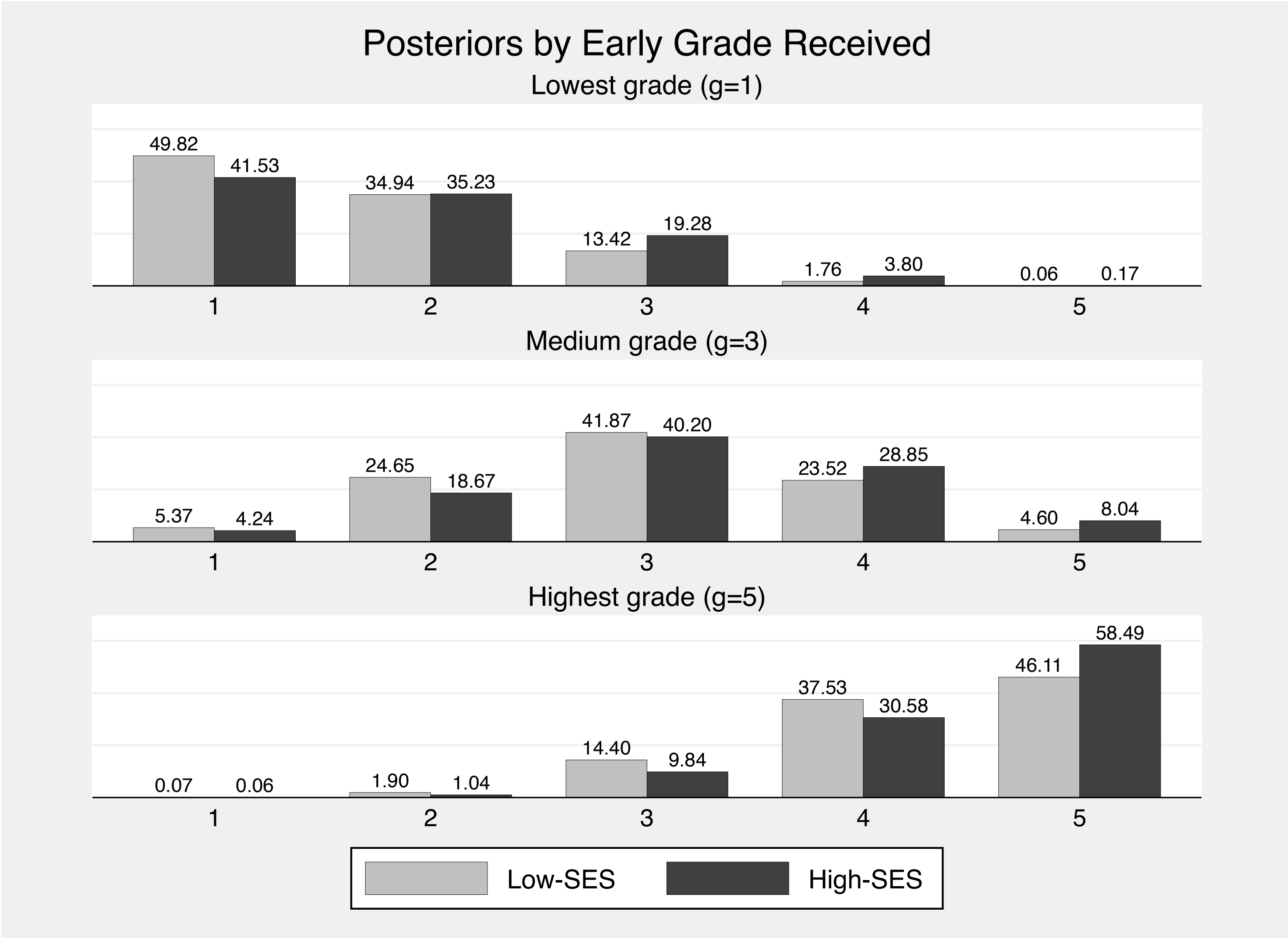


Figure A.10: Comparison of posterior distributions in t_2 after observing grades for low- and high-SES students

Note: The Figure compares posterior distributions for low- and high-SES students in t_2, after students observe respectively lowest, medium, and top grades. Lowest (highest) grades lead to stronger updates for low (high) SES students.

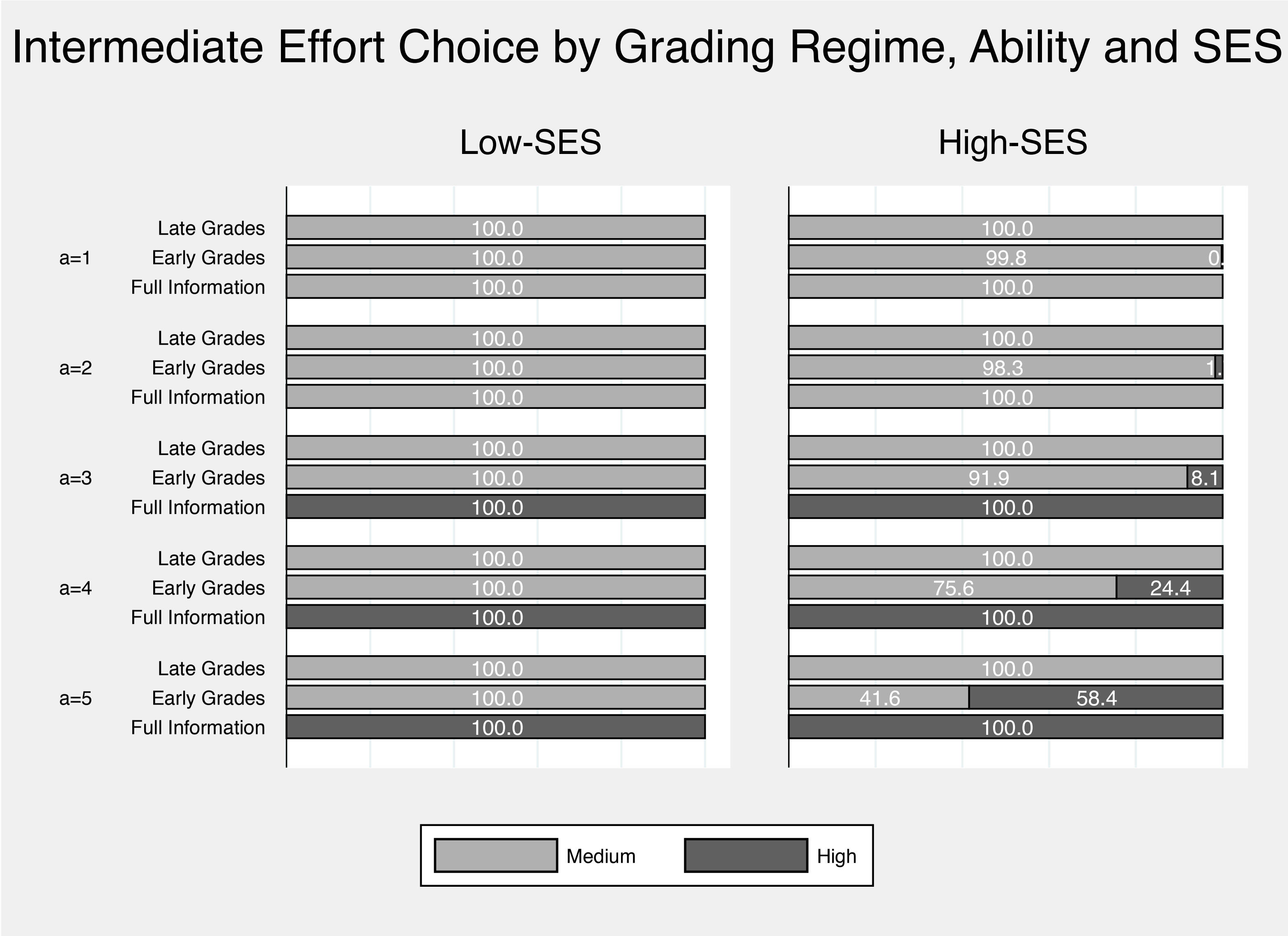


Figure A.11: Intermediate effort choice by ability and SES for different grading regimes

Note: The Figure plots effort distributions in middle compulsory school. Results are presented by ability level. For students with ability level above 3, it is optimal to follow academic paths. For those with lower ability, vocational high school is the optimal choice. SES affects students's priors about ability, and thus optimal choices.

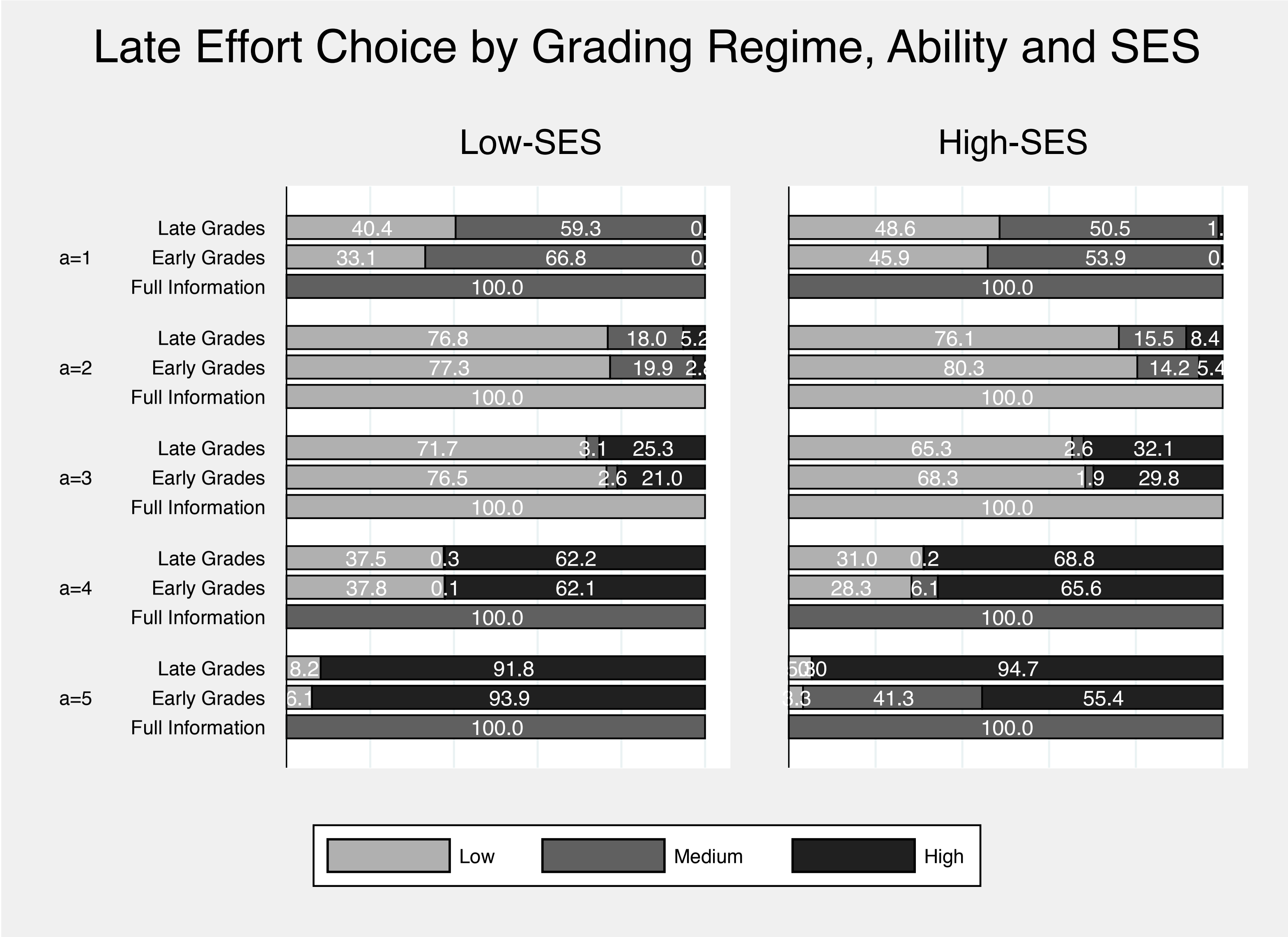


Figure A.12: Late effort choice by ability and SES for different grading regimes

Note: The Figure plots effort distributions in late compulsory school. Results are presented by ability level. For students with ability level above 3, it is optimal to follow academic paths. For those with lower ability, vocational high school is the optimal choice. SES affects students's priors about ability, and thus optimal choices.

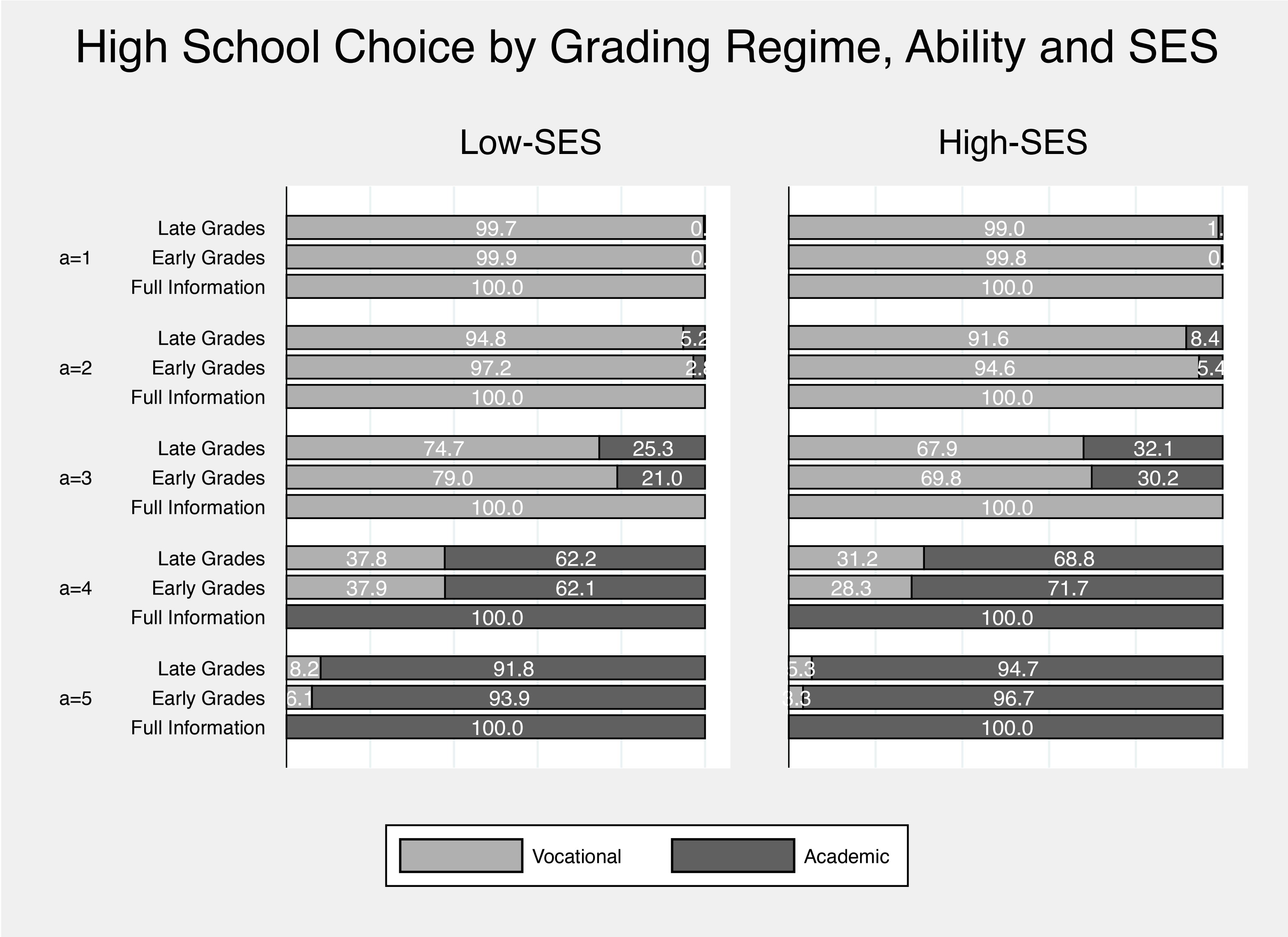


Figure A.13: High school choice by ability and SES for different grading regimes

Note: The Figure plots high school choice distributions. Results are presented by ability level. For students with ability level above 3, it is optimal to follow academic paths. For those with lower ability, vocational high school is the optimal choice. SES affects students's priors about ability, and thus optimal choices.

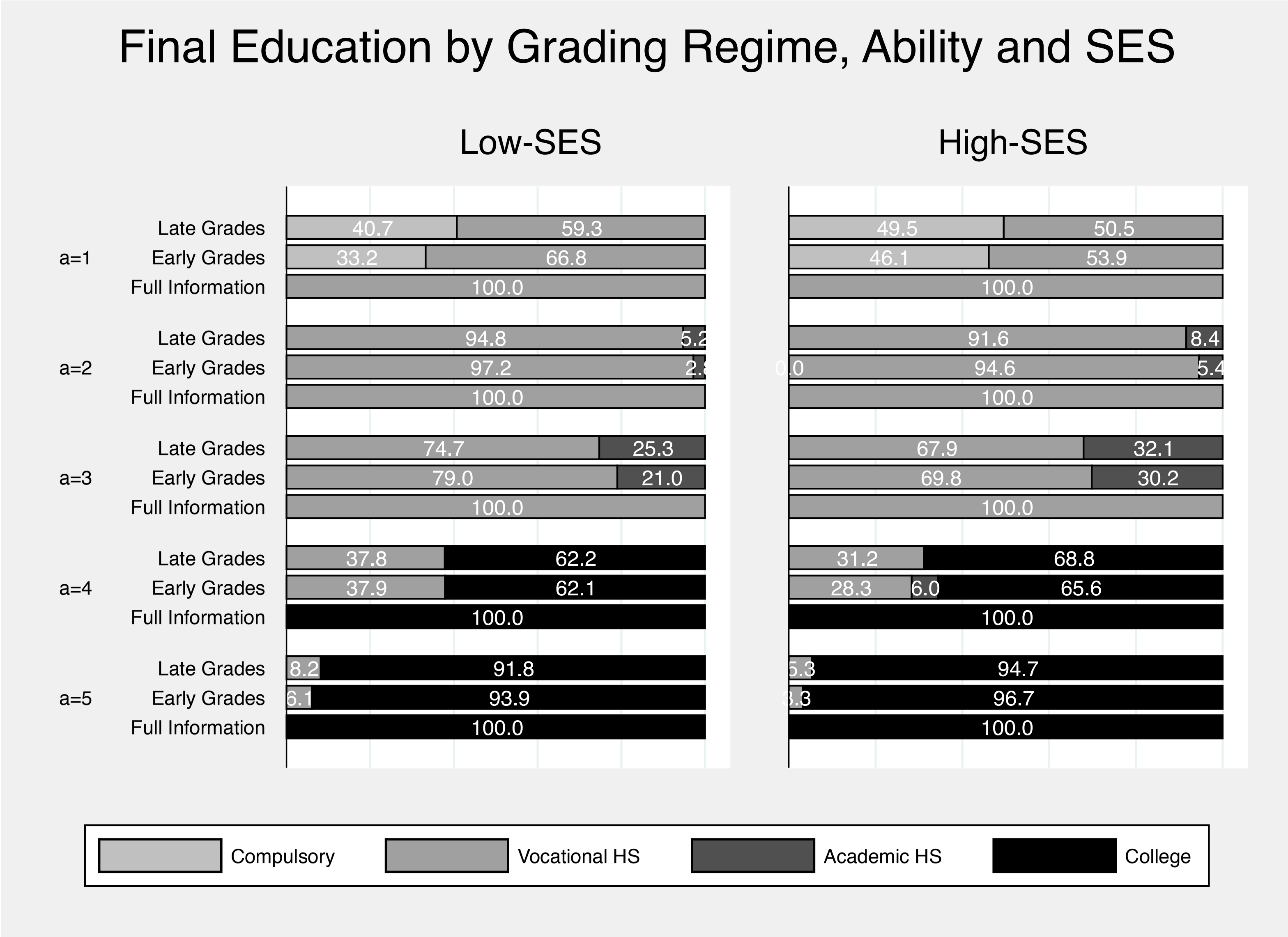


Figure A.14: Final education by ability and SES for different grading regimes

Note: The Figure plots final education distributions. Results are presented by ability level. For students with ability level above 3, it is optimal to follow academic paths. For those with lower ability, vocational high school is the optimal choice. SES affects students's priors about ability, and thus optimal choices.

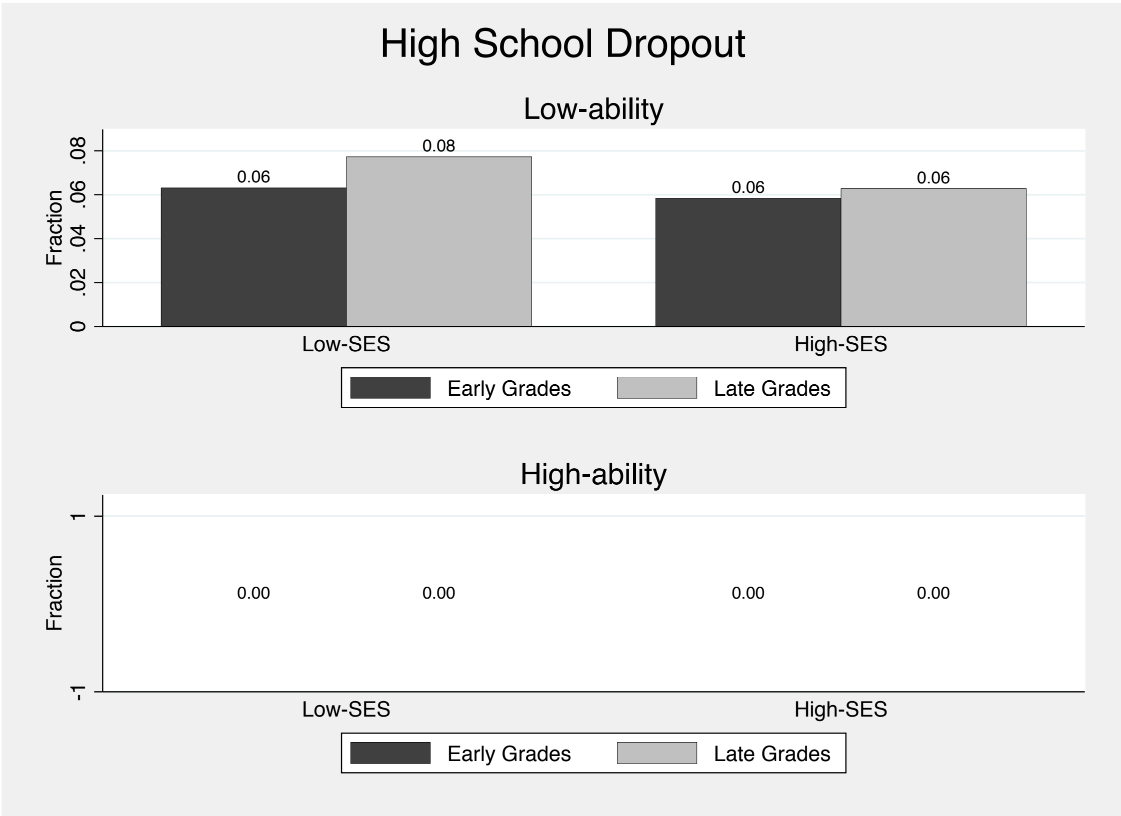


Figure A.15: High school dropout by grading regime

Note: The Figure plots high school dropout rates under early or late grade assignment. Notice that with full information dropout is never optimal. High-ability students never drop out of high school, due to their high levels of ability.

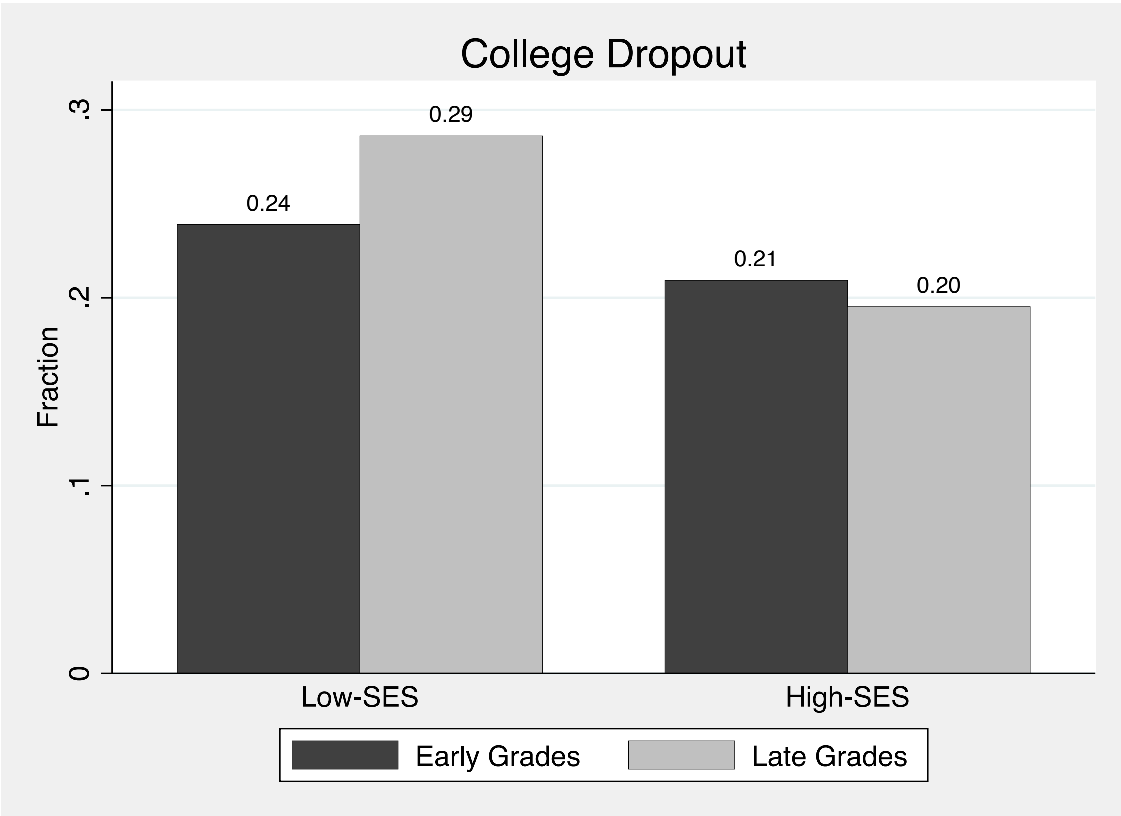


Figure A.16: College dropout by grading regime

Note: The Figure plots college dropout rates under early or late grade assignment. Notice that with full information dropout is never optimal.

A.4 Model and Institutional Setup

Table A.8 compares the model to the institutional setup in terms of choices, selection and information. While I designed the model around the institutional setup, there still are some differences.

Table A.8: Model and empirical setup

	Model	**Empirical setup**
Early Compulsory	choose e_1, no grades	choose effort in s.y. 1-5, no grades
Middle Compulsory	choose e_2, (no) grades	choose effort in s.y. 6 (no) grades
Late Compulsory	choose e_3, grades	choose effort/courses in s.y. 7-9, grades
High school Selection	$\bar{k}^{E_3}$	GPA, course choices
College Selection	E_3, $\bar{k}^{E_4}$	Academic HS, GPA, quotas

First, students in the sample have an additional choice with respect to the model: type of course in late compulsory school. As advanced courses are more difficult than general ones, this type of choice can be interpreted as an effort choice. At the same these choices are relevant for admission into academic high school. Table A.6 confirms this empirically: choosing advanced math in grade 9 and having higher grades substantially affects the probability of admission to the preferred high school choice. The magnitudes of the coefficients are lower bounds, as I am including students who only apply to vocational tracks.[38]

Second, in the model I assume that students need to meet absolute knowledge thresholds to complete college. In my setup a quota system is in place: an increase in college enrollment could in principle affect the admission threshold for all students. These general equilibrium effects are not captured in my model, where the number of students who can complete college is a function of the ability distribution. Öckert (2002) finds that the difference in years of education between students screened out and admitted at college in the early 80s in Sweden is about 0.6. The difference reduces to just 0.20 years when comparing students with similar number of admission credits.[39] This in turn is a good approximation to the marginal change in admission requirements that might be triggered by a reshuffling of the pool seeking college admission after grades are assigned. Given the size of the change results would likely not change significantly allowing for the general equilibrium effect.

[38] I have no information on the track the student applied for.

[39] See Öckert (2010), published version of the IFAU working paper.

B Descriptives

B.1 Definition of Ability and SES

In this section I discuss how I measure ability and SES, and describe how I discretize them to match the model.

Students took during the spring term of school year 6 a battery of three standardized ability tests: a test of *verbal ability,* requiring to find the opposite to a word among a list of four alternatives; a test of *inductive ability,* requiring to complete a number series of 6 terms with two more numbers; a test of *spatial ability,* requiring to find the three-dimensional representation of a two-dimensional picture that can be folded. The tests taken by the two cohorts are exactly the same, and the distributions look similar over time (see Figure B.1).

Students had respectively 15, 27 and 22 seconds to answer each section of the test, assuming they wasted no time at all in the test. The fast pace of the test adds to the quality of the ability measures: Borghans et al. (2008) show that reducing the time available for completing intelligence tests reduces differences in effort between students with different non-cognitive traits.

I create a standardized aggregate index of ability from the z-scores of inductive and verbal ability. I label high-ability those students who scored at least at the 60th percentile of the ability distribution. Consistently with the model, the cutoff roughly corresponds to the median ability of students who attained college education.[40]

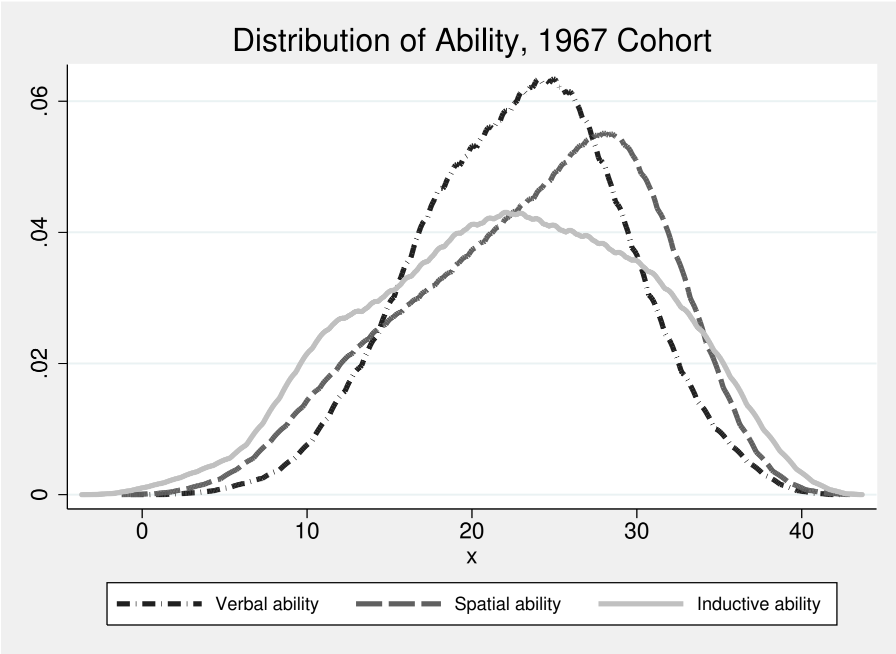
When performing the normalization at the cohort level, ability measures turn out to be 5% of a standard deviation higher in the treatment group, with respect to the control group. For the 1972 cohort there is no such difference. While the main treatment is grade assignment in school year 6, in principle there might be differences in grade assignment also in school year 3, when students were age 10 (see discussion in Section 2.3). Ability could thus have been affected by the treatment. However the literature (e.g., Heckman et al., 2007) reports that cognitive ability should be stable by that age. The ability measures were taken in May of 1981 for the 1967 cohort, quite close to the final tests used for grade assignment. It is possible that test taking behavior was instead affected in the treated municipalities. Students may have thought that the standardized tests were relevant to the final grades, or may have put more effort in the tests simply because they were affected by the more competitive environment.[41] This is consistent with results from the literature on the effect of non-cognitive traits on test taking behavior (Borghans et al.,

[40] I leave out of the index spatial ability, as it poorly correlates with academic choices or outcomes.
[41] Jalava et al. (2015) show that rank-based grading positively affects effort during tests

2008). In order to have more consistent ability measures I thus normalize ability at the cohort-treatment level, and basically use a measure of relative ability in the analysis. This avoids any problem of endogeneity or differential reporting caused by grade assignment.

Extensive investigation of which SES measure is most predictive of education choice shows that parental education strongly predicts children's education choices. Parental income is less predictive of education choice. Measures based on parental occupation yield results similar to parental education. My preferred measure of SES is based on parental education. Occupation-based measures are more difficult to discretize into dummies, and are recorded in my data using a definition that slightly changes between cohorts.[42] Finally, I consider high-SES those students who have at least one parent with academic high school (about 40% of my sample).

[42] Results do not change that much when using the alternative SES definition



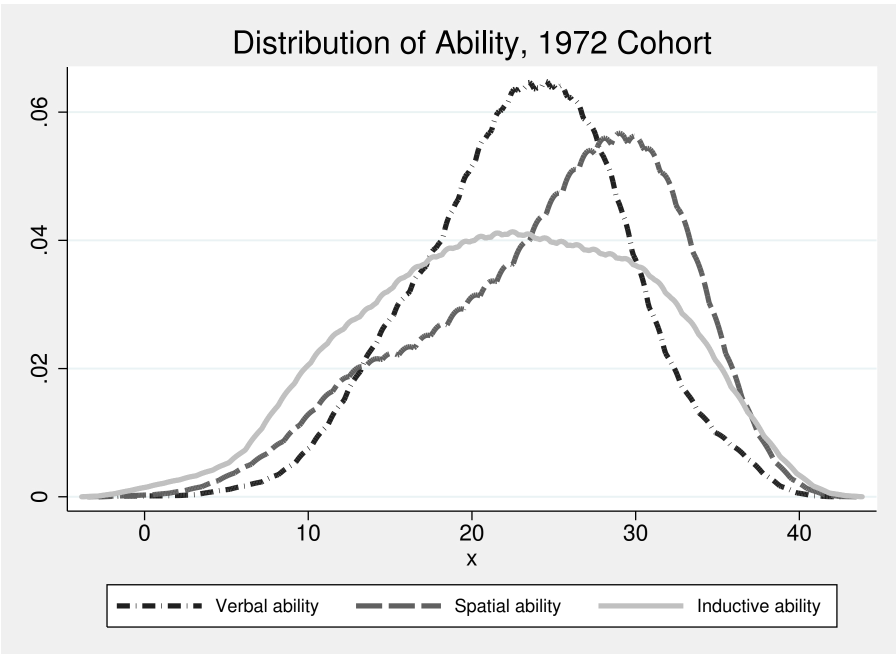


Figure B.1: Absolute ability distributions

Note: Ability measures are taken from tests administered in school year 6. The tests are the same for both cohorts.

B.2 Education Choices, Grades and Outcomes

In the following I provide descriptive evidence on students' choices and educational attainment, the main outcomes in the empirical analysis.

Table B.1: Education choices and outcomes by ability and SES, cohort 1967

	Low-ability		**High-ability**	
	Low-SES	**High-SES**	**Low-SES**	**High-SES**
Compulsory:				
Adv Math (s.y. 7)	0.56	0.73	0.92	0.97
Adv Math (s.y. 8)	0.44	0.62	0.87	0.95
Adv Math (s.y. 9)	0.32	0.51	0.75	0.88
Adv English (s.y. 7)	0.57	0.77	0.93	0.98
Adv English (s.y. 8)	0.53	0.74	0.90	0.97
Adv English (s.y. 9)	0.47	0.66	0.86	0.95
Non Compulsory:				
HS Enroll	0.85	0.92	0.93	0.98
Academic HS Enroll	0.22	0.45	0.60	0.82
HS Dropout	0.19	0.12	0.10	0.09
Attain Academic HS	0.19	0.39	0.56	0.75
Attain College	0.19	0.39	0.48	0.72

Each variable is a dummy. College enrollment is defined as enrolling into a tertiary education program lasting at least 2 years.

Table B.1 shows that high-ability and high-SES students are more likely to make choices consistent with an academic education path. This pattern suggests that the Swedish education system (grading, tracking and funding) successfully managed to allocate the most skilled students to higher education levels, but SES remained a relevant factor in the process, possibly distorting the efficient allocation of skill to human capital. While these differences are less marked for high-ability students up to compulsory school, low-SES students are 20 p.p. less likely to choose an academic education, independently of ability. As there might still be differences in ability between SES categories, in Table B.2 I report coefficients for the differences in choices between high- and low-SES students,

Table B.2: Differences in choices by SES controlling for ability, cohort 1967

	Low-ability	**High-ability**
Compulsory:		
Adv Math (s.y. 7)	0.11	0.04
Adv Math (s.y. 8)	0.12	0.07
Adv Math (s.y. 9)	0.13	0.10
Adv English (s.y. 7)	0.14	0.04
Adv English (s.y. 8)	0.15	0.06
Adv English (s.y. 9)	0.14	0.07
High School:		
HS Enroll	0.06	0.04
Academic Enroll	0.19	0.18
HS Dropout	–0.05	–0.00
Attain Academic HS	0.17	0.16
College:		
College	0.17	0.21

Each variable is a dummy. College enrollment is defined as enrolling into a tertiary education program lasting at least 2 years.

controlling for ability. The picture does not change that much, but it appears that the differences among low ability students are in part due to low-SES students having less ability. I still confirm that low-SES students are much less likely (a 18 p.p difference) to choose an academic education path than their high-SES counterparts.

To understand whether differences in education choices and educational attainment are related to school performance, I report in Table B.3 average grades by ability and SES. Grades are consistently higher for high-ability and high-SES students. However grade differences between high- and low-SES students with similar ability levels are not so big. When considering grades in school year 6, and Swedish grades in school year 9 (which do not reflect course choice), differences are at most $\frac{1}{3}$th of a grade. Table B.4 reports grade differences by SES controlling for ability. The picture remains similar: grade differences among students with similar ability levels are at most $\frac{1}{4}$th of a grade. This suggests that SES plays a fundamental role in education choice in Sweden, potentially reflecting different motivation and preferences for education among students.

Table B.3: Grades by ability and SES, cohort 1967

	Low-ability		**High-ability**	
	Low-SES	**High-SES**	**Low-SES**	**High-SES**
Swedish Grade (s.y. 6)	2.64	2.93	3.67	3.97
Swedish Grade (s.y. 7)	2.61	2.86	3.40	3.63
Swedish Grade (s.y. 8)	2.64	2.88	3.46	3.72
Swedish Grade (s.y. 9)	2.69	3.01	3.56	3.89
Math Grade (s.y. 6)	2.66	2.97	3.80	3.98
Math Grade (s.y. 7)	2.64	2.78	3.38	3.61
Math Grade (s.y. 8)	2.69	2.86	3.37	3.61
Math Grade (s.y. 9)	2.83	3.00	3.51	3.72
English Grade (s.y. 6)	2.64	2.98	3.65	4.02
English Grade (s.y. 7)	2.69	2.84	3.33	3.63
English Grade (s.y. 8)	2.68	2.86	3.37	3.63
English Grade (s.y. 9)	2.80	3.06	3.45	3.76

Grades are expressed on a 1-5 norm-referenced scale. Math and English grades in s.y. 8 and 9 pool together advanced and general courses.

Table B.5 shows that many students switch courses over time. Students are more likely to switch from academic to general courses than the opposite, and there are more switches in math.[43] Switches from one type of course to the other can be interpreted as revision of choice, and imply that students do not have full information over own ability (or knowledge). It is also interesting to see that the switching behavior continues through all grades: most of the students switch in school year 8, but some also switch in the last year. This is consistent with students revising some sort of prior, with the updating process continuing over time. In Table B.4 I show how SES affects the choice of switching in compulsory school from an advanced to a general course. When comparing the choices of low- and high-SES students with the same grades and ability, I find that SES still influences switching choices. This suggests that grading information might affect differently students from different socioeconomic backgrounds.

[43]This is consistent with research finding that students tend to be overly optimistic about own ability/preparation in higher education (see Stinebrickner & Stinebrickner, 2012; Zafar, 2011)

Table B.4: Differences in grades by SES controlling for ability, cohort 1967

	Low-ability	**High-ability**
Swedish Grade (s.y. 6)	0.18	0.21
Swedish Grade (s.y. 7)	0.16	0.15
Swedish Grade (s.y. 8)	0.15	0.17
Swedish Grade (s.y. 9)	0.22	0.24
Math Grade (s.y. 6)	0.18	0.11
Math Grade (s.y. 7)	0.07	0.14
Math Grade (s.y. 8)	0.11	0.16
Math Grade (s.y. 9)	0.10	0.14
English Grade (s.y. 6)	0.22	0.27
English Grade (s.y. 7)	0.07	0.21
English Grade (s.y. 8)	0.11	0.16
English Grade (s.y. 9)	0.20	0.21

Grades are expressed on a 1-5 norm-referenced scale. Math and English grades in s.y. 8 and 9 pool together advanced and general courses.

Table B.5: Fraction of students switching courses, 1967 cohort

	Grade 8	**Grade 9**
Math:		
Switches to gen choice	0.12	0.17
Switches to adv choice	0.04	0.01
English:		
Switches to gen choice	0.06	0.08
Switches to adv choice	0.07	0.03

Switches are conditional on previous year's course choice.

Table B.6: Impact of SES on the stability of course choices, 1967 Cohort

	Sticks to adv Math	Sticks to adv Eng
Regressors of interest:		
High-SES	0.06***	0.02*
	(0.01)	(0.01)
Controls:		
Grade (s.y. 8)	–0.00	–0.05***
	(0.01)	(0.01)
Grade (s.y. 7)	0.21***	0.18***
	(0.01)	(0.01)
Standardized verbal ability	0.02**	0.04***
	(0.01)	(0.01)
Standardized inductive ability	0.02**	0.00
	(0.01)	(0.01)
Standardized spatial ability	0.04***	0.02***
	(0.01)	(0.01)
E[Y]	0.73	0.86
R^2	0.27	0.21
Observations	5532	5653

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Outcome: the student sticks to the advanced course choice made in s.y. 7. The SES division is based on parental education. Standard errors clustered at the class level.

B.3 Treated and Control Municipalities

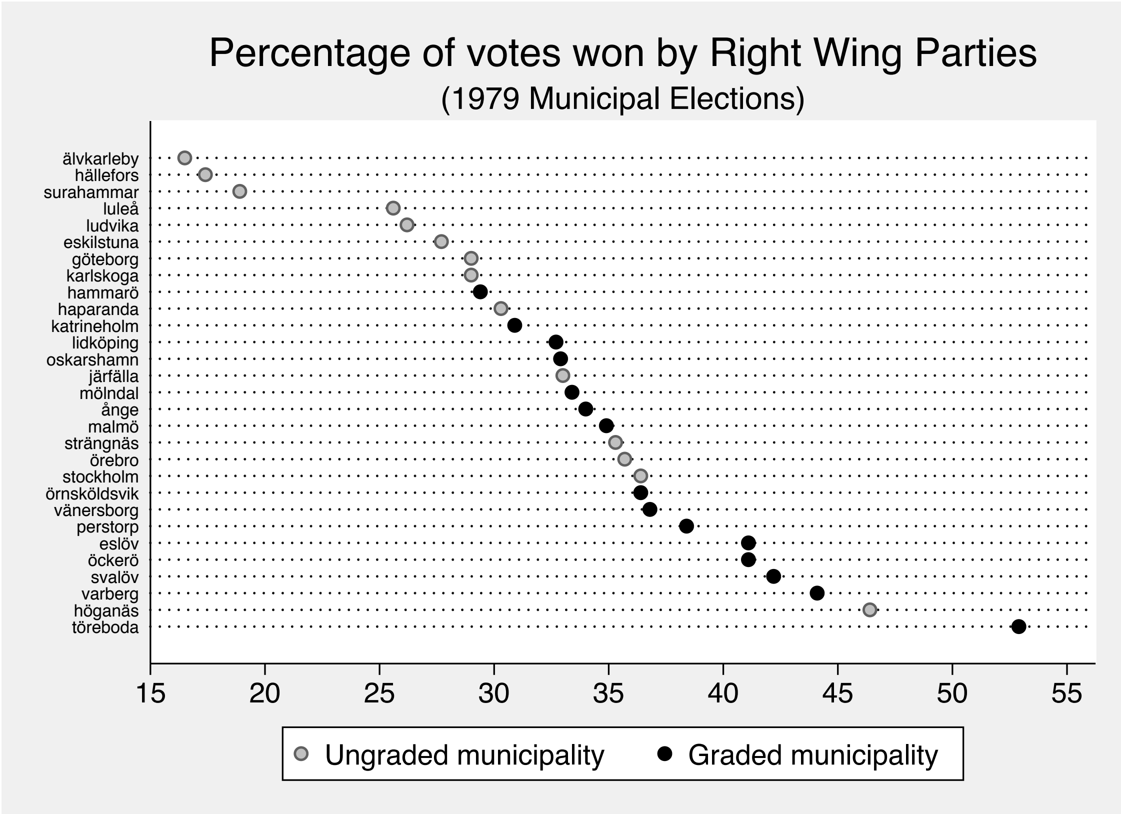


Figure B.2: Vote share of right-wing parties in municipal elections by treatment status

Note: This Figure plots the aggregated vote share of right-wing parties in the 1979 municipal elections. Municipalities assigning early grades had a higher share of right-wing voters.

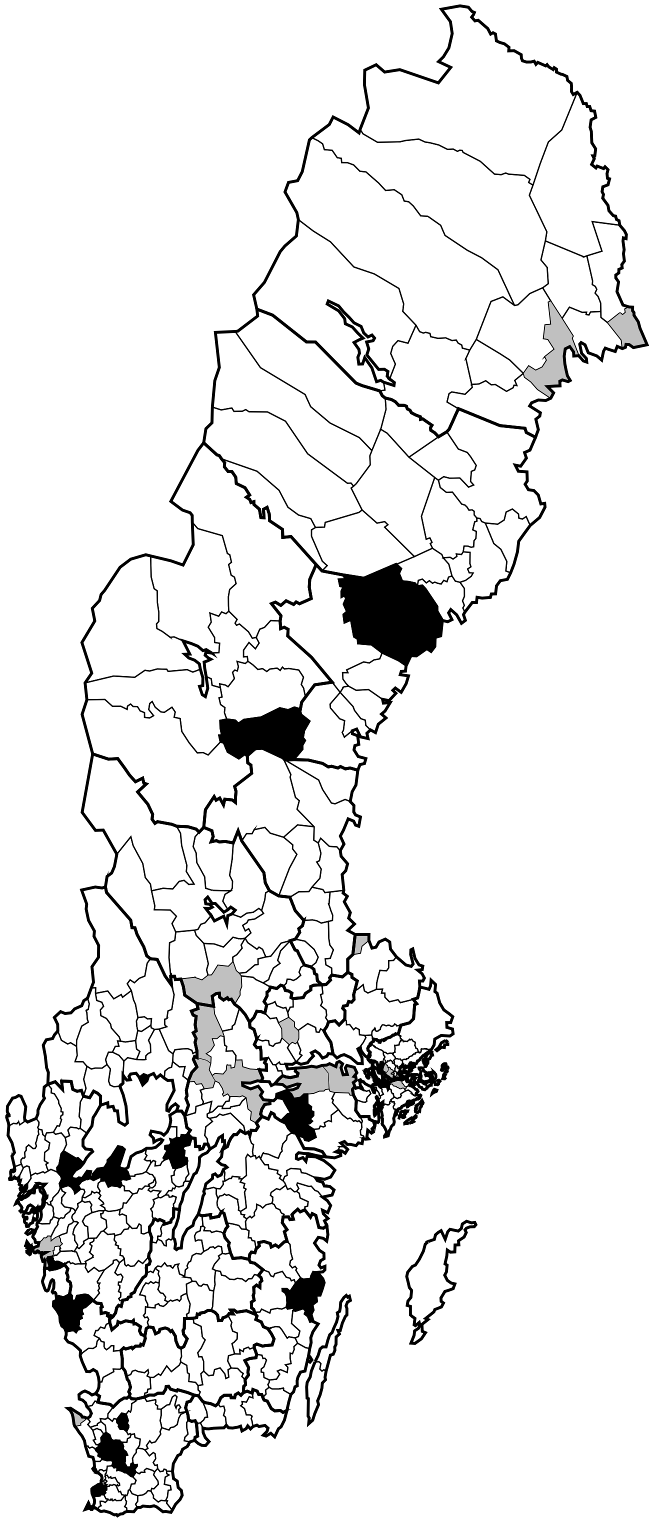


Figure B.3: Sampled municipalities

Note: Municipalities assigning early grades before the reform are colored in black

Table B.7: Differences in students background
by treatment status, 1967 cohort

	Graded (sd)	Ungraded (sd)	Difference (sd)
Female	0.48	0.50	-0.02
	(0.62)	(0.55)	(0.01)
Birth year	66.98	66.97	0.02*
	(0.25)	(0.24)	(0.01)
Foreign born	0.02	0.05	-0.03***
	(0.14)	(0.24)	(0.01)
Both parents not Nordic	0.01	0.02	-0.00
	(0.11)	(0.13)	(0.01)
Verbal ability	23.22	22.70	0.52*
	(7.60)	(6.60)	(0.28)
Inductive ability	22.85	21.78	1.07***
	(10.25)	(8.92)	(0.38)
Spatial ability	23.82	23.41	0.41
	(8.94)	(7.91)	(0.28)
Kindergarten	0.91	0.93	-0.02
	(0.34)	(0.28)	(0.03)
Quiet home environment	0.95	0.95	-0.00
	(0.28)	(0.25)	(0.01)
Switched Class (G6-G9)	0.06	0.08	-0.03***
	(0.27)	(0.30)	(0.01)
Special Education	0.12	0.14	-0.02
	(0.41)	(0.39)	(0.02)
Changes of teacher	0.59	0.53	0.05
	(1.20)	(1.00)	(0.14)
Hours absent in grade 6	7.89	6.82	1.07
	(11.94)	(8.23)	(1.26)
Class size	23.64	23.27	0.37
	(6.64)	(5.18)	(0.75)

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

Standard errors clustered at the municipal level. Ability measures are on a 0-40 scale.

Table B.8: Differences in parental SES
by treatment status, 1967 cohort

	Graded (sd)	Ungraded (sd)	Difference (sd)
Disposable family income (age 11-16)	243.69	253.82	-10.13**
	(101.46)	(84.98)	(4.87)
High Income	0.47	0.54	-0.07**
	(0.62)	(0.55)	(0.03)
High Education	0.36	0.44	-0.08**
	(0.60)	(0.54)	(0.03)
High Income/Educ	0.59	0.68	-0.09***
	(0.62)	(0.51)	(0.03)
High SES	0.41	0.46	-0.05
	(0.61)	(0.55)	(0.04)
Parents: non-skilled workers, goods	0.10	0.09	0.01
	(0.39)	(0.33)	(0.02)
Parents: non-skilled workers, service	0.11	0.09	0.02**
	(0.39)	(0.31)	(0.01)
Parents: skilled workers, goods	0.17	0.19	-0.02
	(0.47)	(0.43)	(0.02)
Parents: skilled workers, service	0.02	0.02	-0.00
	(0.18)	(0.16)	(0.01)
Parents: lower non-manual ii	0.04	0.05	-0.00
	(0.25)	(0.22)	(0.01)
Parents: lower non-manual i	0.09	0.09	0.00
	(0.35)	(0.30)	(0.01)
Parents: intermediate-level non-manual	0.19	0.23	-0.03**
	(0.49)	(0.46)	(0.01)
Parents: higher civil servants and senior salaried	0.11	0.13	-0.01
	(0.39)	(0.36)	(0.02)
Parents: independent professionals	0.00	0.00	-0.00**
	(0.03)	(0.06)	(0.00)
Parents: entrepreneur	0.11	0.10	0.01
	(0.39)	(0.33)	(0.01)
Parents: farmer	0.05	0.02	0.04***
	(0.30)	(0.16)	(0.01)

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

Standard errors clustered at the municipal level. Income in 1000 kr, measured when the student was 11-16. Occupation variables are taken from the 1980 Census.

Table B.9: Differences in father background by treatment status, 1967 cohort

	Graded (sd)	Ungraded (sd)	Difference (sd)
Father not Nordic	0.03	0.03	-0.01
	(0.19)	(0.20)	(0.01)
Married father	0.81	0.77	0.04***
	(0.49)	(0.46)	(0.01)
Father SES, 1 (low) to 3 (high)	1.73	1.75	-0.02
	(0.87)	(0.79)	(0.05)
Divorced father	0.15	0.18	-0.03**
	(0.45)	(0.42)	(0.01)
Father educ: compulsory school or less	0.47	0.41	0.06**
	(0.62)	(0.54)	(0.03)
Father educ: high school	0.38	0.41	-0.04**
	(0.60)	(0.54)	(0.02)
Father educ: college or more	0.16	0.18	-0.02
	(0.45)	(0.42)	(0.02)
Father: in the labor force	0.90	0.90	0.00
	(0.36)	(0.33)	(0.01)
Father: unemployed	0.04	0.04	0.00
	(0.24)	(0.20)	(0.00)

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

Standard errors clustered at the municipal level.

Table B.10: Differences in mother background
by treatment status, 1967 cohort

	Graded (sd)	Ungraded (sd)	Difference (sd)
Mother not Nordic	0.04	0.05	-0.01
	(0.21)	(0.23)	(0.01)
Married mother	0.77	0.72	0.04***
	(0.52)	(0.49)	(0.02)
Divorced mother	0.16	0.19	-0.04**
	(0.45)	(0.43)	(0.01)
Mother SES, 1 (low) to 3 (high)	1.55	1.62	-0.06
	(0.76)	(0.68)	(0.05)
Mother educ: compulsory school or less	0.42	0.41	0.01
	(0.61)	(0.54)	(0.03)
Mother educ: high school	0.41	0.41	0.00
	(0.62)	(0.54)	(0.02)
Mother educ: college or more	0.17	0.18	-0.01
	(0.47)	(0.42)	(0.02)
Mother: in the labor force	0.91	0.90	0.01
	(0.36)	(0.34)	(0.01)
Mother: unemployed	0.04	0.03	0.01*
	(0.25)	(0.19)	(0.00)

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

Standard errors clustered at the municipal level.

Table B.11: Differences in father occupation
by treatment status, 1967 cohort

	Graded (sd)	Ungraded (sd)	Difference (sd)
Father: non-skilled workers, goods	0.12	0.12	0.01
	(0.42)	(0.35)	(0.02)
Father: non-skilled workers, service	0.10	0.08	0.03***
	(0.38)	(0.29)	(0.01)
Father: skilled workers, goods	0.21	0.23	-0.02
	(0.50)	(0.46)	(0.02)
Father: skilled workers, service	0.01	0.01	-0.00
	(0.09)	(0.10)	(0.00)
Father: lower non-manual ii	0.03	0.02	0.01
	(0.21)	(0.17)	(0.00)
Father: lower non-manual i	0.07	0.07	0.00
	(0.32)	(0.29)	(0.01)
Father: intermediate-level non-manual	0.18	0.22	-0.04**
	(0.47)	(0.45)	(0.02)
Father: higher civil servants and senior salaried	0.11	0.12	-0.02
	(0.38)	(0.36)	(0.02)
Father: independent professionals	0.00	0.00	-0.00*
	(0.03)	(0.05)	(0.00)
Father: entrepreneur	0.11	0.10	0.01
	(0.39)	(0.32)	(0.01)
Father: farmer	0.06	0.02	0.03**
	(0.30)	(0.17)	(0.01)

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

Standard errors clustered at the municipal level. Occupation variables are taken from the 1980 Census.

Table B.12: Differences in mother occupation by treatment status, 1967 cohort

	Graded (sd)	Ungraded (sd)	Difference (sd)
Mother: non-skilled workers, goods	0.05	0.06	-0.00
	(0.27)	(0.26)	(0.01)
Mother: non-skilled workers, service	0.37	0.36	0.02
	(0.60)	(0.53)	(0.03)
Mother: skilled workers, goods	0.01	0.02	-0.01
	(0.14)	(0.15)	(0.01)
Mother: skilled workers, service	0.06	0.06	0.00
	(0.31)	(0.25)	(0.01)
Mother: lower non-manual ii	0.11	0.14	-0.03*
	(0.39)	(0.38)	(0.01)
Mother: lower non-manual i	0.09	0.11	-0.02
	(0.34)	(0.33)	(0.01)
Mother: intermediate-level non-manual	0.16	0.16	-0.00
	(0.47)	(0.40)	(0.01)
Mother: higher civil servants and senior salaried	0.04	0.04	0.00
	(0.25)	(0.21)	(0.01)
Mother: independent professionals	0.00	0.00	-0.00*
	(0.00)	(0.03)	(0.00)
Mother: entrepreneur	0.05	0.04	0.00
	(0.28)	(0.23)	(0.01)
Mother: farmer	0.05	0.02	0.03***
	(0.28)	(0.15)	(0.01)

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

Standard errors clustered at the municipal level. Occupation variables are taken from the 1980 Census.

Table B.13: Differences in parents choice protocols by treatment status, 1967 cohort

	Graded (sd)	Ungraded (sd)	Difference (sd)
Electives chosen for: child ability	0.62 (0.60)	0.61 (0.53)	0.02 (0.02)
Electives chosen for: child preferences	0.37 (0.61)	0.39 (0.53)	-0.02 (0.02)
Electives chosen for: more choice in HS	0.37 (0.61)	0.38 (0.53)	-0.00 (0.02)
Electives chosen for: entrance requirements	0.24 (0.53)	0.26 (0.48)	-0.01 (0.01)
Electives chosen for: teacher suggestion	0.04 (0.24)	0.07 (0.28)	-0.03*** (0.01)
Electives chosen for: classmates choice	0.01 (0.12)	0.01 (0.13)	-0.00 (0.00)
Electives chosen for: other	0.01 (0.12)	0.01 (0.12)	-0.00 (0.00)

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

Standard errors clustered at the municipal level. Variables represent agreement with the statement on a 0-1 scale.

Table B.14: Differences in parental school preferences by treatment status, 1967 cohort

	Graded (sd)	Ungraded (sd)	Difference (sd)
School priority: Teach Math and Swedish	8.21	8.06	0.14**
	(1.73)	(1.63)	(0.06)
School priority: Develop collaborative skills	6.64	6.70	-0.06
	(2.26)	(2.02)	(0.09)
School priority: Teach other subjects	6.03	6.01	0.02
	(2.35)	(1.99)	(0.07)
School priority: Develop critical thinking	5.95	6.17	-0.22**
	(2.62)	(2.33)	(0.10)
School priority: Teach Foreign languages	5.46	5.61	-0.16
	(2.57)	(2.19)	(0.14)
School priority: Inform about working life	4.53	4.64	-0.11
	(2.36)	(2.20)	(0.11)
School priority: Teach children to obey adults	3.65	3.42	0.23*
	(2.93)	(2.49)	(0.13)
School priority: Cope in a competitive society	3.57	3.62	-0.05
	(2.58)	(2.23)	(0.08)
School priority: Select for higher education	1.79	1.80	-0.01
	(1.84)	(1.66)	(0.07)

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

Standard errors clustered at the municipal level. Variables are on a 1-9 scale (9 = top priority).

C Refutability Tests

C.1 Tests for Parallel Trends

The following Figures show tests for parallel trends between treatment and control municipalities in determinants of education and educational attainment. For each outcome I plot in the upper panel trends for the grading and non-grading municipalities that are part of my sample. The two dashed lines mark the years in which the 1967 and 1972 cohorts were in school year 6: the tests close to this period are thus more relevant. In the lower panel I show coefficient and 95% confidence interval from difference in differences placebo regressions. In the regressions I control for a linear trend, and run tests over a 5-year window centered on the year marked in the picture. This way I test precisely the assumption that underlies my specification: over a 5-year window there should be no differential trend in education (or related variables) between treated and control municipalities.[44]

In Figure C.1 and Table C.1, I test for parallel trends in the aggregate vote share of right wing parties in municipal elections, held every 4 years in Sweden. It is reassuring to see that the differences in vote share, which can be considered as the "treatment assignment," are quite stable over time. In Figures C.2 to C.4 I consider educational attainment in the population aged 38-74, corresponding to cohorts who completed their education before the first grading reform, likely including also the parents of the students in the sample. This avoids picking up any effect of the reform. Trends appear to be parallel for all education levels, while there seems to be some catching up on the part of the graded municipalities in high school attainment. That coefficient however is small and marginally significant only in 1989, after the period I consider.[45] Figures C.5 to C.7 consider flows in educational attainment for cohorts born 1969 onwards. These cohorts were all studying under the reformed school system, and thus they started getting grades in school year 8. The first of the pictures plots the fraction of students who graduate from a 2-year high school, which up to the early 90s corresponds to vocational high school.[46] The fraction of students who completes vocational education is higher in grading municipalities, but the diff-in-diff coefficients in the tests are all close to 0, implying that differences remained

[44]Difference in differences is functional form dependent (Lechner, 2011), and the functional form assumed for the trend should be consistent with the data (Mora & Reggio, 2012). In my analysis I can only control for a linear trend, but this should not be problematic as I use cohorts just 5 years apart.

[45]Notice that administrative education data is only available starting from 1985.

[46]A reform in the 90s increased the length of vocational training to three years.

stable over time. This is one of the key education variables I will be using as an outcome in my analysis, hence it is particularly reassuring to see that the test passes. A similar picture emerges for high school attainment and college.[47]

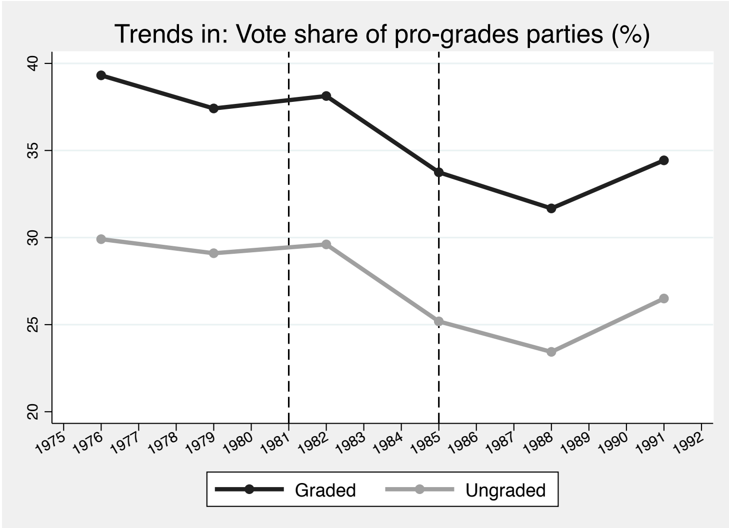


Figure C.1: Test for parallel trends in treatment assignment

Note: The Figure plots aggregate vote share of right-wing parties (in general favoring early grade assignment) in municipal elections, held every 4 years in Sweden, for grading and non-grading municipalities in the sample. Dashed vertical lines mark the period in which the two cohorts in the sample were in school year 6.

Table C.1: Test for parallel trends in pro-grade vote share

	Vote share of pro-grades parties (%)
Graded x Year	–0.05
	(0.08)
Graded	111.89
	(160.44)
Year	–24.62
	(25.11)
Mean	35.23
R^2	0.32

Standard errors clustered at the municipality level.
Note: * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$

[47]Statistics Sweden was changing the classification of education in the last part of the panel, a thus there are breaks in the trends. A dashed line marks the first and last year in which estimates are affected by the break, which may lead to spurious rejections in the tests.

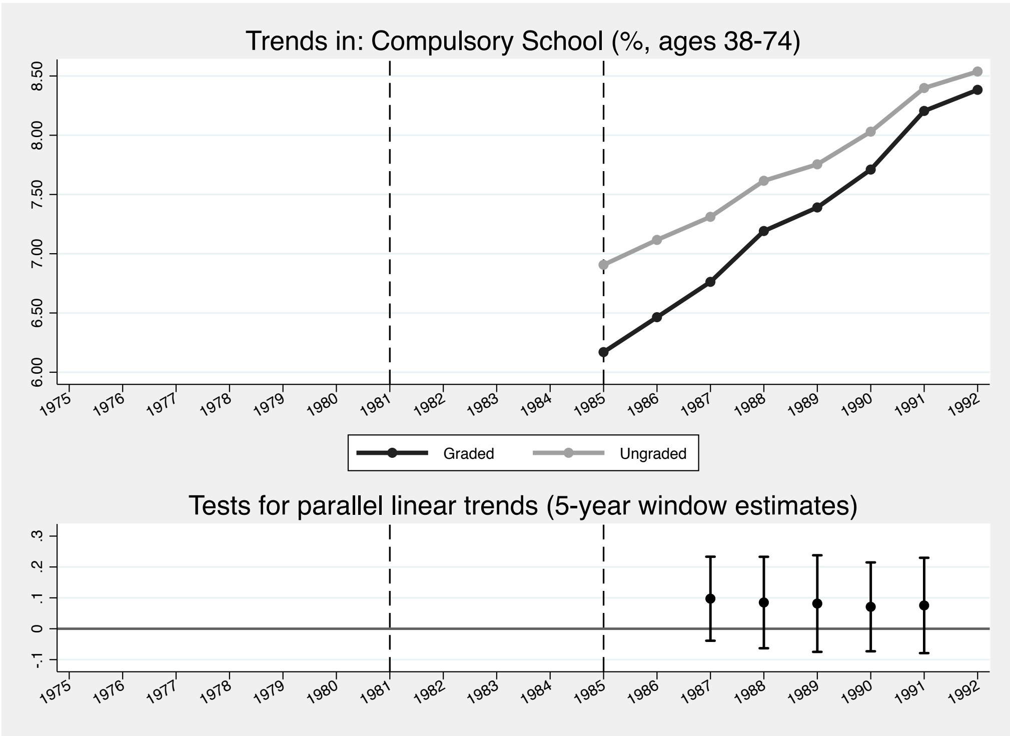


Figure C.2: Test for parallel trends in education: Compulsory school (graded cohorts)

Note: The Figure plots in the upper panel trends in 9-year-compulsory-school attainment for people aged 38-74, who studied before the initial reform, for grading and non-grading municipalities in the sample. Dashed vertical lines mark the period in which the two cohorts in the sample were studying. The lower panel presents coefficient and 95% confidence interval from diff-in-diff placebo regressions. Tests are run over a 5-year window centered on the marked year, and control for a linear trend.

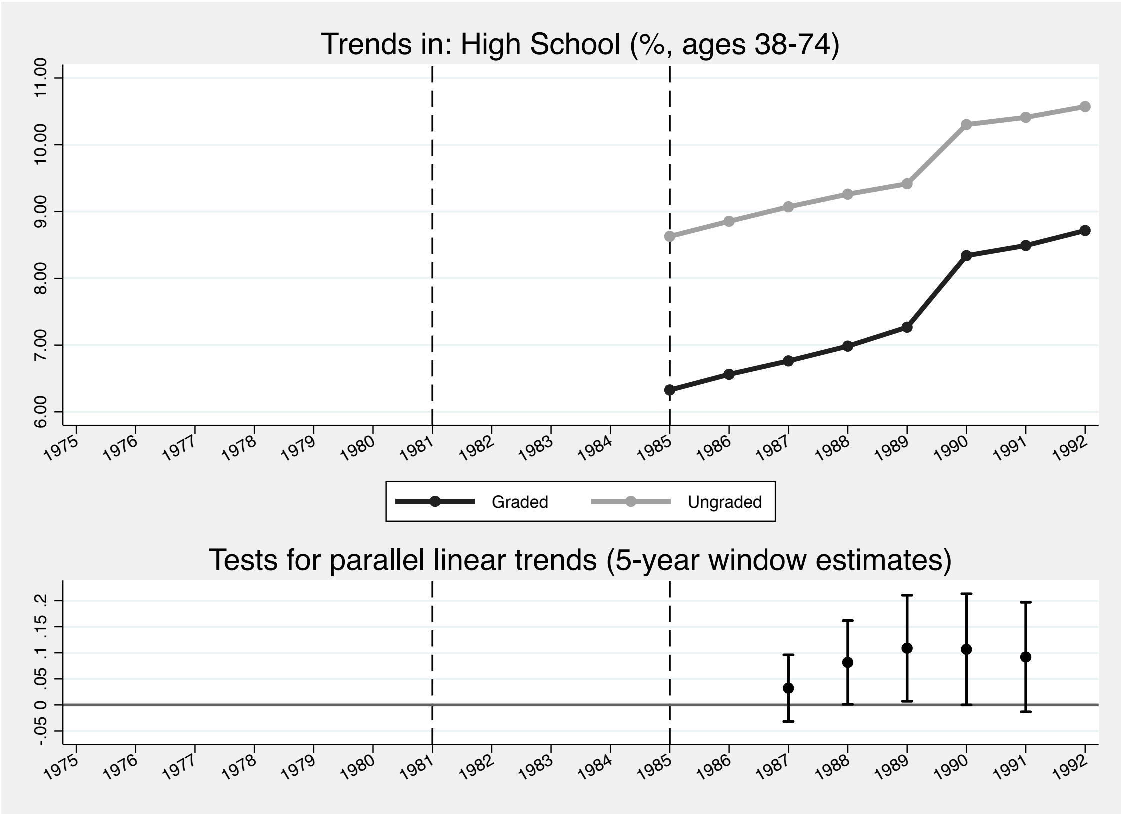


Figure C.3: Test for parallel trends in education: HS (graded cohorts)

Note: The Figure plots in the upper panel trends in high-school attainment for people aged 38-74, who studied before the initial reform, for grading and non-grading municipalities in the sample. Dashed vertical lines mark the period in which the two cohorts in the sample were studying. The lower panel presents coefficient and 95% confidence interval from diff-in-diff placebo regressions. Tests are run over a 5-year window centered on the marked year, and control for a linear trend.

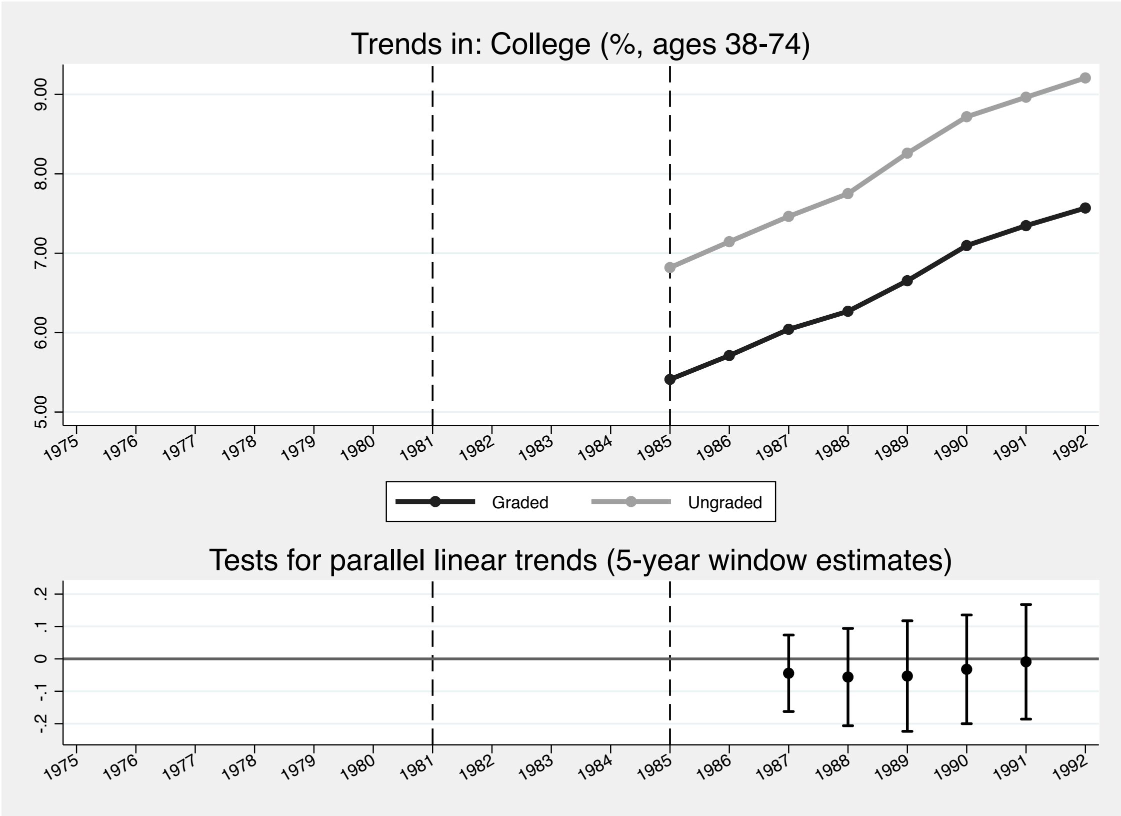


Figure C.4: Test for parallel trends in education: College (graded cohorts)

Note: The Figure plots in the upper panel trends in college attainment for people aged 38-74, who studied before the initial reform, for grading and non-grading municipalities in the sample. Dashed vertical lines mark the period in which the two cohorts in the sample were studying. The lower panel presents coefficient and 95% confidence interval from diff-in-diff placebo regressions. Tests are run over a 5-year window centered on the marked year, and control for a linear trend.

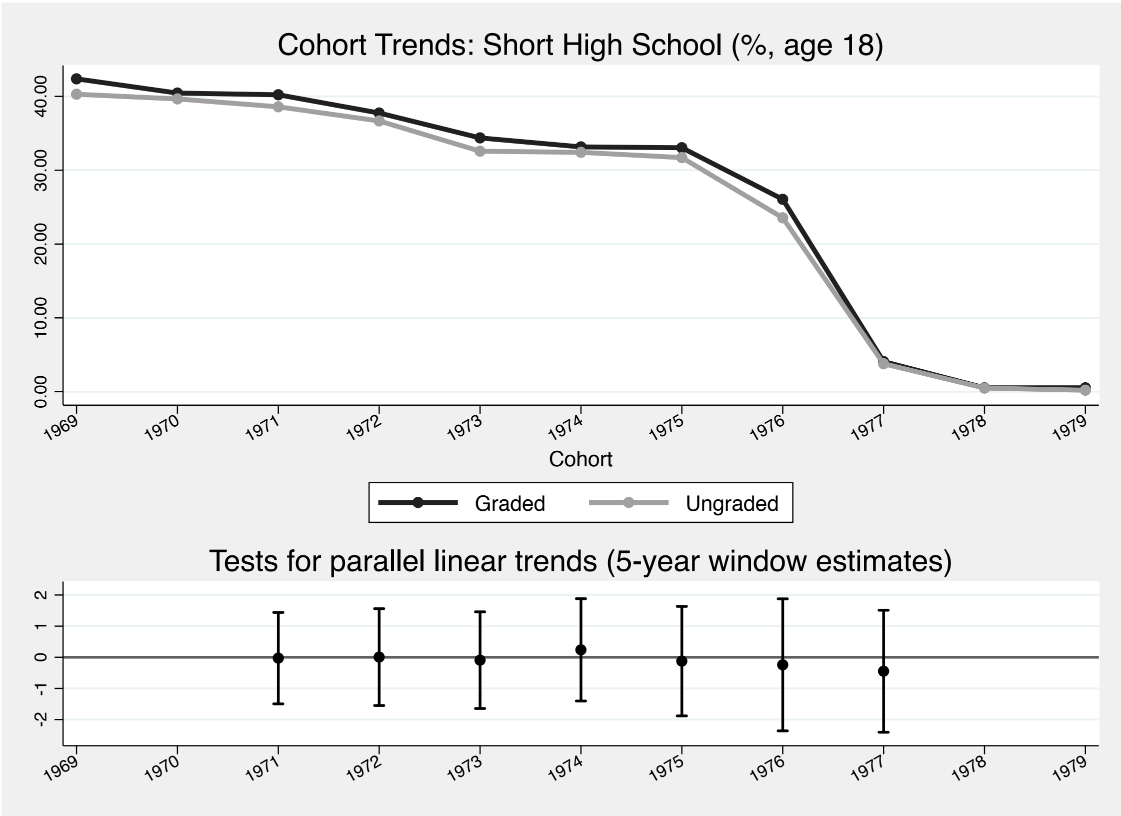


Figure C.5: Test for parallel trends in education flows: Short HS (ungraded cohorts)

Note: The Figure plots in the upper panel trends in vocational high school attainment for cohorts who studied when early grades were abolished, for grading and non-grading municipalities in the sample. The lower panel presents coefficient and 95% confidence interval from diff-in-diff placebo regressions. Tests are run over a 5-year window centered on the marked year, and control for a linear trend.

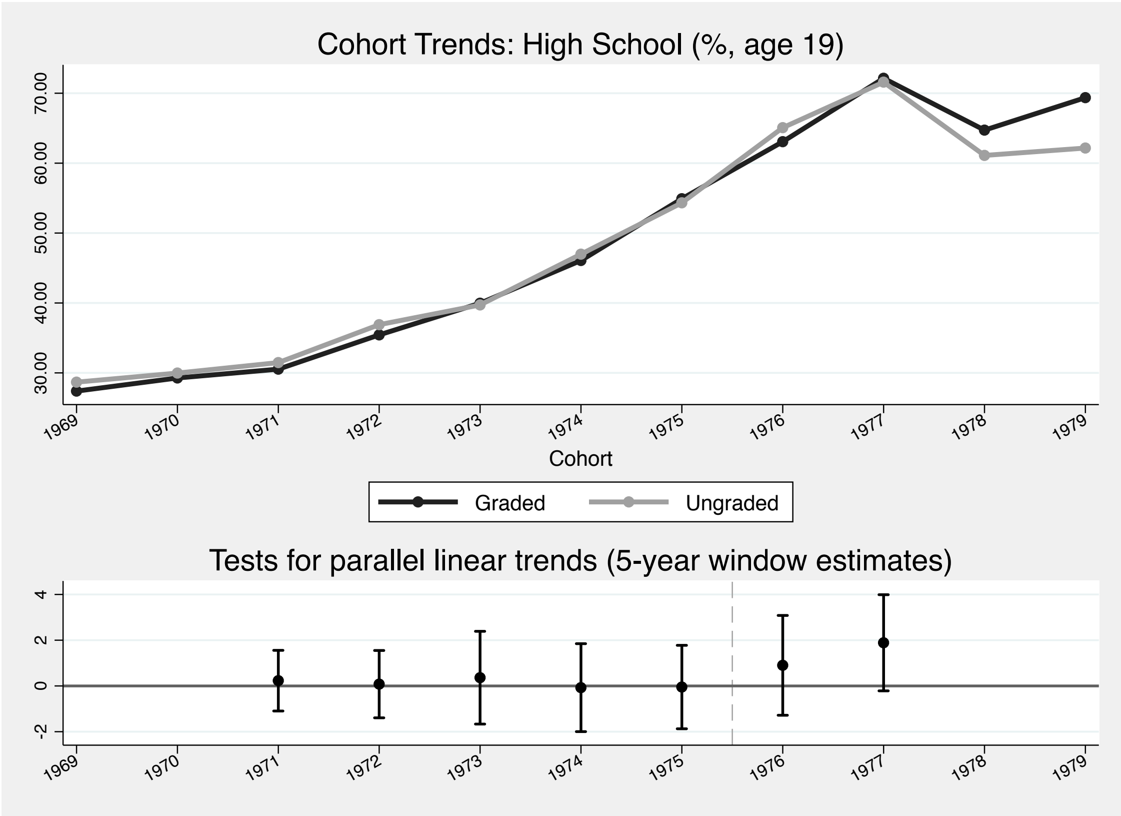


Figure C.6: Test for parallel trends in education flows: HS (ungraded cohorts)

Note: The Figure plots in the upper panel trends in high school attainment for cohorts who studied when early grades were abolished, for grading and non-grading municipalities in the sample. The lower panel presents coefficient and 95% confidence interval from diff-in-diff placebo regressions. Tests are run over a 5-year window centered on the marked year, and control for a linear trend. The dashed line marks a break in the time series due to reclassification of education by Statistics Sweden.

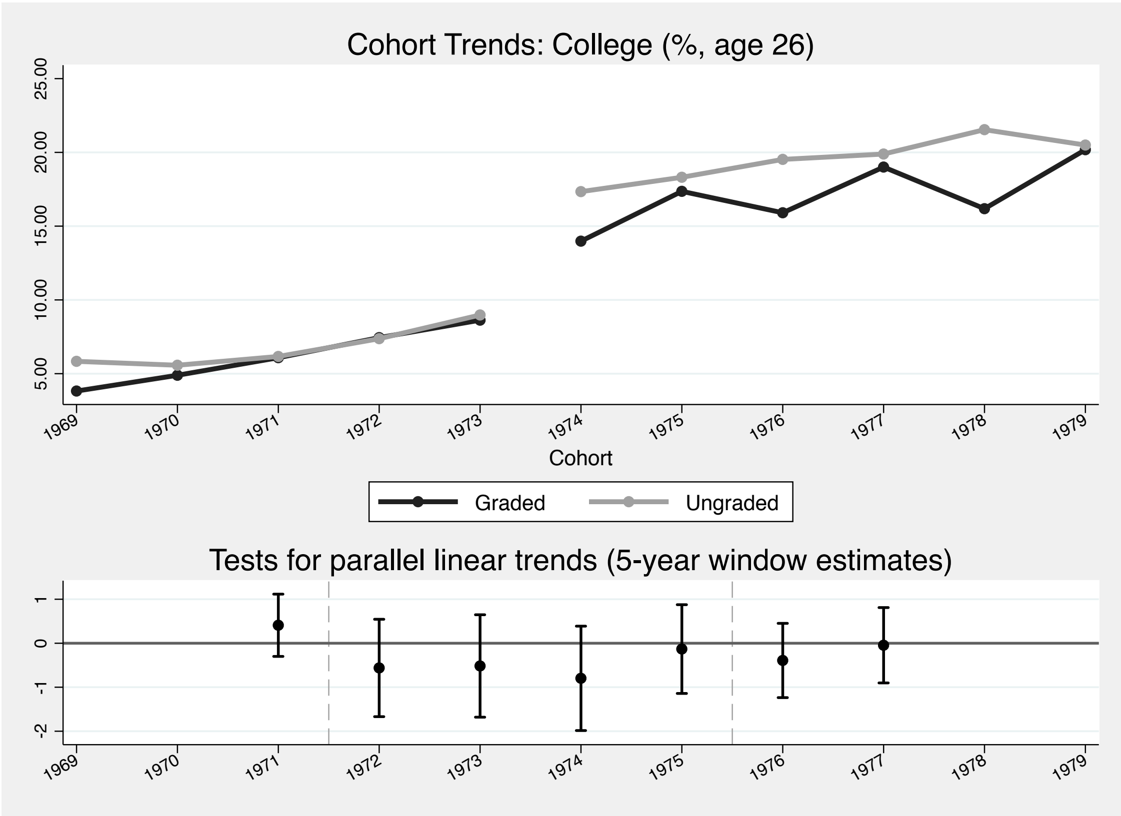


Figure C.7: Test for parallel trends in education flows: College (ungraded cohorts)

Note: The Figure plots in the upper panel trends in college attainment for cohorts who studied when early grades were abolished, for grading and non-grading municipalities in the sample. The lower panel presents coefficient and 95% confidence interval from diff-in-diff placebo regressions. Tests are run over a 5-year window centered on the marked year, and control for a linear trend. Estimates within the dashed lines are affected by a break in the time series due to reclassification of education by Statistics Sweden.

C.2 Tests for Differential Response and Compositional Change

Intelligence and SES data is missing for 18% of my sample, but attrition does not change over time between grading and non-grading municipalities (see Table C.3). It is thus possible to explore heterogeneous effects by ability and SES. There appear to be no issues for the surveys taken in grades 6 and 10, but parental surveys display differential attrition (see Table C.2). Thus I can not use variables from these surveys in the final specification. Finally, among the standardized tests that end-of-the-year grades are based upon, only the Swedish test does not exhibit differential attrition. So this will be the only standardized test I will be using as an outcome.

In Tables C.3 to C.8 I test for differential compositional change in the two sets of municipalities for a large set of pre-treatment variables. As I run the tests for many outcomes and for both the whole sample and the individual ability-SES cells I am likely to find spurious rejections. I thus comment on how the tests perform on average.[48] All demographic and school-level variables pass the tests (see Tables C.4 and C.5). The placebo tests for relative verbal and inductive ability fail in some cases within cell, but are by definition 0 in the sample since I normalized ability at the treatment cohort level.[49] The cross-sectional differences in marriage and divorce rates found for parents seem to persist over time (Tables C.6 and C.7). Both income (Table C.5) and a broad measure of parental education (SES, in Table C.3) pass the tests. Finally most of the occupational categories (Table C.8) pass the tests for compositional change, confirming that the cross sectional differences in occupation remained constant over time. When looking at parental educational attainment, it appears that the fraction of students with college educated mothers (Table C.6) and fathers (Table C.7) increases less in early grading municipalities with respect to the late grading municipalities. These differences are consistent with the two tests that fail for compositional change in occupation: the share of parents involved in non-manual occupations increased less over time in early grading municipalities.

[48] In some cases the tests pass in the sample, but not within ability-SES cell. This could be due to compositional change that I find in the ability measures used for the cell. In my analysis I thus always control for ability.

[49] See the discussion in Appendix B.1.

Table C.2: Tests for differential response:
Summary of difference in differences estimates

Outcome:		**Low-ability**		**High-ability**	
	All Sample	**Low-SES**	**High-SES**	**Low-SES**	**High-SES**
Student survey	-0.03	0.01	0.01	0.01	0.01
(school year 6)	(0.73)	(0.50)	(0.16)	(0.83)	(0.56)
	[0.90]	[0.99]	[0.99]	[0.99]	[0.99]
Student survey	0.02	0.02	0.04*	0.05**	0.03
(school year 10)	(0.26)	(0.37)	(0.05)	(0.02)	(0.19)
	[0.76]	[0.69]	[0.77]	[0.83]	[0.88]
Parent survey	-0.06	-0.01	-0.06**	-0.01	-0.05***
	(0.15)	(0.67)	(0.05)	(0.81)	(0.01)
	[0.74]	[0.71]	[0.78]	[0.82]	[0.87]
English Test	0.13*	0.10	0.09	0.17*	0.14
(school year 8)	(0.08)	(0.20)	(0.31)	(0.09)	(0.13)
	[0.24]	[0.27]	[0.21]	[0.20]	[0.18]
Swedish Test	0.08	0.10	0.07	0.12	0.07
(school year 9)	(0.45)	(0.34)	(0.56)	(0.34)	(0.52)
	[0.33]	[0.34]	[0.32]	[0.30]	[0.30]
Math Test	-0.26**	-0.29***	-0.26**	-0.28***	-0.31***
(school year 9)	(0.02)	(0.00)	(0.02)	(0.01)	(0.00)
	[0.27]	[0.30]	[0.25]	[0.23]	[0.22]

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

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets.

Table C.3: Tests for compositional change: SES and ability. Summary of diff-in-diff estimates

	All Sample
No SES or ability data	0.03
	(0.64)
	[0.18]
High-SES	0.00
	(0.84)
	[0.44]
Low-ability Low-SES	-0.00
	(0.87)
	[0.38]
Low-ability High-SES	-0.04*
	(0.07)
	[0.21]
High-ability Low-SES	0.01
	(0.55)
	[0.17]
High-ability High-SES	0.03**
	(0.04)
	[0.23]

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

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets.

Table C.4: Tests for compositional change: Demographics.
Summary of difference in differences estimates

Outcome:		**Low-ability**		**High-ability**	
	All Sample	**Low-SES**	**High-SES**	**Low-SES**	**High-SES**
Female	-0.02	-0.03	-0.04	0.01	0.04
	(0.52)	(0.50)	(0.19)	(0.71)	(0.42)
	[0.49]	[0.47]	[0.48]	[0.52]	[0.50]
Birth year	0.01	0.01	-0.01	0.01	0.02**
	(0.53)	(0.36)	(0.74)	(0.19)	(0.03)
	[1,969.39]	[1,969.20]	[1,969.65]	[1,969.18]	[1,969.63]
Verbal ability	0.01	0.02	-0.09*	-0.10	-0.06
	(0.85)	(0.54)	(0.07)	(0.13)	(0.32)
	[-0.00]	[-0.53]	[-0.33]	[0.60]	[0.77]
Inductive ability	0.01	0.00	0.02	-0.07	-0.12***
	(0.86)	(0.95)	(0.60)	(0.23)	(0.00)
	[0.00]	[-0.53]	[-0.41]	[0.70]	[0.75]
Spatial ability	0.00	0.02	-0.06	-0.08	-0.05
	(0.91)	(0.47)	(0.13)	(0.10)	(0.25)
	[0.00]	[-0.33]	[-0.15]	[0.34]	[0.45]
Special Education	-0.01	-0.01	-0.04	0.02	0.02
	(0.68)	(0.72)	(0.44)	(0.30)	(0.36)
	[0.16]	[0.25]	[0.21]	[0.03]	[0.03]

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

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets.

Table C.5: Tests for compositional change: School and SES.
Summary of difference in differences estimates

Outcome:		**Low-ability**		**High-ability**	
	All Sample	**Low-SES**	**High-SES**	**Low-SES**	**High-SES**
Switched Class	-0.00	0.01	-0.02	0.01	-0.00
	(0.98)	(0.75)	(0.41)	(0.78)	(1.00)
	[0.12]	[0.09]	[0.14]	[0.07]	[0.12]
Changes of teacher	0.03	0.04	0.14	0.03	0.09
	(0.80)	(0.75)	(0.37)	(0.81)	(0.54)
	[0.55]	[0.59]	[0.53]	[0.52]	[0.46]
Hours absent (s.y. 6)	-0.79	-2.65	2.26	-0.16	2.13
	(0.86)	(0.55)	(0.59)	(1.00)	(0.65)
	[26.19]	[26.16]	[26.40]	[24.15]	[25.98]
Class size	0.44	0.16	1.08	0.24	0.60
	(0.50)	(0.89)	(0.31)	(0.71)	(0.40)
	[21.39]	[21.14]	[21.54]	[21.62]	[21.58]
Parents not Nordic	0.00	-0.02**	0.02	0.00	0.01
	(0.86)	(0.03)	(0.11)	(0.90)	(0.44)
	[0.02]	[0.02]	[0.02]	[0.01]	[0.01]
Quiet home envir.	-0.00	-0.02	-0.00	0.01	0.02
	(0.75)	(0.46)	(0.87)	(0.35)	(0.27)
	[0.95]	[0.93]	[0.95]	[0.96]	[0.96]
Family income	0.42	0.30	0.26	0.30	-0.02
	(0.36)	(0.46)	(0.66)	(0.68)	(0.98)
	[272.37]	[245.29]	[297.31]	[248.11]	[316.44]

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

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets. Family income measured when the child is 11-16.

Table C.6: Tests for compositional change: Mother.
Summary of difference in differences estimates

Outcome:		**Low-ability**		**High-ability**	
	All Sample	**Low-SES**	**High-SES**	**Low-SES**	**High-SES**
Not Nordic	0.01	-0.01	0.01	0.00	0.02
	(0.43)	(0.21)	(0.49)	(0.87)	(0.11)
	[0.05]	[0.05]	[0.05]	[0.03]	[0.04]
Married	-0.01	-0.01	0.03	0.00	-0.04
	(0.52)	(0.54)	(0.27)	(0.95)	(0.20)
	[0.74]	[0.72]	[0.75]	[0.74]	[0.79]
Divorced	-0.01	-0.02	-0.04*	-0.03**	0.01
	(0.44)	(0.40)	(0.07)	(0.04)	(0.78)
	[0.18]	[0.17]	[0.19]	[0.15]	[0.17]
Compulsory or less	-0.03	-0.01	-0.07***	-0.02	-0.07***
	(0.14)	(0.76)	(0.00)	(0.62)	(0.00)
	[0.35]	[0.53]	[0.16]	[0.48]	[0.11]
High school	-0.01	-0.02	0.04*	-0.02	0.04
	(0.82)	(0.64)	(0.10)	(0.61)	(0.42)
	[0.44]	[0.47]	[0.43]	[0.52]	[0.36]
College or more	0.04*	0.03**	0.03	0.03*	0.02
	(0.06)	(0.05)	(0.28)	(0.05)	(0.69)
	[0.21]	[0.00]	[0.41]	[0.00]	[0.53]
In the labor force	-0.00	0.00	0.02	0.01	-0.03
	(0.89)	(0.91)	(0.19)	(0.65)	(0.34)
	[0.89]	[0.86]	[0.93]	[0.88]	[0.95]
Unemployed	0.00	0.01	0.01	-0.00	0.00
	(0.47)	(0.26)	(0.37)	(0.99)	(0.81)
	[0.05]	[0.06]	[0.04]	[0.05]	[0.04]

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

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets.

Table C.7: Tests for Compositional Change: Father.
Summary of difference in difference estimates

Outcome:		**Low-ability**		**High-ability**	
	All Sample	**Low-SES**	**High-SES**	**Low-SES**	**High-SES**
Not Nordic	0.01	-0.01	0.02*	0.00	0.01
	(0.53)	(0.32)	(0.10)	(0.69)	(0.37)
	[0.03]	[0.04]	[0.03]	[0.02]	[0.03]
Married	0.00	-0.00	0.00	-0.01	-0.02
	(0.78)	(0.98)	(0.88)	(0.53)	(0.30)
	[0.77]	[0.75]	[0.78]	[0.79]	[0.82]
Divorced	-0.01	0.01	-0.02	-0.00	-0.00
	(0.57)	(0.74)	(0.16)	(0.91)	(0.85)
	[0.17]	[0.17]	[0.17]	[0.14]	[0.15]
Compulsory or less	-0.01	-0.01	-0.01	0.03	-0.05**
	(0.32)	(0.64)	(0.78)	(0.39)	(0.03)
	[0.41]	[0.67]	[0.13]	[0.63]	[0.09]
High school	-0.02	-0.02	0.01	-0.06*	0.02
	(0.20)	(0.50)	(0.67)	(0.09)	(0.57)
	[0.40]	[0.33]	[0.51]	[0.37]	[0.44]
College or more	0.04***	0.03**	-0.00	0.03**	0.03
	(0.00)	(0.04)	(0.87)	(0.04)	(0.22)
	[0.19]	[0.00]	[0.36]	[0.00]	[0.47]
In the labor force	-0.01	0.01	-0.02	-0.03	-0.02*
	(0.24)	(0.70)	(0.33)	(0.32)	(0.06)
	[0.88]	[0.85]	[0.91]	[0.88]	[0.93]
Unemployed	0.00	-0.00	-0.01	0.02	0.04**
	(0.60)	(0.91)	(0.32)	(0.19)	(0.01)
	[0.04]	[0.05]	[0.04]	[0.04]	[0.04]

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

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets.

Table C.8: Tests for Compositional Change:
Best of parent occupation.
Summary of difference in difference estimates

	Diff-in-Diff (p-value)
Parents: non-skilled workers, goods	-0.01 (0.69)
Parents: non-skilled workers, service	0.01 (0.46)
Parents: skilled workers, goods	-0.06*** (0.00)
Parents: skilled workers, service	0.01 (0.17)
Parents: lower non-manual ii	0.00 (0.96)
Parents: lower non-manual i	0.03*** (0.00)
Parents: intermediate-level non-manual	0.02 (0.45)
Parents: higher civil servants and senior salaried	0.02 (0.21)
Parents: independent professionals	-0.00* (0.09)
Parents: entrepreneur	-0.01 (0.42)
Parents: farmer	-0.00 (0.60)

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

Standard errors clustered at the municipal level. Occupation variables are taken from Census 1980.