European Economic Review 187 (2026) 105372

Contents lists available at ScienceDirect

European Economic Review

journal homepage: www.elsevier.com/locate/eer

School quality beyond test scores: The role of schools in shaping educational outcomes[☆]

Annalisa Loviglio *

Department of Economics, University of Bologna, Italy
IZA@LISER, Luxembourg

ARTICLE INFO

JEL classification:
I20
J24
C35

Keywords:
School value-added
Educational choices
Attainment
Retention

ABSTRACT

I study how schools affect student performance and attainment throughout secondary education. I develop a dynamic structural model of cognitive skill accumulation and educational decisions in which students have imperfect information about their ability and learn about it from academic evaluations. The estimation relies on administrative data covering the universe of students in public middle schools in Barcelona (Spain).

I use the estimated model to simulate the educational outcomes that each student would achieve in every school in the sample. The results show that school quality cannot be adequately summarized by any single ranking: students with differing characteristics and ability levels benefit from different school inputs, and the school environment is particularly important for the educational attainment of students from disadvantaged backgrounds or with lower ability. In particular, the schools that would maximize these students' final test scores are often not those that maximize their probability of graduating or pursuing further education. These findings suggest that evaluating schools using only standardized assessments may fail to identify the schools that best support disadvantaged students.

I then evaluate a counterfactual policy requiring students to remain in school through the final grade and take the final assessment rather than only until they turn sixteen. The policy substantially increases graduation rates, especially among disadvantaged students, while having only modest effects on the skill composition of high school entrants.

1. Introduction

Higher educational attainment is associated with better labor market outcomes, as well as improved health and life satisfaction.[1] Furthermore, socioeconomic background is often the main determinant of educational attainment. Therefore, providing inclusive

☆ I am grateful to Caterina Calsamiglia and Joan Llull for their guidance and continuous support during my doctoral studies. I would also like to thank Manuel Arellano, Peter Arcidiacono, Chiara Binelli, Massimilano Bratti, Ana Rute Cardoso, Olivier De Groote, Davide Dragone, Anna Houstecka, Arnaud Maurel, Michele Pellizzari, Lavinia Piemontese, and participants at the Applied Working Group at UAB for their useful comments along different stages of the project. I am grateful to the staff at the Departament d'Ensenyament and IDESCAT, and in particular to Xavier Corominas and Miquel Delgado, for their help in processing the data. I acknowledge financial support from the La Caixa-Severo Ochoa Program for Centers of Excellence in R&D of the Spanish Ministry of Economy and Competitiveness.

* Correspondence to: Department of Economics, University of Bologna, Italy.
E-mail address: annalisa.loviglio@unibo.it.

[1] For the OECD as a whole, the employment rate is 85% for tertiary-educated adults, 76% for adults who have completed high school, and less than 60% for those who have not. Moreover, adults of working age with a tertiary degree earn 54% more than those with only an upper secondary education, while those with less than an upper secondary education earn 22% less (OECD, 2018). Those with high literacy skills and a high level of education are 55% more likely to report being in good health than those with low literacy skills and a low level of education. 92% of tertiary-educated adults were satisfied with their life in 2015, as compared to 83% with lower educational attainment (OECD, 2016).

https://doi.org/10.1016/j.euroecorev.2026.105372
Received 14 May 2025; Received in revised form 4 May 2026; Accepted 5 May 2026
Available online 9 May 2026
0014-2921/

and high quality education that improves outcomes for all students, particularly the most disadvantaged, is a primary goal for policy makers all around the world. To address this issue, many countries have implemented school accountability measures for the purpose of monitoring school quality, deciding on corrective actions, and in some cases assigning funding. A prime example was the "No Child Left Behind" program in the US. In practice, school quality is often measured using nationwide tests, under the assumption that a school's capacity to enhance students' test scores is highly indicative of its ability to improve their overall educational outcomes.[2] In this work, I show that this assumption is questionable. In fact, I show that although school rankings based on standardized assessments are fairly successful in measuring a school's value added in terms of performance, they fail to capture a school's contribution to educational attainment, particularly for students from less favorable socioeconomic backgrounds. Unfortunately, there is no easy solution: other readily available summary measures at the school level, such as the graduation rate, do not adequately capture the school contribution to the attainment of students with differing characteristics.

Specifically, I study cognitive skills development and educational choices of students enrolled in public middle schools in Barcelona (Spain), a setting in which 17% of students do not complete basic education and 34% do not enroll in high school. While this study focuses on a specific city, it addresses a widespread issue in developed countries, using administrative data that is likely accessible to policy makers at local or national levels. According to Eurostat, in 2022 approximately 15% of young adults in the EU are early school leavers, with rates ranging from 5% in Ireland to 24% in Spain, and experience considerably worse labor market outcomes than their peers with higher qualifications.[3] Secondary education attainment is higher in the USA, but many students experience school disruption, leading to GED credentials rather than traditional high school diplomas. While the GED is formally equivalent to a diploma, several studies have questioned this equivalence, showing that it has limited returns in the labor market and may distort social statistics on high school completion rates (e.g. Heckman et al., 2011). This disparity is particularly pronounced among students from lower socioeconomic backgrounds, who are much more likely to experience school disruption.[4]

The existing literature demonstrates that success in life is determined by more than just cognitive skills. Interventions aimed at improving a broader set of skills show impressive long-term returns and contribute to closing gaps due to socioeconomic background (e.g. Cunha et al., 2006; Cunha and Heckman, 2008; Cunha et al., 2010). These results underscore the importance of moving beyond test scores alone in the debate surrounding school quality. Children are left behind not only when they receive low scores on a standardized test, but also when their school environment fails to develop both their cognitive and non-cognitive skills sufficiently and motivate them to pursue further studies. In this respect, secondary education is a crucial stage, because for the first time students can decide whether they want to continue their education or leave. In fact, while in most countries basic education is compulsory, students can legally drop out when they reach a certain age (not necessarily upon completing a particular grade). Moreover, in many European countries, upon completing lower secondary education, students choose whether to enroll in upper secondary academic education, which will provide them with access to university.[5] Students at risk of dropping out may benefit more from attending a lower-ranked school where they feel comfortable and are able to earn a diploma than a higher-ranked one from which they are likely to drop out. Similarly, the decision to continue into upper secondary or tertiary education may depend on previous performance, but also on non-cognitive factors such as a student's motivation or family support. The school environment may play an important role, substantially affecting a student's desire to acquire further education.

Educational trajectories unfold over several periods and involve a sequence of interrelated decisions. Skills accumulate over time, past performance affects future opportunities through retention and graduation outcomes, and students may progressively learn about their own ability from school evaluations, which in turn may affect their educational choices. A dynamic framework is therefore necessary to capture how these elements jointly influence schooling decisions and attainment. To this end, I develop and estimate a dynamic model of cognitive skills accumulation and schooling decisions throughout lower secondary education of students enrolled in heterogeneous schools. In each period, students acquire new skills and decide whether to continue their education or drop out. Upon graduation, they choose whether to enroll in academic upper secondary education, which consists of two additional years of high school that provide access to university. Importantly, before the decision of whether to drop out, a student can be required to repeat a year in order to stay in school, referred to as "retention". Retention may raise skills in the following period, but it also increases the time required to graduate and can therefore alter subsequent educational decisions. Skills growth depends on innate ability, individual characteristics (particularly parental background), and school environment. Students do not observe their own innate ability and therefore have imperfect information about their current skill level; instead, they progressively learn about it from the evaluations they receive in school through Bayesian updating. Their utility from schooling depends on their perceived cognitive skills, individual characteristics, and school environment. Additionally, the model allows the flow utility to vary with retention status.

[2] The No Child Left Behind Act of 2021 (replaced by the Every Student Succeeds Act in 2015) requires public schools to administer an annual statewide standardized test. If a school repeatedly shows poor results, then various steps are taken to improve its performance. Since 1992, the UK has published the so-called "school league tables" which summarize the average GCSE results of state-funded secondary schools. Underperforming schools face various types of sanctions (Leckie and Goldstein, 2017).

[3] Data from "Education and training outcomes", available online at Eurostat. Importantly, only 56% of these early school leavers were employed, as opposed to 77% of high school graduates and 83% of tertiary education holders in the same age-range.

[4] Approximately 12% of high school credentials in a year are issued through The General Educational Development (GED) subject-based test, that can be taken by high school dropouts. According to a survey of the National Center of Education Statistics (Table219.62a in the Digest of Education Statistics), 17% of 9th graders in the lowest quintile of socioeconomic status leave school for at least a month in the following two years, as opposed to 3% of students in the highest quintile.

[5] For instance, students in Italy choose what high school to attend after completing grade 8, in France after completing grade 9, and in Spain after completing grade 10.

The school environment is modeled through a vector of peer characteristics and a vector of school inputs. School inputs are parameters that capture school heterogeneity; they are estimated within the model and can be regarded as school values added along multiple dimensions. First, schools differ in the way they contribute to the accumulation of cognitive skills for given peer quality. Second, they have different grading and retention policies, such that they vary in the probability of retaining students with given cognitive skills and individual characteristics. Third, they directly influence students' educational choices in various ways. The primary advantage of this structural approach is that it facilitates disentangling these channels based on the sequence of students' decisions and test scores. A second advantage is that the model explicitly incorporates imperfect information about ability, making it possible to study how it affects educational outcomes and to assess their importance relative to school environments and institutional constraints.

The analysis is based on administrative data for the universe of students attending public middle schools in Barcelona during the period 2009–2015. As common in many European countries, nationwide tests are administered at the end of primary school and at the end of lower secondary education, but in the latter case several students have dropped out before taking the test. Moreover, given the compulsory education laws in Spain, all students spend at least some time in lower secondary education and are evaluated by their teachers at least once, although grading policies vary across schools. The structure of the model makes it possible to combine the signals provided by these different evaluations, even if they are not directly comparable across schools, and to account for the self-selection of students who take the final standardized test.

The model is estimated using an approach that builds on James (2011) and Arcidiacono et al. (2025). I first estimate the production function of cognitive skills (in particular, the grade equations and the variance of unobserved ability), and individual beliefs over time using an Expectation-Maximization algorithm for computational convenience. I then estimate logit equations for the retention events. Finally, I estimate the parameters that govern the sequence of students' choices using maximum likelihood.

The results show that individual innate ability, parental background, and school environment are all important determinants of cognitive skills. For instance, the effect on skill growth of attending a school in the top quartile rather than in the bottom one is comparable to that of having a highly educated father rather than a low-educated one. Other things being equal, better parental education somewhat decreases the probability of retention and somewhat increases the probability of enrolling in high school. However, most of the effect of parental background on student attainment is due to its large contribution to cognitive skill levels. Conversely, schools have a sizable direct effect on educational choices beyond their indirect effect by way of cognitive skills. For instance, being in a school in the 75th percentile with respect to the distribution of the relevant school input rather than in the 25th percentile increases the utility from continuing one's education as much as an increase in cognitive skills of 0.5 s.d. Schools also vary widely in their retention policies: this is particularly important in view of the estimation results showing that retention negatively affects utility, especially for students with higher skill level.

Using the estimated model, I simulate the educational outcomes that each student in the sample would experience in each of Barcelona's schools. This exercise allows me to compare schools based on the skill levels and educational attainment that the same student would achieve if enrolled there. The results show that school rankings based on cognitive skills are relatively similar across students and closely aligned with rankings based on observed test scores. By contrast, school rankings based on educational attainment — such as graduation or enrollment in academic upper secondary education — vary substantially across students and are only weakly correlated with aggregate measures such as school graduation rates. Differences in attainment across schools are particularly large for students from disadvantaged backgrounds. For instance, the graduation probability of a representative student with average ability and low-educated parents ranges from 55% to 90% across schools, whereas it is always above 90% for a similar student with highly educated parents. Among students with highly educated parents, sizable variation in attainment across schools arises mainly for those with low innate ability. These findings suggest that standard performance-based accountability measures may capture the most relevant differences across schools for advantaged students, but fail to reflect dimensions of school quality that are crucial for the educational attainment of disadvantaged students.

Additional counterfactual simulations assess the role of uncertainty about ability in shaping educational outcomes. Comparing the baseline model with alternative scenarios in which students and school personnel perfectly observe ability shows that uncertainty can affect individual decisions but has a relatively modest aggregate effect on graduation and high school enrollment in the sample. Differences in school environments and institutional constraints play a larger role in explaining educational attainment.

Finally, I use the estimated model to evaluate a counterfactual policy requiring students to remain enrolled in lower secondary education until the final grade, rather than allowing them to leave school once they reach the age of 16. The simulations indicate that many students who currently drop out would successfully pass the final exams if required to remain enrolled. As a result, the graduation rate would increase by about 10 percentage points overall and by 18 percentage points among students with low-educated parents. These findings underscore the importance for policymakers aiming to enhance educational attainment of carefully considering the interplay between existing regulations and student decision-making processes.

Related literature

The paper relates to several strands of the literature, particularly school quality, human capital development, and decision making. School accountability requires reliable measures of quality of individual schools (Kane and Staiger, 2002; Allen and Burgess, 2013; Angrist et al., 2017), or of school types (e.g. charter *vs* traditional public schools, as in Dobbie et al., 2011; Abdulkadiroğlu et al., 2011). This is typically accomplished using a value added approach, i.e. estimating the net effect of attending a given institution on a relevant outcome.[6] Test scores have been the most widely used outcome to capture quality, under the assumptions that performance in school measures cognitive skills and that they are positively correlated with desirable outcomes in subsequent educational stages and in the labor market. The current analysis uses a comparable value added approach, but it explicitly models school inputs along multiple dimensions and studies the effect of attending a school with that combination of inputs on several outcomes. The results underscore the importance of moving beyond performance and rankings to study school quality.

Other strands of the literature have shown that cognitive skills alone do not explain educational choices and attainment that matter for future life outcomes. On the one hand, the literature on human capital development has established that the return on non-cognitive skills is no lower than that on cognitive skills and that the former may not be captured well by using test scores (for one of the seminal works, see Heckman and Rubinstein, 2001). On the other hand, the literature on decision making in Education shows that educational choices vary substantially among individuals with identical prior performances, for a multitude of reasons, ranging from differences in beliefs with respect to the return on each choice to differences in consumption value.[7] These studies focus primarily on individual traits and preferences, and on differences by gender and socioeconomic background.[8] The current analysis contributes to incorporate their findings within the research on the effect of school quality. Indeed, it seems plausible that the school environment, like the home environment, may substantially contribute to non-cognitive skill development, taste formation, and provision of information regarding the returns on education. A structural model of cognitive skills development, retention, and choices makes it possible to first estimate the effect of school environment on cognitive skills, and then to estimate its direct effect on a given educational choice.

The current analysis differs from other contributions (Deming et al., 2014; Angrist et al., 2016) that study the effect of attending a given school on measures of attainment, such as high school graduation, college enrollment, and college persistence. They use those outcomes (rather than test scores) as alternative measures of human capital, to confirm that the gain found by previous studies is not due to a "teaching to the test" attitude, but rather an actual improvement in skills. Their analyses cannot assess whether improvements in graduation rate or college enrollment are due only to the improvement in cognitive skills as measured by standardized tests or to other factors as well. The advantage of my approach is that it allows to disentangle effect of schools on attainment by way of cognitive skills and that by way of other channels, and to examine their interaction. The results are consistent with contemporaneous studies that show that a school's impact on high-stakes tests is weakly related to educational attainment and other important life outcomes, such as crime or teen motherhood (Beuermann et al., 2022), while a school's impact on non-cognitive skills can be important in explaining them (Jackson et al., 2020). Their findings also suggest that relying solely on one-dimensional indicators based on test scores is inadequate for assessing school quality. Beuermann et al. (2022) explore the relationship between parents' preferences and school effects, while my focus lies in understanding the quality of the school system from a policymaker's perspective.

A related literature studies teacher value-added across multiple dimensions of student development. For example, Jackson (2018) and Petek and Pope (2023) estimate teacher value added not only on test scores but also on measures capturing other dimensions of student development, such as behavior and non-cognitive skills, and show that exposure to teachers with high value added along these dimensions affects students' later educational outcomes. Interestingly, they find that value added on test scores is weakly correlated with value added on non-cognitive skills. While these studies focus on teacher-level contributions, my analysis examines how school-level environments — combining multiple educational inputs — shape students' skill formation and educational attainment.

Another related literature studies how parental beliefs and investments interact with school and peer environments in the process of skill formation. For example, Kinsler and Pavan (2021) show that parental beliefs about children's ability are affected by the local school environment, while Agostinelli et al. (2026) develop a dynamic model in which parental investments respond to peer environments, and the two jointly determine the evolution of children's skills over time. In contrast, the present framework does not model parents as a separate decision-making agent. Students' choices can instead be interpreted as reflecting family decisions when parental involvement is present, while parental background is captured in a reduced-form way through observed characteristics such as parental education. This approach allows the analysis to focus on how differences in school environments affect students' skill accumulation and educational choices. It also reflects the use of population-wide administrative records, which make it possible to study school-level environments in the full student population and make the framework potentially applicable to other settings, but do not contain direct measures of parental beliefs or investments.

[6] For a survey, see Angrist et al. (2022). Previous studies also investigated the determinants of school value added. For instance Dobbie et al. (2013) and Angrist et al. (2013) identifies practices such as increased instructional time, high-dosage tutoring, and high expectations which are responsible for the success of some kind of charter schools, while Fryer (2014) shows that some of these practices can be successfully exported to public schools.

[7] Several papers have looked at the choice of college major in the US (for a survey, see Altonji et al. (2015)). Avery et al. (2006) and Hoxby and Avery (2012) study the role of financial constraints and information in the applications to selective colleges of high-achieving low-income students. Arcidiacono (2004) finds that individual preferences for particular majors in college or in the workplace is the main reason for ability sorting. Zafar (2013) and Wiswall and Zafar (2014) find that while expected earnings and perceived ability are a significant determinant of the major chosen, heterogeneous tastes are the dominant factor. On the other hand, Wiswall and Zafar (2015) find that college students are misinformed to a great extent about population earnings and revise their earnings beliefs in response to new information. Kinsler and Pavan (2015), Bordon and Fu (2015), Hastings et al. (2013) exploit Chilean data to study choice of college and major.

[8] Recent examples include Belfield et al. (2020), who collect survey data on students' motives in pursuing upper secondary and tertiary education in the UK. They find that beliefs about future consumption values play a more important role than beliefs about monetary benefits and costs, and differences in the perceived consumption value across gender and socioeconomic groups can account for a sizable proportion of corresponding gaps in students' intentions to pursue further education. Giustinelli and Pavoni (2017) study high school track choice in Italy and find that children from less advantaged families display lower initial perceived knowledge and acquire information at a slower pace.

I find that the school environment is particularly relevant to educational choices and attainment in the case of students with low-educated parents or low prior cognitive skills. Dearden et al. (2011) question the use of a single measure of value added based on performance in order to assess school effectiveness by showing that a school's effect can differ according to prior ability levels. The current analysis shows that even if the performance of every student is affected similarly by a given school environment, its effect on educational attainment may vary according to background and ability.

Finally, the current study contributes to the debate on the effectiveness of grade retention. Even though retention is a common practice in many countries, the empirical literature provides mixed evidence of its effectiveness in improving student performance (Allen et al., 2009; Fruehwirth et al., 2016). The results discussed here suggest that it can improve test scores at the end of middle school (at the cost of longer time spent in education); however, it has a negative effect on a student's consumption value of schooling. The net result is a large increase in drop out rates among retained students and a lower probability of enrollment in high school.[9] Interestingly, the gap is larger for students with a relatively high level of cognitive skills who would not be at risk of retention in more lenient schools. These results are consistent with contemporaneous findings from Portugal in Borghesan et al. (2022): they show that, on average, retention leads to higher test scores, but it also significantly increases dropout.

The remainder of this paper is organized as follows. Section 2 provides background on the Spanish education system, describes the data, and discusses descriptive statistics. Section 3 describes the model and Section 4 summarizes the estimation procedure. Section 5 presents the estimation results. Section 6 discusses simulations and counterfactual analysis, with focus on outcomes of students with low parental background. Section 7 concludes.

2. Data

I employ administrative data on the universe of students who began attending one of 47 public middle school in Barcelona (Spain) in 2009 or in 2010. I exploit various data sources to collect detailed information on enrollment, school progression, performance, and sociodemographic characteristics. This section provides some background on the school system, describes the data sources, and discusses the descriptive statistics.

2.1. The education system

Basic education in Spain is divided into two stages: primary school (corresponding to ISCED level 10, primary education) and middle school (corresponding to ISCED level 24, general lower secondary education). Normally, students spend 6 years in primary education, followed by 4 years in middle school. All students begin primary school in the year in which they turn 6 years old. Retention during primary education is uncommon and, by law, it can occur no more than once. Thus, students usually start middle school in the year in which they turn 12, or at most one year later. In middle school, students can be retained at most twice, but not in the same grade. While repeating a grade during lower secondary education is fairly common, repeating twice is rare. About 20% of students in public schools in Barcelona are retained once, but only 3% are retained twice. Thus, students can graduate in the year in which they turn 16 or later, depending on their retention history.[10] Education is compulsory until age 16, not until completion of lower secondary education. Given that retention is common, there are students who turn 16 well before their potential graduation date.

After successfully completing lower secondary education, students can enroll in high school for two more years (corresponding to ISCED level 34, general upper secondary education). They can also choose vocational training, which, however, does not provide direct access to tertiary education after completion. Alternatively, they can choose not to continue their formal education. Students who do not complete lower secondary education, either because they leave or because they failed too many times, can access some basic vocational training.[11]

About 60% of students attend a public middle school. All public schools are largely homogeneous in infrastructure, curricula, funding per pupil, limits on class size, and teacher assignment.[12] On the other hand, schools have extensive autonomy in deciding how to evaluate student performance and whether to advance them to the next grade.

Families have limited options when choosing a middle school for their children. In fact, every primary school is affiliated with one or more middle schools and students from an affiliated primary school have priority if the middle school is oversubscribed. Other priority criteria such as distance between school and home are also used, if necessary. The structure of the application process creates a major incentive to designate the school for which the student has the highest priority as first choice, because students who are not admitted to their first choice lose their priority for other schools. For instance, 92% of families in Barcelona applied to an affiliated middle school in 2009 and 88% to the closest one. As a result, 93% of families got the school that they ranked first.

[9] Jacob and Lefgren (2009) and Cockx et al. (2017) find that retention has an adverse effect on the probability of graduating from high school (in the US and in Flanders, Belgium, respectively).

[10] For primary education: Decret 142/2007, issued on June, 26 (in DOGC núm. 4915 - 26/6/2007). For secondary education: Decret 143/2007, issued on June, 26 (in DOGC núm. 4915 - 26/6/2007).

[11] I do not observe in the data whether they enroll in vocational training.

[12] Schools in Barcelona have from 1 to 5 classes per grade, depending on the school size. From first to third grade the curriculum is identical for all classes, although they may be taught by different teachers. In fourth grade, all students study core subjects such as Mathematics, Spanish, Catalan, and English. They can also choose to attend a limited number of elective subjects, whose evaluations are not part of this study. Further details on the allocation of students to classes are provided in Appendix G.

While lower and upper secondary education are two separate stages, they are usually located in the same school ("Instituto de Educación Secundaria"). Thus, the principal is the same for both and classes take place on the same premises. Typically, there are different teachers in each, although some might switch between them over time. All public schools offer two main tracks in upper secondary education, namely Science and Humanities, and admittance is guaranteed to all students who graduate middle school. Arts, is offered as a third track in a small number of schools but there is no guaranteed admission in this case. About 4% of the students in Barcelona choose Arts, 43% choose Humanities, and 53% choose Science. About 95% of the students in Science and Humanities stay in the same school for both upper and lower secondary education.

2.2. Variables

The panel used for the analysis consists of 5140 students in 47 middle schools in Barcelona who began their lower secondary education in either September 2009 or September 2010. They are observed from their last year of primary education (2008/2009 or 2009/2010) to their last year of lower secondary education (up to 6 years later). Moreover, for those who successfully graduate, it is observed whether or not they enroll in academic upper secondary education subsequently. Data on school progression and performance and data on socioeconomic background are taken from several administrative data sources, including the Catalan Ministry of Education and the national census. Data sources are described in more detail in Appendix B.1, while the data cleaning is discussed in Appendix B.2. This subsection introduces the variables used in the empirical analysis.

2.2.1. Sociodemographic characteristics

Parental education. Mother's and father's education take three values: Low (at most lower secondary education), Average (upper secondary education), and High (tertiary education). Two dummies are included in order to capture missing information for the mother or the father. In some cases, I use three categories of parental background, which are created by averaging the mother's and father's education: "Low" if both parents have at most lower secondary education, "High" if one of them has tertiary education and the other has at least upper secondary education, and "Average" in the remaining cases.[13]

Other demographic characteristics. The data includes dummies for gender and immigrant, and two variables for the exact age of the student. More specifically, the dummy "retained in primary school" takes a value of 1 if the child turns 13 rather than 12 in the first year of middle school.[14] The other variable is the day of birth, rescaled to the interval 0–1, such that 1 is January 1st and 0 is December 31st. Furthermore, students in the sample belong to one of two cohorts: those who began lower secondary education in 2009, and those who began in 2010.

Application to middle school. The dummy "school not first choice" takes a value of 1 if the middle school that the student attends was not the one ranked first when they submitted their application list. As explained in Section 2.1, the "choice" of middle school is only apparent in Barcelona: the overwhelming majority of households apply for their "safe option" (the closest affiliated middle school) and typically secure it. Nonetheless, I include a control for those who did not adhere to this strategy and ended up in a different school than the one ranked first.[15]

2.2.2. Neighborhood quality

I create an index of neighborhood quality based on average gross income and on proportion of highly educated mothers and fathers in the postal code of the student's home address.[16] This index is the first component of a principal component analysis. It explains 92% of the variance and loads similarly on the three factors.[17] The index is normalized to have mean 0 and standard deviation 1.

[13] If information is missing for either parent, then only the education level of the other is used.

[14] As discussed in Calsamiglia and Loviglio (2020), rules for enrollment in primary education were inflexible for the cohorts under analysis. Virtually all students enrolled in the year in which they turn 6. Therefore it is reasonable to assume that 13-year-old students were retained during primary school.

[15] Households in this category may either have a strong preference for a different school and decide to take the risk, or they may lack understanding of how the system works. I cannot further disentangle their reasons. I conducted the analysis described in the paper including a larger set of controls from the application data (including proxies for distance and other priorities). Given that the results remain unchanged, I prefer to use the more parsimonious approach and include only one dummy variable.

[16] About 96% of the students in the sample live in Barcelona, which is divided into 42 postal code areas. The remaining 4% come from nearby municipalities. Average income at the postal code level is available for large municipalities (Badalona, Hospitalet de Llobregat, and Sabadell) and at the town level for the few smaller municipalities that appear in the sample.

To the best of my knowledge, information on the education of residents is not available at the postal code level. Instead, I use enrollment data in Catalonia to compute the proportions of college educated mothers and fathers among the parents of children aged 6 to 15 living in each postal code area in the relevant years. While these measures are only a proxy for the proportion of college graduates in the overall population, they refer to the group of adults that is likely to be more relevant for teens enrolled in school.

[17] I replicated the analysis introducing the three variables separately. The results are very similar and the coefficients of the other regressors remain virtually unchanged. Therefore, I preferred the more parsimonious specification.

2.2.3. Performance, retention, and choices

External evaluations. Students undergo an externally administered region-wide assessment in Math, Catalan and Spanish at the end of primary education and at the end of lower secondary education. Scores are normalized to have mean 0 and standard deviation 1 in the full sample.[18]

Internal evaluations. End-of-the-year evaluations are carried out by teachers in each subject. The final grade takes in account the results of several tests administered during the year. For comparability with the region-wide exams, I use the average of the grades in Math, Spanish, and Catalan, normalized to have mean 0 and standard deviation 1 in the full sample.

Retention. At the end of the year, students either successfully advance to the next grade or are retained, namely they have to repeat the same grade if they wish to stay in school. In the analysis, dummies for end-of-the-year retention are aggregated in various ways in order to define variables over a longer time span, such as, for example, being retained before a given time period.

Dropout/staying in school. Students are classified as dropout in a given year if they are not recorded in any school in Catalonia, according to an enrollment database covering all schools in the region. Variables for staying in school are defined symmetrically.[19]

Enrollment in high school. Students are considered enrolled in high school, if they are recorded in any high school in Catalonia up to two years after graduation.

2.2.4. School variables

The empirical analysis uses a vector of dummies to identify the school in which a student is enrolled. The data makes it possible to identify a student's class in a given year, and therefore peer variables can be computed at the class level. More specifically, peer variables are defined as the average value among all the peers in the class, including students who do not belong to the sample used in the analysis (e.g., because they were retained and belong to an older cohort).[20]

Two peer variables are used in the main analysis: the proportion of females and an index of peer quality. The latter combines information on the proportion of college graduates among the students' mothers and fathers, the proportion of immigrant students, the average test score at the end of primary education, and the share of students who took that test.[21] The index is the first component of a PCA and explains 73% of the variance. The loadings are very similar in magnitude ranging from 0.42 to 0.47. The proportion of immigrant peers has a negative sign while the others are positive.[22] Finally, the two peer variables are standardized to have mean 0 and standard deviation 1 in the sample.[23]

To estimate the model of educational choices, I also need to define peer variables for counterfactual situations that are not observed in the data. For example, I must define the peers that a retained student who dropped out would have had if they had instead remained in school. To do so, I use the average peer variables observed for classmates who were also retained but did not drop out, and therefore repeated the level and joined the following cohort.[24]

2.3. Descriptive statistics

About 17% of the students in the sample do not graduate middle school. Most of them leave school voluntarily before completing lower secondary education: 8.6% drop out as soon as possible, while the rest remain for one or more additional years after reaching the legal age to drop out. About 66% of the initial pool of students eventually enroll in high school (78% of those who graduated).

As shown in Table 1, there is significant variation in the descriptive statistics across subgroups of the population. Children with low-educated parents have much lower test scores when they start middle school, they are more likely to drop out at 16, and only 70% of them complete lower secondary education (as compared to 95% of children with highly educated parents). Moreover, those who graduate have on average lower test scores (−0.38 s.d. versus 0.59 s.d. for students with highly educated parents) and

[18] I predict evaluations at the end of secondary education for students who did not take them and include the predicted values in the standardization. This approach ensures that the full sample of students is used for both rescaling, and makes average observed scores qualitatively easier to compare across periods (for instance, to check if there is positive sorting of students in taking the second test). A similar approach is used for internal evaluations.

[19] Importantly, a student who change school is observed and not considered a dropout. Only 5% of students in this sample change school; further details are provided in Appendix B.2. Students who move to a different region or country would be missclassified as dropout; I do not address this potential issue in the analysis, as it is likely to be exceedingly minor.

[20] Defining peers at the class level rather than at the school level appears to be preferable because in the Spanish system students in the same class are exposed to the same teachers and the same contents, and are always together in school. Moreover, this allows peer variables to vary both over time and within school. Given the limited number of cohorts this is a desirable feature. Class allocation is further discussed in Appendix G.

[21] The proportion of students for whom I could not retrieve test score is also a proxy for recent immigrants who may have limited knowledge of the local languages and therefore may have been exempted from the test or may have completed primary education in another country.

[22] The correlation between peer quality and proportion of females is virtually 0. When the proportion of females is included in the PCA and two (rotated) components are extracted, the second component is almost identical to the proportion of female.

[23] I replicated the analysis using each of the peer regressors separately. Overall, the results are fully aligned although some parameters are not precisely estimated. Therefore, I prefer the more parsimonious specification.

[24] In most cases, students stay in the same class throughout middle school, except for those who are retained or drop out. Students who are retained repeat the grade and therefore join the following cohort in the next year. In some cases, they may also be assigned to a different class (for example if another class has more free places). Thus, counterfactual peer variables are a weighted average of the possible groups of peers that a retained student would have had in the counterfactual scenario. In the very few cases in which there are no peers in the relevant situation, such as for example if no one in the class repeated the year, I use peers at the school level.

Table 1
Descriptive statistics by subgroups of the population.

	N	%	Test score at entry	final	Drop out at 16	Attainment grad.	high sch.	Peer quality	Neighbor. quality	School 1st choice
ALL	5140	1.000	0.002	0.160	0.086	0.832	0.656	0.000	−0.000	0.925
low parental edu.	1315	0.256	−0.547	−0.383	0.155	0.697	0.430	−0.514	−0.388	0.919
avg parental edu.	2029	0.395	−0.081	0.043	0.097	0.812	0.621	−0.061	−0.046	0.916
high parental edu.	1796	0.349	0.497	0.575	0.022	0.954	0.860	0.445	0.336	0.940
male	2663	0.518	−0.026	0.175	0.096	0.802	0.602	−0.024	−0.001	0.922
female	2477	0.482	0.031	0.144	0.075	0.865	0.713	0.026	0.001	0.929
Spanish	4353	0.847	0.111	0.252	0.064	0.865	0.692	0.112	0.035	0.931
immigrant	787	0.153	−0.606	−0.484	0.207	0.652	0.456	−0.619	−0.196	0.895
regular	3710	0.722	0.277	0.320	0.032	0.968	0.832	0.158	0.084	0.932
retained in primary	417	0.081	−0.858	−0.784	0.230	0.499	0.266	−0.546	−0.166	0.897
retained in grade 1–3	794	0.154	−0.730	−0.493	0.286	0.406	0.140	−0.394	−0.242	0.903
retained in grade 4	219	0.043	−0.380	−0.445	0.000	0.708	0.283	−0.204	−0.236	0.950

Note. The table reports summary statistics for the sample of students used to estimate the structural model described in the paper. It consists of students who enrolled in a public middle school in Barcelona (Spain) in 2009 or in 2010. "Test score at entry" (final) measures results in a region-wide test administered at the end of primary school (middle school), "School 1st choice" takes value 1 if the student is enrolled in the preferred school.

are less likely to pursue further studies. Only 43% of the initial pool of students with low-educated parents enroll in high school, as compared with 86% of students with highly educated parents. Importantly, Fig. 1 shows that educational outcomes vary even among students who have similar performance at the beginning of lower secondary education, as measured by the standardized assessment at the end of primary education ("test score at entry" in the figure). Panel (a) plots the average graduation rate for each decile of test score at entry for students with low-educated parents (blue dots) and those with highly educated parents (red dots). The latter have much higher graduation rate in every decile. For example, the graduation rate in the second decile is already 90% for students with highly educated parents, as opposed to only 70% for students with low-educated parents. Similarly, panel (b) shows that students with highly educated parents have substantially higher odds of continuing into high school in each decile and, if anything, differences are larger among students whose test score at entry is above average. The other panels suggest that there are several channels contributing to these differences. Students with low-educated parents are more likely to fail a grade for any level of test score at entry (panel (c)). For example, more than 20% of students with above average scores at the end of primary education repeat a grade, while the proportion is negligible for those with highly educated parents. Moreover, their test scores diverge during lower secondary education: students with low-educated parents do worse according to both internal and external evaluations (panels (d) and (e), respectively). Interestingly, they are also less likely to enroll in high school after graduation in every decile of final test score besides the first one (panel (f)).

According to Table 1, there is also a large gap between students from an immigrant background and natives: the former show lower performance and have lower probabilities of completing middle school and enrolling in high school. Boys and girls have on average similar performance both at the beginning and at the end of middle school. However, there are salient differences in their educational attainment: boys are more likely to drop out as soon as possible, 20% of them do not complete middle school (as compared to 14% of girls), and only 60% enroll in high school (as compared to 71% of girls). About 93% of students attend the middle school that was ranked first in their application list. The rate is very high among all the subgroups.

Disadvantaged students are more likely to have classmates from a similar background and to live in a neighborhood with a lower socioeconomic status. The index of peer quality is about 0.5 s.d. less for students with low-educated parents, and the index of neighborhood quality is lower by about 0.4 s.d. Meanwhile, boys and girls have similar peers and live in similar neighborhoods.

The last four rows of Table 1 give the descriptive statistics by retention status. Students are grouped into four categories: those who were never retained before leaving middle school; those who were retained in primary school (8% of the sample); those who were retained for the first time in middle school before reaching the last grade (15%); those who were retained for the first time in the last grade (4%). Students who were already behind before turning 16 are the most likely to drop out early, especially if they were retained during secondary education: 30% of them immediately drop out, while only 3% of non-retained students do so. Moreover, less than half of them graduate middle school and very few enroll in high school. Students who repeat the last grade are less likely to graduate and enroll in high school, but they have better odds than students retained at an earlier stage.

Table 2 describes the distribution of incoming students' characteristics and their outcomes by school percentile and shows that schools are quite different both in the types of students they teach and in the outcomes they produce. For instance, for a school at the 75th percentile, incoming students have primary school test scores that are on average 0.5 s.d. better than those in the school at the 25th percentile. Moreover, the interquantile range (IQR) of the proportion of students with highly educated parents is 30 percentage points. The IQR of the final test score is almost 0.6 s.d. and that for enrollment in high school is 21 p.p. On the other hand, in all schools only a small share of students had a different school as their first choice (even the school at the 25th percentile was ranked first in the application form by 90% of its students).

Fig. 2 provides further evidence of the large variation in outcomes across schools. It plots the average test score at the end of lower secondary education on the x-axis and the proportion of students that graduated (left) or the share of students that enroll in high school (right) on the y-axis. Performance and attainment at the school level are positively correlated.[25] Fig. 3 shows the same

[25] The correlation in the left panel is 0.59, while that in the right panel is 0.72 . Weighted correlations by school size are 0.57 and 0.74, respectively.

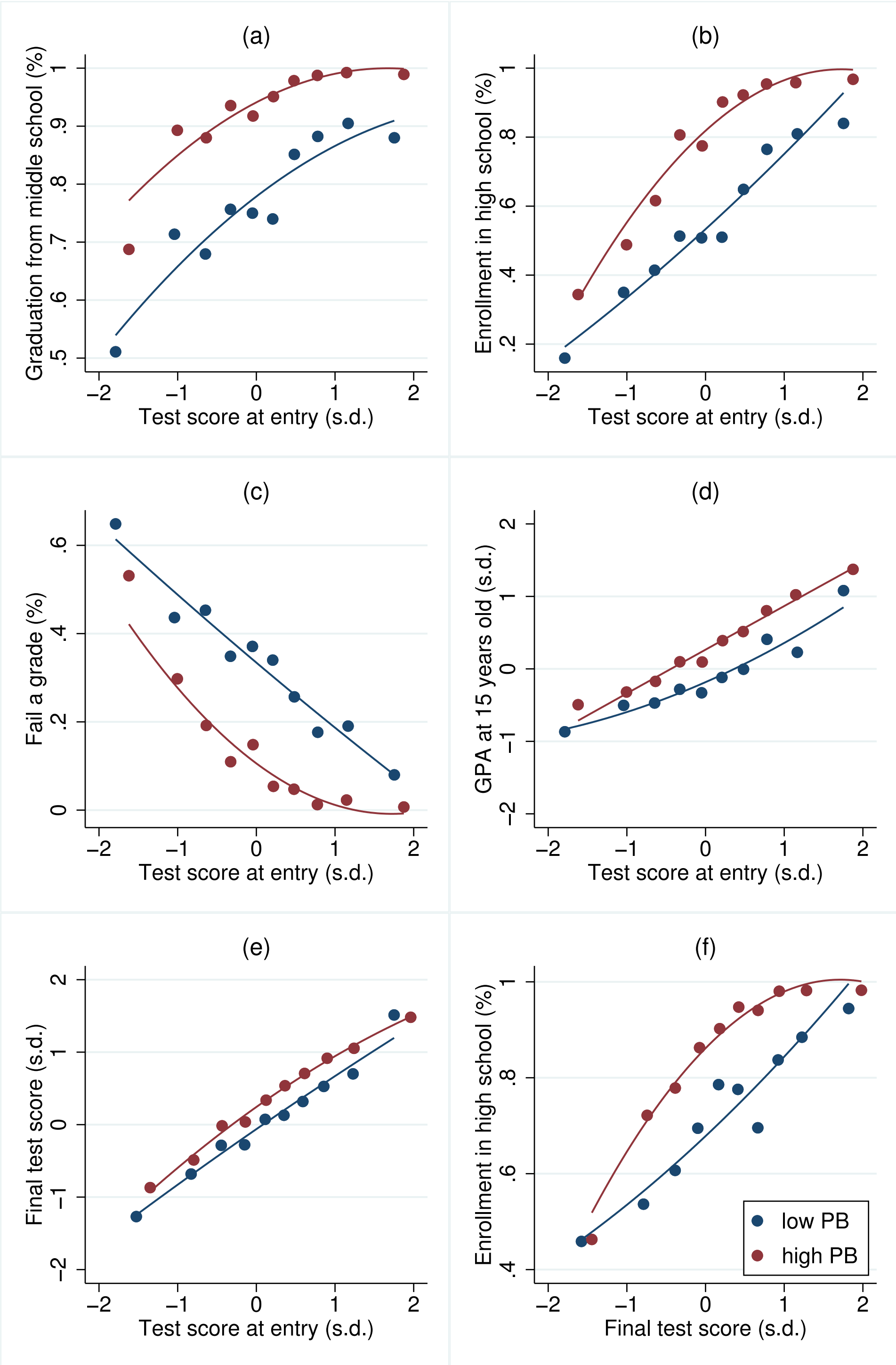

Fig. 1. Educational outcomes by parental background.
Note. Statistics in panels (a) to (d) are computed using all students in the sample. Statistics in panels (e) and (f) use the sub-sample of students who graduate from lower secondary education. In panel (a) to (e) the population is allocated in deciles by test score at the end of primary education; in panel (f) the deciles are computed using the test score in the last grade of lower secondary education. The dots plots the average value in each decile of the variable displayed in the *y*-axis by parental education. The continuous lines are quadratic fits of the data.

Table 2
Descriptive statistics by schools.

	Highly edu. parents	Spanish	Test score at entry	Graduate	High school	Test score final	Neighbor. quality	School 1st choice
p10	0.04	0.65	−0.73	0.68	0.45	−0.58	−1.09	0.72
p25	0.14	0.77	−0.33	0.73	0.52	−0.23	−0.74	0.90
median	0.31	0.82	−0.04	0.81	0.63	0.09	−0.07	0.95
p75	0.44	0.91	0.20	0.89	0.73	0.36	0.52	0.98
p90	0.56	0.95	0.39	0.94	0.81	0.52	0.79	1.00

Note. This table reports summary statistics for the 47 public middle schools in Barcelona which are used to estimate the structural model discussed in this paper.

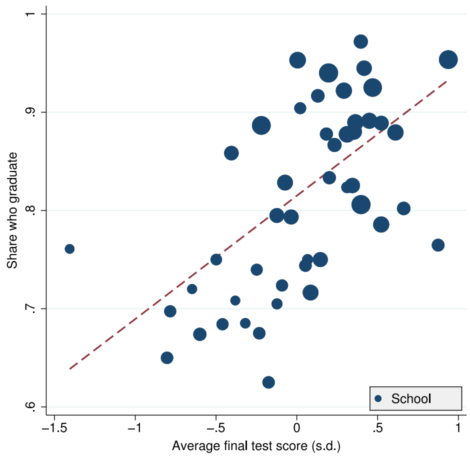

Fig. 2. Educational outcomes by school.
Note. Each dot plots average statistics at the school level. The dot size is proportional to the school size. The dashed line is a linear fit.

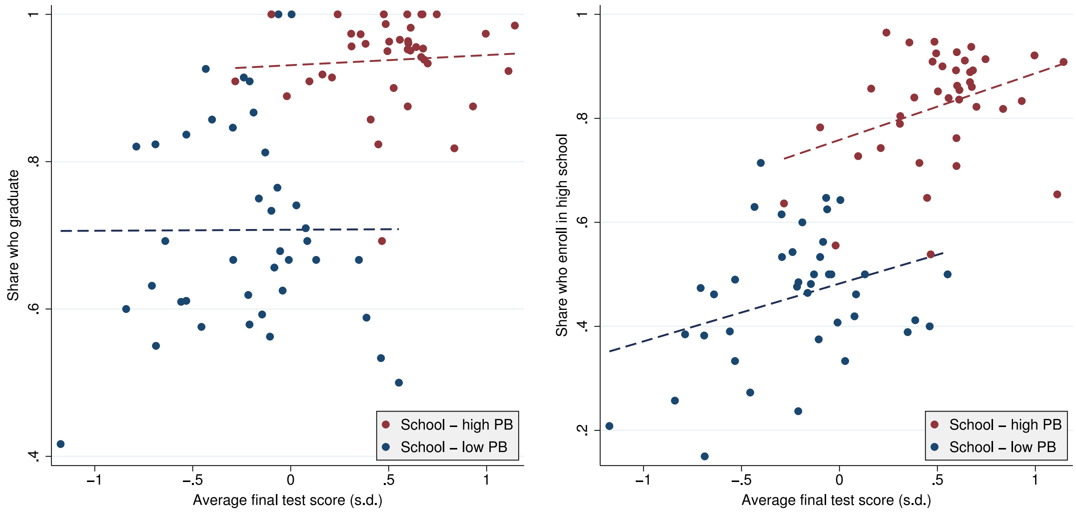

Fig. 3. Educational outcomes by school and parental background.
Note. Each dot plots average statistics at the school level for students with low parental background (blue dot) or high parental background (red dot). Dashed lines are linear fits.

outcomes by parental background. It can be seen that variation in attainment is particularly large for students with low-educated parents. Moreover, within-group correlation between test scores and attainment appear to be lower than the overall correlation. This is confirmed in Fig. A.8 in the appendix, which shows that the correlation decreases considerably once parental education is controlled for. This descriptive evidence suggests that average final test score at the school level is a poor proxy for the educational attainment of students with low parental background.

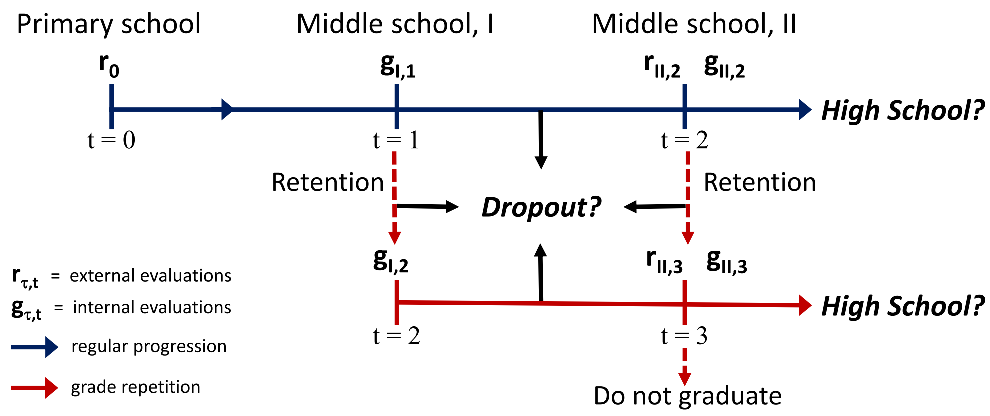

Fig. 4. Timeline of the model.

3. Model

3.1. Overview

A model is constructed to explain cognitive skills development and educational choices among students enrolled in secondary education. The latter include whether to stay in school after the legal age to drop out, and whether to pursue further academic education. While in school, students might fail a grade and have to repeat it, which may affect their incentive to continue, especially because it prolongs the time required to graduate.

Cognitive skill accumulation depends on previous skill level and on contemporaneous inputs: individual characteristics (including ability h which is unknown to both the student and the econometrician) and school environment. While in school, students are evaluated and use the results to infer their level of ability, and corresponding cognitive skills. There are two type of evaluations: 1. standardized test, which are the same across schools; 2. internal grades, which have school-specific components.

Retention is probabilistic and depends on a student's cognitive skills, individual characteristics and school environment. Students are assumed to be forward-looking and make choices that maximize the expected present discounted value of utility. Their flow utility at each point in time depends on what they believe to be their cognitive skills, on their individual characteristics, on the school environment, and on their retention history.

The model is constructed to reflect the Spanish education system with some unavoidable simplifications. The timeline is summarized in Fig. 4, while Table 3 summarizes the most frequently used variables and parameters.

3.1.1. A student's progression in school

Timeline. At time $t = 0$, student i is evaluated by means of a nation wide test ($r_{0,i}$), completes primary school, and begin middle school s. Lower secondary education covers two stages: level I and level II. Each level takes one time period, but students may be retained once, during either level I or level II. In that case, they have to spend an additional period in the same level, unless they drop out.[26]

At time $t = 1$, students finish the first period in school, undergo an internal evaluation ($g_{\mathrm{I},i1}$) and are informed as to whether they can continue on to level II ($\mathrm{fail}_{\mathrm{I},i1}$). Education is no longer compulsory after $t = 1$, and therefore students decide whether to stay in school or drop out:

- Retained students who do not drop out repeat level I. At time $t = 2$, they again undergo an internal evaluations ($g_{\mathrm{I},i2}$) and can advance with certainty to level II.
- Not-retained students who do not drop out continue to level II. At time $t = 2$, they undergo an internal evaluation and a nationwide external evaluation ($g_{\mathrm{II},i2}$ and $r_{\mathrm{II},i2}$). They are also informed as to whether they have successfully completed lower secondary education or have been retained ($\mathrm{fail}_{\mathrm{II},i2}$).

Thus, at time $t = 2$, students who have not successfully completed lower secondary education decide whether to drop out or to stay in level II. If they do not drop out, then at time $t = 3$ they undergo internal and external evaluations ($g_{\mathrm{II},i3}$ and $r_{\mathrm{II},i3}$), and are informed as to whether or not they are graduating ($\mathrm{fail}_{\mathrm{II},i3}$). If they do not graduate at $t = 3$, they must leave school without obtaining a diploma. Students who successfully complete lower secondary education – whether at $t = 2$ or at $t = 3$ – decide whether to continue on to high school.

[26] Given that students do not take any decision in the first years of lower secondary education, the model collapses them in one level. In the estimation, students are retained in level I if they are retained in first, second or third grade of lower secondary education.

Table 3
Most frequently used symbols.

Symbol	Explanation	Level of variation	Source
x_i	vector of time-invariant observed student characteristics	individual	data
ret_{it}	vector with history of retention up to time t	individual	data
$r_{\tau,it}$	region-wide test score in level τ at time t	individual	data
$g_{\tau,it}$	internal evaluation in level τ at time t	individual	data
h_i	innate ability	individual	estimated
$C_{i,\tau t}$	cognitive skills at time t in period τ	individual	estimated
$fail_{\tau,it}$	retention in level τ at time t	individual	data
$p_{ic(i,t)t}$	vector of peers' mean characteristics	class	data
$\mathcal{A}$	school input on cognitive skills	school	estimated
$\mathcal{I}$	school leniency (input on retention)	school	estimated
$\mathcal{M}_M$	school input on choice of staying ("taste" for education)	school	estimated
$\mathcal{M}_H$	school input on choice of high school	school	estimated

Observable student characteristics. Student characteristics x_i (which directly affects skills, retention, and choices) include gender, citizenship, parental education, year and day of birth, cohort, an index of neighborhood quality, and a dummy for not attending one's first-choice school (as described in Section 2.2).

Unobserved ability. Cognitive ability h is an innate trait, randomly distributed in the population according to a normal distribution $\mathcal{N}(0,\sigma)$. Students' cognitive ability h_i is uncorrelated with their individual characteristics x_i. Students do not know their level of ability, but they know the distribution in the population. When they receive new (normally distributed) signals about their level of ability, they update their beliefs and formulate a posterior distribution. Ability is one of the determinants of cognitive skills; intuitively, it captures the portion not explained by individual and family characteristics.

3.1.2. The role of schools

The school environment affects a student's cognitive skills, progression and choices through a vector of peer characteristics $p_{ic(i,t)t}$, where $c(i,t)$ denotes the class attended by student i at time t, and a vector of school inputs $\mathcal{S}$. As described in Section 2.2, when bringing the model to the data, two variables capture peer characteristics: an index of peer quality, based on classmates' previous performance and family background, and the proportion of females. To account for non-linearity in peer effects, a third-degree polynomial in each variable is used. Furthermore, each school is characterized by a vector of four school inputs $\mathcal{S} = (\mathcal{A}, \mathcal{I}, \mathcal{M}_M, \mathcal{M}_A)$, which are estimated using the model and can be interpreted as measures of school value added, above and beyond peer quality in that school. To simplify the notation, the subscripts i and s are omitted from the school inputs.

$\mathcal{A}$ is the school's contribution to cognitive skill development during lower secondary education (as described in Section 3.2). $\mathcal{I}$ is the school grading policy, or in other words the school's level of leniency in grading and deciding whether to retain. It has a direct effect on internal evaluations (see 3.2.3), retention and graduation (see 3.3), and the flow utility from school attendance (see 3.4.1). While the model does not explicitly incorporate a student's effort, the direct effect of $\mathcal{I}$ on utility captures the fact that students may have higher utility in schools in which they need to invest less effort in order to advance to the next grade or achieve a given evaluation score. $\mathcal{M}_M$ and $\mathcal{M}_A$ are the school's inputs that affect the choice of whether to drop out of middle school and whether to continue on to high school, respectively (see 3.4.1). In other words, they are the school's contribution to "tastes" or "motivation" for education, beyond the school's effect on skills and the grading policy in place.

To illustrate, schools with high teacher quality can be thought of as having high $\mathcal{A}$, while schools with lenient grading standards can be thought of as having high $\mathcal{I}$. Inclusive schools that provide students at risk with psychological support to avoid dropping out may have high $\mathcal{M}_M$. Schools hosting orientation events and emphasizing the benefits of further studies may exhibit high $\mathcal{M}_A$.

Importantly, the total contribution of the school environment on the observed outcomes depends on the interplay between its various components. For instance, the school environment directly affects the choice of enrolling in upper secondary education through its effect on tastes, and indirectly, because it affects cognitive skill development (higher skill levels may increase the utility from education) and the student's progression through lower secondary education (by way of retention policy and the choice whether to drop out).

3.2. Cognitive skills formation and evaluations

While in school, a student accumulates cognitive skills. As detailed in the following subsections, these skills are not directly observable to either the student or the econometrician; instead, both only observe noisy measures, namely the yearly evaluations. Upon receiving new evaluations, students update their beliefs about their current skill level.

3.2.1. Skills

The creation of cognitive skills is a cumulative process. Their level in a given time period depends on the level previously achieved and on contemporaneous determinants. Student i starts secondary education with skills $C_{i,0}$, gains $C_{i,\mathrm{I}}$ at the end of the first level, and $C_{i,\mathrm{II}}$ at the end of the second level. When they repeat a level, the most recent attainment of skills replaces what was attained in the previous time period. $C_{i,\tau t}$ denotes the cognitive skills of student i in period τ at time t:

$$C_{i,0} = z_i' b_0 + m_0 h_i \tag{1}$$

$$C_{i,\tau t} = a_\tau C_{i,\tau-1} + z_{it}' b_\tau + m_\tau h_i, \qquad (\tau, t) \in \{(1,1),(1,2),(2,2),(2,3)\} \tag{2}$$

In each period, the contemporaneous determinants are ability $h_i \sim \mathcal{N}(0,\sigma)$ (which is unobserved), and individual and school variables. Since h_i is unknown to the student, C_i is likewise unobserved. At time $t = 0$, its observed determinants are the individual time-invariant characteristics x_i and primary school effects $\mathcal{P}_i$. Starting from $t = 1$, $z_{it}' b_\tau$ is a linear function of x_i, a dummy rep_{it} (which takes a value of 1 if the level is being repeated for the second time), peer variables $p_{ic(i,t)t}$, school input $\mathcal{A}$, and interaction between school input and certain individual characteristics.[27] In practice, interactions between school input and gender, nationality and parental education dummies are included. This specification allows a school to have differential effects on students according to their characteristics while remaining parsimonious with respect the number of parameters to estimate.[28]

3.2.2. External evaluations

The nationwide test score at time t in level τ is an unbiased measure of cognitive skills, that is, an affine transformation plus an exogenous normally distributed error:

$$r_{\tau,it} = o_\tau + \lambda_\tau C_{i,\tau t} + \epsilon_{r_\tau,it}, \qquad \epsilon_{r_\tau,it} \sim \mathcal{N}(0, \rho_{r_\tau}). \tag{4}$$

The nationwide test is administered only at the end of primary education and in level II. Therefore, all students observe:

$$r_{0,i} = C_{i,0} + \epsilon_{r_0,i}, \qquad \epsilon_{r_0,i} \sim \mathcal{N}(0, \rho_{r_0}), \tag{5}$$

and those who stay in school long enough also receive:

$$r_{\mathrm{II},it} = o_{\mathrm{II}} + \lambda_{\mathrm{II}} C_{i,\mathrm{II}t} + \epsilon_{r_{\mathrm{II}},it}, \qquad \epsilon_{r_{\mathrm{II}},it} \sim \mathcal{N}(0, \rho_{r_{\mathrm{II}}}), \tag{6}$$

with $t = 2$ or $t = 3$. Note that in period 0 the parameters (o_0, λ_0) have been normalized to (0,1).

3.2.3. Internal evaluations

At the end of each period in secondary education, students receive an evaluation from their teachers. Given that exams are designed and graded internally, teachers' biases or comparison with peers may affect the assigned score. Moreover, schools may have different grading policies, such that they may design and administer more or less difficult tests and are more or less lenient when grading. In other words, children with the same level of underlying cognitive skills may expect to receive different evaluations depending on their characteristics, peers, and the school in which they are enrolled. Finally, as in the case of nationwide test scores, there is an exogenous normally distributed error:

$$g_{\tau,it} = v_\tau + \mu_\tau C_{i,\tau t} + z_{it}' \gamma_\tau + \epsilon_{g_\tau,it}, \qquad \epsilon_{g_\tau,it} \sim \mathcal{N}(0, \rho_{g_\tau}). \tag{7}$$

Note that in principle all of the contemporary observed determinants of cognitive skills can be a source of discrepancy between internal and external evaluations. In particular, $z_{it}' \gamma_\tau$ include differences in grading policies across schools, namely the school input $\mathcal{I}$: the higher $\mathcal{I}$ is, the more lenient is the school. Conversely, unobserved ability h_i only affects evaluations through skills.

Finally, the error terms of internal and external evaluations may have different variances. This would be the case, for instance, if internal evaluations were more precise (even if "biased") because they average several tests administered during the year.

3.2.4. Final grade equations and parameters to estimate

The scale factors $\lambda_\tau, o_\tau, \mu_\tau, v_\tau$ in the grade equations, the shares α_τ, and the coefficients β_τ cannot be identified separately. Therefore, I will not be able to disentangle the contemporary effect of time-invariant characteristics, but only their cumulative effect. Moreover, skill parameters in level I and level II are identified "up to scale", that is, meaningful comparisons of skill values

[27] With some abuse of notation, I define $z_{it}' b_\tau$ for student i enrolled in s at time t in level τ as follow:

$$z_{it}' b_\tau = \sum_{j=1}^{J} b_{\tau,x_j} x_{ij} + \mathcal{A} + \sum_{j=1}^{J'} b_{\tau,sx_j} (\mathcal{A} x_{ij}) + b_{\tau,rep} rep_{it} + \sum_{k=1}^{K} b_{\tau,p_k} p_{ic(i,t)tk}, \tag{3}$$

where J is the number of individual characteristics, J' is the number of interactions ($J' \le J$), and K is the number of peer characteristics. For clarity, I always use $z_{it}' b_\tau$ or similarly defined terms in the remainder of this section.

[28] A specification that estimates different school inputs according to a student's characteristics would be overly demanding in the current setting.

can be done within a level, but not across levels. Finally, a necessary assumption for identification is that school and teachers' grading policy is constant across levels, i.e. $\gamma_{\text{I}} = \gamma_{\text{II}} = \gamma$.[29]

The final grade equations – which I bring to the data to estimate the parameters – are obtained by substituting the definition of cognitive skills (2) in the evaluation equations defined in (4) and (7), and redefining some of the coefficients:

$$r_{0,i} = z_{i0}'\beta_0 + h_i + \epsilon_{r_0,i} = R_{0,i} + h_i + \epsilon_{r_0,i} \tag{8}$$

$$g_{\text{I},it} = \nu_{\text{I}} + z_{it}'(\beta_{\text{I}} + \gamma) + \kappa_{\text{I}} I_{0,i} + \mu_{\text{I}} h_i + \epsilon_{g_{\text{I}},it} = G_{\text{I},it} + \mu_{\text{I}} h_i + \epsilon_{g_{\text{I}},it} \tag{9}$$

$$r_{\text{II},it} = o_{\text{II}} + z_{it}'\beta_{\text{II}} + \kappa_{\text{II}} I_{\text{I},i} + \lambda_{\text{II}} h_i + \epsilon_{r_{\text{II}},it} = R_{\text{II},it} + \lambda_{\text{II}} h_i + \epsilon_{r_{\text{II}},it} \tag{10}$$

$$g_{\text{II},it} = \nu_{\text{II}} + \mu(z_{it}'\beta_{\text{II}} + \kappa_{\text{II}} I_{\text{I},i}) + z_{it}'\gamma + \mu_{\text{II}} h_i + \epsilon_{g_{\text{II}},it} = G_{\text{II},it} + \mu_{\text{II}} h_i + \epsilon_{g_{\text{II}},it}, \tag{11}$$

where $I_{\tau-1,i}$ is the portion of previous cognitive skills that is due to time-varying observed covariates.[30] The coefficients in β_τ capture the cumulative effects of time invariant regressors, and the innovation of time-varying regressors. Furthermore, to ease the notation I define $\mu = \frac{\mu_{\text{II}}}{\lambda_{\text{II}}}$. $R_{\tau,it}$ and $G_{\tau,it}$ denote grades net of ability and idiosyncratic shock, namely the component of grades that is known *ex ante* to the student.

As a matter of notation, herein I use $g_{\tau,it}$ ($r_{\tau,it}$) for internal (external) evaluations at time t in period τ. I denote g_{it} (r_{it}) to be the evaluation at time t, while abstracting from the level, and $g_{\tau,i}$ ($r_{\tau,i}$) to be the last evaluation in period τ, while abstracting from time.

3.2.5. *Signals and posterior distribution*

I assume that students know the parameters that govern cognitive skills and grading, but do not observe h_i, and therefore they do not know $C_{i,\tau t}$ exactly at any point in time. Given that h_i enters the skill equation linearly and it is the only unobserved determinant, there is a one-to-one correspondence between beliefs about ability and beliefs about skills. Students infer signals on h_i from evaluations. Specifically, the signals are grades "netted out" of the observed components $R_{\tau,it}$ or $G_{\tau,it}$. From Eqs. (8)–(11), it follows that signals are given by:

$$s(r_{\tau,it}) = \frac{1}{\lambda_\tau}\left(r_{\tau,it} - R_{\tau,it}\right) = h_i + \frac{1}{\lambda_\tau}\epsilon_{r_\tau,it} \tag{12}$$

$$s(g_{\tau,it}) = \frac{1}{\mu_\tau}\left(g_{\tau,it} - G_{\tau,it}\right) = h_i + \frac{1}{\mu_\tau}\epsilon_{g_\tau,it}. \tag{13}$$

After receiving one or more signals, students update their beliefs about their ability – and therefore about their level of cognitive skills – using standard Bayesian learning. Given a prior distribution at time t with mean m_{it-1} and variance ω_{it-1}, students compute the posterior mean m_{it} as a weighted average of the prior and the signals, with weights given by the precision of ability's prior distribution and signals' distributions.[31]

To be more concrete, all students observe $r_{0,i}$ before starting middle school. Given that ability h_i has mean 0 and variance σ in the population, the signal $s(r_{0,i})$ has mean 0 and variance $v_{r0} = \sigma + \rho_{r_0}$. Thus, a student i who received grade $r_{0,i} = \hat{r}_{0,i}$ and therefore observes signal $s(r_{0,i}) = \hat{s}(r_{0,i})$ enters middle school with beliefs about their ability characterized by

$$m_{i0} = \frac{\rho_{r_0}^{-1}\hat{s}(r_{0,i})}{\sigma^{-1} + \rho_{r_0}^{-1}} \tag{14}$$

$$\omega_0 = \frac{1}{\sigma^{-1} + \rho_{r_0}^{-1}} \tag{15}$$

At time $t = 1$ the student observes the signal $s(g_{\text{I},i1}) = \hat{s}(g_{\text{I},i1})$. The student assumes that the signal follows a normal distribution with mean m_{i0} and variance $v_{g1} = \omega_0 + \frac{1}{\mu_{\text{I}}^2}\rho_{g_{\text{I}}}$ and update beliefs accordingly:

$$m_{i1} = \frac{\omega_0^{-1} m_{i0} + \mu_{\text{I}}^2 \rho_{g_{\text{I}}}^{-1}\hat{s}(g_{\text{I},i1})}{\omega_0^{-1} + \mu_{\text{I}}^2 \rho_{g_{\text{I}}}^{-1}} \tag{16}$$

$$\omega_1 = \frac{1}{\omega_0^{-1} + \mu_{\text{I}}^2 \rho_{g_{\text{I}}}^{-1}} \tag{17}$$

While all students observe $r_{0,i}$ and $g_{\text{I},i1}$, signals at time $t = 2$ and $t = 3$ depend on their choices and on whether they are retained; in all cases, the updating process follows exactly the same logic. For instance, a student promoted to level II at time $t = 2$ observes

[29] With regard to teachers' grading policy, the assumption is necessary because external evaluations are observed only in level II. The fact that parameters are identified "up to scale" is a consequence of the fact that evaluations are not fully comparable across levels. For instance, getting a 8/10 in first grade, require a different set of skills than getting the same score in the last grade. A student who scores a 8/10 in first grade, would probably fail a test in the last grade, because they have not yet acquired the necessary knowledge.

[30] For instance, using the previous notation, $I_{i,\text{I}} = p_{ic(i,1)1}' b_{p\text{I}} + \alpha_{\text{I}} \mathcal{P}_i$, for a student who did not repeat level I.

[31] See DeGroot (1970) for the general theory of the Bayesian updating. The precision of a normal distribution is defined as the inverse of its variance; using precision rather than variance is convenient because it simplifies the formulas (e.g. Holmström, 1999)

signals $s(r_{\text{II},i2}) = \hat{s}(r_{\text{II},i2})$ and $s(g_{\text{II},i2}) = \hat{s}(g_{\text{II},i2})$ with variance $v_{r2} = \omega_2 + \frac{1}{\lambda_{\text{II}}^2}\rho_{r_{\text{II}}}$ and $v_{g2} = \omega_1 + \frac{1}{\mu_{\text{II}}^2}\rho_{g_{\text{II}}}$ respectively. The posterior distribution of the student's ability is given by[32]:

$$m_{i2} = \frac{\omega_1^{-1} m_{i1} + \lambda_{\text{II}}^2 \rho_{r_{\text{II}}}^{-1} \hat{s}(r_{\text{II},i2}) + \mu_{\text{II}}^2 \rho_{g_{\text{II}}}^{-1} \hat{s}(g_{\text{II},i2})}{\omega_1^{-1} + \lambda_{\text{II}}^2 \rho_{r_{\text{II}}}^{-1} + \mu_{\text{II}}^2 \rho_{g_{\text{II}}}^{-1}} \tag{18}$$

$$\omega_{i2} = \frac{1}{\omega_1^{-1} + \lambda_{\text{II}}^2 \rho_{r_{\text{II}}}^{-1} + \mu_{\text{II}}^2 \rho_{g_{\text{II}}}^{-1}} \tag{19}$$

As a matter of notation, I use $\text{E}_{i,t}(C_\tau)$ to denote a student's belief at time t about their cognitive skills in level τ; similarly $\text{E}_{i,t}(h)$ denotes the belief at time t about innate ability. Moreover, $\psi_{it}(h)$ denotes the posterior distribution after observing signals from time 0 to time t; $\boldsymbol{\psi}_i(h)$ denotes the final posterior distribution using all available signals and $\text{E}_i(h)$ the final posterior about ability.

3.3. Retention and graduation

The retention and graduation events are treated as probabilistic. At time $t = 1$, everyone is at risk of retention: students either fail and can repeat level I or continue on to level II. At time $t = 2$, students who are in level II for a year either fail (and can repeat level II) or graduate. Similarly, at time $t = 3$, students who repeated a level in the past either fail (and leave school) or they graduate.

I assume that the conditional probability takes a logit form:

$$\Pr(\text{fail}_{\tau,it} = 1 | w_{it}) = \frac{\exp(w_{it}'\zeta_\tau)}{1 + \exp(w_{it}'\zeta_\tau)}, \qquad (\tau, t) \in \{(\text{I}, 1), (\text{II}, 2), (\text{II}, 3)\} \tag{20}$$

Students who have already repeated level I advance to the next level with certainty. For ease of notation, I extend the definition to include the case of $\Pr(\text{fail}_{\text{I},i2}) \equiv 0$. Moreover, I will sometimes refer to the "probability of graduation" in level II, defined as $\Pr(\text{grad}_{it}) = 1 - \Pr(\text{fail}_{\text{II},it})$.

Vector w_{it} includes beliefs $\text{E}_{i,t-1}(C_\tau)$, individual characteristics x_i, peer characteristics $p_{ic(i,t)t}$ and the school leniency input $\mathcal{I}$.

This specification accounts for the absence of a deterministic retention criterion, allowing schools to adopt more or less lenient promotion standards. While grades do not mechanically determine retention or graduation, they are strong predictors of both outcomes in the data. As a result, more lenient grading practices reduce the likelihood that students fall below promotion thresholds. I therefore assume that a school's degree of leniency in retention is proportional to its degree of leniency in grading, and include $\mathcal{I}$ as a regressor.[33] Students are assumed to know the parameters and form expectations over their probability of graduation using (20).

While prior beliefs enter the retention probability, $C_{i,\tau}$ or equivalently h_i does not. This assumption appears to be reasonable since the school personnel do not know the student's true h_i either when deciding on retention, though they can form a belief about it by observing the student's performance over time, exactly as the student does.[34]

3.4. Educational choices

3.4.1. Flow utilities

As shown in the flow chart in Fig. 4, students do not make an educational choice at $t = 0$; from $t \geq 1$ onward, in each period they face a binary decision between remaining enrolled and leaving education, which in the final period corresponds to not enrolling in high school.

Students receive a flow payoff for each period they spend in school. The payoff depends on beliefs about the level of their cognitive skills at the beginning of the period, on individual characteristics x_i, the history of retention ret_{it}, peer characteristics $p_{ic(i,t)t}$ and school inputs $\mathcal{I}$ and $\mathcal{M}_\mathcal{M}$ or $\mathcal{M}_\mathcal{A}$ (in the case of lower secondary education or upper secondary education respectively).

The flow payoff per period in lower secondary education for individual i attending level τ at time t is given by:

$$\begin{aligned} U_{it}^M &= \phi_{M,r}\text{E}_{i,t}(C_\tau) + \text{ret}_{M,it}'\theta_{M,r} + x_i'\theta_{M,x} + p_{ic(i,t)t}'\theta_{M,p} + \theta_{M,\mathcal{I}}\mathcal{I} + \mathcal{M}_M + \varepsilon_{it} = \\ &= \phi_{M,r}\text{E}_{i,t}(C_\tau) + y_{it}'\theta_M + \varepsilon_{it}. \end{aligned} \tag{21}$$

The coefficients of the covariates capture all of the motivational and non-cognitive factors that affect the student's choice, in addition to their (perceived) level of cognitive skills. Moreover, the specification allows for differences among students according to retention history. In fact, the vector $\text{ret}_{it}' = (stI2_{it}, stII3_{it}, ftII3_{it})$ includes three mutually exclusive dummies to capture all of the possible combinations of time and retention. The baseline category is students who advance to level II at time $t = 1$. $stI2$ takes a value of 1 if the student failed at $t = 1$ and has to repeat level I at $t = 2$. $stII3$ takes a value of 1 if the student has to repeat level

[32] Note that at $t \leq 1$, ω_t is the same for all students, whereas in subsequent periods the index i is needed because the posterior variance depends on the student's retention status.

[33] In principle, a school-specific measure of leniency in retention could be identified separately from grading leniency. I explored this approach in an earlier version of the paper; however, the two measures turned out to be very highly correlated, leading me to adopt the more parsimonious specification used here.

[34] Furthermore, including $C_{i,\tau}$ in the model would be cumbersome, because students could then learn about their ability through the realization of the event. In fact, $\text{fail}_{\tau,it}$ would be a signal with binary value and a non-normal distribution. As a consequence, the individual posterior distribution $\psi_{i1}(h)$ would not have a normal distribution. I follow (Arcidiacono et al., 2025) in avoiding this complication.

II at $t = 3$. $ftII3$ takes a value of 1 for a student who repeated the first level at $t = 2$ and can advance to level II for the first time at $t = 3$. The value of the coefficient $\phi_{M,r}$ depends on the history of retention and as a result the level of skills may have differing effects on the flow utility, depending on the retention history.[35] For brevity, I will sometimes use $y_{it}'\theta_M$ for the total effect of the inputs other than the skill level.

I assume that $\mathcal{M}_M$ and $\mathcal{I}$ are orthogonal, to be able to separately estimate them from the data.[36]

The flow utility for the choice of whether to enroll in high school has a similar formulation:

$$\begin{aligned} U_{it}^A &= \phi_{A,r}\mathrm{E}_{i,t}(C_{\mathrm{II}}) + \mathrm{ret}_{A,it}'\theta_{A,r} + x_i'\theta_{A,x} + p_{ic(i,t)t}'\theta_{A,p} + \theta_{A,\mathcal{I}}\mathcal{I} + \mathcal{M}_A + \varepsilon_{i,t} \\ &= \phi_{A,r}\mathrm{E}_{i,t}(C_{\mathrm{II}}) + y_{it}'\theta_A + \varepsilon_{it}, \end{aligned} \tag{22}$$

where $\mathrm{ret}_{A,it}' = (rII_{it}, rI_{it})$, with $rII = 1$ if the student repeated the level II, and $rI = 1$ if they repeat level I. As above, the value of $\phi_{A,r}$ depends on the retention status, and $\mathcal{M}_A \perp \mathcal{I}$.[37]

Given the binary nature of the choice in each period, the payoff of the outside option is normalized to 0 and the errors $\varepsilon_{i,t}$ are assumed to be logistic and i.i.d. This normalization implies that the binary choice problem can be written in terms of $\max\{0, U_{it}\}$ rather than explicitly including alternative-specific preference shocks.

3.4.2. Decision making

Students make an educational decision in each period, while taking into account their flow utility and expected future utility. Individuals are assumed to be forward-looking and choose the sequence of actions that yields the highest expected value of utility.

The one-period discount factor is δ. I use u_{it} to denote utility at time t. It is important to recall that students know all of the present and future covariates described in the previous subsection, while signals and shocks to preferences are random variables. Furthermore, the utility of the outside option is normalized to 0 in each period.

At the end of lower secondary education. For those who graduated, the utility of pursuing further education is simply the flow utility in Eq. (22) with $t = 2$ if retention never took place or $t = 3$ if the student was retained in either level I or II. Therefore:

$$u_{it}|(\mathrm{grad}_{it} = 1) = \max\left\{0,\ U_{it}^A\right\} = \max\left\{0,\ v_{it}^A + \varepsilon_{it}\right\}, \tag{23}$$

where v_{it}^A is the utility just before observing the realization of the random shock to preferences ε_{it}.

During lower secondary education. At $t = 2$, those who are still in school but have not yet graduated, again make the choice of whether to drop out, knowing that if they stay in school they will graduate with some probability and may gain access to upper secondary education.

$$\begin{aligned} u_{i2}|(\mathrm{grad}_{i2} = 0) &= \max\left\{0,\ U_{i2}^M + \delta\Pr(\mathrm{grad}_{i3} = 1)\mathrm{E}_{i,2}(u_{i3}|\mathrm{grad}_{i3} = 1)\right\} = \\ &= \max\left\{0,\ v_{i2}^M + \varepsilon_{i2}\right\}. \end{aligned} \tag{24}$$

At $t = 1$, students make their first choice of whether to drop out. They face different problems depending on the level that they will be in if they stay in school. Those who are progressing regularly know that if they stay in school during the next period they will graduate with some probability or the may have to repeat level II. Conversely, those who have to repeat level I anticipate that they will advance to level II with certainty in two periods if they stay, and then graduate with some probability. Thus,

$$\begin{aligned} u_{i1}|(\mathrm{fail}_{I,i1} = 0) = \max\Big\{ & 0,\ U_{i1}^M|(\mathrm{fail}_{I,i1} = 0) + \\ & + \delta\Big(\Pr(\mathrm{grad}_{i2} = 1)\mathrm{E}_{i,1}(u_{i2}|\mathrm{grad}_{i2} = 1) + (1 - \Pr(\mathrm{grad}_{i2} = 1))\mathrm{E}_{i,1}(u_{i2}|\mathrm{grad}_{i2} = 0)\Big)\Big\} \\ & = \max\left\{0,\ v_{i1}^M|(\mathrm{fail}_{I,i1} = 0) + \varepsilon_{i1}\right\}, \end{aligned} \tag{25}$$

$$\begin{aligned} u_{i1}|(\mathrm{fail}_{I,i1} = 1) &= \max\left\{0,\ U_{i1}^M|(\mathrm{fail}_{I,i1} = 1) + \delta\mathrm{E}_{i,1}(u_{i2}|\mathrm{grad}_{i2} = 0)\right\} = \\ &= \max\left\{0,\ v_{i1}^M|(\mathrm{fail}_{I,i1} = 1) + \varepsilon_{i1}\right\}. \end{aligned} \tag{26}$$

3.5. Identification

As is common for this type of dynamic discrete choice models (e.g., Rust, 1987; Arcidiacono et al., 2025), identification of the flow utility parameters relies on the distributional assumptions imposed on the idiosyncratic shocks, the normalization of the outside option, and the discount factor δ, which is set to 0.95.[38]

---

35 In practice, the specification includes interaction terms between $\mathrm{E}_{i,t}(C_\tau)$ and the dummies ret_{it}. Therefore, $\phi_{M,r}$ can take 4 different values.

36 The total school input to utility can be defined as $\mathcal{T}_{\mathcal{M}} = \theta_{M,\mathcal{I}}\mathcal{I} + \mathcal{M}_M$. It measures the overall school contribution to the educational decision beyond the school's effect through skills or retention. $\mathcal{M}_M$ is the component of the contribution that is unrelated to the school's grading policy.

37 This formulation makes the implicit assumption that students take the grading policy $\mathcal{I}$ in middle school as a proxy for the grading policy in high school. It is worth recalling that students usually attend upper secondary education in the same school and under the same principal. Whether $\mathcal{I}$ has a significant effect on utility is an empirical question that will be answered by estimating the coefficient $\theta_{A,\mathcal{I}}$.

38 I replicated the estimation using other values for δ in the interval $[0.9, 1)$ and the results remained virtually unchanged.

The assumption that a student's ability affects educational choices only by way of its effect on cognitive skills is necessary in order to consistently estimate the parameters that govern skills and evaluations. Indeed, evaluations from time $t = 2$ onward are observed only for individuals who choose to continue their education, raising a potential selection issue. However, econometricians can make the same inference as students do when they use the observed evaluations as signals for their own ability. Thus, they can compute the beliefs $\mathrm{E}(h)$ about ability and control for $\mathrm{E}(h)$ in order to consistently estimate the coefficients of the grade equations.[39] On the other hand, the noise of previous test scores, which is used as a signals to infer skills, only affects educational choices, but not skill development nor new test scores. This can be thought of as an exclusion restriction.

It is worthwhile to grasp the intuition of how the variation in the data is used to estimate the parameters of the grade equations. Assume for now that posterior beliefs $\mathrm{E}_i(h)$ have been computed for each student. Consider the following regression for external evaluations in level II at time $t \in \{2, 3\}$, which mimics Eq. (10) in Section 3.2.4, except that time varying variables for previous levels are directly included as regressors:

$$r_{\mathrm{II},it} = o_{\mathrm{II}} + z'_{it}\beta_{\mathrm{II}} + z'_{\mathrm{I},i}\widetilde{\kappa}_{\mathrm{II}} + \lambda_{\mathrm{II}}\mathrm{E}_i(h) + \widetilde{\epsilon}_{r_{\mathrm{II}},it}, \tag{27}$$

where $z_{\mathrm{I},i}\widetilde{\kappa}_{\mathrm{II}} = \kappa_{\mathrm{II}} I_{\mathrm{I},i}$ and $z_{\mathrm{I},i}$ is the vector of time varying regressors for levels I and 0. Furthermore, $\widetilde{\epsilon}_{r_{\mathrm{II}},it} = \lambda_{\mathrm{II}}(h_i - \mathrm{E}_i(h)) + \epsilon_{r_{\mathrm{II}},it}$. Under the assumption that educational choices depend only on the student's belief about their ability, the error $\widetilde{\epsilon}_{r_{\mathrm{II}},it}$ is uncorrelated with the regressors (i.e. $(h_i - \mathrm{E}_i(h))$ is white noise) and therefore OLS can consistently estimate the parameters $o_{\mathrm{II}}, \beta_{\mathrm{II}}, \lambda_{\mathrm{II}}$.

Similarly, OLS can consistently estimate the reduced form parameters of the following regression (based on Eq. (11)):

$$g_{\mathrm{II},it} = \nu_{\mathrm{II}} + z'_{it}(\mu\beta_{\mathrm{II}} + \gamma) + z'_{\mathrm{I},i}\mu\widetilde{\kappa}_{\mathrm{II}} + \mu_{\mathrm{II}}\mathrm{E}_i(h) + \widetilde{\epsilon}_{g_{\mathrm{II}},it}, \tag{28}$$

and using the previous estimates of β_{II} and λ_{II} one can retrieve estimates of μ and γ.

Having estimated γ, one can retrieve the structural parameters β_I and μ_I by OLS:

$$g_{\mathrm{I},it} = \nu_{\mathrm{I}} + z'_{it}(\beta_{\mathrm{I}} + \gamma) + z'_{0,i}\widetilde{\kappa}_{\mathrm{I}} + \mu_{\mathrm{I}}\mathrm{E}_i(h) + \widetilde{\epsilon}_{g_{\mathrm{I}},it}, \tag{29}$$

where $z_{0,i}\widetilde{\kappa}_{\mathrm{I}} = \kappa_{\mathrm{I}} I_{0,i}$ ($z_{0,i}$ are time varying regressors from level 0). Finally, OLS applied to:

$$r_{0,i} = z'_{i0}\beta_0 + h_i + \widetilde{\epsilon}_{r_0,i} \tag{30}$$

makes it possible to consistently estimate β_0 given that there is no selection at time 0. It is then possible to estimate $I_{0,i}$ and $I_{\mathrm{I},i}$ and to retrieve the parameters κ_{I} and κ_{II}.

Up to this point, the identification of the parameters relied on the simplifying assumption that belief $\mathrm{E}_i(h)$ has already been computed. However, in order to perform the Bayesian updating one should know the variance σ of ability h and the variances of the errors in the grade equations ($\rho_{r_0}, \rho_{g_{\mathrm{I}}}, \rho_{g_{\mathrm{II}}}, \rho_{r_{\mathrm{II}}}$). Those parameters are identified from past signals and in particular the covariance between evaluations carried out in the same period or in different periods. In particular, σ can be inferred from the covariance between the residuals of $g_{\mathrm{II},it}$ and those of $r_{\mathrm{II},it}$, which are obtained by regressing the evaluation scores on the observable covariates. The variance of each type of residual is a linear function of σ and the variance of the relevant error, and therefore the latter can be retrieved after estimating σ.[40] Therefore, in practice, the econometrician must estimate both the parameters of the grade equations and the variances of ability and signals in a way that accounts for their mutual dependence, since posterior beliefs depend on the variance of ability and signal noise.

3.6. Modeling choices and limitations

The model abstracts from permanent unobserved heterogeneity that may be observed by students but unobserved by the econometrician. Persistent unobserved traits — such as motivation, preferences for schooling, or non-cognitive skills — may in principle affect both academic performance and educational choices. Appendix C illustrates these issues using a highly simplified version of the model. Here, the main implications are briefly summarized.

First, when permanent unobserved heterogeneity is present and known to students, students can net it out from observed evaluations when forming beliefs about their skills. In contrast, an econometrician who ignores such heterogeneity conflates ability with permanent unobserved traits when inferring skills from observed outcomes. As shown in Appendix C, this leads the econometrician to overstate the precision of the signals and to understate posterior uncertainty. As a result, inferred beliefs about skills are less dispersed and more tightly concentrated around the population mean than the beliefs held by students.

Second, if permanent unobserved traits affect the utility from schooling in addition to their effect on cognitive skills, they enter the econometrician's utility specification as part of the unobserved error term. Because these traits are also reflected in inferred beliefs about skills, the resulting error term is correlated with beliefs about skills. In this case, the estimated contribution of cognitive skills to utility may partly capture the effect of omitted permanent traits.

Third, if permanent unobserved heterogeneity is correlated with observed covariates, estimated coefficients in both grade and choice equations reflect this correlation. For instance, if students from more advantaged backgrounds derive higher enjoyment from schooling, this may be absorbed by the estimated coefficients on parental education in the choice equations. More generally,

[39] Intuitively, $\mathrm{E}(h)$ is a good "proxy" for unobserved ability h because $h - \mathrm{E}(h)$ is purely white noise, and therefore uncorrelated with the other covariates.

[40] More precisely, $\mathrm{Cov}(g_{\mathrm{II},it}, r_{\mathrm{II},it} | z_{it}, z_{\mathrm{I},i}) = \mu\lambda_{\mathrm{II}}^2\sigma$. For instance, $\mathrm{Var}(r_{\mathrm{II},it}) = \lambda_{\mathrm{II}}^2\sigma + \rho^r_{\mathrm{II}}$.

the socio-demographic characteristics included in the model can be interpreted as a "reduced-form" way of summarizing the effects of persistent traits correlated with observed characteristics. Of particular relevance for this paper, if permanent unobserved heterogeneity is correlated with middle school assignment, estimated school-level effects may reflect both causal school inputs and selection on persistent unobserved characteristics. This concern rests on the same identifying assumption adopted for ability, namely that both ability and other persistent unobserved traits are uncorrelated with school assignment conditional on observed covariates. In practice, the estimation includes a rich set of individual characteristics and primary school fixed effects, which should mitigate potential selection concerns.

In related work, such as (Arcidiacono et al., 2025), permanent unobserved heterogeneity is explicitly incorporated and identified through an auxiliary measurement system, which relies on additional information beyond administrative records, including tests and survey-based measures of schooling preferences, behaviors, and expectations. In the present setting, the available data consist exclusively of administrative records covering the universe of students, such as grades, test scores, retention outcomes, and enrollment decisions, and do not include independent measurements that would allow one to separately identify persistent unobserved heterogeneity of this kind. Importantly, the use of population-wide administrative data is central to the analysis, as the model is designed to study school-level effects, which would be difficult to assess using survey data that typically cover only selected samples of students. The resulting framework reflects a trade-off between incorporating rich auxiliary measurements and exploiting comprehensive administrative data at the school level; while the model is estimated in a specific institutional setting, its core structure relies only on standard administrative inputs that are routinely available across education systems, and could in principle be adapted to other contexts without requiring setting-specific survey data.

4. Estimation

This section summarizes the main steps of the estimation procedure. Further details are provided in Appendix D. Define $d_i = (d_{it})_t$ ($t \in \{1,2,3\}$) to be the vector of choices of student i, $\text{fail}_i = (\text{fail}_{it})_t$ to be the vector of retention/graduation events, and $o_i = (o_{it})_t$ to be the vector of evaluations observed by i.[41] The student makes $T_d \in \{1,2,3\}$ decisions, receives $T_{\text{f}} \in \{1,2,3\}$ notifications of retention/graduation, and observes signals in $T_d + 1$ periods. More specifically, they receive T_d internal evaluations and $T_r \geq 1$ external evaluations. For instance, consider a student who is retained in level I, stays in school one more period and then drops out. They make two choices, i.e. $d_i = (1,0)$. At $t = 1$ they are retained, and at $t = 2$ they are promoted to level II, such that $\text{fail}_i = (1,0)$. They are evaluated externally at the end of primary school, and internally in level I at $t = 1$ and $t = 2$, i.e. $o_i = (r_{0,i}, g_{\text{I},i1}, g_{\text{I},i2})$.

Recall that $\boldsymbol{\phi}$ is the pdf of ability $h \sim \mathcal{N}(0,\sigma)$. Omitting the dependence on observable characteristics for ease of notation, the individual likelihood is given by:

$$L_i = L(d_i, \text{fail}_i, o_i) = L(d_{i1}, \dots, d_{iT_d}, \text{fail}_{i1}, \dots, \text{fail}_{iT_{\text{f}}}, g_{i1}, \dots, g_{iT_d}, r_{i0}, \dots, r_{iT_r}) \tag{31}$$

Moreover, $L(d_i, \text{fail}_i, o_i) = \int L(d_i, \text{fail}_i, o_i|h)\boldsymbol{\phi}(h)dh$ and therefore:

$$\begin{aligned} L_i = & \int L(r_{i0}|h)L(\text{fail}_{i1}|h, r_{i0})L(g_{i1}|h, r_{i0})L(d_{i1}|h, r_{i0}, g_{i1}) \cdots L(d_{iT_d}|h, o_i, d_{i1}, \dots, d_{it_d-1})\, \boldsymbol{\phi}(h)dh = \\ & \Big(L(d_{i1}|r_{i0}, g_{i1}) \cdots L(d_{iT_d}|o_i, d_{i1}, \dots, d_{iT_d-1})\Big) \times \Big(L(\text{fail}_{i1}|r_{i0}) \cdots L(\text{fail}_{iT_{\text{f}}}|o_i, d_{i1}, \dots, d_{iT_d-1})\Big) \times \\ & \times \int L(o_{iT_d}|h, d_{i1}, \dots, r_{i0}, \dots) \cdots L(r_{i0}|h)\, \boldsymbol{\phi}(h)dh \end{aligned} \tag{32}$$

where the second equality follows from the fact that choices and retention/graduation probabilities depend on h only through students' beliefs, i.e. through the signals inferred from the evaluations. Thus, the loglikelihood is separable into three parts (choices, retention probabilities, and evaluations) that can be estimated sequentially:

$$\log L_i = \log L_{i,d} + \log L_{i,\text{fail}} + \log L_{i,o} \tag{33}$$

Estimation of $\log L_{i,o}$. Maximizing the likelihood $\log L_{i,o}$ would be computationally costly due to the integration of h. Following James (2011) and Arcidiacono et al. (2025), I use an Expectation-Maximization (EM) algorithm to overcome this issue. In a nutshell, it iterates two steps – the E-step and the M-step – until convergence. In the E-step of the kth iteration, the observed evaluation scores and the parameters estimated in the previous $k-1$ iteration are used to compute the signals, estimate a posterior distribution $\boldsymbol{\psi}_i^k(h)$ of each student's ability and update the estimate of the variance σ^k. The M-step maximizes the log-likelihood computed using individual posterior distributions and estimates a new vector of parameters. Once the parameters have been estimated, I can compute beliefs about cognitive skills for each student at any point in time, and then use them as covariates in the subsequent stages.

Estimation of $\log L_{i,fail}$. I estimate the logit model for retention using the beliefs about cognitive skills and the previously found school leniency. Using the estimated parameters, I then compute the probability of repeating level I at time $t = 1$ and the probability of graduation in the subsequent time periods. These probabilities are then used in the next step.

[41] o_{it} is a vector containing one evaluation at $t = 0$ and in level I, and up to two evaluations in level II. Recall that I use $g_{\tau,it}$ ($r_{\tau,it}$) for internal (external) evaluation at time t in level τ. I denote g_{it} (r_{it}) as the evaluation at time t, while abstracting from the level, and $g_{\tau,i}$ ($r_{\tau,i}$) as the last evaluation in level τ, while abstracting from time.

Estimation of $\log L_{i,d}$. Given the beliefs and probabilities computed in the previous steps, I maximize the total log-likelihood in order to estimate the parameters, following Rust (1987). Because the choice is binary in each period and the outside option is normalized to 0, the idiosyncratic shocks to preferences are standard logistic. As a result, the likelihood contribution of educational choices, $\log L_{i,d}$, is based on logit choice probabilities; however, it does not reduce to a simple product of logit likelihoods because expected utilities depend on beliefs that evolve over time. In fact, students use their beliefs about their skills in the flow utility, since they anticipate that if they stay in school they will receive new signals and therefore modify their beliefs. Thus, the computation of their expected utility for a given choice at time t requires integrating over all the signals that they may receive from $t+1$ onward.[42]

Standard errors. Standard errors are estimated using a bootstrap procedure with 200 replications. Let N_s be the number of students in the sample enrolled in school s. In each replication, and for each school s, N_s individuals are sampled with replacements.[43] An alternative approach would be to not impose that the number of students in the school remain constant across iterations, and would simply draw with replacement N students from the total sample at each iteration. I estimated standard errors using this alternative approach as well and results were extremely similar (in particular the significance of the coefficients remained unchanged throughout).

5. Empirical results

This section summarizes the main findings regarding the estimated parameters, which will be used in the subsequent section to conduct simulations and assess the variation in student outcomes across schools. Appendix E provides a more detailed discussion of the results.[44]

Cognitive skills. About half of the total variance of cognitive skills can be attributed to unobserved ability, while the other half stems from observed individual characteristics and school environment (Table A-10 in Appendix E). Evaluations are informative, leading students to acquire accurate beliefs about their ability, as evidenced by the rapid reduction in posterior individual variance over time. In fact it diminishes from 0.17 at time $t = 0$, to 0.07 at time $t = 1$, and further to 0.03 at time $t = 2$. This suggests that students' choices in the later stages would likely remain unchanged even if they were to directly observe their exact skill level rather than relying on their beliefs. However, uncertainty may still influence decisions in earlier periods, when beliefs are less precise.[45] Among the observed determinants, parental education is the most important contributor to cognitive skill accumulation. Having both parents with tertiary degrees, as opposed to only basic education, is associated with an improvement in skills by approximately 1 s.d. (Table A-11 in Appendix E). The school environment also plays a crucial role: having higher-quality peers increases skill level, and there is substantial variation in school inputs. Specifically, the difference between a "good" school at the 75th percentile and a "bad" school at the 25th percentile (i.e., the interquantile range) amounts to almost 0.4 s.d., a magnitude consistent with previous results in the literature (e.g. Angrist et al., 2017; Abdulkadiroğlu et al., 2011). Although the model allows for interactions between school inputs and individual characteristics, they appear to be unimportant: school effects on skills are roughly homogeneous across individual characteristics.

Educational choices. Cognitive skills – or, more precisely, beliefs about them – are an important determinant of the choices made (Table A-13 in Appendix E). *Ceteris paribus*, girls are less likely than boys to drop out and more likely to enroll in high school after graduation. Students with a highly educated father are more likely to choose to enroll in high school; however, overall, the direct effects of parental background appear less relevant than the indirect effects through skills. Conversely, the school environment also has a substantial direct impact. For example, attending a school at the 75th percentile of the distribution of school inputs, rather than the 25th, affects the decision to pursue further education almost as much as having an additional 0.5 s.d. in cognitive skill level. Additionally, peer quality is associated with a modest decline in the flow utility derived from staying in middle school, perhaps due to ranking concerns.

Retention. Unsurprisingly, higher cognitive skills decrease the probability of retention (Table A-12 in Appendix E). However, other factors also contribute to it. Girls have a lower probability of retention regardless of skill level. Moreover, having highly educated parents, especially a father, is also associated with a lower probability. Higher quality peers increase the probability of retention at any given level of skills, consistent with teachers "grading on a curve".[46] Furthermore, students in more lenient schools are less likely to be retained and the difference is large: increasing school leniency by 1 s.d. lowers the probability of retention by as much as an increase of 0.3 s.d. in cognitive skills. Repeating a level for a second time has a positive and sizable effect on cognitive skills (up to 0.5 s.d. in level II, as shown in Table A-11), thus having a positive indirect effect on educational choices. However, retention also has a substantial negative direct effect on educational choices, particularly enrollment in high school (Table A-13 in Appendix E). More specifically, retained students exhibit a flatter flow utility in cognitive skill level, resulting in a larger gap between retained

42 This is the most computationally costly part of the maximum likelihood estimation and is performed using Gauss–Hermite quadrature.

43 For two schools, in very few iterations the random sample did not includes any dropouts. In those instances, I resampled again the observations for the school. This happened in less than 4% of the iterations for one of the school and in less than 0.4% for the other.

44 Specifically, the estimated coefficients are presented in Tables A-10 and A-11 (skills and evaluations), Table A-12 (logit for retention and graduation), and Table A-13 (choices)

45 As explained in Section 3, the posterior individual variance is the updated variance of the student's belief after receiving one or more new signals. To assess the impact of uncertainty about one's ability on educational outcomes, in Section 6.1.3 I analyze a counterfactual scenario where uncertainty is removed.

46 Calsamiglia and Loviglio (2019) documents that in Catalonia teachers' evaluations are deflated in the presence of better-quality peers.

Table 4
Average school input by parental education.

School input		Parental education		Difference
Symbol	Explanation	Low	High	p-value
$\mathcal{A}$	Input on cognitive skills	0.019	−0.032	0.150
$\mathcal{I}$	Input on retention (leniency)	0.014	0.021	0.847
$\mathcal{M}_M$	Input on choice of staying	−0.007	−0.069	0.098
$\mathcal{M}_A$	Input on choice of high school	0.280	−0.228	2.26e−45

Note. Each row refers to a school input, whose symbol and description are reported in the first two columns. The following columns report the average value of each input for students with low-educated parents and for students with highly educated parents, as well as the p-value of a t-test for the difference in means. The estimated school inputs are standardized to have mean 0 and standard deviation 1 in the sample.

and non-retained students at higher levels of cognitive skills compared to lower levels. Regarding the choice of whether to continue on to high school, the value of the flow utility for retained students relative to their outside option is significantly lower than that for non-retained students at any given level of cognitive skills, as illustrated in Fig. A.9 in the appendix. This trade-off between a positive effect on performance and a negative effect on attainment aligns with findings in Borghesan et al. (2022).

School environment and parental background. As discussed earlier, the school environment has a sizable effect on cognitive skills, on the logit function that determines promotion and retention, and on the flow utility of educational choices. The overall effect of the school environment on attainment depends on the interplay between all of these channels. To understand this interplay, in the next Section 1 will simulate the outcomes for various groups of students when exposed to different school environments. In particular, I will focus on differences by parental education. As noted in Section 2.3, students with highly educated parents have on average higher quality peers than students with low-educated parents. However, it is important to recall that all the schools in the sample feature a diverse mix of students with varying parental education levels. Moreover, according to estimation results, students with low and high parental background are, on average, exposed to similar school inputs. As shown in Table 4, the average grading policy $\mathcal{I}$ is nearly identical for the two groups. While the school input on skills $\mathcal{A}$ and the school input on utility from remaining in education $\mathcal{M}_M$ are slightly larger for students with low-educated parents, the differences are small in magnitude.[47] The only exception is the school input on the utility from high school: students with low-educated parents are more likely to attend schools that encourage high school enrollment, resulting in a significantly larger values of $\mathcal{M}_A$ in their case.

Students with high and low parental education can be found in all schools in Barcelona, and those schools vary widely in the inputs they provide. Hence, it is crucial to investigate how the impact of the school environment on student outcomes varies with students' background.

5.1. Fit of the model

To assess the fit of the model, I simulate the choices and outcomes of each individual in the sample, using the structural parameter estimates presented above. More specifically, I create 1000 copies of each individual at time 0 (according to their time invariant characteristics, the primary school they attended, and the middle school they are attending). For each of them, I draw ability and shocks to evaluation scores, preferences, and retention events, using the estimated distributions. I can then compute their outcomes, and in particular their cognitive skills and choices.

Table 5 reports empirical frequencies (in the "data" columns) and frequencies predicted using the model (in the "model" columns) for the following events: the choice to stay in school at time $t = 1$, graduation, enrollment in high school, retention in level I, retention in level II. Frequencies are computed over the entire sample at time $t = 1$. Table A.7 in the appendix reports similar statistics calculated only for the subsample of the initial population who reached the relevant stage for the event to take place (e.g. enrollment in high school is calculated for the subsample of students who completed middle school). The first line of each table presents frequencies for the overall sample, while the subsequent rows present that same information by subgroups (such as parental education, gender, etc.). The predicted choice of whether to stay in school, graduation rate, and choice of whether to enroll in high school are very close to the actual ones, both in the overall sample and by subgroup. The retention rate in level I is also almost identical to the actual one, while the retention rate in level II is slightly higher.

I next investigate how evaluations and events simulated by the model replicate the patterns observed in the data. For instance, Fig. A.10 in the appendix plots the proportion of students who choose to stay in school at $t = 1$ by test score quantile at $t = 0$. Fig. A.11 in the appendix plots their enrollment in high school, again by test score quantile. The model predictions successfully mimic the empirical outcomes. The other evaluations and choices exhibit similar patterns.

[47] They are significant at the 10% only for $\mathcal{M}_M$.

Table 5
Fit of the model.

	Stay at $t = 1$		Graduate		High school		Retained in I		Retained in II	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
ALL	91.44	91.97	83.23	84.21	65.56	65.43	23.56	23.50	4.26	5.84
male	90.42	90.44	80.17	80.76	60.23	59.61	26.96	27.14	4.84	6.62
female	92.53	93.61	86.52	87.91	71.30	71.70	19.90	19.58	3.63	4.99
Spanish	93.64	93.76	86.49	87.27	69.17	68.78	18.91	19.13	4.11	5.64
immigrant	79.29	82.07	65.18	67.28	45.62	46.95	49.30	47.65	5.08	6.92
low parental edu.	84.49	85.04	69.73	70.04	43.04	42.74	43.19	42.32	6.62	8.59
avg parental edu.	90.29	91.41	81.17	83.13	62.10	62.48	25.68	25.60	4.88	6.65
high parental edu.	97.83	97.67	95.43	95.80	85.97	85.39	6.79	7.34	1.84	2.91
below median peers	86.64	88.51	75.88	77.77	53.40	54.36	33.24	32.21	5.07	6.78
above median peers	94.89	94.46	88.50	88.83	74.30	73.38	16.61	17.24	3.68	5.16

Note. Data frequencies are computed from the sample of students used in the estimation. Model frequencies are constructed using 1000 simulations of the structural model for each individual in the estimation sample, where an "individual" is defined by a set of observable initial characteristics.

6. Model implications and counterfactual analyses

This section uses the estimated structural model to study how schools affect students differently and how education policies interact with these mechanisms. I first use the model to characterize heterogeneity in school contributions to students' educational outcomes by simulating the outcomes that students with different characteristics and ability would experience across schools, and by examining how information frictions — arising from imperfect knowledge of ability — affect these outcomes. I then analyze a counterfactual policy that makes education compulsory until the last grade of lower secondary school, highlighting how its effects depend on the interaction between student characteristics, school environment, and information frictions.

6.1. Counterfactual school assignments

The purpose of this section is to characterize how school contributions to student outcomes vary across students. I focus on differences by parental background and ability, as estimated by the model, and on the role played by information frictions.

Using the estimated parameters of the model, I simulate the educational outcomes that a student with given observable characteristics and innate ability would experience if enrolled in each middle school in Barcelona. For each school, I simulate choices and outcomes under repeated draws of preference, evaluations, and retention shocks, and compute expected outcomes by averaging across simulations.[48] I focus on three outcomes: cognitive skills at the end of lower secondary education C_{II}, the ex ante probability of graduating from lower secondary education, and the ex ante probability of enrolling in academic upper secondary education.[49] The first outcome is a measure of performance, while the other two are measures of attainment (for brevity, I will refer to them as "graduation" and "high school enrollment", respectively).

For a given student, these simulations induce a ranking of schools that is specific to that student for each outcome. Moreover, to assess the magnitude of school contributions for each outcome, I summarize the dispersion of simulated outcomes across schools using the interquantile range (IQR), defined as the difference between the 75th and 25th percentiles of the outcome distribution across schools.

6.1.1. Aggregate patterns in school contributions

A natural question is whether such individualized school rankings generated by the structural model provide valuable information, or whether aggregate statistics from the data would be sufficient to approximate outcomes for the majority of students. However, as detailed in Appendix F.1.1, this is not the case. Rankings based on performance outcomes are relatively stable across students in the sample and closely resemble rankings obtained from external evaluations. In contrast, individualized rankings based on attainment outcomes vary substantially across students and are only weakly correlated not only with performance-based rankings, but also with aggregate attainment measures observed in the data, such as school-level graduation or dropout rates. As a result, a one-dimensional ranking based on test scores is insufficient to capture how schools affect students' educational trajectories, and the model is necessary to jointly account for performance and attainment at the individual level.

I then use the model to assess how the magnitude of school contributions varies across students. Across the full sample, cognitive skills at the end of lower secondary education exhibit substantial dispersion across schools, with an average IQR of approximately 0.4 standard deviations. Importantly, the magnitude of this dispersion is similar across parental education groups, indicating that the school environment matters for performance for all students. In contrast, dispersion in attainment outcomes is much larger for students with low-educated parents than for those with highly educated parents, implying that school environment is particularly

[48] Peer characteristics are set at the average value for each school and level.

[49] Skills at the end of lower secondary education are computed using the occurrences in which the student reaches level II. This occurs in at least some iterations for all students in all schools.

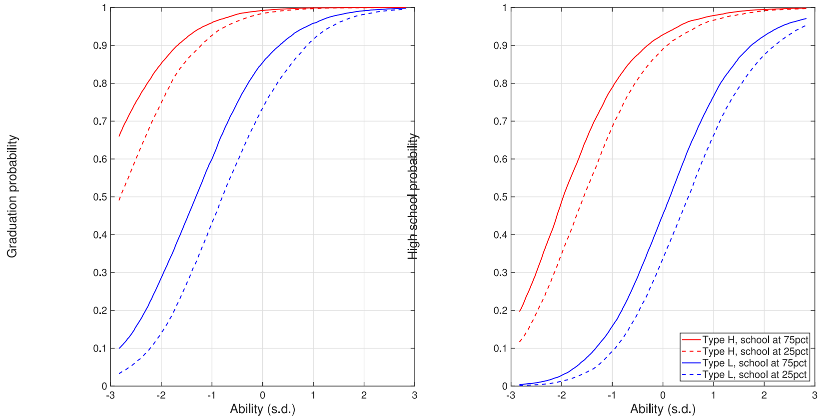

Fig. 5. Attainment of type L and type H students by ability.
Note. The left figure plots the graduation probability (y-axis) by ability level (x-axis) for four representative student types: student with highly educated parents enrolled in a school with high graduation probability (continuous red line), student with highly educated parents enrolled in a school with low graduation probability (dotted red line), student with low-educated parents enrolled in a school with high graduation probability (continuous blue line), student with low-educated parents enrolled in a school with low graduation probability (dotted blue line). The school with high (low) graduation probability is the one at the 75th percentile (25th percentile) of the school ranking by graduation probability. Each probability is the average of 10000 simulations for a given student type at a given level of ability. The individual characteristics used are described in Appendix F.2. The right figure has the same structure, but the outcome used is the probability of enrolling in high school.

relevant for disadvantaged students. For instance, the IQR of graduation probabilities exceeds 9 percentage points for 70% of students with low-educated parents and 16 percentage points for 30% of them. Conversely, among students with highly educated parents, the IQR is below 3 percentage points for 70% of students and below 0.5 percentage points for 30% of them. Patterns for academic high school enrollment are similar, although dispersion across schools remains non-negligible for a sizable subset of students with highly educated parents as well. Detailed results are reported in Appendix F.1.2.

6.1.2. School contributions by student type

Students who differ by parental education may also differ along other dimensions, such as immigrant status or neighborhood characteristics. To isolate the role of parental education, I therefore consider two representative student types who are identical in all observable and unobservable characteristics except parental education and, for each type, analyze how school contributions vary with ability.

Type L has low-educated parents (both parents attained at most lower secondary education), while type H has highly educated parents (both parents completed tertiary education). For each type, I simulate outcomes in every middle school for values of innate ability ranging from three standard deviations below to three standard deviations above the mean.[50] Importantly, in all simulations type L and type H are exposed to exactly the same school environment, which makes it possible to quantify how school contributions vary by student type and ability. Further details on the simulation design, including the individual characteristics used, are reported in Appendix F.2.

Fig. 5 plots graduation and high school enrollment probabilities for type L and type H students as a function of ability, comparing outcomes in schools at the 75th and 25th percentiles of the corresponding outcome distribution.[51] The left panel shows that school contributions to graduation are largest for type L students with below-average ability: for students with ability around one standard deviation below the mean, attending a school at the 75th percentile rather than the 25th percentile increases graduation probability by up to 18 percentage points. The IQR remains large (12 p.p.) for average ability and gradually decreases towards 0 as ability increases. Conversely, the school environment has a sizable effect on type H's probability of graduation only at very low ability levels.

The right panel shows that the probability of enrolling in high school is most sensitive to the school environment for two groups of students: type H students with below-average ability and type L students with average or above-average ability. For these groups, attending the school at the 75th percentile rather than the 25th percentile increases the probability of enrolling in high school by up to 14 percentage points, corresponding to an improvement of more than 30%.

To further illustrate these patterns, I next focus on a zoom-in for two representative students with average ability ($h = 0$), denoted L_0 and H_0. Fig. 6 plots, for each school in the sample, expected cognitive skills at the end of lower secondary education and the

[50] The graphs show standardized values in order to facilitate the interpretation of the results. The estimated variance of ability is 0.28 (see E.1), thus I use values ranging from about −1.5 to about 1.5 in the simulation.
[51] Fig. A.12 in the appendix shows the interquantile ranges.

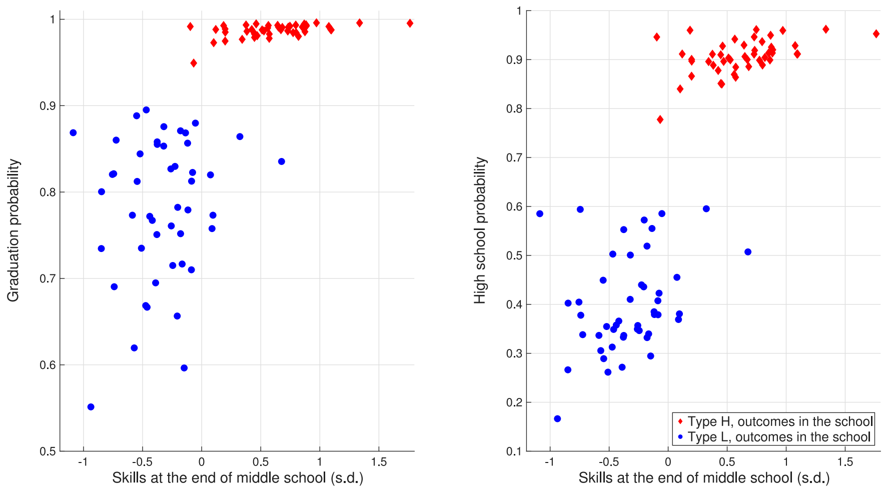

Fig. 6. Outcomes of type L_0 and type H_0 by school.

Note. The left figure plots graduation probability (y-axis) and skills at the end of lower secondary education (x-axis) in each school of the sample for two representative students: Type L_0 (blue dots) has parents with primary education, type H_0 (red diamonds) has parents with tertiary education. The average peer characteristics in each school are used. Frequencies are constructed using 10000 simulations of the structural model for each type and every school in the sample. Individual characteristics used are described in Appendix F.2. The right figure has exactly the same structure, but the outcome on the y-axis is the probability of enrolling in high school.

corresponding probability of graduation (left panel) and high school enrollment (right panel) for the two types. While school-to-school variation in performance is similar for L_0 and H_0, attainment outcomes differ markedly across types. For H_0, graduation probabilities are above 94% in all schools and close to one in schools with above-average expected skills, and high school enrollment exceeds 80% in all but one school. In contrast, L_0 exhibits much larger dispersion across schools: graduation probabilities range from about 55% to 90%, and the probability of enrolling in high school ranges from below 20% to about 60%.

The figure visually summarizes the main findings discussed above. School contributions to performance and attainment are positively but not strongly correlated: many schools that generate only average skill gains ensure higher attainment than schools with much larger effects on performance. This distinction is particularly relevant for disadvantaged students. While performance differences across schools convey most of the relevant information for students with highly educated parents, multiple outcome dimensions are necessary to characterize school contributions for students with low-educated parents.

6.1.3. Information frictions and educational outcomes

As discussed in Section 5, evaluations are informative and students progressively acquire more accurate beliefs about their ability, as reflected in the decline of posterior variance over time. This suggests that uncertainty about ability is gradually resolved. Nonetheless, temporary misperceptions about ability may still affect educational choices and outcomes, particularly in the early stages of middle school. This subsection quantifies the extent to which such information frictions influence educational trajectories.

Uncertainty affects educational outcomes through two channels. The first operates directly through students' choices. Students who hold pessimistic beliefs about their ability may choose to leave education even when they would have continued had they known their true ability. Conversely, students who overestimate their ability may choose to remain enrolled even though the outside option would have been more attractive under full information. While noise in evaluation signals is idiosyncratic and therefore uncorrelated with school contributions to choices, school effects shift this decision margin. Conditional on holding pessimistic beliefs, students are less likely to drop out when attending a school with stronger school effects on educational choices, and more likely to leave education when these effects are weaker.

The second channel operates through institutional decisions. School personnel observe the same noisy evaluations and therefore face similar uncertainty about students' underlying ability. As a result, uncertainty also affects their retention decisions. In particular, pessimistic signals increase the likelihood that a student is retained, and may also increase the probability of failing the final assessment at time $t = 3$ for students who have previously been retained. These two channels interact because, all else equal, retention is more likely when signals underestimate the student's true ability. Consequently, retained students tend to hold more pessimistic beliefs about their skills, which may further reduce their propensity to remain in education, over and above any direct disutility of retention.

To assess the relevance of uncertainty for educational outcomes and the importance of these channels, I compare outcomes in the baseline scenario with two counterfactual simulations.[52] In the first simulation, ability is perfectly known to the student but not to school personnel; therefore uncertainty affects retention events but not students' choices. In the second simulation, ability is observed by both students and school personnel, so uncertainty is completely eliminated.

Fig. 7 compares graduation rates (left panel) and high school enrollment rates (right panel) across schools in the baseline and counterfactual scenarios. Table A.8 in the Appendix shows the corresponding average statistics for the full sample and by ability level (above or below zero). All schools are affected in a similar way. When uncertainty is fully removed, the overall graduation rate is slightly lower, with differences typically between 1.5 and 3 percentage points. On average, eliminating uncertainty reduces graduation probabilities for students with below-average ability and increases them for students with above-average ability, although the negative effect for low-ability students is somewhat larger.[53]

High school enrollment rates are even closer across scenarios, with differences ranging from approximately -1.5 to 1.5 percentage points. Results are qualitatively similar when ability is known only to students, although the differences are smaller in magnitude, suggesting that the retention channel explains roughly half of the overall effect of uncertainty on outcomes.

Overall, in the simulation with fully observed ability, 6.9% of students change their graduation outcome and 8% change their high school enrollment decision. The corresponding figures are 3.4% and 5% when ability is known only to students. This indicates that uncertainty affects the educational trajectories of a minority of students, although its effects may still be important in some cases.

Appendix F.3 focuses on the representative student L_0, who has low-educated parents and average ability. The analysis shows that eliminating uncertainty would slightly improve this student's graduation probability in almost all schools, typically by 1–2 percentage points, while the probability of enrolling in high school remains largely unaffected. The appendix also explores the interaction between retention and uncertainty. Consistent with the mechanism described above, retained students tend to hold beliefs about their ability that are slightly below their true ability. As a result, even abstracting from the direct disutility of retention, retained students are somewhat more likely to drop out. By contrast, for students who remain enrolled and graduate at time $t = 3$, both perceived and actual ability tend to be higher than those observed at time $t = 2$. Consequently, the large differences in high school enrollment between students graduating at different times appear to be primarily driven by differences in preferences rather than by beliefs.

Overall, while uncertainty may influence the decisions of some students with low parental background, its quantitative impact is relatively modest. For comparison, as discussed in Section 6.1.2, the representative student considered here would increase his graduation probability by about 12 percentage points when moving from a school at the 25th percentile of the school distribution to one at the 75th percentile.

6.2. A counterfactual policy for mandatory education

In Spain lower secondary education is often thought of as compulsory, given that students normally turn 16 years old during their last year in middle school. However the previous analyses made it clear that in practice the legal dropout age is not necessarily aligned with the end of lower secondary education. I therefore use the model to explore the effects of a counterfactual policy, in which school attendance is mandatory until a given level, rather than a given age. More specifically, students are required to stay until the end of level II, repeating a grade if necessary; they are not allowed to leave at $t = 1$ or $t = 2$, unless they graduate. Those who fail at $t = 3$ fulfill the legal requirement, but, as in the baseline model, cannot enroll in upper secondary education.[54]

This counterfactual directly builds on the mechanisms identified above, in particular the role of school-level retention policies in influencing graduation. Schools differ in leniency, and these differences affect attainment because retention matters for students' subsequent dropout decisions. Under the counterfactual policy, this channel is largely switched off: while school leniency may still affect retention, it can no longer affect whether students leave school before reaching the final grade. The policy therefore mechanically compresses differences across schools in graduation that operate through retention-induced dropout.

It is important to acknowledge that the policy change might prompt schools to respond by somehow adjusting their inputs. However, this goes beyond the scope of the exercise, which treats school inputs as given. Nonetheless, the results illustrate what would happen in the short run and can be taken as a benchmark for the outcomes absent any school adjustment.[55]

[52] To perform the simulations, I replicate 1000 times the vector of observable characteristics of each student enrolled in a given school and draw ability and idiosyncratic shocks from their estimated distributions. This procedure generates a sufficiently large simulated sample to compute outcomes at the school level while preserving the distribution of observable characteristics and the proportion of students across schools observed in the data.

[53] This asymmetry reflects the nonlinear relationship between ability and the probabilities of retention and dropout implied by the dynamic decision problem. For students whose expected utilities place them close to the margin between staying in school and leaving, small shifts in perceived ability can substantially affect outcomes. A similar mechanism operates for students near the retention threshold. Lower-ability students are more likely to lie near these margins and, uncertainty tends to smooth their perceived ability toward zero. This increases the utility of remaining in education and shifts the retention decision toward promotion. Conversely, high-ability students are less likely to lie close to these margins; therefore symmetric shifts in perceived ability have a smaller effect on their probability of leaving school or being retained.

[54] To perform the simulations, I replicate 1000 times the vector of observable characteristics of each student enrolled in a given school and draw ability and idiosyncratic shocks from their estimated distributions. This procedure generates a sufficiently large simulated sample to compute outcomes at the school level while preserving the distribution of observable characteristics and the proportion of students across schools observed in the data.

[55] On the other hand, I allow peer effects to vary in the counterfactual simulation. More specifically, peer quality usually worsens. In fact, in the baseline case, peers in level II are somewhat positively selected, given that many students drop out before graduation. Under the counterfactual policy, all students reach level II, and therefore a student in a given school can expect the same peer quality in both level I and level II.

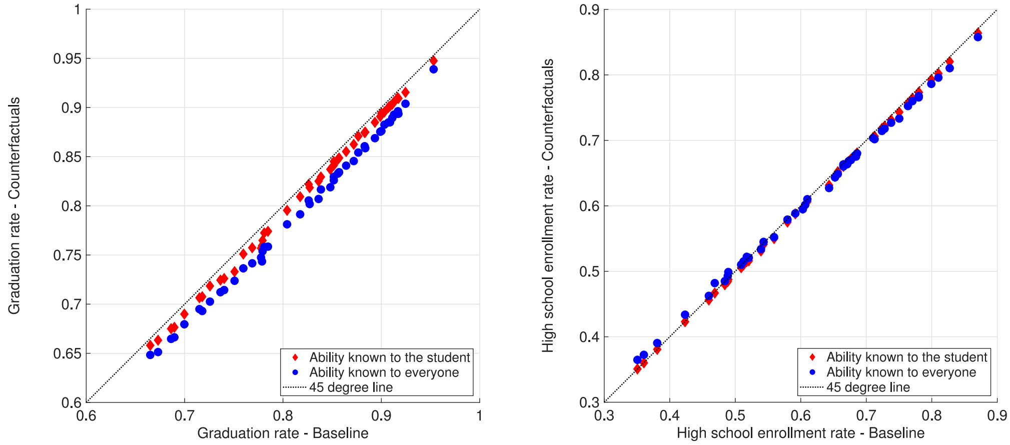

Fig. 7. Outcomes with and without uncertainty.
Note. The left figure contrasts the graduation rate in each school of the sample simulated using the baseline model (x-axis) and counterfactual models (y-axis). In the first counterfactual model (red diamonds) ability h is known only to the student. In the second counterfactual model (blue dots) ability h is known to both the student and the school personnel. Frequencies are constructed using 1000 simulations for each individual (i.e. vector of observable characteristics) in the sample. The right figure has the same structure, but the outcome plotted is the enrollment rate in high school.

Table 6
Counterfactual: changing mandatory education rules.

(a)Average outcomes in the population

	All		Low p.e.		Avg p.e.		High p.e.	
	(b)	(c)	(b)	(c)	(b)	(c)	(b)	(c)
graduate	0.842	0.939	0.699	0.874	0.831	0.936	0.958	0.988
enroll in high school	0.654	0.690	0.427	0.488	0.626	0.667	0.854	0.865
never dropout	0.870	1.000	0.754	1.000	0.862	1.000	0.965	1.000
graduate| no drop.	0.967	0.939	0.928	0.874	0.964	0.936	0.992	0.988
enroll in h.s.|grad.	0.777	0.736	0.611	0.558	0.752	0.712	0.891	0.876
C_{II}| grad.	0.255	0.123	−0.366	−0.512	0.098	−0.009	0.741	0.678
C_{II}| h.s.	0.465	0.388	−0.107	−0.199	0.273	0.208	0.833	0.788

(b)Expected outcomes for three student types

	Type L_0		Type M_0		Type H_0	
	(b)	(c)	(b)	(c)	(b)	(c)
graduate	0.774	0.935	0.928	0.981	0.990	0.999
enroll in high school	0.397	0.449	0.737	0.762	0.927	0.931
never dropout	0.815	1.000	0.944	1.000	0.991	1.000
graduate| no drop.	0.949	0.935	0.983	0.981	0.999	0.999
enroll in h.s.|grad.	0.514	0.481	0.794	0.777	0.936	0.932
C_{II}| grad.	−0.395	−0.400	0.286	0.264	0.833	0.809
C_{II}| h.s.	−0.413	−0.419	0.278	0.256	0.832	0.808

Note. Outcomes in columns (b) ("baseline") are computed using the estimated parameters of the model, while columns (c) ("counterfactual") contain results of a counterfactual simulation in which students are not allowed to drop out: they have to either graduate at $t = 2$ or stay in school until time $t = 3$. In panel a) values are the average of 1000 simulations of the structural model for each individual in the sample; the table shows the average in the sample and by subgroup of parental education (low, average, high). In panel (b) values are the average of 100000 simulations for each student type with given characteristics. More specifically, L_0, M_0 and H_0 have parents with primary, upper secondary, and tertiary education respectively. They are assigned the mean school environments for students with low, average, and high parental background respectively. Their ability is set at the average value of 0. Other characteristics are detailed in Appendix F.2.

Panel (a) of Table 6 summarizes average outcomes in the population and by subgroups of parental education. For comparison, panel (b) replicates the exercise for three student types: L_0, M_0, H_0. As before, individual characteristics other than parental education are kept constant and ability is set at its average value of 0. Each type is assigned the average school environment for students with similar parental education.

In the baseline scenario, 87% of students remain enrolled until the last possible period (i.e., never drop out), and almost 97% of them eventually graduate at either $t = 2$ or $t = 3$. This indicates that students who choose to stay in school are extremely likely to pass the final assessment. A natural concern with the counterfactual policy is that it would force students to remain in school even though they would ultimately be unable to obtain a diploma. The simulations suggest that this concern is limited: under the counterfactual, 94% of students graduate. Thus, while some students who are forced to stay still fail to complete lower secondary education, the overwhelming majority would successfully graduate when required to remain in school. The increase is particularly large for students with low-educated parents (+17 p.p.), but it is also sizable for students with average parental background (+10 p.p.). Thanks to the large increase in the pool of middle school graduates, overall enrollment in high school increases by almost 4 p.p. (6 p.p. in the case of students with low-educated parents), while enrollment conditional on graduation slightly decreases, since students at the margin of dropping out are not very likely to pursue further academic education if they eventually graduate. The results for the representative student types are consistent with these patterns in the general population.

Making education compulsory up to a specific level rather than a specific age has only a modest effect on the skill composition of students who enroll in high school. The average skill level of high school students decreases by approximately 0.07 s.d., suggesting that the policy would have limited effects on the next educational stage through changes in peer composition. Among middle school graduates, the decline in average skills is slightly larger, about 0.13 s.d., but still modest in magnitude. Overall, these results provide evidence that many early leavers would be able to complete lower secondary education — and in some cases continue to upper secondary education — under a policy that requires students to remain in school until the end of final level.

To assess whether the results of the policy counterfactual are affected by uncertainty about ability, Table A.9 in the Appendix replicates the analysis after eliminating uncertainty from both the baseline scenario and the mandatory schooling scenario. As discussed in Section 6.1.3, uncertainty about ability plays a relatively limited role in shaping students' educational trajectories. Consistent with this finding, the results are very similar to those obtained in the baseline model with uncertainty. In particular, graduation increases substantially under mandatory schooling – especially among students with low-educated parents – and high school enrollment rises modestly. These patterns confirm that, in this setting, school environment and institutional rules play a larger role than uncertainty in determining students' educational outcomes.

7. Conclusions

Suppose that a policy maker would like to identify the most "successful" schools, in order to learn their best practices and apply them in other schools. The school inputs that affect cognitive skills, as identified using the model of this study or other comparable measures of value added, would allow them to rank schools based on their ability to improve performance, as measured by a nationwide test. However, attending one of the top ranked schools would not be of benefit to every type of student, if those schools are not committed to helping every student succeed. Indeed, while attending such schools potentially raises cognitive skills, this is not accomplished if the student drops out. Moreover, in a different school they might have graduated and acquired a stronger motivation to pursue further education, which would likely lead to better outcomes in the labor market. The findings of this paper indicate that using school value added on performance as a proxy for "school quality" is not a benign assumption. In the best case, it might be a useful simplification for students with favorable socioeconomic conditions and high ability since they are very likely to pursue further academic education regardless of their current school environment. On the other hand, school environment is a crucial determinant of educational attainment in the case of students from a less favorable socioeconomic background, not only through its contribution to cognitive skills development, but also through its impact on students' educational choices and school-level retention policies. As shown in the counterfactual scenario in which dropping out is eliminated, the overwhelming majority of students would successfully graduate if they remained in school. Under the current rules, however, many of them decide to drop out. Evaluating school effectiveness using only performance may lead to conclusions that do not benefit disadvantaged students: a policy maker whose goal is to improve educational outcomes for all should not ignore other dimensions, such as educational attainment.

The methodological approach proposed here makes it possible to disentangle a school's impact on attainment through its effects on skills, educational decisions, and grading policies, relying exclusively on administrative data that are usually available at local or central levels. Future research should endeavor to open the "black box" of school inputs and understand the mechanisms that lead to the differentiation across schools. Furthermore, the results highlight the importance of studying how seemingly homogeneous rules are implemented in decentralized institutions. In fact, while the public schools in the sample share the same general set of rules for grading, retention, and graduation, differences in implementation have long-lasting effects on student outcomes.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used ChatGPT (OpenAI) to assist with proofreading, language editing, and formatting of tables and figures. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Table A.7
Fit of the model (conditional)

	graduate at $t = 2$		graduate at $t = 3$		enroll in hs	
	Data	Model	Data	Model	Data	Model
ALL	94.25	92.05	83.43	85.27	78.78	77.71
male	93.17	90.46	80.30	82.67	75.13	73.81
female	95.32	93.57	87.57	88.80	82.41	81.55
Spanish	94.83	92.77	83.07	85.37	79.97	78.81
immigrant	88.54	85.58	84.30	84.97	69.98	69.78
low parental edu.	87.66	83.82	84.70	82.41	61.72	61.02
avg parental edu.	93.15	90.62	82.42	85.96	76.50	75.16
high parental edu.	98.01	96.81	82.69	90.98	90.08	89.13
below median peers	91.97	89.36	84.51	85.22	70.37	69.90
above median peers	95.52	93.58	82.11	85.33	83.95	82.61

Note. Data frequencies are computed from the sample of students used in the estimation. Model frequencies are constructed using 1000 simulations of the structural model for each individual in the estimation sample, where an "individual" is defined by a set of observable initial characteristics. Statistics are conditional on reaching the relevant level, for instance graduation at $t = 2$ is conditional on being promoted at $t = 1$ and choosing to stay in education.

Table A.8
Students' outcomes when ability is known.

(a) Average outcomes in the population

	All			$h < 0$			$h \geq 0$		
	(b)	(pu)	(nu)	(b)	(pu)	(nu)	(b)	(pu)	(nu)
retained at $t = 1$	0.235	0.235	0.263	0.285	0.285	0.387	0.185	0.185	0.139
dropout at $t = 1$	0.081	0.093	0.096	0.132	0.164	0.171	0.029	0.022	0.021
graduate	0.842	0.833	0.819	0.745	0.719	0.682	0.938	0.947	0.955
enroll in high school	0.654	0.649	0.649	0.497	0.472	0.452	0.812	0.826	0.846
$C_{II}\|$ grad.	0.255	0.280	0.320	−0.361	−0.324	−0.272	0.743	0.739	0.741

(b) Proportion of students who changed outcome

	h known to the student			h known to everyone		
	All	(b) better	(pu) better	All	(b) better	(nu) better
retained at $t = 1$	0.000	0.000	0.000	0.096	0.062	0.034
dropout at $t = 1$	0.036	0.024	0.012	0.040	0.028	0.012
graduate	0.034	0.022	0.013	0.069	0.046	0.023
enroll in high school	0.050	0.027	0.022	0.080	0.043	0.037

Note. Panel (a). Values in columns (b) are the average of 1000 simulations of the structural model for each individual in the sample ("baseline" simulation). Values in columns (pu) and (nu) are the average of 1000 simulations of counterfactual models for each individual in the sample, Specifically in (pu) ability h is known by the student but not by the school personnel ("partial uncertainty"). In (nu) ability h is perfectly observed ("no uncertainty"). Columns "All" contains average outcomes in the whole populations, columns "$h < 0$" contain average among the subsample of students with ability below 0, columns "$h \geq 0$" contain average among the subsample of students with ability above 0. *Panel (b).* For each outcome, the table shows the share of simulations in which the student has a different outcome in the baseline scenario and in the scenario with reduced uncertainty (pu) or without uncertainty (nu). Columns "All" count all differences, columns "(b) better" include cases in which the outcome is more favorable in the baseline scenario, columns "(pu) better" and "(nu) better" include cases in which the outcome is less favorable in the baseline scenario.

Appendix A

A.1. Additional tables

See Tables A.7–A.9.

A.2. Additional figures

See Figs. A.8–A.12.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.euroecorev.2026.105372.

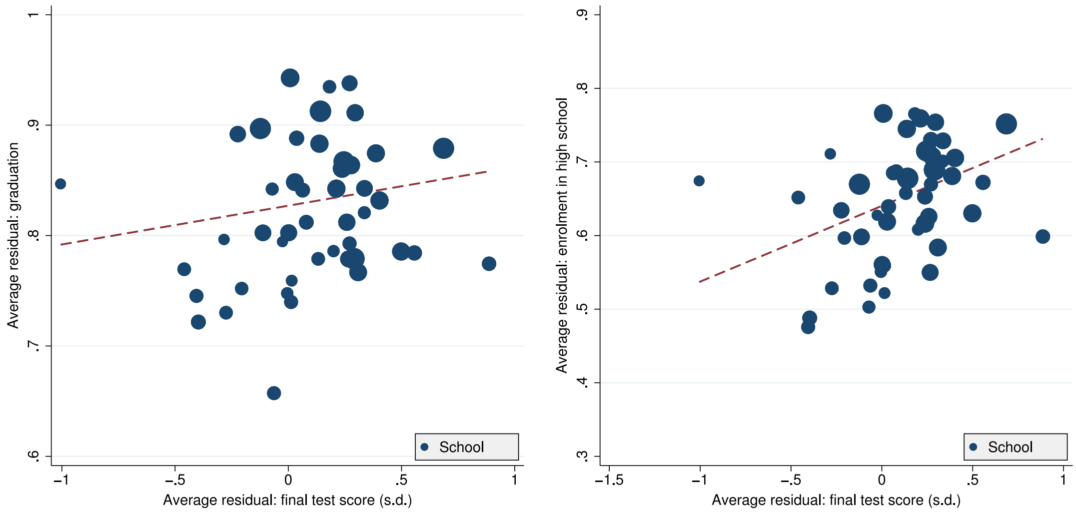

Fig. A.8. Educational outcomes by school.
Note. Each dot plots average residuals at the school level from a regression on a vector of dummies for mother and father education (the sample mean is added back). The dot size is proportional to the school size. The dashed line is a linear fit. Correlations are 0.19 (left) and 0.34 (right); correlations weighted by school size are 0.16 and 0.39 respectively.

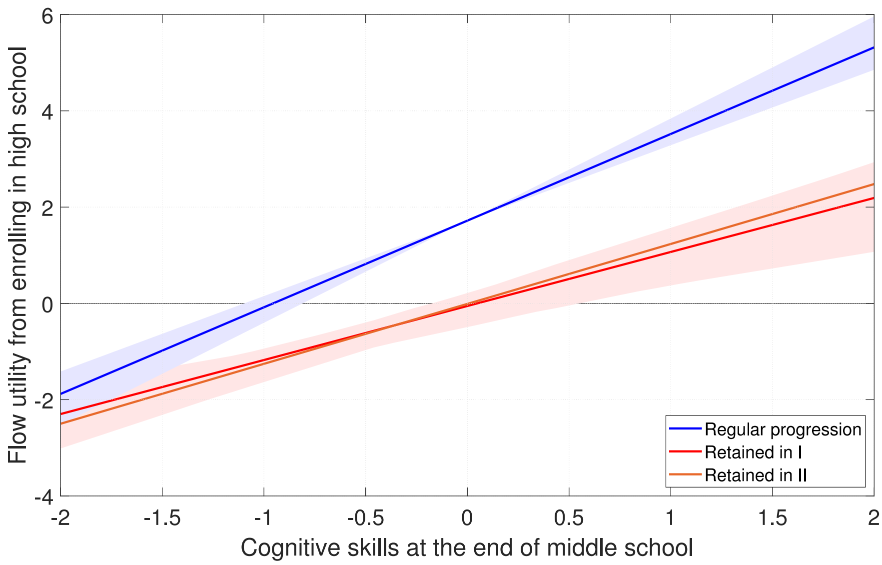

Fig. A.9. Effect of cognitive skills on high school enrollment by retention status.
Note. The figure plots the effect of beliefs about cognitive skills on the utility from the choice of enrolling in high school for three types of students: those who completed middle school regularly at time $t = 2$ (blue line), those who graduated at time $t = 3$ because they were retained at time$t = 1$ and repeat level I (red line), those who graduated at time $t = 3$ because they were retained at time $t = 2$ and repeat level II (orange line). Shaded areas are 95% confidence intervals. The confidence interval for "Retained in II" is not plotted for clarity, it largely overlaps with the one for "Retained in I".

Table A.9
Counterfactual: changing mandatory education rules - without uncertainty.

(a) Average outcomes in the population

	All		Low p.e.		Avg p.e.		High p.e.	
	(b)	(c)	(b)	(c)	(b)	(c)	(b)	(c)
graduate	0.819	0.922	0.676	0.852	0.804	0.917	0.940	0.979
enroll in high school	0.649	0.685	0.435	0.493	0.620	0.662	0.838	0.853
never dropout	0.850	1.000	0.732	1.000	0.838	1.000	0.951	1.000
graduate| no drop.	0.963	0.922	0.923	0.852	0.959	0.917	0.989	0.979
enroll in h.s.|grad.	0.793	0.743	0.644	0.579	0.771	0.721	0.892	0.871
C_{II}| grad.	0.320	0.160	−0.288	−0.476	0.174	0.033	0.780	0.699
C_{II}| h.s.	0.523	0.437	−0.030	−0.138	0.350	0.271	0.878	0.825

(b) Expected outcomes for three student types

	Type L_0		Type M_0		Type H_0	
	(b)	(c)	(b)	(c)	(b)	(c)
graduate	0.795	0.945	0.949	0.986	0.994	0.999
enroll in high school	0.400	0.459	0.768	0.793	0.936	0.941
never dropout	0.832	1.000	0.960	1.000	0.999	0.999
graduate| no drop.	0.955	0.945	0.988	0.986	0.999	0.999
enroll in h.s.|grad.	0.504	0.485	0.810	0.805	0.942	0.942
C_{II}| grad.	−0.397	−0.393	0.296	0.297	0.820	0.820
C_{II}| h.s.	−0.412	−0.408	0.291	0.292	0.819	0.819

Note. This table replicates Table 6 using a counterfactual model in which ability h is perfectly observed. Columns (b) ("baseline") summarize outcomes of a simulation in which students are allowed to dropout at time $t = 1$ and $t = 2$. Columns (c) ("counterfactual") contain results of a simulation in which students are not allowed to drop out: they have to either graduate at $t = 2$ or stay in school until time $t = 3$. In panel a) values are the average of 1000 simulations of the structural model for each individual in the sample; the table shows the average in the sample and by subgroup of parental education (low, average, high). In panel b) values are the average of 100000 simulations for each student type with given characteristics. More specifically, L_0, M_0 and H_0 have parents with primary, upper secondary, and tertiary education respectively. They are assigned the mean school environments for students with low, average, and high parental background respectively. Their ability is set at the average value of 0. Other characteristics are detailed in Appendix F.2.

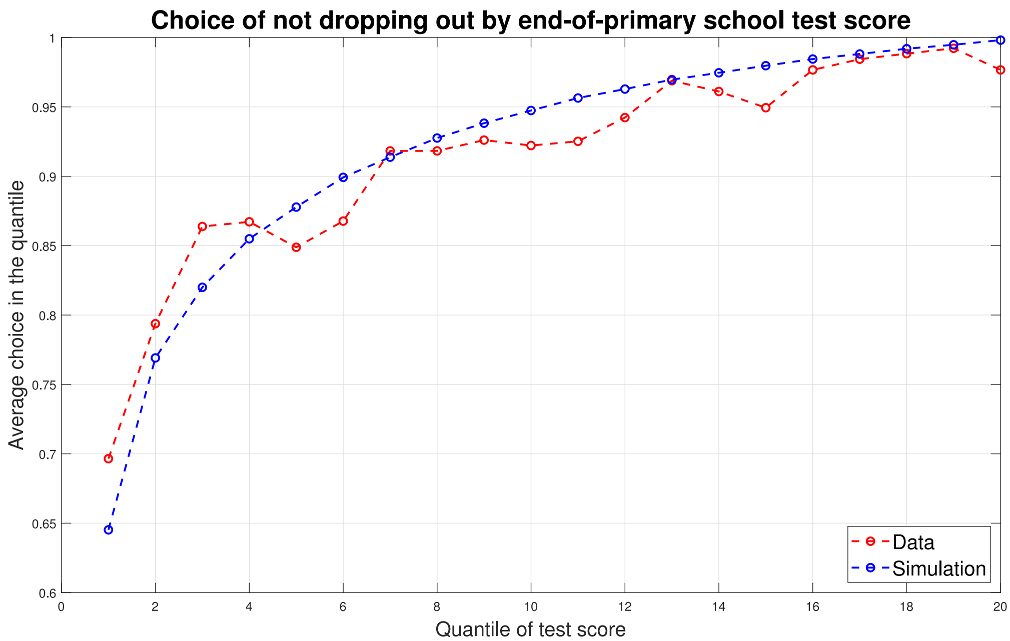

Fig. A.10. Fit of the model (i).
Note. The figure plots the share of students who choose to stay in school at time $t = 1$ by quantile of their test score at the end of primary school. Sample average from the real data are in red, while results of the simulation performed using the estimated parameters of the model are in blue.

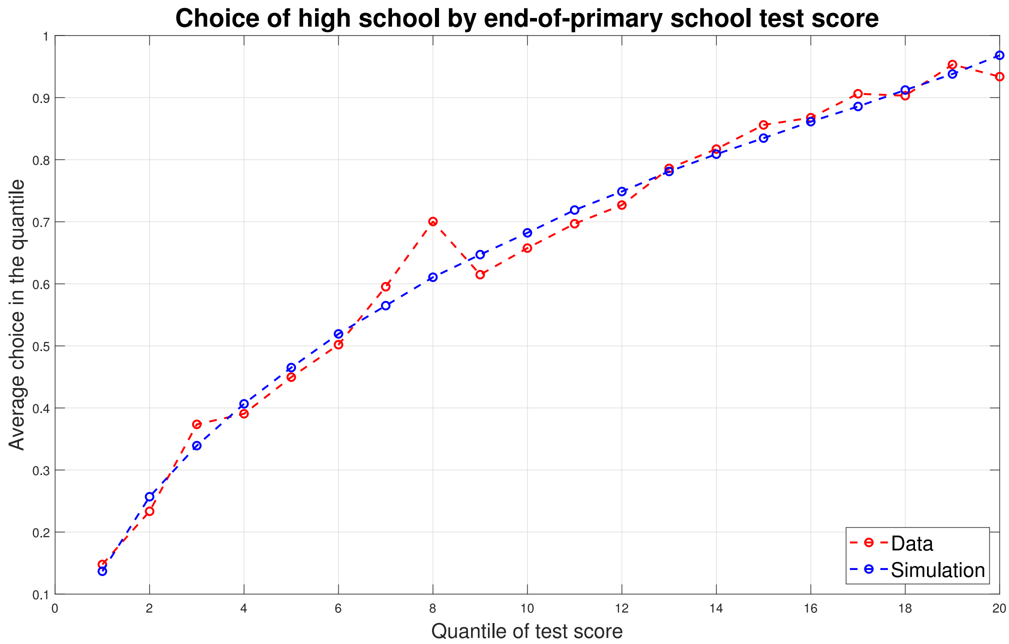

Fig. A.11. Fit of the model (ii).
Note. The figure plots the share of students who enroll in high school by quantile of their test score at the end of primary school. Sample average from the real data are in red, while results of the simulation performed using the estimated parameters of the model are in blue.

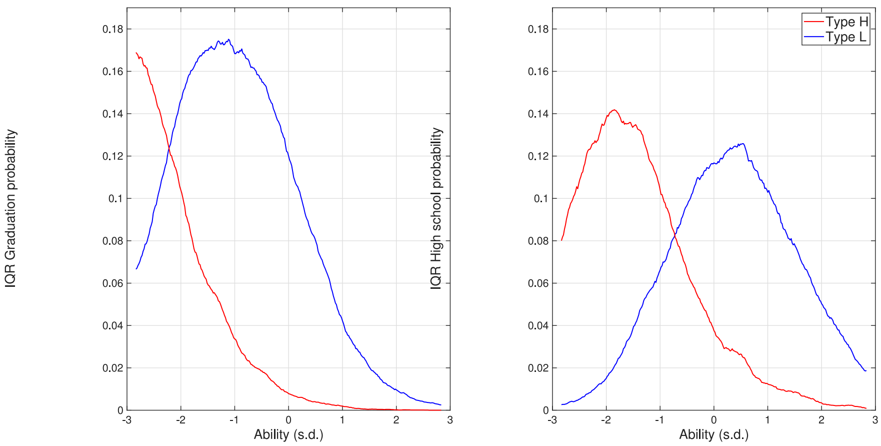

Fig. A.12. Attainment of type L and type H students by ability.
Note. The left figure plots the interquantile range (IQR) of graduation probability (y-axis) by ability level (x-axis) for representative students with highly educated parents (red line) and low-educated parents (blue line). The IQR is the difference in graduation probability between the school at the 75th percentile and the school at the 25th percentile. Each probability is the average of 10000 simulations for a given student type at a given level of ability. Other characteristics are detailed in Appendix F.2. The right figure has the same structure, but the outcome used is the probability of enrolling in high school.

Data availability

The data that has been used is confidential.

References

Abdulkadiroğlu, A., Angrist, J.D., Dynarski, S.M., Kane, T.J., Pathak, P.A., 2011. Accountability and flexibility in public schools: Evidence from boston's charters and pilots. Q. J. Econ. 126 (2), 699.

Agostinelli, F., Doepke, M., Sorrenti, G., Zilibotti, F., 2026. It takes a village: The economics of parenting with neighborhood and peer effects. J. Political Econ. 134 (1), 313–365.
Allen, R., Burgess, S., 2013. Evaluating the provision of school performance information for school choice. Econ. Educ. Rev. 34 (C), 175–190.
Allen, C.S., Chen, Q., Willson, V.L., Hughes, J.N., 2009. Quality of research design moderates effects of grade retention on achievement: A meta-analytic, multilevel analysis. Educ. Eval. Policy Anal. 31 (4), 480–499, PMID: 20717492.
Altonji, J.G., Arcidiacono, P., Maurel, A., 2015. The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects. NBER Working Papers 21655, National Bureau of Economic Research, Inc.
Angrist, J.D., Cohodes, S.R., Dynarski, S.M., Pathak, P.A., Walters, C.R., 2016. Stand and deliver: Effects of boston's charter high schools on college preparation, entry, and choice. J. Labor Econ. 34 (2), 275–318.
Angrist, J.D., Hull, P.D., Pathak, P.A., Walters, C.R., 2017. Leveraging lotteries for school value-added: Testing and estimation*. Q. J. Econ. 132 (2), 871–919.
Angrist, J., Hull, P., Walters, C.R., 2022. Methods for measuring school effectiveness. Working Paper 30803, National Bureau of Economic Research.
Angrist, J.D., Pathak, P.A., Walters, C.R., 2013. Explaining charter school effectiveness. Am. Econ. J.: Appl. Econ. 5 (4), 1–27.
Arcidiacono, P., 2004. Ability sorting and the returns to college major. J. Econometrics 121 (1–2), 343–375.
Arcidiacono, P., Aucejo, E., Maurel, A., Ransom, T., 2025. College attrition and the dynamics of information revelation. J. Political Econ. 133 (1), 53–110.
Avery, C., Hoxby, C., Jackson, C., Burek, K., Pope, G., Raman, M., 2006. Cost Should Be No Barrier: An Evaluation of the First Year of Harvard's Financial Aid Initiative. NBER Working Papers 12029, National Bureau of Economic Research, Inc.
Belfield, C., Boneva, T., Rauh, C., Shaw, J., 2020. What drives enrolment gaps in further education? The role of beliefs in sequential schooling decisions. Economica 87 (346), 490–529.
Beuermann, D.W., Jackson, C.K., Navarro-Sola, L., Pardo, F., 2022. What is a Good School, and Can Parents Tell? Evidence on the Multidimensionality of School Output. Rev. Econ. Stud. 90 (1), 65–101.
Bordon, P., Fu, C., 2015. College-major choice to college-then-major choice. Rev. Econ. Stud. 82 (4), 1247–1288.
Borghesan, E., Reis, H., Todd, P.E., 2022. Learning Through Repetition? A Dynamic Evaluation of Grade Retention in Portugal. PIER Working Paper 22-030.
Calsamiglia, C., Loviglio, A., 2019. Grading on a curve: When having good peers is not good. Econom. Educ. Rev. 73, 101916.
Calsamiglia, C., Loviglio, A., 2020. Maturity and school outcomes in an inflexible system: evidence from catalonia. SERIEs 11 (1), 1–49.
Cockx, B., Picchio, M., Baert, S., 2017. Modeling the Effects of Grade Retention in High School. Discussion Paper Series 148, Global Labor Organization (GLO).
Cunha, F., Heckman, J.J., 2008. Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. J. Hum. Resour. 43 (4).
Cunha, F., Heckman, J.J., Lochner, L., 2006. Interpreting the Evidence on Life Cycle Skill Formation. In: Hanushek, E., Welch, F. (Eds.), In: Handbook of the Economics of Education, vol. 1, Elsevier, pp. 697–812.
Cunha, F., Heckman, J.J., Schennach, S.M., 2010. Estimating the Technology of Cognitive and Noncognitive Skill Formation. Econometrica 78 (3), 883–931.
Dearden, L., Micklewright, J., Vignoles, A., 2011. The effectiveness of english secondary schools for pupils of different ability levels. Fisc. Stud. 32 (2), 225–244.
DeGroot, M.H., 1970. Optimal Statistical Decisions. McGraw Hill.
Deming, D.J., Hastings, J.S., Kane, T.J., Staiger, D.O., 2014. School choice, school quality, and postsecondary attainment. Am. Econ. Rev. 104 (3), 991–1013.
Dobbie, W., Fryer, J., Roland, G., 2011. Are high-quality schools enough to increase achievement among the poor? evidence from the harlem children's zone. Am. Econ. J.: Appl. Econ. 3 (3), 158–187.
Dobbie, W., Fryer, J., Roland, G., 2013. Getting beneath the veil of effective schools: Evidence from new york city. Am. Econ. J.: Appl. Econ. 5 (4), 28–60.
Fruehwirth, J.C., Navarro, S., Takahashi, Y., 2016. How the timing of grade retention affects outcomes: Identification and estimation of time-varying treatment effects. J. Labor Econ. 34 (4), 979–1021.
Fryer, Jr., R.G., 2014. Injecting charter school best practices into traditional public schools: Evidence from field experiments. Q. J. Econ. 129 (3), 1355–1407.
Giustinelli, P., Pavoni, N., 2017. The evolution of awareness and belief ambiguity in the process of high school track choice. Rev. Econ. Dynamics 25, 93–120, Special Issue on Human Capital and Inequality.
Hastings, J.S., Neilson, C.A., Zimmerman, S.D., 2013. Are Some Degrees Worth More than Others? Evidence from college admission cutoffs in Chile. NBER Working Papers 19241, National Bureau of Economic Research, Inc..
Heckman, J.J., Humphries, J.E., Mader, N.S., 2011. The GED. In: Hanushek, E., Machin, S., Woessman, L. (Eds.), In: Handbook of the Economics of Education, 3, Elsevier, pp. 423–483.
Heckman, J.J., Rubinstein, Y., 2001. The importance of noncognitive skills: Lessons from the GED testing program. Am. Econ. Rev. 91 (2), 145–149.
Holmström, B., 1999. Managerial incentive problems: A dynamic perspective. Rev. Econ. Stud. 66 (1), 169–182.
Hoxby, C.M., Avery, C., 2012. The Missing "One-Offs": The Hidden Supply of High-Achieving, Low Income Students.
Jackson, C.K., 2018. What do test scores miss? The importance of teacher effects on non–test score outcomes. J. Political Econ. 126 (5), 2072–2107.
Jackson, C.K., Porter, S.C., Easton, J.Q., Blanchard, A., Kiguel, S., 2020. School effects on socioemotional development, school-based arrests, and educational attainment. Am. Econ. Rev.: Insights 2 (4), 491–508.
Jacob, B.A., Lefgren, L., 2009. The effect of grade retention on high school completion. Am. Econ. J.: Appl. Econ. 1 (3), 33–58.
James, J., 2011. Ability matching and occupational choice. (Working Paper 11-25).
Kane, T.J., Staiger, D.O., 2002. The promise and pitfalls of using imprecise school accountability measures. J. Econ. Perspect. 16 (4), 91–114.
Kinsler, J., Pavan, R., 2015. The specificity of general human capital: Evidence from college major choice. J. Labor Econ. 33 (4), 933–972.
Kinsler, J., Pavan, R., 2021. Local distortions in parental beliefs over child skill. J. Political Econ. 129 (1), 81–100.
Leckie, G., Goldstein, H., 2017. The evolution of school league tables in England 1992–2016: 'contextual value-added', 'expected progress' and 'progress 8'. Br. Educ. Res. J. 43 (2), 193–212.
OECD, 2016. Education at a Glance 2016: OECD Indicators. Oecd publishing, Paris.
OECD, 2018. Education at a Glance 2018: OECD Indicators. Oecd publishing, Paris.
Petek, N., Pope, N.G., 2023. The multidimensional impact of teachers on students. J. Political Econ. 131 (4), 1057–1107.
Rust, J., 1987. Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. Econometrica 55 (5), 999–1033.
Wiswall, M., Zafar, B., 2014. Determinants of college major choice: Identification using an information experiment. Rev. Econ. Stud..
Wiswall, M., Zafar, B., 2015. How Do College Students Respond to Public Information about Earnings? J. Hum. Cap. 9 (2), 117–169.
Zafar, B., 2013. College Major Choice and the Gender Gap. J. Hum. Resour. 48 (3), 545–595.