Università degli studi di Torino

Dip. di Culture, politica e società

Università degli studi di Milano

Dip. di Scienze sociali e politiche

PhD PROGRAM
SOCIOLOGY AND METHODOLOGY OF SOCIAL RESEARCH – 32nd cohort
SETTORE SCIENTIFICO SPS/07 or 08

THE POLICY OF GRADE REPETITION

Determinants and Consequences of a Contested Educational Practice

Doctoral dissertation by

Guido Salza

Supervisor: Prof. Dalit Contini

Director of Doctoral Program: Prof. Mauro Barisione

Contents

Acknowledgements

Consistent with the spirit of this thesis, I should thank my parents' occupation and educational attainment, or the many books on the shelves in my family house while I was a child.

Luckily enough, even accounting for all the relevant variables that determined this outcome would miss the non-quantifiable role of the individuals that filled up of sympathy and kindness this journey.

The present thesis, from his genesis to its final draft, arises from the continuous intellectual exchange with Professor Dalit Contini. Work in contact with her passion and meticulousness has allowed me to make my first steps inside the academic world. The highly ambitious standards she asks of herself and her collaborators have made me grow enormously as a researcher.

My sincere gratitude also goes to Carlo Barone, as my tutor in OSC-SciencesPo and thesis referee. I had many chances to talk about the thesis and receive feedback: his advice has crucially improved the theoretical consistency of this work.

I wish to express a similar kind of gratitude to two talented young researchers, Camilla Borgna and Marco Romito. They are again examples of how this job can be passionately accomplished. In particular, I would like to thank Camilla for her insightful and punctual feedback as a referee.

Many other scholars have gifted me with insights, criticisms, and enriching discussions. Among many others, I would like here to cherish the insightful comments that came to me at different stages of the work from Gianluca Argentin, Moris Triventi, and Enrico Rettore.

My sincere gratitude also goes to the scholars and colleagues I met in my stay in Paris. I had the precious opportunity to extensively present my work and receive feedback before Lidia Panico, Anne Solaz and the whole Unité 9 ad INED. Moreover, I was warmly welcomed by the ex-pat community at INED: thanks all of you!

The SOMET-32 people (Nicole, Roberto, Alberto, Beatrice, Robin, Alessandro, Caterina, Anouck, Francesca, Silvia and Marta) have played a central and unforgettable

role in this PhD. Among them, a specific word of gratitude goes to my friend Alberto: through his generous company (and great cooking skills), he taught me a lot about this job. More in general, as a collective and as distinct individuals, I own the SOMET-32 people very much, and I wish every one of us a happy future, any future you like. In contrast to the ideology of individual competition fostered in academia, I feel that the solidarity that has bound us together in the first place has never been broken.

A special thank is to my cherished mate Matteo ‘Peo’ Cavallaro, for me a constant (and never-resting) intellectual reference in the last decade.

I also take the opportunity to genuinely thank all my friends who in these years have been asking me back their tax money from my salary. They have continuously and lovingly remembered me not to take all this matter too seriously.

It is likewise difficult to imagine my survival in the writing process without the tender counterbalance of Giulia. It never felt so right to take some time off and remember that I am made of many other things besides this job.

Finally, well beyond the limited event of this PhD, my most profound gratitude goes to my brother Claudio, my cousin Davide, my lifelong friend Josh and my sister Lula. My sincerest thanks go to my mother Patrizia, who always knows when I need help, and Toni; to my loyal readers, my father Paolo, and Gabriella.

All the people mentioned above, and all the others I care for but are not directly involved, unconditionally deserve the co-authorship of this thesis.

Introduction

"Reform ideas win or lose according to the way they resonate with a particular social context, attract or repeal particular constituencies, and respond to the social problems that are seen as most salient at the time".
David F. Labaree 2010, p.163

Policymakers and the public are increasingly interested in the trends of educational outcomes, and commitment to supporting and strengthening aggregate learning levels is growing (OECD 2015; UN 2012; EU 2018).

Even in Italy, many have sought to mobilise public opinion around this task. For instance, in 2017, a few hundred university teachers published an open letter to the Ministry of Education. In the letter, they denounced the worrying levels of academic preparation among enrolled students, including the 'too many young people [who] write badly in Italian, read little and find it hard to express themselves orally'. The signers called for 'a school that is demanding, that verifies the learning and can be more effective in teaching, otherwise, neither the generous commitment of many outstanding teachers nor the acquisition of new methodologies will suffice'.[1]

If a lack of rigorous exams and strict selection in schools are responsible for the decline in aggregate learning outcomes, then an attractive solution may be the practice of grade repetition, which denies students progression in their educational career for an additional year.

Recently, a prominent Italian newspaper published an article which states that '[i]n the Italy of low economic growth, there is a voice that has actually grown at a sustained rate: the students promoted at the end of the year'. The author argues that the positive trend in promotions nonetheless 'does not seem to result from an improvement in knowledge on a large scale, given the low levels of learning in reading and mathematics, but from the continuous loosening of the pupil assessment system…often transversal to the governments in office'.[2]

[1] Retrieved on September 2019 from http://gruppodifirenze2.blogspot.com/
[2] Bruno and Tucci 2019, 'La scuola non boccia quasi più (e non per merito degli studenti)', Il Sole 24 Ore; retrieved in September 2019 from https://www.ilsole24ore.com

The present doctoral thesis investigates grade repetition as a selection and recovery device. It places particular emphasis on its implications through the broader mechanisms of educational inequalities and social stratification. Socially underprivileged pupils are likely to be over-represented among poor performers (cf. Chapter 1). In this regard, the Organisation for Economic Cooperation and Development (OECD; 2015) has reported that a 15-year-old student 'with certain characteristics is more likely to have repeated a grade than other students' (p.164). The likelihood of grade repetition is significantly higher among ethnic minorities, migrants, families with low socioeconomic status and residents of rural areas.

In the past, underprivileged children have had to blame themselves for their academic failures. However, the unprecedented amount of empirical evidence that has accumulated over the last decades reveals that one's social position drastically influences academic preferences. However, it is still unclear how grade repetition relates to rooted educational inequalities, and the extent to which it is a useful device to assist underachievers is uncertain (cf. Chapter 2). On the one hand, advocates of grade repetition claim that it could give students the necessary additional time to re-align their competencies and knowledge before proceeding in their educational careers. Many teachers, deans and parents favour the practice of grade repetition and use it as leverage to elicit effort from students, to 'protect' underperforming students from overly challenging curricula in the next grade (Shepard and Smith 1989) or even to incorporate another academic year to allow for emotional or psychological maturation.

It is not rare to hear stories from teachers about the individual successes of retained students (Shepard 1989; Marxoux and Crahay 2008). For example, José Claveria, the dean of one of the most prominent private schools in Milan, recently intervened in the debate over grade repetition by sharing that he 'saw a mature, happy man, with a fulfilled life, say to a fourteen-year-old boy who flunked: "I was flunked twice as a boy". That boy's lowered eyes suddenly lit up. With those words, he did not deny a school failure but made it an occasion for good. It is nothing but a lost year!'.[3]

[3] Claveria 2019 'Fermare alle superiori: this is not the question', Avvenire; retrieved in September 2019 from https://www.avvenire.it

Supposedly, the 'lost year' allows children to get back on track before it is too late. 'After grade repetition - an Italian middle school teacher claim in response to a negative commentary on the subject - many retained pupils found their environment by changing the school (or simply the class) ambience, and they recovered the lost year abundantly, much more than they would have been able to do if they had not repeated the year. This is often the purpose and rationale of the grade repetition'.[4]

'In the secondary schools - continues the above-mentioned secondary school teacher - we often find ourselves in front of people without a family substratum, and they after us have no other way to have a minimum of civic education…But we, the teachers, and the whole school system, if we promote those who do not deserve, we are defeated and do not do our job well'.[5]

However, there is no shortage of critical voices. Several international institutions have firmly advocated for a limitation to grade repetition. The OECD has observed that grade repetition is more prevalent in school systems in which students achieve lower-range scores on the Programme for International Student Assessment (PISA) assessment or school attendance is less equitable (OECD 2015).

The United Nations Educational, Scientific and Cultural Organization (UNESCO 2012) has recommended 'ensur[ing] transitions (*i.e.* promotions, editor's note) to reduce or target the practice of repetition' (p.56). They have even suggested that the automatic progression of students through grade levels could represent a more cost-effective policy.

In following this advice, national educational systems should plan to gradually limit grade repetition to the end of each cycle rather than each academic year (ibidem). At the European level, the Education and Training Monitor (2018) has recommended that countries reduce the rate of grade repetition to improve their aggregate human capital outcomes (E.U. Commission 2018).[6]

[4] Retrieved on September 2019 at https://www.avvenire.it/opinioni/pagine/sempre-vero-che-chi-boccia-perde-nella-scuola-primaria-sempre-vero

[5] *Ibidem.*

[6] In the Italian context, Andrea Gavosto, the director of the Giovanni Agnelli Foundation, a leading private institution that publishes each year a ranking of the Italian upper secondary schools, recently claimed that 'in the regions where grade repetition rate is the lowest, the levels of achievement [measured by standardised tests, ed.] are the highest', and hence concluded that grade repetition is not a useful nor an efficient educational practice (Andrea Gavosto 2013 'Ma bocciare serve davvero?', La Stampa; retrived in September 2019 from https://www.fondazioneagnelli.it). These standpoints often seem to overlook the

Many educators do not favour grade repetition. The Italian pedagogue Daniele Novara has maintained that teachers should 'analyse mistakes without punishing…and evaluate progress, not failures. [Therefore,] it is necessary to get rid of all coercive practices, which have no scientific reasons'.[7] Another Italian educational scholar, Raffaele Mantegazza, has publicly opposed the 'educational massacre' that is apparent from the considerable rate of grade repetition in the first year of high school.[8]

On these bases, this work seeks to organise the relevant existing literature and add to scientific knowledge of the understudied Italian context (cf. Chapter 3). In Italy, grade repetition is by far the principal factor in the variation in student ages within grades (PISA 2015). In this country, 15.1% of students reported that they had repeated at least one year by the end of the upper-secondary school; for comparison, the OECD average was 11.3% (*ibidem*). Repetition is concentrated at the upper-secondary level, although more than 5% of students reported that they had already faced a grade repetition at the lower-secondary level.[9]

Teachers have the most influence on the decision of whether a student will repeat a grade, although they must transmit this resolution to parents in advance. Academic performance, behaviour and the rate of absence from school are the formal criteria that govern this decision (EURYDICE 2010). Although grade repetition is a widely accepted educational practice, it remains substantially understudied in Italy. Consequently, there is a lack of appropriate data. Italy has produced neither a specific official report on the topic nor a longitudinal study of students' careers at the national or local level.

The present work intends to answer the three following empirical questions:

(1) What are the main determinants of grade repetition? How important are socioeconomic and migration background once the analysis accounts for past academic performance?

crucial contribution of well-rooted educational inequalities on academic performance. For example, a low grade-repetition rate in a school might correspond to a socially and culturally privileged student-body, which in facts has little to do with any implemented educational practice.

[7] Retrieved on September 2019 from https://www.orizzontescuola.it

[8] Retrieved on September 2019 from https://www.repubblica.it

[9] Notice that in Italy retention at primary level is barely inexistent, while it is quite diffuse in other countries (cf. Chapter 2 and 3).

This work examines the determinants of grade repetition by statistically modelling the probability of being subject to grade repetition at the end of ninth grade (cf. Chapter 5). At this stage, the analysis intends to clarify the influence of a comprehensive set of risk factors while emphasising the significance of socioeconomic and migration backgrounds. The rich array of measurements for academic performance before grade repetition allows for a credible assessment of the enduring influence of parental resources on the occurrence of grade repetition.

(2) Which are the academic careers of retained students? Are there significant disparities between social groups?

A detailed examination compares ninth-grade repeaters with distinct socioeconomic and migration characteristics and estimates further increases in educational inequalities (cf. Chapter 7). The results highlight that regardless of their social background, repeaters have a dramatically higher chance of experiencing further difficulties in their subsequent careers. Moreover, careers after grade repetition also diverge according to social and, to a smaller extent, migration backgrounds.

(3) What is the causal impact of grade repetition on the students' subsequent academic careers? Is there heterogeneity of effect in grade repetition across social groups?

Identification of the impact of grade repetition on subsequent educational careers is accomplished *via* two distinct matching strategies. In essence, this method facilitates a comparison of the educational outcomes of retained and non-retained students by constructing a sample of 'treated' and 'controls' with highly similar characteristics. The underlying assumption is that, depending on student characteristics (including a rich array of performance measurements) as well as current and prior school characteristics, the grade repetition assignment can be considered random. According to the results, grade repetition not only generates disappointing outcomes for all parties but also entrenches the well-rooted lines of social stratification within schools, eventually contributing to the social gap in academic performance. Although the investigation in Chapter 8 is preliminary, and the reader should thus receive the results with caution, this concluding empirical chapter contributes to the debate over the educational usefulness of grade repetition.

The empirical examination utilised a vast longitudinal database of students that provided information about their academic performance and socioeconomic and migration status (cf. Chapter 4). The primary empirical source of this work was the set of administrative data about individual educational careers that were collected by the National Register of Students (Anagrafe Nazionale degli Studenti). The records refer to the population of upper-secondary students in the last four scholastic years (from 2013-14 to the present) and three most populous Northern Italian regions (Lombardy, Piedmont and Veneto) with almost 600,000 individual students. The National Register of Students gathers information related to end-of-year marks, progressions and grade repetitions, changes of school, transitions in grades and withdrawal rates of high school students.

Data from the register were integrated with those collected by the National Institute for the Evaluation of the Educational and Formative System, which administers standardised tests to the entire student population and receives information about their socioeconomic status and attitudes towards school. The richness of the data represents an innovative development in research on grade repetition in Italy, as it accounts for performance differentials before grade repetition, which in turn allows for more accurate identification of both socioeconomic risk factors and academic consequences.[10]

The theoretical and empirical material that was gathered and synthesised for this dissertation contributes significantly to the debate over the contested practice of grade repetition. In this regard, the present thesis could support the daily work of educators, principals and other workers in educational institutions. As its main contribution, it aims to raise awareness of the interrelation of grade repetition with several social processes that reflect deeply rooted structural inequalities.

The present thesis also intends to add to the educational stratification literature with a systematisation of grade repetition literature as well as a vast, novel empirical analysis of the selection, consequences and impact of this educational practice.

[10] The analysis makes use of two broad families of methodological tools (cf. Chapter 5): multilevel regression methods, which serve to model grade repetition's occurrence and consequences, and matching strategies for causal identification, which assist to estimates the impact of grade repetition on the subsequent academic careers of repeaters.

Chapter 1: Left behind. The social determinants of underachievement

"C'est là la première cause de la non-démocratisation: l'influence du milieu familial sur le développement de l'enfant et, par suite, sa réussite scolaire. […] C'est là la seconde cause de la non-démocratisation: même à égalité de notes, la chance pour l'enfant d'entrer en sixième est en relation avec sa condition sociale."

Girard and Bastide 1963, p. 437-439

1.1 Introduction

Grade repetition strikes underachiever students. With the adjective 'underachiever', I define a kid who fails to meet (*or fails to show to meet*) the learning demands set by schools for a specific grade. On this subject, the primary interest is to unveil the mechanisms behind scholastic underachievement, and their link with grade repetition.

Underachievement could result from both biological and environmental factors. In this respect, there is a vast consensus over the fact that parental (economic, cultural and social) resources strongly influence academic performance. Within the social structure, people experience different constraints in the (economic, cultural and social) educational investments. Underachievement could also arise from modest or poorly informed academic preferences.

On top of the gaps driven by parental resources, ethnicity and migration status strongly affect the distribution of academic performance and educational attainments (Schmid 2001). Migrants and ethnic minorities experience persistent difficulties in the learning process, together with higher failure and dropout rate (Dronkers 2011). Moreover, immigrant students show relational problems with classmates and teachers more frequently (Agirdag et al. 2012).

Of course, the stake affects the outcome: from the perspective of the students, the struggle to reach the learning threshold depends on the threshold itself. Learning thresholds change over grades or tracks and, within general learning guidelines, teachers

have a substantial degree of arbitrariness. For example, learning thresholds change from one school to another. This fact impacts the likelihood to reach the learning threshold and, consequently, to incur in grade repetition. Eventually, thresholds may also change from one teacher to another: teachers' expectations and evaluation practices are essential elements in a student's achievement.

Against this background, the present chapter reviews the research on the socioeconomic determinants of academic underachievement. The next pages display as follows.

The chapter begins with the discussion of the channels easing the transmission of the family's resources concerning the formation and the development of academic ability. As the reader should readily recognise, academic ability is a pivotal element in grade repetition's occurrence. In this course, the review takes into account academic ability as a multidimensional concept, which has both cognitive and non-cognitive components. Then, the reader will find a systematic display of various forms of resources concurring in the pivotal process of ability transmission.

As mentioned above, social groups do not just differ in academic ability, but also their academic preferences. This fact is of interest for this work: differences in academic preferences may lead to a gap in grade repetition probabilities across social groups. To date, there is no theoretical model of academic preferences that can claim its primacy. Much of the effect of family over academic choices still, in facts, unexplained. While the bunch of empirical studies tried to validate the relative risk aversion hypothesis (cf section 1.3.2), they delivered a much-blurred picture.

Somehow to get out of the corner, the last decade witnessed a renewed effort in the estimation of schools and teachers' role teachers in perpetuating inequalities. In this perspective, teachers may directly contribute to an unfair distribution of grade repetition among social groups, beyond the unequal distribution of both academic ability and preferences. A review of the latest developments in this field concludes the chapter.

1.2 Academic performance

Contemporary educational systems enhance and select skills and competences. One of the goals is to sort individuals into distinct social occupations and positions

(Schizzerotto and Barone 2006). However, the dispute is open on which individual traits best guarantee scholastic success.

1.2.1 Cognitive and non-cognitive skills

On the one hand, schools widen kids' cognitive skills by providing them with formal instructions. A minimum set of cognitive skills should include vocabulary, reading comprehension, spelling, capitalisation, math concepts, problem-solving and computation, as well as abstract or mechanical reasoning, visualisation, clerical checking, specific academic knowledge and measure of memory (Farkas 2003). As neoclassic economists argue, cognitive skills associated with productivity are valuable in the labour market. Tasks in economic production require specific content of cognitive skills. Each year of education improves the individual cognitive skills level. This link explains the positive correlation between educational attainment and earnings.

1.2.1.1 The contribution of Bowles and Gintis

Bowles and Gintis (1976) were among the first scholars to show that schooling does not just provide instructions on cognitive skills, which are far from fully explaining the association between the individual educational attainment and economic success.[1] As the authors point out, 'a substantial portion of the returns to schooling are generated by effects or correlates of schooling substantially unrelated to the skills measured on the available tests' (Bowles and Gintis 2000, p.6).[2]

What complements the contribution of schooling on future earnings? According to Bowles and Gintis, schooling fosters specific non-cognitive skills, i.e. behavioural traits that are valuable to employers (1976). The specific behavioural traits rewarded in

[1] Bowles and Gintis' ground-breaking book was primarily animated by the effort to disentangle scholastic and labour market success from immutable genetic factors. The work added to a radical reconceptualization of the notion of ability. Their goal was to call in question the 'technocratic-meritocratic' conception of education, which assumes that schooling provides an objective sorting mechanism for it assigns individuals to unequal economic positions according to their cognitive level. The technocratic-meritocratic approach complements\integrates a theory of inequality based on hereditary differentials in individual ability (seen as intelligence) and socially rooted in social classes (in authors' words the 'poor are dumb' theory). Against this idea, Bowles and Gintis brightly show that the association between length of education and economic success cannot be explained solely by the level of cognitive achievements. The interested reader on early research on non-cognitive skills can refer to Crockett (1962); Elder (1968); Duncan et al. (1972); Sewell and Hauser (1975); Jencks (1979).

[2] They showed that, once the model accounts for cognitive levels, a large portion of income variation remained unexplained in their data.

schools vary across cycles and tracks. Lower cycles and vocational tracks favour traits like obedience and dependability. Instead, teachers in academic tracks and university professors encourage students to develop their creativity and independence (or even aggressiveness).

Bowles and Gintis draw from this hypothesis a Marxist theory of the role of the school on the social stratification. Under capitalism, the educational systems obey a 'correspondence principle'. In brief, the authors believe that in schools, pupils face a correspondence between the instructions they receive, and the behavioural traits required to fill specific class positions.

For example, working-class kids attending vocational tracks shape their behaviour to fit their future task in the production lines. Vocational schools reward obedience while discouraging autonomous initiative. In contrast, academic tracks and universities encourage creativity and even arrogance and aggressivity when they raise affluent kids. Elite schools foster the necessary behavioural traits to endorse future generation's leadership.

In recent papers, while the political emphasis had softened, the authors still have maintained the centrality of non-cognitive skills (Bowles and Gintis 2000, 2002; Bowles et al. 2001a, 2001b), and Bowles and Gintis have framed the correspondence principle in a principal-agent framework. In the interaction between the employer and the employee, employees face high cost in exerting effort from workers for the exact amount of worker's effort is not exogenously contractible. Hence, employees are more likely to choose those applicants that show incentive enhancing preferences. These preferences are '[…] valuable to the employer, and though they are not capacities in the sense that they appear in a production function, they may nonetheless be rewarded by profit-maximizing employers facing a competitive labour market' (Bowles et al. 2000a, p.155). A better orientation to the future and a sense of personal efficacy are examples of incentive enhancing preferences. The first raises the perceived value of keeping the job, while the second directly influences the marginal subjective benefits in exerting effort.[3]

[3] Behavioural skills learnt in schools' functions as a signal to the employer in an incomplete information transaction. Their theory adds on post-Walrasian economics. See Bowles and Gintis 1990.

According to this framework, schooling equips individuals with various incentive enhancing preferences that, in a different field, signal a willingness to exert effort in the workplaces.

To conclude on Bowles and Gintis' contribution, the role of education is twofold. On the one hand, schooling imparts the necessary technical and cognitive skills to meet jobs' tasks. On the other, it also provides future workers of behavioural skills valuable to employers.

1.2.2 Categories of non-cognitive skills

Since the 70s, Bowles and Gintis' innovative line of research has flourished in a broad investigation on non-cognitive skills. Dalton and colleagues (Rosen et al. 2010) define non-cognitive skills as 'academically and occupationally relevant skills and traits that are not specifically intellectual or analytically in nature' (p. 1). Dalton classifies non-cognitive skills in six broad and non-mutually exclusive categories: motivation, effort, self-regulated learning, self-efficacy, academic self-concept, anti- and pro-social behaviour, and coping and resilience.

1.2.2.1 Motivation

Motivation sketchily conceived is the desire to achieve something; in the field of education, it appears as 'achievement motivation' as the desire to obtain academic success. Notably, the reasons behind motivation can be either instrumental (e.g. a cost-benefit calculation) or affective (e.g. a status goal). The two underlying sources of motivation are somehow oppositional: internal interest or enjoyment counterposed to the value of external rewards. Scholars who stress this opposition focus on intrinsic *versus* extrinsic reward of action (Rauber 2007; Covington 2000; Weiner 1987). Others (Wigfield and Eccles 2000) add on intrinsic\extrinsic motives to develop the so-called Expectancy-Value Theory, by which the likelihood of success itself determines the value that agents attach to success. According to this theory, agents are more prone to engage in positive behaviours to reach a determined goal. When the attribution of this behaviour is internal, the sense of self-efficacy increases.

The Achievement Goal Theory (Pintrich 2000) explicitly refers to the educational environment as it focuses on educational goals. In this respect, educational goals can be

either performance or mastery related. Performance goals push students to show their merit to judging others (e.g. to teachers), whereas mastery goals encourage kids to master skill- or subject-related competencies. According to the Achievement theory, mastery goals are more engaging than performance goals, and hence they provide higher learning levels.

1.2.2.2 Effort

The effort is the behavioural manifestation of engagement defined as kids' 'energises, enthusiastic, emotionally positive, cognitively focused interactions with academic activities' (Kindermann 2007, p.1186 in Rosen et al. 2010). Scholars have identified two crucial empirical dimensions of effort, namely, its degree and specificity. The degree refers to the amount of energy exerted and it measures the procedural (simple compliance) or the substantial (genuine commitment) engagement with the task. Scholars detect the lack of (procedural) effort by a range of non-compliance behaviours, such as school non-attendance, delays, homework non-completion and class daydreaming.

1.2.2.3 Self-regulated Learning

Although partially cognitive, self-regulated learning has a fundamental social-behavioural character: it signals the ability of a student to 'evaluate tasks, review the strategies available to them for accomplishing the tasks; apply themselves to completing the tasks; monitor the effectiveness of their strategies; and […] revise their model for approaching similar tasks in the future' (Rosen et al. 2010, p.69). Volition is a fundamental element in self-regulated learning. Self-regulated learning is not easily distinguishable from other non-cognitive traits (e.g. motivation), although, differently to motivation, it develops with age (Rauber 2007).

1.2.2.4 Self-efficacy

As initially defined by Bandura (2010), 'perceived self-efficacy is defined as people's beliefs about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives' (p.1). Self-efficacy has been a cornerstone in social psychology. Bandura (*ibidem*) describes four ways to acquire self-efficacy: through mastery or successful experiences; through the observation of similar others reaching success; through social persuasion, intended as the boost from others;

finally, the alteration of the mood and the limitation of adverse reactions ease self-efficacy perception.

1.2.2.5 Academic self-concept

Academic self-concept emerges in the interaction with the academic environment, as the self-perceived academic ability (O'Mara et al. 2006). Assessing the causal path that relates academic self-concept and the broader effect of the interaction with significant others is challenging. For this reason, many scholars reckon some degree of recursive circularity between the two.

1.2.2.6 Coping and Resilience

Copying and resilience refer to the set of behavioural responses and the individual strategies that people put in practice in costly-resources and investment-demanding environments. While coping embraces the intentional responses to solve specific situations, the concept of resilience points at the adaptation skills of an individual. As Masten (2001) maintains, coping requires a set of skills, while resilience refers to the successful exploitation of these skills.

1.2.2.7 Anti-social and pro-social behaviour

Physical and non-physical act of violence, aggression and dismissive behaviours, such as ignoring and refusal to cooperate, fall in the category of anti-social behaviour. At the other pole of the spectrum of behaviours, pro-social behaviours include positive social feeling or desire of others' inclusiveness. Pro-social behaviour correlates with many positive educational outcomes (literacy comprehension, school completion, friendships, peer acceptance, and occupational status).

To sum up, the ability is multidimensional as it embraces both mere cognitive skills and complex behavioural skills. However, how does ability develop? To what extent are socioeconomic factors likely to influence ability formation and scholastic performance?

1.2.3 'Poor are dumb' theories

A set of hypotheses attribute the performance gap across social groups to a supposed socio-cultural handicap experienced by poor or disadvantaged people. For

example, Lewis (1966) described the culture of poverty as ‘a subculture of Western society with its structure. [..] A way of life handed on from generation to generation along family lines’ (p.19).

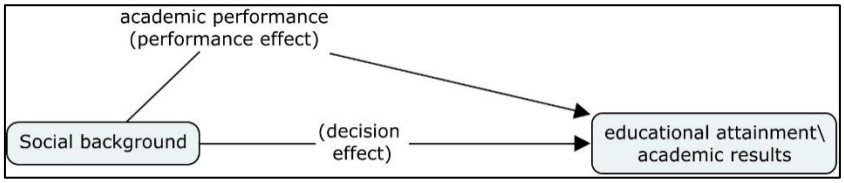
The culture of poverty is not a mere lack of resources or opportunities but ‘a ready-made set of solutions for human problems which serves a significant adaptive function’ (*ibidem*). Over and beyond the structural constraints that have produced the unfortunate circumstances, the culture of poverty was a real obstacle for social change: ‘By the time slum children are age six or seven, they have usually absorbed the basic values and attitudes of their subculture and are not psychologically geared to take full advantage of changing conditions or increased opportunities which may occur in their lifetime’ (p.21). The underlying assumption is that low socio-economic status families suffer a socio-cultural handicap (Blanchard and Cayouette-Remblière 2016).

‘Poor people are dumb’ theories have been widely criticised as by-defect-explanations that renounce to advance any hypothesis upon the mechanisms of transmissions of inequalities (Bautier et al. 2000). While they exonerate academic institutions, the socio-cultural handicap assumes high behavioural and preference homogeneity among individuals belonging to the low socioeconomic positions. As a matter of facts, the poor are dumb theories neglect the poor the use of rational reasoning.

1.2.4 Mechanisms of intergenerational transmission of ability

In contrast to socio-cultural handicap hypotheses, a significant amount of evidence suggests that social background’s influence on educational attainment is twofold (Figure 1.1). On the one hand, social background exerts an indirect influence via academic performance. Thus, on average, students with lower (higher) background achieve poorer (better) academic results. On the other, social background direct influences attainment and students with lower (higher) social background could be more (less) prone to give up studying. Literature refers to the first as primary (or performance) effect, whereas it labels the second as secondary (or performance) effect (Boudon 1973).

Figure 1.1 The twofold influence of social background on educational attainments.



1.2.4.1 The contribution of Boudon

Scholars have acknowledged the twofold role of family's background since the early Forties.[4] However, the fortune of the concepts of primary and secondary effects in the scientific community is owed to Boudon's *L'inegalité des chances* (1973) -published in English the following year as 'Education, Opportunity, and Social Inequality'.

Boudon speculates on the individual actions that generate the aggregate outcomes. People, Boudon claims, act in a structured social space, in which they struggle to take up scarce social positions and labour market occupation. The educational qualifications are the strategic tools by which the individuals grant themselves better chances to reach the high-status social position or high-rank occupations.

Boudon develops two main propositions to explain the emergence of educational inequality. The first is that the class structure in the broad society shapes the distribution of chances of success for each educational step. The second is that, at any specific turning point in educational paths, educational choices are functions of one's social position, perceived costs, expected advantages and risks. Academic performance strongly impacts on expected risks: the lower it is, the higher is the risk to invest in education.[5]

Scholars in Sociology have often disregarded the identification of generative mechanisms of primary effects. A notorious example is the 1997 Breen and Goldthorpe's quotation in which they prefer 'not take up the vexed and complex question of the extent

[4] See: Girard, Alain, and Henri Bastide. "La stratification sociale et la démocratisation de l'enseignement." *Population* 18.3 (1963): 435-472.

[5] Here it is worthy of quoting Boudon's words: « Pour un fils de cadre supérieur, atteindre le niveau du baccalauréat plutôt que celui de la fin du premier cycle du secondaire est plus "utile" que pour un fils d'ouvrier. Dans le premier cas, le risque de démotion (sociale – ed.) est élevé. Dans le second cas, la promotion (sociale – ed.) peut être assurée même avec un niveau d'instruction plus bas. De même, il est évident qu'il est plus coûteux pour le second individu que pour le premier d'atteindre le niveau le plus élevé plutôt que le plus bas. » (ibidem, p.110).

to which they are genetic, psychological or cultural in character' (Breen and Golthorpe 1997, p. 277). However, the very concept of performance effects tells very little on the causes or mechanisms that put the social background influence on performance to work (Nash 2003).

1.2.5 Generative mechanisms of primary effects

The last decade has witnessed some crucial steps towards a comprehensive theory of inequality should explicitly address the generative mechanisms of performance effect.

Primary effects arise from a complex interaction between the educational institutions and household resources (broadly understood). Jackson (2013) has identified six major classes of mechanisms: genetic, health and nutrition, sibship size, the home environment, psychological mechanisms and cultural biases in school.[6]

1.2.5.1 Genetic endowments

Undoubtedly, there is growing interest in the role of genetic endowments in determining educational attainments. This section recalls the main contributions briefly in this area of research.

Genetics have cyclically acquired legitimacy to the detriment of the environment as the leading cause of social inequalities. In the debate over schooling democratisation, the tension between biological and sociological explanations has heated up each time inclusion policies have failed in considerably eroding educational inequalities.

One among the fiercest critics of the educational inclusion, Arthur R. Jensen, bluntly affirmed in 1969: 'The compensatory education in the US has been tried and it has failed' (Jensen 1969, p.2). Jensen remarked that the impact on students' IQ of educational interventions had been minimal and this fact should cast severe doubts to the so-called 'deprivation hypothesis'.

While the deprivation hypothesis blames the social and economic realm for the gaps in educational attainments, Jensen believed that the failure of educational reforms

[6] For the limited scope of this work, I avoid discussing mechanisms related to health, nutrition and sibship size; instead, psychological mechanisms are intimately related to the variety of students' reactions to the school and class setting, and I will discuss their implication for the stratification of scholastic attainments in Section XX of the present chapter.

rooted in the genetic distribution of ability. Therefore, socioeconomic and ethnic gaps in educational attainment are biologically-driven.[7] Although environmental factors might act as a threshold in extreme cases of deprivation, they can at most depress individual innate potential.[8]

Jensen has been the first to borrow the term heritability from zoology and botany (Feldman et al. 2000). In genetics, heritability is a specific phenotype's proportion of variation due to vertical transmission, as opposed to the proportion of variance due to the environment.

Until the middle '70s, heritability in human beings had been studied mostly on small samples of monozygotic twins, as they share the total genetic variation.

The distinguished Italian scholar, Luigi Luca Cavalli-Sforza, in a 1973 paper with Marcus Feldman started to point out that small samples of monozygotic twins were likely to provide biased estimations. For example, the estimation of IQ's heritability in twins was nearly 80% (Cavalli-Sforza and Feldman 1973). The larger the sample size, the much moderate effect of genetic endowments (Cloninger et al. 1979).[9]

Cavalli-Sforza and Feldman claim that, in humans, genes are just one of the channels through which offspring inherits traits as IQ from parents and that cultural transmission, besides genetic endowments, must play a role in the transmission of intelligence. In human heritability, both parental phenotypes and parental cultural values are involved.[10] In this respect, IQ's heritability can be decomposed in three main

[7] For example, Jensen writes that it is "not unreasonable hypothesis that genetic factors are strongly implicated in the average Negro-White intelligence difference" (ibidem, p.82).

[8] More recently in 1994, Herrnstein and Murray resume this conception. The two authors claim that intelligence (measured by the US Armed Force Qualification Test- AFQT- at age 15\23) is by far the most important predictor for academic and economic success (Herrnstein and Murray 1994; Dunifon and Duncan 1998). They find a strong association between the score in the AFQT test and income, high school dropout, unemployment, children out of wedlock, have ever been on welfare, had a low-birthweight or low IQ score child. The book has received many critics. Heckman (1995) casts doubts upon Herrnstein and Murray's intelligence measurement, which ignores 15 to 23 years of environmental influences. Koreman and Winship (in Arrow et al. 2000) point out the inadequacy of their parental SES measurement, which is highly correlated with AFQT, and the fact that Herrnstein and Murray neglect the effect of school on income.

[9] Being the heritability a statistical construct which by no means refers to the actual causal link between genes and phenotypes, it is severely sensitive to its statistical specification.

[10] Additionally, the authors claim that models of heritability should also consider the effect of assortative mating, that is, the tendency of human beings to mate with phenotypically or socially similar individuals.

components: genetically transmitted, culturally transmitted and non-transmitted environmental (Feldman et al. 2000).

The availability of DNA data in the last decade has exponentially increased, encouraging a deeper understanding of the dynamics between genes and specific behavioural traits. Recent work exploits the whole human genome and looks for statistical associations between regions of DNA and behavioural phenotypes. Since each identified region (termed 'locus') conveys a minimal influence over a specific phenotype, analysists compute phenotype-specific polygenic scores, that are, the sum of the traits-associated alleles across genetic loci (Tropf et al. 2017). Up to date, polygenic scores have produced promising results in the field of reproductive behaviour. Meanwhile, the potential genetic contribution over educational attainment remains somewhat obscure.[11]

On the subject, Paige Harden et al. (2019- preprint) find a moderate genetic association with maths courses' completion on a USA sample 3635 young individuals.

Nevertheless, they get a substantial mediating effect of both parental socio-economic status and high school socio-economic composition. The authors also find that privileged schools can better sustain genetic potential (cumulative advantages) and that high parental socio-economic status compensates for lower polygenic score (resources substitutions). The two mechanisms of cumulative advantages and resources substitution contribute to bias STEM attainments in favour of privileged kids.

To sum up, contrary to the early works in this field, the latest evidence calls for a renewed primacy of the environmental and social factors, which severely enhance or suppress the (somewhat limited) influence of genetic endowments on human social and complex behaviours (such educational performance and investments).

If biological molecules explain parental influence to a limited extent, a complete 'family resource theory' (Farkas 2000) should instead embrace the whole set household assets that are likely to impact offspring's performance. The next sections review the three main classes of family resources, namely, economic, cultural and social.[12]

[11] For further reading on this topic visit the publication page of the Sociogenome project at this link: http://www.sociogenome.com//publications/

[12] In accomplishing this, the review limits to those parental practices and resources that influence the ability and hence exert a primary effect on academic performance. Eventually, the last section of this

1.2.5.2 Economic resources

The economic resources that parents can mobilise in the nurturing exert a substantial influence on children's academic performance. Parents' economic investments such as private schooling, extracurricular activities, and investments in educational enrichment goods, positively modify the learning environment and contribute to children's educational development (Covay and Carbonaro 2010).

Differentials in parental investments are likely to produce performance gaps already at early stages in children's life (Farkas and Beron 2004; Klooserman 2011), and particularly in oral vocabulary knowledge (Farkas and Beron 2004). For example, Kornrich (2016) underlines the decisive role of pre-school activities and enrichment goods and activities.

Parental social class severely impact on credit constraints when investing in education. Kaushal et al. (2011) show that when income increases, educational expenditures (e.g. in items of educational enrichment and time-use) grow consequently. Spells of parental unemployment negatively impact on children's attainments (Levin 2011; Micklewright et al. 1990).

Households income trends also affect academic performance. As Heckman and colleagues argue, changes in investments price alter the dynamics of skills formation: 'investment in any form of capital entails forward-looking behaviour comparing current costs to future returns' (in Arrow & Bowles 2000, p.231).

Families that experience harsh labour market conditions are less likely to provide for educational investments (Sherman and Harris 2012). Schneider et al. (2018) find that the higher the income inequality at state-level, the wider the gap in parental financial investment.

Two empirical studies from the US have suggested that the gap in educational investments has widened in the last decades. Kornrich (2016) claims that, between 1972 and 2010, higher-income families had increased their spending, while little had changed at the bottom of the income distribution. Alike, Reardon (2011) shows that, between

chapter (devoted to decisions and preferences) completes the task by including parental influences in determining secondary effects.

1976-2001, the gap in educational investments between the rich and poor is 30%/40% larger in the latest cohorts. The only exception would be the investments in early childhood, which seems to have narrowed down (Bassok et al. 2016).

In the US, other scholars have explicitly addressed the impact of the Great Recession on educational investments. Ananat et al. (2011) advance that community-level job losses undermine standardised test score achievements. The effect is more massive for older and average performing children. Both structural (economic) and psychological (stress and spillover effects) mechanisms might have provoked the performance drop. Lunn and Kornrich (2018) analysed the changes in the household's educational expenditure between 2005 and 2011, and they claim that pre-existent income gaps have consistently widened.

1.2.5.3 Cultural resources

Academic performance is also the product of cultural endowments. Lamont et Lareau (1988) define cultural capital as 'high-status cultural signals (attitudes, preferences, formal knowledge, behaviours, goods and credentials) used for social and cultural exclusion' (p.156). The high-status cultural signals are historically and geographically determined and, as such, they are likely to evolve and change over time.

Unlike economic capital, cultural capital is not 'immediately and directly convertible into money' (Bourdieu 1986, p.47). According to Bourdieu (*ibidem*), cultural capital's uneven distribution between classes explains the 'specific profits which children from the different classes and class fractions can obtain in the academic market'.

Many scholars are uncomfortable with the notion of cultural capital (and with that intimately linked to the *habitus*). This ambiguity might derive from the various functions that cultural capital plays in Bourdieu's thought. As Lamont and Lareau (1988) remark, 'cultural capital is alternatively an informal academic standard, a class attribute, a basis for social selection, and a resource for power which is salient as an indicator/basis of class position' (p.156). In the educational field, cultural capital performs at once as a resource, but also as a criterion of selection and elimination, and as a preferences' structuring factor. This section addresses cultural capital to its limited role as a resource.

Bourdieu identifies cultural capital in three primary forms (1986), namely, objectified, embodied and institutionalised. In its objectified form, cultural capital is directly visible and consumable. The objectified cultural capital consists of cultural goods, such as books and paintings. Notice that the property of cultural goods is not alone sufficient, for the bearer must appropriately master them.

The cultural capital in its embodied form is, for its part, acquired with significant long-lasting effort, and it is context-specific. Bourdieu calls *habitus* the device that encloses and put to work the embodied cultural capital. The *habitus* is, essentially, a set of behavioural predispositions that gives the 'feeling' for the game within a social field.

Drawing upon Bourdieu, Swidler (1986) has argued that *habitus* coincides with culture, to the extent it shapes the way people organise their action and their strategies to coupe within a specific structural frame. Rather than determining individuals' goals, the culture (or the *habitus*) gives them a toolkit of actions. Under this light, the culture is a nothing else than a repertoire of behaviours from which actors select and construct *ad hoc* lines of actions. Social groups act differently within the same structural setting by drawing their strategy from separated cultures.[13]

Eventually, the institutional form of cultural capital assures the preservation and transmission of accumulated cultural capital. Educational qualifications distinguish people from autodidacts. Schooling offers a 'conventional, constant, legally guaranteed value concerning culture and social alchemy' (p.50).

Many empirical studies have tried to estimate the effect of cultural capital on academic performance. Ashaffenburg and Maas (1997) have investigated the role of highbrow culture on scholastic success in the US. The effect is substantial, although it has recently declined. Sullivan (2001) has found a significant contribution of parental cultural capital to the UK General Certificate of Secondary Education score. Cultural capital is an essential factor in determining educational attainments across OECD

[13] Swildler (1986) presents two models of cultural influence. In settled environments or lives, culture and structural circumstances reinforce each other, and it is hard to disentangle the two, to such an extent that the observer might assume cultural values determine actors' action. In uncertain environments or lives, culture is less able to sustain strategies: ouverte ideologies govern action, but the structural opportunities select the new fitting line of actions (strategies).

countries, although it only partially explains the link between social class and school success (Barone 2006).

1.2.5.4 Social resources

Families endow offspring of a set of social relations. Social resources consist 'in some aspects of the social structure, and they facilitate certain actions of individuals who are within the structure' (Coleman 1994, p.302).

Social capital is not fully fungible by an individual and, alike cultural capital, is valuable in respect of specific activities. Social capital inhabits the structures of relations and hence does not reside in (neither is owned by) individuals.

Parents and schools, pictured as a network of relations, are likely to contribute to the accumulation of the child's social capital (Helliwell and Putnam 1999).

1.3 Academic choices

On average, low socio-economic status families are more likely to make less ambitious choices. Secondary (or decision) effects embrace all those mechanisms that link social class to educational attainment through academic choices on top of achievement differences. Nevertheless, the identification of such mechanisms is the object of a long-lasting scientific dispute.

1.3.1 The individual function of educational decision

A minimum model for individual educational choices should acknowledge three fundamental parameters: the academic performance, the costs and the socioeconomic profitability of each alternative decisions (Gambetta 1987; Erikson and Jonsson 1996; Breen and Goldthorpe 1999; Barone 2019).

Individual academic performance influences academic choices to the extent that it defines the expected chances of success. The unequal distribution of academic performance across social classes consequently defines the unequal distribution of chances of academic success.

Costs are both direct (*e.g.* school fees) and indirect (*e.g.* books and opportunity costs). Families face different budget constraints, which limit their setting of educational opportunities and contribute to educational stratification.

The third pivotal parameter of socioeconomic profitability represents people's evaluation of the educational alternatives in light of the entailed socioeconomic risks and gains.

While this broad definition allows identifying a common epistemological ground by according legitimacy to individual agency, it also accommodates a plurality of different (and often incompatible) hypotheses.[14]

The exact determination of the micro-level mechanisms that determine the socioeconomic profitability (*i.e.* the exact 'function' of educational decisions) remains a source of dispute within scholars.

1.3.2 Socioeconomic profitability: a struggle for definition

Neoclassical economists usually define socioeconomic profitability in expected monetary returns of education. In their view, families can accurately forecast income gains from each specific educational choice (Becker 1964; Blundell 2001; Checchi et al. 2006, 2010; Cameron and Heckman 1998; Carneiro and Heckman 2003).

[14] At a careful look, the results derived by assuming complete economic rationality are of little difference to those of 'pushed-from-behind' view (Gambetta 1987). Causal rather than intentional forces lead choices; both preferences and constraints are exogenous. People are expected to obey to standardised procedures of market-type egoistic evaluation while facing any decision of their lives. As Elster (1983) underlines, while push-from-behind theories accord no importance to preferences, neoclassical economics tend to consider them as given and fixed. Differences in the course of actions are changes in the opportunity set given mutable constraints, with no change occurring in the ranking of preferences. (See Elster 1984, 1993 for a classic review of the issue). During the last decades, the accumulation of controversial experimental evidence called into question the neoclassical paradigm of the rational choices (for a review refer to Bowles 2004). Latest developments in (behavioural) economics assimilate a more psychosociological view of individual choices. Within this field of research, post-Walrasian accounts (Bowles and Gintis 1990, 2000) allows both preferences and constraints to emerge from the set of actions endogenously. For example, prospect theory's advocates (Kahneman and Tversky 2000) advance four crucial adjustments to the conventional expected utility paradigm. First, that people overweight the importance of losses, and hence do not act accordingly to the standard expected utility paradigm. Breen and Goldthorpe acknowledge these developments. Second, the way choices are framed impact on preference ranking. Third, they weight alternatives with substantive and subjective measures. Fourth, their model deals with situation-dependent preferences. Preferences emerge from the setting of the action, *i.e.*, individual may rank alternatives differently according to the situation they experience. Such widely observed preferences contradict the standard version of expected utility theory, which would claim this kind of preferences to be inconsistent and, hence, irrational.

However, as noted by Akerlof and Kranton (2002), educational decisions are indicative of social decisions. Sharing this insight, sociologists have favoured the investigation of the link between educational choices and social reproduction. Collins argued that educational attainment is the outcome of 'efforts of competing status groups to monopolise or dominate jobs by imposing their cultural standards on selection processes' (1971, p.1002).

Boudon (1974) is of the same position when he claims that people's struggle for scarce social positions rests upon formal schooling. Educational credentials are a crucial tool through which elites secure their position in socio-economic structures. In Boudon's view, parents aim at minimising their offspring's downward mobility risk (or, put the other way around, parents aim at maximising their offspring's chances to secure their actual social position). Breen and Goldthorpe tried to formalise this hypothesis within a marginalist framework (1997; Breen and Yaish 2006).[15]

Their crucial assumption is that families are risk-averse when it comes to educational investments. Given that the risk of social demotion is relative to each specific situation in the social structure, the higher a kid starts, the higher the qualification she\he needs to secure her\his social position; the higher the risk of downward mobility, the more compelling the necessity to attain prestigious educational titles. The result is that other things being equal (*i.e.* costs and academic performance at time *t*), a working-class kid would be happy with a less ambitious choice at time *t+1*, whereas a middle-class kid would need to undertake more ambitious paths if he or she does not want to drop in the social hierarchy.[16]

Breen and Yashi (2006) try to reconcile the relative aversion hypothesis with Kahneman and Twersky's prospect theory paradigm (1979). Following this paradigm, they suggest that individual academic preferences shape around a reference threshold. This 'reference states' influences the way people look at social mobility (distances might be long or short, depending on the actors' perspective). Moreover, Breen and Yashi's

[15] Breen and Goldthorpe look at the marginal utility attached to a specific educational choice differs across social groups: rich people prefer to invest more in education, i.e., make more ambitious choices if compared to needy families.

[16] The model includes other minor assumptions. E.g., Breen and Goldthorpe also assume that losses have negative impact on the probability to get access to higher social classes, i.e., staying at school and failing increases the odds to end up in the lowest social strata.

formalisation complements the loss aversion hypothesis, for which losses weigh more than gains in individual decisions.

1.3.3 Inconclusive empirical evidence

Up to date, the empirical evidence for the Breen and Goldthorpe hypothesis is mostly indirect (Barone et al. 2018). As Breen and Yashi put it, 'to operationalise the assumed common (among students from all class origins) beliefs about the returns to educational attainments. […] it turns out that our results depend on which particular source of information about the class returns to different levels and types of educational attainment we use' (p.27).

Therefore, both Breen and Yaish (2006) and Davies et al. (2002) have tested the model exploiting indirect measurements of socio-economic profitability (or returns) of education and find favourable results.

Instead, Van de Werfhorst and Hofstede (2007) have investigated the effect of class maintenance and upward mobility concerns on next school choice and schooling ambitions by the age of thirty in five secondary schools in Amsterdam.

Their results show that children are equally concerned about downward mobility, regardless of their background. Moreover, downward mobility concerns do not exert any significant effect on scholastic performance, which is instead determined by parental cultural capital. The same is true for the association between RRA and the next school choice. In partial support of the Breen and Goldthorpe hypothesis, the results show that students who are the most worried about downward mobility also display the highest educational ambitions.

Stocke (2007) has found mixed evidence by testing Breen and Goldthorpe's predictions on a longitudinal dataset of German kids. The author measures the impact of a subjective measure of the importance of status maintenance on the secondary school choice. As it turns out, the perceived chances of success exert a much stronger effect than the status maintenance motives.

The author admits that, in his data, Breen and Goldthorpe have failed to explain the direct effect of parental social class, which remains substantially unexplained when

the model accounts for status maintenance motives. In the author's words, although their hypothesis 'is valid in assuming the parents' desire to avoid status demotion to be a relevant determinant, […] the theory falls short of explaining the secondary effects of class in educational decisions' (*ibidem,* p.516).

Gabay-Egozi et al. (2009) have tested the status maintenance hypothesis in curriculum choices in Tel Aviv secondary schools. The authors have found that, while the marginal utility and perceived chances of success played a role, status maintenance motives were not influential in the curriculum choice.

In line with the previous research, Barone et al. (2018) have found that status maintenance role in tertiary education decisions has quite limited importance among Italian students. Relying on a longitudinal dataset of students from different Italian high schools, the authors measured university aspirations and, in a subsequent collection, the actual enrolment decisions. They have shown that, while secondary effects explain approximately two-thirds of the social background differentials in higher education enrolment, both income and status maintenance motives only explain them to a limited extent.

Overall, the accumulated direct empirical evidence casts consistent doubts upon the explanatory power of status maintenance motive in explaining decision effects. Differences in educational choices are still mostly unexplained, and there is substantial room for alternative causal mechanisms. For example, cultural reproduction theories still offer competitive and compelling explanations (*e.g.* Bourdieu's *habitus*).

Another source of inequalities is, however, represented by the role of schooling in influencing educational preferences. The next and last section of the present chapter reviews the evidence on this subject.

1.4 Setting the threshold

As recalled previously, Jackson (2011) encompasses the cultural biases exhibited by schools within the plethora of mechanisms responsible for the primary effects. In other

words, the effect of the academic environment exert varies according to children's social background.[17]

Conflict theorists first called in question the content of the scholastic instructions and the system of evaluation as socially biased. Bourdieu and Passeron (1970) and Bourdieu (1977; 1984) analyse the school content concerning the mechanisms of social allocation. According to their seminal contribution, the school acts as a conservative force as it pushes individuals to acquire a specific world's view.

The world's view propagated by the school is a direct translation of the values of the dominant social classes. As such, schooling is a form of symbolic violence because it promotes the recognition of its legitimacy through the misrecognition of the relations of power that itself entails. In Bourdieu's view, educational institutions exclude the students whose habitus is incompatible with the dominant explicit and implicit knowledge. The distance from the academic culture and the different disposition to acquire it is at the root of social selection in schools.

The heterogeneity in behavioural instructions across schools is at the core of the groundbreaking analysis by Bowles and Gintis (1976). Schooling fosters differentiated cognitive and behavioural schemes, for it channels students into social positions.

In Collins (1979), the values transmitted by the curriculum tend to justify and preserves the dominant groups' structural position. Similarly, the concept of the 'hidden curriculum' (Apple 1990) underlines the social selection over students' familiarity with the implicit knowledge transmitted by the academic curriculum.

Within a more institutionalist view, the academic curriculum is also a critical factor in the building of shared knowledge and understanding (McEneaney and Meyer 2000). With the massification of the school system, academic curricula face contradictory goals: on the one hand, they must accommodate the universal expansion of education; on the other, they must not stop to legitimate the social selection and the differentiation.

[17] Nonetheless, the collocation of factors related to schooling's contribution to scholastic achievements is not clear-cut (Barone et al. 2018). Schooling exerts an effect on ability (performance effect) that, in turn, acts on perceived chances of success and opportunity costs and, ultimately, it mediates subsequent choices (decision effect).

1.4.1 The institutional setting

Curriculum tracking and ability grouping are key structural factors in determining future learning outcomes (Carbonaro 2005). For example, students' effort significantly differs between academic and non-academic tracks. Even controlling for academic performance, non-academic tracks are likely to harm students' subsequent achievements (Farkas 2011). Gamoran and Mare (1989) have reached similar conclusions by focusing on maths scores, the probability of graduating and the probability to drop out. The authors claim that tracking reinforces prior gaps in achievements. Others (Hallinan et al. 2001; Tach and Farkas 2006) have collected convincing shreds of evidence on the impact on the achievement of ability grouping.

Bol and van de Werfhorst (2013) highlight the existence of a trade-off between labour market allocation and equality of opportunities. While tracking enhances a better labour market match, it accentuates the role of parental resources in educational attainments.

The literature consistently shows that early tracking brings about higher levels of inequalities in attainments (Geven, Batruch & van de Werfhorst 2018). By reviewing the international literature, Van de Werfhorst and Mijs (2010) concludes that highly horizontally stratified educational systems contribute to the dispersion of student test scores and educational opportunities. When central examinations guide the allocation of students within tracks, the social background effect on attainments is, however, alleviated (Bol et al. 2014). Across time, socioeconomic inequalities (among eight-graders) more strongly narrow in countries that have switched to more comprehensive (less tracked) systems and the reduction has been more pronounced at the top of the achievement distribution (van de Werfhorst 2018).

The tools of evaluation may additionally reinforce achievement gaps across social classes (Autin et al. 2018). The fundamental tension between education and selection, typical of modern educational systems, is reflected in the system of assessments, which is likely to emphasise one of the two functions.

'Normative' assessments display individual merit and rank students within the class, enforce the comparison with others, and usually consist of alphabetic or numeric

scale. In contrast, ‘formative' assessments are detailed commentaries that give a student useful feedbacks to allow her\him to overcome eventual difficulties.

While it is shared beliefs (Autin et al. 2015), especially among underprivileged groups (McCoy and Major 2007), that a normative grading system alone fulfils equity and fairness standards, formative assessments have appeared to be more equitable. Indeed, by relying on experimental data, Autin et al. (2018) show that normative assessment tends to artificially increase the social achievement gap by encouraging the goal of selection on the side of the evaluator\teacher. The authors conclude that the ‘institutional practices of assessment constrain the way individuals in a position of evaluator behave toward students from the lower or higher social class’ (p.3). [18]

1.4.2 Teachers’ effect

Teachers are essential inputs in achievement. They are also likely to contribute to shaping attitudes towards educational attainment. The literature has approached the teachers’ effect in two main ways (Blanchard and Cayouette-Remblière 2016).

On the one hand, scholars have investigated the impact of the character (and, supposedly, the quality) of the interaction teacher-students on students’ academic performance. On the other, scholars have stressed the crucial role of both educational practices and shared ways of thinking at the institutional level in determining academic performance.

1.4.3 Expectations

A rich literature has pointed out that expectations on the side of teachers are in general accurate, that is, they correctly predict future performance (Jussim et al. 2009; Egan & Archer 1985; Good 1987; Pedulla, Airasian, & Madaus 1980).

[18] Refer to Kohn (1999; 2011) for a comprehensive review of the critical literature on grading practice. According to the author, grades harm the intrinsic interest of students in the subject itself and in learning itself. Normative grading is likely to enlarge social inequalities in academic achievements. For example, Croizet and Claire (1998) claim that when tests explicitly address the cognitive ability, stereotyped groups show apprehension and underperform. Additionally, grading undermines the relationship with the teacher and between students.

In turn, teachers' expectations are also likely to influence students' performance directly. For example, Cayouette-Rembliére (2016) has found that people, even many years after, usually can recollect teachers' judgements at the time they were students.

Teachers' expectations act on performance in at least three ways. First, teachers can modify the way they teach according to their expectations toward students. Second, low expectations on the side of teachers might stimulate emotional responses that directly harm performance or jeopardise students' identification with educational environments (Steel 1997). Third, students are likely to comply with teachers' expectations and perform accordingly (Ferguson 2003).

The last two mechanisms resemble Merton's self-fulfilling prophecy. Quoting Merton (1948): 'men respond also to the meaning this situation has for them, not only to its objective features. Public definitions of a situation (prophecies or predictions) become an integral part of the situation and thus affect subsequent developments. The prophecy led to its own fulfilment' (p. 194).

Rosenthal and Jacobson (1968) have referred to the self-confirming expectations on the side of teachers as the Pygmalion effect.[19]

In their classic experiment, Rosenthal and Jacobson randomly assign students in a class to two groups, and they instruct the teachers that the (randomly pick) students in the one group had an above-average IQ. At a one-year distance, the two authors showed that, regardless of their actual IQ score, students in that group had better results compared to their peers. Rosenthal and Jacobson conclude that teachers' expectations have a substantial effect on students' performance.

The school's activities that the most serve to a logic of selection (*e.g.* recommendations, decisions over grade repetition, and admissions) are particularly sensitive to expectations, and teachers' decisions necessarily suffer from a degree of uncertainty (Darmon 2012). The teachers' power of 'divination' is the search for the 'academic energy' of students and it goes beyond the observed performance (*ibidem*, 'divination professorale' in the original French text). When class, gender, racial or ethnic

[19] In the Metamorphoses, Ovid tells the myth of Pygmalion, a sculptor who falls in love with a statue he modelled. By persuading himself that its masterpiece was a real woman, Pygmalion eventually sees it coming to life.

biases are in place, teachers' action can have long-lasting impacts in reinforcing inequalities.

The empirical investigation on biases in teachers' expectations related to students' socioeconomic status has produced consistent findings (Geven et al. 2018). Irrespectively to the actual academic performance, teachers generally evaluate students from disadvantaged background less positively than their wealthier peers.[20]

Expectation biases consistently influence teacher track recommendations in Europe. Teacher track recommendations under-estimate low-status kids in France (Duru-Bellat 2002), Germany (Ditton et al. 2005), Netherlands (Timmermans 2016) and Italy (Abbiati et al. 2017).

In Italy, Romito (2014) carried out an ethnography in three Italian middle schools and has shown that teachers tend to discourage low-SES kids from enrolling into academic tracks disproportionally. The interviewed teachers believe that their underprivileged students would not fit in academic tracks and that the costly effort would be overwhelming.

The (consistent) gap in evaluations is generally small. Nonetheless, teachers might hold prejudiced perceptions of students' ability (Geven et al. 2018). For example, Boone et al. 2013 have shown that teachers in Flanders emphasise specific characteristics that are typically weaker in low-SES kids, such as self-reliance, planning capacity and punctuality.

Empirical results are instead ambiguous in the case of ethnicity and immigrant students in the European context (Geven et al. 2018). Some experimental studies in Germany (Sprietsma 2013; Glock & Karbach 2015), and Russia (Raisa and Alieva 2016) found different teachers' expectations among ethnic groups. On the contrary, an experiment on Dutch teachers did not reveal any significant bias (van Ewijk 2010).

Observational studies have produced controversial findings as well. By comparing standardised test scores and teachers marks in Italy, Triventi (2019) shows a systematic

[20] The chapter does not deal with the influence of gender on teachers' expectation. For the interested reader, Carlana (2017) finds that gender-biased teachers tend to have a significant impact on female STEM achievements and self-confidence.

downward evaluation when teachers deal with immigrant students. In Sweden, conditionally to national test results, Lindahl (2016) find no evidence of ethnic bias on the side of teachers.

As pointed out by Geven et al. (2018), the sources of the observed inconsistency might be related to the higher level of uncertainty when dealing with ethnic minorities. Uncertainty may lead to a loss in the accuracy of teachers' recommendations and evaluations.

However, ethnicity plays a significant role in non-European contexts. In the USA, Ferguson (2003) has shown that biases in teachers expectations sustain and enlarge the Black-White gap score. Similarly, Gershenson et al. (2015) observe that non-black teachers hold downward expectation toward black students. In New Zealand, Meissel et al. (2017) find that ability grouping decisions penalise ethnic minorities.

1.4.4 Discrimination

Teachers may hold biased expectations towards stigmatised social groups. Stigma is a complex process that joints social labelling and social discrimination (Link and Phelan 2001). Stigma occurs when, first, an observable character among individuals becomes socially salient and, secondly, people starts to discriminate the individuals with that specific characteristic; eventually, discrimination of specific characters must provoke in its bearer's status loss.

Stigma is situation-dependent (e.g., they vary across contexts) and imposes a separation between the bearers of the stereotyped label ('them') and the others ('us').

Neo-classical economics has investigated taste-based discrimination in the labour market (Becker 1957). Taste-based discrimination emerges when employers display productivity-unrelated preferences towards employees. The 'taste for discrimination' consists in the willingness 'to pay something, either directly or in the form of a reduced income, to be associated with some persons instead of others' (ibidem, p.4).

A well-known variation to Becker's model is the statistical discrimination theory developed by Arrow (1972). Arrow claims that employers may infer workers productivity by relying on observable characteristics (such as social class, gender, or ethnicity) when they face informational incompleteness.

For example, if it is the common belief that, on average, women display a lower level of productivity in comparison to men, an employer that has no clue about the individual productivity of a specific female worker may offer her a low salary. Alike, if a teacher believes that girls perform worst in maths than boys, that teacher might opt to assign higher notes to male kids.

Statistical discrimination does not necessarily imply status loss and, hence, does not automatically entails stigma. The social labelling may be perfectly ‘value-free', being the mere consequence of common rooted beliefs.

Let us linger a few more words on this last remark. Both taste and statistical discrimination models do not entail any political challenge. The first predicts a natural disappearance of productivity-unrelated differences due to the push of market competition, while the second is expected to phan out as soon as new pieces of information update the beliefs of employers (or teachers).

At a closer look, both Becker and Arrow neglect the possibility that discrimination may be a fundamental need of the capitalistic society in which economic disparities and social privileges are the rules rather than the exception (Cavallaro 2017). As even Arrow admits: ‘It is certainly a common view that in some sense racial discrimination is a device by which the whites in the aggregate gain at the expense of the blacks. Hence, the whole problem is to be interpreted as an exploitative relation. There is a stable relation here; the values inherent in discrimination uphold a structure that is profitable to those holding those values’ (Arrow 1972, p.26).

Discrimination studies have explored two categories of bias, namely, implicit and explicit biases (Geven et al. 2018). While explicit biases are deliberately held by actors and hence easy to trace, implicit biases are behind the consciousness of the subjects.[21]

In Germany, Wenz et al. (2017) find stereotypes in a small sample of secondary teachers. When asked about group averages in academic achievements, the teachers were openly negative towards immigrant and low socioeconomic students. Nonetheless, the

[21] Implicit biases are usually measured by the strength of the association between social categories and semantic contents, or positive\negative implicit attitudes (*ibidem*).

perceptions expressed by teachers in the sample were not particularly inaccurate and almost in line with the national recorded averages.

Always in Germany, Dunkake and Schuchart (2015) asked pre-service teachers to fill an online questionnaire. Their study reveals that aggressive, lazy and undisciplined behaviours, as well as less ambitious self-perception, were much more frequently associated with socially disadvantaged students. Interestingly, subjects would have addressed disruptive behaviours publicly when dealing with lower-class students; while the same behaviour would have received more discrete and private feedback from the fictitious teachers in the case of higher-class students.

As regards implicit bias, disadvantaged students, as compared to their peers, seems to be rather frequently the target of negative (implicit) attitudes (Geven et al. 2018). For example, Pit-ten Cate and Glock (2018) find that a group of pre-service teachers in a German university show negative attitudes towards immigrant and low-background students, while they could not find any explicit negative beliefs towards these students' profiles.

1.4.5 Class and school composition

In the present review, a few words on the class and school composition are useful. Scholars have investigated the effect of class and school composition on teachers' decisions reaching no conclusive findings (Geven et al. 2018).

Scholars have explored two distinct mechanisms. On the one hand, teachers may evaluate an individual student about her\his class or school mates. Marsh et al. (2008) find the so-called 'big fish little pond' effect, that is, students are more likely to receive higher marks when the average in their class is low.

On the other hand, teachers might hold 'generalised' expectations; that is, her\his classmates might define their evaluation over a singular student. In other words, the average expectation influences the evaluation of everyone in the class, irrespectively from her relative ranking in the group. For instance, a higher-performing student as compared to her classmate would receive the average mark.

1.5 Conclusions

Grade repetition is substantially a tale of underachievement, and the present chapter has aimed at reconstructing underachievement's social determinants of underachievement.

The chapter has reviewed the principal mechanisms related to the intergenerational transmission of ability. Academic performance, in its sophisticated and multifaced phenomenology, is strongly influenced by parental resources; as such, it is a crucial factor in the primary gap between kids coming from different social *milieus*. The resources families might mobilise several in kind: economic, cultural and social.

Additionally, when constrained by their limited resources, families make different choices according to their social class and status. On this subject, the chapter has reviewed the main theories which aim at explaining variations in academic decisions and investments, regardless of the individual academic performance.

Over and above the family, I have pointed out that there is an autonomous influence of scholastic institutions. The school and teacher effects are the focus of emerging literature aiming to add a tile in our knowledge of the causal role of education in social stratification complex dynamics. Growing empirical evidence suggests that expectations might be biased towards low-SES and migrant kids.

The next chapter systematises the specialistic literature on grade repetition. The main goal will be that of gathering the available evidence on inequalities in the selection, and in the benefits and disadvantages of grade repetition.

Chapter 2: Grade Repetition, a contested educational practice

"In the last few years, standards for student promotion have become a major topic of discussion among those concerned about public education. Educators, parents, and citizens in general have become worried about the large number of students who are not mastering grade-level basic skills. Increasingly, critics are arguing that one way to solve this problem is to demand that students demonstrate a minimum level of competence before being promoted to the next grade."

Labaree (1984, p.67)

2.1 Introduction

Contemporary educational systems concur in two (partially irreconcilable) fundamental functions, namely those of education and selection (*cf.* among others: Schizzerotto and Barone 2006). The function of education consists of equipping all individuals with a basic set of knowledge and skills, shared culture and shared values. Meanwhile, the function of selection sorts students into future social positions through selection based on knowledge and skills that are valuable for the functioning of the society and the economy (cf. Chapter 1). The function of education favours group learning, which allows for accommodating and instructing a large body of students. In contrast, the function of selection privileges individual selection and develops tools to screen pupils for their ability.

Labaree (1984) has defined the merit promotion as an institutional arrangement that resolves the tension between education and the selection functions in favour of the latter. The logic of merit promotion is that 'each student is retained or advanced a grade based solely on his or her proven ability as measured against a fixed achievement standard' (p.70). At specific turning points – usually, when grades are issued but sometimes at the end of each cycle – students must demonstrate that they meet the respective promotional standards, which are the minimum threshold for the learning target that is set by schools.[1]

[1] The compartmentalisation of the educational systems in grades serves both of the logics (Labaree 1984). On the one hand, it allows for equalising the contents of education at each step; on the other hand, it formally certifies the educational level of the individual.

The logic of merit promotion supports grade repetition (GR), which is a practice that denies students progression in their educational careers for one year (Jackson 1975, Bropy et al. 2006), as the optimal approach to cope with students who do not meet promotional standards. The goal of the present chapter is to gather and organise the available empirical literature on grade repetition as a peculiar educational and organisational practice.

Several international institutions have advocated for limitation of grade repetition. For example, the Organisation for Economic Co-operation and Development (OECD) has peremptorily affirmed that research has 'mainly found a negative effect of grade repetition on academic achievement' (OECD 2015, p.161). Furthermore, the United Nations Educational, Scientific and Cultural Organization (UNESCO 2012) has recommended 'ensur[ing] transitions (*i.e.* promotions, editor's note) to reduce or target the practice of repetition' (p.56). The UN organisation for education has even suggested that the automatic progression of students through grade levels could represent a more cost-effective policy. Hence, educational systems should plan a gradual transition from grade repetition by grade to grade repetition by educational level or cycle (ibidem). The OECD has observed that grade repetition is more prevalent in school systems in which students achieve lower-range scores on the Programme for International Student Assessment (PISA) assessment or school attendance is less equitable. In the European context, the Education and Training Monitor 2018 has recommended that countries reduce the rate of grade repetition to improve their human capital aggregate outcomes (EU Commission 2018).

Opponents of grade repetition have advanced four broad categories of criticism. The first relates to social stratification. A common argument is that grade repetition reinforces segregation patterns in gender, ethnicity and migration backgrounds, social class and geographical contexts. The second criticism concerns the economic cost of grade repetition, which is significant at both the individual and aggregate levels. The third criticism assumes a pedagogical perspective in arguing that grade repetition might pose harmful consequences for students' perceived self-efficacy and social and academic adjustment. The fourth and final criticism focuses on the impact of grade repetition on future achievements. This chapter carefully addresses the debate on the effects of grade repetition, which is essential to the research on this topic. Circumstantial evidence

suggests that grade repetition corresponds with higher dropout rates and lower scholastic performance. However, the evaluation of the causal impact of grade repetition remains an open empirical inquiry. Several scholars have recommended caution, and recent developments in grade repetition research reflect a less direct, non-trivial or even beneficial impact.

Most importantly, the practice of grade repetition finds the favour with many educators and parents, who apply it to, for example, elicit effort from students or 'protect' them from excessively challenging curricula in the next grade (Shepard and Smith 1989). Grade repetition could indeed offer an efficient motivational scheme to motivate students. Moreover, when students fail in school, grade repetition could provide a sufficient remedy by granting them the necessary additional time to re-align competencies and knowledge before proceeding in their educational careers.

Two countries have significantly contributed to the literature on the impact of grade repetition: the US and, to a lesser extent, France. This chapter traces the debates in these two countries. The only definitive fact that emerges from the empirical literature on the impact of grade repetition is that the available data, chosen methods and institutional context of grade repetition systematically and severely influence its outcomes. The main obstacle is to convincingly distinguish between previous students' characteristics and the effect of grade repetition. Most strategies for the identification of causal effects are weakly motivated or inconsistent.

Above all, the institutional level heavily influences the use and features of grade repetition. As indicated above, grade repetition is widespread when merit promotion is prevalent. The radical alternative to merit promotion (and grade repetition) is social promotion. In contrast to merit promotion, social promotion emphasises the function of education and the logic of group learning. It also preserves same-age classes through the virtually automatic progression of pupils through grades; regardless of individual performance, pupils advance together with their peers. The assumption is that the achievement of the students will converge in the long term.

The organisational stratification of the educational system influences both merit promotion and social promotion. Vertical and horizontal differentiation within educational systems mediates the effects of both merit and social promotion. Vertical

differentiation precludes or discourages grade repetition at some educational levels. In certain countries, merit standards govern progression between cycles, yet social promotion is the rule within each educational level (i.e. from one grade to another within the same cycle, social promotion is applied). For instance, in French and Italian elementary schools, horizontal differentiation is a standard tool to mitigate social and merit promotion alternatives.

Such differentiation consists of tracking or streaming. Tracks provide different educational instructions and goals, whereas streams groups students by ability within the same subject. Within each track or stream, students advance through either merit or social promotion. When merit promotion is in place, promotional standards vary between tracks (Labaree 1984).

The primary goal of the present chapter is to collect and organise the available evidence on the benefits and disadvantages of grade repetition (GR, henceforth). It is interested in potential biases and inequalities within the process of selection and considers the consequences of GR. Regarding the impact evaluation of GR, the chapter applies a country perspective.

Two ad hoc sections provide a reconstruction of the U.S. and French debates concerning the effects of GR. The second part of the chapter instead accounts for the influence of institutional level by conducting a descriptive analysis of the use of GR within various institutional European educational systems.

2.2 The sources of the debate

Grade repetition has attracted increasing attention in OECD reports. The OECD has compared ex-post retained versus never-retained students in several dimensions (OECD 2015).[2] Unsurprisingly, the two groups differed primarily in their academic proficiency. An increase of 100 points in PISA mathematics and reading tests decreased the likelihood of having repeated a grade by 43% and 34%, respectively. Moreover, students who had repeated a grade were more likely to have skipped school days or

[2] The OECD measures academic proficiency at age 15 through the PISA test. Additionally, they collect students' background information. Students answer whether (and at which level) they have ever experienced one or more grade repetitions.

arrived late within the two weeks preceding the interview. They also exhibited lower levels of interest in scholastic achievement goals.

2.2.1 Social stratification

Nonetheless, under-achievement and behavioural conduct are not the only factors that relate to the past occurrence of GR (OECD 2015; UNESCO 2012). Even when controlling for PISA's test, 'a student with certain characteristics is more likely to have repeated a grade than other students' (ibidem, p.164). For instance, boys are more likely to be overage in a class or to repeat a grade, although the gap is generally small. Furthermore, pupils who belong to ethnic minorities or have a background of migration are more likely to have experienced GR. The odds of having repeated a grade significantly favour low socioeconomic status students in respect of high socioeconomic status students. Finally, children from poor or rural areas are more likely to have repeated a grade.

However, ex-post comparisons lack any insight into pre-existing differences in academic performance. This evidence does not disentangle prior (under)achievement from GR; as such, it cannot clarify whether migrant and underprivileged students would remain overrepresented among repeaters when accounting for their academic performance. The present section reviews the scarce empirical works that have attempted to overcome this severe limitation. It includes only studies that were able to account for prior-to-GR academic performance.

As anticipated, the research has established a link between GR's probabilities and long-term underachievement. Alexander et al. (2003), among others, have found that GR correlated with poor school adjustment, which entails the process of adapting to the role of a student and the school environment. To-be-retained students exhibited weak academic performance was frequently absent and did not conform easily to behavioural demands. In contrast, to-be-promoted students regularly attended class, engaged in non-problematic conduct and were popular with other students.

There is only weak and scarce empirical evidence of an influence of gender on the probability of GR. McCoy and Reynolds (1999) have analysed GR risk factors among 1,164 low-income minority pupils who enrolled in the Chicago school district. Their data

indicate that male gender was the second-strongest predictor of GR after school performance (and among low-income and minority pupils). This result was not unexpected since behavioural problems are significantly more common among boys (Meisels and Liaw 1993).[3]

While gender has a marginal role in the probabilities of GR, the family has a much more distinct relevance. In this regard, Jimerson (1997) has eloquently stated that 'the major factor that distinguishes a retained child from a comparable peer may be a parent characteristic rather than a child characteristic' (p.6).

First, in considering ethnicity, it is notable that most of the empirical evidence on the ethnicity factor derives from the US, although the research is likely to provide a general indication. Halligan (2000) has emphasised that the racial gap in GR probabilities is consistent across U.S. states.

Cosden et al. (1993) have investigated GR patterns of ethnic groups in kindergarten. They observed that Hispanic pupils were consistently more likely to experience GR compared to their Anglo peers. Economic constraints have a crucial role in the limited investment in pre-school education by Latin families. Moreover, even though Latino parents want their children to succeed in school, limited information about school mechanisms jeopardises their ability to support them in this endeavour.[4]

In another study, Dauber et al. (1993) have reported that African American children had a higher probability of GR from the first to the fourth grades. Other evidence (Hughes et al. 2005) has highlighted the relation between teachers' expectations, academic performance and students' ethnicity (cf. the following chapter). Specifically, the results reveal that teachers had lower expectations for their relationships with African American parents and kids. Teachers might react to a child's behaviour within the classroom, parenting practices, communication styles and educational beliefs. Rodney et al. (1999) have identified three main factors that contributed to GR among a sample of

[3] Ikeda and Garcia (2013) are not included in the text because they did not have a prior-to-retention performance measure in comparing countries using PISA. They found that the gender coefficient became non-significant when the model considered scholastic performance.

[4] By following Latino students at risk of retention, Willson and Hughes (2006) have found that the children's age and a lower degree of parental support directly contributed to the probability of GR.

African American students: the number of suspensions from school reported behavioural conduct and a reported lack of discipline.

Evidence from non-U.S. contexts is scarce. Klapproth et al. (2015) have found that having a migrant background significantly predicted a higher GR probability in Luxemburgish schools. Meanwhile, Krohne et al. (2004) have demonstrated that migrant pupils in Germany were substantially more likely to experience GR in elementary school. They also found a significant interaction between being male and having a migrant background that influenced the probability of GR later in the scholastic career. In the French context, Paul and Troncin (2004) have reported that the effect of being migrant became non-significant once the model controlled for social class.

Social class is a critical factor for GR probabilities. In this regard, Abidin et al. (1971) were among the first researchers to underline the importance of parental socioeconomic status for GR probability. Through a later study of the association between health, social factors and GR in the U.S. context, Byrd and Weitzman (1994) have illustrated that poverty and low maternal education were relevant predictors of GR in kindergarten. Persistent dependence on welfare significantly heightened the probability of GR even when other factors, such as cognitive ability, were equal (Guo et al. 1996). In research by Jimerson et al. (1997), parental involvement in school significantly affected the probability of GR.

Moreover, Alexander et al. (2003) have identified significant variations between retained and promoted students in terms of socioeconomic status and parents' educational background. Among repeaters, one out of two fathers was a high school dropout; meanwhile, never-retained students had economically advantaged parents. Notably, only one study (Dauber et al. 1993) has yielded divergent results. Through this research, which relied on a longitudinal study of student careers in the U.S. city of Baltimore, the authors found that parental education became non-significant once the model accounted for academic performance.

In a study of France in the 1980s, 30% of farmworkers' children were grade repeaters in the first grade, while this figure was just 2.9% among children of professionals and managers (Seibel 1984). Twenty years later, having a father in a professional or managerial occupation or a mother with a tertiary degree reduces the

chance of GR by 10% and 11.4%, respectively (Paul and Troncin 2004). In Luxemburg, low-SES students were more often subject to GR, even when controlling for prior-to-GR ability (Klapproth et al. 2015). Gonzàlez-Betancor and Lòpez-Puig (2016) have reported that retained Spanish children were more likely to belong to low-SES families, while limited financial means characterised retained students in Portugal (Pereira and Reis 2014). Hunt (2008) has noted a higher share of GR in rural areas and suburbs. Finally, Paul and Troncin (2004) have observed significant geographical variations in France, which are possibly due to the unequal distribution of resources and the differences in educational beliefs and practices. No research has addressed the Italian context.

2.2.2 Economic costs of GR

Grade repetition incurs considerable costs at both the family and collective levels. Specifically, families need to cope with a full additional year of education (N'tchougan-Sonou 2001), which implies both direct and indirect costs. Direct costs refer to enrolment fees and learning materials, while indirect costs relate to the postponement of labour market entry, which delays a salary. Additionally, retained students lose one year in the wage premium that is acquired by the diploma (Alexander et al. 2003). For educational systems, imposing GR on students generates consistent costs (Manacorda 2008; Alexander et al. 2003). The OECD (2018) has computed the expenditure per student by dividing the state cash outflow for each level of education by the number of full-time enrolled students. The European Union average cost was estimated to be $9,465 per student each year. Notably, core education services, such as teaching costs, represent 86% of the total amount (i.e. $8,843).[5]

2.2.3 Grade repetition as a pedagogical practice

According to Shepard (1989), scholars have long overlooked teachers' beliefs about GR. Teachers have a crucial role in decision-making about GR in virtually all educational systems, as they decide to have a student repeat a grade based on their

[5] The additional costs for the Italian educational system can be roughly computed. In this country, the annual spending per student is equal to $8,452 USD. In 2018, the share of retained students in high school was 7.1% (around 189,040 students). Assuming that all of them will graduate from high school, the Italian educational system is projected to spend an additional $1.5 billion USD.

confidence in the effectiveness of GR and their expectations of the student's chance of recovery (Hong and Yu 2008; Bonvin, Bless, and Schuepbach 2008).

Nevertheless, they vary in their judgement of the several factors that should lead to repetition (Paul, J J 1997; Tomchin and Impara 1992). Teacher expectations of students' future outcomes represent a crucial issue in this thesis. Chapter 1 has underlined that expectations affect the ability of students. Later, Chapter 3 explains that the Italian law refers directly to teachers' expectations as a critical factor in guiding decision-making about GR.

Empirical findings on teachers' GR-related beliefs are rare and highly localised. In the Swiss city of Geneva, Marxoux and Crahay (2008) conducted an extensive study of such beliefs. They have indicated that the contribution of scholastic achievement to the GR decision is relatively small. The process of evaluation is far more complex. In their academic life, pupils must conform to the role of a student (the *métier d'èléve*), which encompasses the autonomy to carry out academic tasks, a positive attitude towards the school realm and the display of effort in the learning process. A certain degree of ambiguity and uncertainty often crucially characterises the decisional processes of teachers. Hence, the evaluation is likely to consider familiar characteristics. Teachers estimate the parent's ability to sustain their child's academic performance. In sum, teachers' perceptions related to both individual characteristics and the educative potential of their significant others (ibidem p.507).

The practice of GR is likely to depend on teachers' pedagogical beliefs in its beneficial effects (Marxoux and Crahay 2008; Crahay et al. 2013). Often, teachers understand repetition to be a second chance for students who bear the potential to get back on track. First, GR offers time to develop maturity. Second, it avoids additional scholastic troubles when there is a challenging environment at home. Third, GR might reassure pupils of their ability as students by reducing the workload. According to Alexander et al. (1989), the practice of GR grounds over a quite linear conception of learning, whereby learning is assumed to be a progressive accumulation of knowledge and skills. Students advance through blocks (i.e. grades). Any deficiency in this pathway

represents an obstacle for future accumulation. Consequently, it is more productive to reduce the pace and address the defect.[6]

2.2.4 Grade repetition as a motivational tool

A strong argument in favour of GR refers to its motivational power as a deterrent for academic failure (Manacorda 2008). Grade repetition might act as a 'Damocles' sword' by exerting a threat effect on students who are at risk of failure (Belot and Vandenberghe 2014). Such effect consists of the anticipation of costs. When opportunity costs are high enough, the threat is sufficient to stimulate low-achieving students and their parents to catch up. However, the existent empirical evidence does not support any strong claim in this respect. Based on a recent change in the GR policy, Belot and Vandenberghe (2014) have found little evidence for the threat effect's effectiveness in Belgium. Meanwhile, Battistin and Schizzerotto (2018) have determined that an increase of the threat of GR in northwest Italy harmed the actual GR rate, especially among students in the non-academic track.

Labaree (1984) has discussed the implication of GR as a motivational scheme at a more theoretical level, and he has urged teachers and administrators to be cautious. For example, prompting motivation can be of minimal use if the motivation itself does not primarily contribute to underachievement; in this regard, other factors, such as social class, are crucial. Educational policies and practices should foremost foster a reduction in the impact of socioeconomic and familial background. Moreover, Labaree has underlined an additional caveat which emerges from the intrinsic tension between the threat effect of GR and the remedial goal. Grade repetition can be a valuable motivational tool since it can evoke high psychological and social costs through its threat effect. Such costs might include unhappiness about separating from classmates or shame regarding the stigma or labels. When the threat is sufficient, students will make the necessary progress to avoid a harmful outcome.

However, in the event of GR occurrence, these anticipations are likely to harm students' academic performance. As such, the threat effect jeopardises the remedial goal of GR. Since teachers want to avoid such adverse outcomes, they should reassure students

[6] Crahay (2007) shows that teachers' positive attitudes towards GR usually goes along with a fixed (rather than constructivist) and one dimensional (rather than multidimensional) conception of intelligence.

about the harmful anticipations of GR, which can undermine its threat effect. The implied conundrum is apparent: when schools (or parents) stress the threat function of GR, they undermine its remedial aim. Thus, there is a trade-off between the remedial and motivational aims of GR whereby the emphasis on the former undercuts GR's motivational power.[7]

2.3 The impact of Grade Repetition

Although many teachers and parents believe in the beneficial effects of GR, rigorous attempts to evaluate its consequences have yielded mixed results. As the next pages demonstrate, data availability, methodological strategies and institutional contexts have largely determined empirical results. In the effort of discussing the role of institutional contexts, this section reconstructs the discourse on GR's impact at the national level. In particular, it considers the evolution of the scientific debate in two essential countries: the US and France. Both countries have experienced intense debate about GR-related issues that have in turn animated the production of a vast body of the empirical literature. While the U.S. case is crucial because of the amount of available empirical literature on the issue in this context, the French case is of more specific interest to this thesis. The French educational system shares many similarities with the Italian system. Moreover, after intense public and scientific deliberation, the legislative framework has opted for a (top-down) limitation to the use of GR.

This section sections aim to reconstruct the heated discussion of the effect of GR in the US. The dispute has developed over the last century. In this context, GR gained early notoriety because of its intimate connection with promotional standards, which refer to the examination targets that institutions require students to meet in order to advance to the next grade. Due to the amount of the bibliographical material, Table 2.1 below provides the reader with a summary.

[7] As Labaree has noted, the same contradiction exists within the family, as parents can read retention as a remedial pedagogical intervention or, on the contrary, a punishment.

Tab. 2.1: Summary of the reviewed studies in the U.S. context

Paper	*Description*	*Matching\conditionally on*	*Main results*
Jackson 1975	Systematic review of 30 empirical studies published between 1911 and 1973. 3 types of studies: (a) comparison between retained and non-retained students; (b) comparison pre- and post-grade repetition; (c) experimental design (3 studies)	Types: (a) loose matching over academic performance; (b) no conditioning over external causes; (c) randomisation	Analysis focuses on achievement and school adjustment. types: (a) detrimental effect of grade repetition; (b) beneficial effect of grade repetition; (c) no statistical or negative
Holmes and Matthews 1984	Meta-analysis of 44 studies (575 individual effect sizes) between 1929 and 1982	Only 18 had matched individuals. Only 17 had matched individuals over at least IQ or achievement test	Both non-matched and matched studies report negative effects of GR
Holmes 1989	Meta-analysis of 63 studies (861 individual effect sizes) between 1929 and 1988	Only 25 had matched individuals	54 studies (21 among the matched studies) found a negative effect
Jimerson 2001	Meta-analysis of 20 papers between 1900 and 1999.	19 had matched subjects. 17 had matched subjects over IQ or prior academic achievement	16 attested a negative or null effect of GR on achievement and adjustment
Eide & Showalter 2001	Panel data to study GR impact on dropout and future labour market entry	IV kindergarten entry dates	Non-statistically significant positive effect of GR
Jimerson 2002a	Longitudinal study on a small group of retained student (N = 58) to study the risk of dropping out after grade repetition	Various measures of achievement, mothers' level of education, socio-emotional and behavioural adjustment	Both mothers' educational status and adjustment differ among the groups (leavers versus non-leavers)
Lorence et al. 2002	Data from the population of elementary student in Texas (1994-1999)	Matched over test score, race, and free lunch receiver	Positive effect of GR, diminishing over the years
Lorence 2006	Re-analysis of Holmes 1989 and Jimerson 2001 + 5 additional studies	Lorence restricts the adequacy for research quality.	"no overwhelming body of […] evidence" against GR, p. 767
Stearn et al. 2007	NELS data (panel) are used to study the correlation between race, GR and dropout	Socio-economic status, school characteristics, and achievement	GR predicts dropout. Among retained students, difference in resources do not fully explain dropout
Wu et al. 2008	Short term GR effect of N = 784 Texan students	Subjects were matched on over 72 variables. 92 matched pairs found	GR was detrimental for the growth in maths skills
Jacob and Lefgren 2009	Administrative data from the Chicago Public schools over three years	Natural experiment (Diff-in-Diff); the introduction of a test-based promotion policy	GR increases dropout risk
Griffith 2010	NELS data used for grouping 878 pairs of comparable students	Matched over prior achievement, race, gender, and SES	Retained had lower reading achievement, and slower reading growth
West & Schwerdt 2012	Administrative state-wide data from Florida following children from 3 to 9 grade	Regression discontinuity design thanks to the introduction of a test-based promotion policy	Short terms gains for retained students in achievement and on the probability of GR in subsequent grade
Hughes et al. 2017	The effect of early GR (primary school) on the probability to dropout in high school. N=538 from a panel in Texas.	Propensity score matching on 65 covariates	Grade repetition led to a sharp increase in dropout rate
Ferreira et al. 2018	Panel data from Columbia district used to measure GR effect on secondary school performance	Natural experiment (Diff-in-Diff); State introduced the threshold of 5% of per-year-grade repetition	Improvements in language but not in maths observed among retained.

2.3.1 The debate in the US

According to Labaree (1984), U.S. public education has been slowly fluctuating between loosening and tightening its promotional standards. In the initial stages of the U.S. public educational system, a strict promotional policy rendered it highly challenging to progress through the educational levels. Therefore, even average-performing students faced a high probability of failure. From the beginning of the 20th century, strict promotional standards were subject to severe critique. Critics questioned the educational system as highly unfair and elitist. Moreover, a high number of retained pupils was exerting severe organisational stress on schools. During this time, many people raised the argument that more flexible promotional standards would have avoided the overload of the system.

2.3.1.1 Early research

In the effort to endorse these claims, sceptics of GR started to collect the first pieces of empirical evidence. In 1901, Superintendent William H. Maxwell published a ground-breaking report on the worrying level of above-the-modal-age children in New York City schools (Coffield and Blommers 1956). In 1907, the Russell Sage Foundation supported an investigation 'to study the problem of the progress of school children through the grades' (Ayres 1908). The goals were to identify children who fail to follow a regular path and to explore related risk factors. Hence, on behalf of the Russell Sage Foundation, Leonard P. Ayres harmonised the administrative records of 20,000 children at 15 Manhattan school.

Ayres noted that in as early as primary school, approximately one-third of students were repeaters. The author further reported that three out of four students had failed at least once by the eighth grade. Late entrance and irregular attendance predicted GR, and girls were less likely than boys to be overage. Ayres observed essential differences between 'races' and nationalities among repeaters. Ultimately, he concluded that grade repetition 'is not at all a problem concerning a few under-developed of feeble-minded children' (p.3). The U.S. courses of study 'are fitted not to the slow child or the average child but the unusually bright one' (p.5). He critically states that 'the average city school system trains its pupils well in the habit of failure' (p.146).

2.3.1.2 Review by Jackson (1975)

Maxwell's report and the Russell Foundation's investigation has stimulated many other studies. In 1975, Gregg B. Jackson published the first systematic review of the literature, which proved influential for many years to come. The author reviewed 30 empirical studies carried out between 1911 and 1973. He categorised this empirical material into three main groups.

The first group compares retained versus promoted students; these studies tend towards a negative evaluation of GR's impact. However, when data provided controls for prior achievement and student characteristics, results often did not indicate any statistically significant difference between retained and non-retained students. Jackson believed that these studies failed to consider contextual variations in performance due to school- and class-specific contexts.

The second body of research compares academic performance before and after GR. The studies found mainly beneficial effects of GR. As Jackson has underlined, the primary shortage of this design is the lack of control for improvements due to causes other than GR.

The third group exploits an experimental design. Surprisingly, three studies (Cook 1941; Farley 1936; Klene and Branson 1929) have randomly assigned similar students to grade GR. Cook (1941) has matched students in terms of age, IQ, achievements and personality traits and retained a randomly selected group. By the end of the semester, no statistically significant differences emerged. Farley (1936) has reached the same conclusion by observing another group of 400 students randomly split in the fourth and fifth grades. Klene and Branson (1929) have noted higher success among promoted students compared to their retained peers. Unfortunately, the authors have not reported any statistical test for corroborating the observed gap. Jackson has concluded that ‘the effect of GR relative to grade promotion is warranted by all the results taken as a whole: there is no reliable body of evidence to indicate that GR is more beneficial than grade promotion for students with serious academic or adjustment difficulties’ (p.627). Eventually, he peremptory affirmed, ‘those educators who retain pupils in a grade do so without valid research evidence’ (ibidem).

2.3.1.3 Meta-analyses by Holmes and Matthews and Flunking Grades

Holmes and Matthews (1984) have expanded on Jackson's work with another systematic review of research results on the topic of GR. However, for this review, Holmes and Matthews collected only studies that reported enough data to allow for a calculation of the effect sizes and whose designs included a comparison group of promoted students.[1] Based on these two criteria, they selected 44 original studies published between 1929 and the early 1980s. The selected studies compare the two groups (retained versus promoted students) in several dimensions: academic achievement (31 out of 44), personal adjustment (21 out of 44), self-concept (9 out of 44), attitude toward school (8 out of 44) and attendance (4 out of 44). The authors computed 575 individual effect sizes.

Holmes and Matthews found an overall negative effect for retained students that was consistent across the dimensions. According to the authors, 'the high degree of consistency between these measures lends credibility to the validity of the findings' (p.17). Holmes and Matthews concluded that 'those who continue to retain pupils at grade level do so despite cumulative research evidence showing the potential for negative effects consistently outweighs positive outcomes' (p.17).

Holmes (1989) has provided an update to Holmes and Matthews (1984) by analysing 19 additional studies selected according to the same criteria. The results are consistent with those of the previous analysis: on average, repeaters performed worse than promoted students and exhibited problematic adjustments (Holmes 1989). Although Holmes also identified nine studies that have reported positive effects, the author underlined that the retained students followed intensive remediation activities.

The above work by Holmes (1989) appears in a pivotal book entitled *Flunking Grades: Research and Policies on Retention (Education Policy Perspectives)* (Shepard and Smith 1989). This collective work gathers contributions from scholars who had reached a critical posture towards the use of GR at that time. As such, it is the U.S. anti-GR manifesto par excellence. The authors have summarised their findings as follows (ibidem, p.215):

[1] The effect size is the difference in outcomes between retained and not retained pupils, divided by the standard deviation of the control group (the promoted students).s

a) Neither school achievement nor personal adjustment benefits from grade repetition.

b) Grade repetition strongly predicts scholastic dropout.

c) Early grade repetition (kindergarten) does not resolve school unpreparedness in pupils.

d) Students experience grade repetition as a punishment and cause of stigmatisation, and it causes them emotional and psychological distress.

e) A positive attitude of teachers towards GR is the result of a lack of any (unlikely) experimental knowledge. Teachers have failed to compare retained students to like promoted students.

f) Grade repetition fails to create more homogenous classes: 'the next year some…teacher will have pupils with a range of ages…with corresponding variations in size, maturity and accomplishments' (p.220).

2.3.1.4 The meta-analysis by Jimerson (2001)

A decade after the publication of *Flunking Grades*, Jimerson (2001) published a meta-analysis of 20 articles printed between 1990 and 1999. With this analysis, Jimerson reinforced the negative discourse that surrounds GR. The author found a wide variation in the variables used to match the groups (ibidem). Besides, 19 out of the 20 studies matched in at least one dimension among IQ, previous academic achievements, socio-emotional adjustment, socioeconomic status and gender. The largest share of studies produced no significant differences between retained students and their corresponding promoted peers. Half of the studies reported that promoted students displayed higher achievement as compared to retained students. Overall, 16 out of the 20 studies attested 'that grade retention is ineffective as an intervention for academic achievement and socioemotional adjustment' (ibidem, p.432). Jimerson has suggested that grade retention alone cannot fully explain post-GR academic failure or success. Still, schools and educators should apply both preventive and remedial intervention strategies to correct student achievement trajectories.

Jimerson et al. (2002a) have extended the work of Jimerson (2001) by examining the link between GR and high school dropout rate. To this end, they analysed 17 empirical studies that model dropout probabilities after GR. The authors maintained that they 'clearly demonstrate that early grade retention is one of the most powerful predictors of later school withdrawal' (p.452).

2.3.1.5 Heterogeneity of the impact of GR in the U.S. context

Few scholars have explored whether the impact of GR is heterogeneous across social groups. In a small-N longitudinal study, Jimerson et al. (2002b) have compared school leavers versus non-school leavers among early retained. For this study, the researchers considered pupils who were retained in first or second grade and followed them through 11th grade. A higher level of education of the mother and a positive attitude of the mother towards education predicted smaller dropout probabilities. In the same manner, among repeaters, several socio-emotional and behavioural measures appeared to be positively correlated with dropout probabilities, aggressive behaviour and low self-esteem.

To explore the heterogeneity of GR's impact, Stearn et al. (2007) have highlighted the effect of the interaction between ethnicity and GR on dropout probabilities. The authors reviewed three main school dropout mechanisms. Finn's frustration–self-esteem model (1989) dictates that unsuccessful outcomes, such as GR, are damaging to the self-esteem of students. When poor academic results undermine their social identities (*sensu* Goffman, cf. Chapter 1), repeaters might engage in social creativity (Kelly 2009; Van Houtte 2012). By rejecting the notion that academic success is essential, they may start to compare the 'in-group' (repeaters) with the 'out-group' (non-repeaters) in new dimensions to induce a reorientation of values away from school.

Considering the engagement with the school realm (Tinto 1993), GR can give rise to alienation and abstention from participation in school life. The stigma of GR is also a threat to students' social capital (Gottfredson et al. 1994) because it weakens social bonds with teachers and classmates. In this context, Stearn et al. (2007) have examined differences in GR's consequences between ethnic groups. They have revealed that differences in resources failed to fully account for the gap in dropout probabilities

between retained and promoted students. The authors have explained these variations by citing differentials in frustration level, engagement factors and social capital exploitation.

2.3.1.6 Methodological shortcomings of the U.S. empirical literature

During the 2000s, the robust consensus regarding the harmfulness of GR began to falter. New data and novel methodologies prompted a wave of studies that have challenged the unidirectional interpretation of GR as a detrimental practice. However, the ideological basis of this paradigm shift emerges from the latest developments in the broader discussion of promotional standards. While U.S. public education shifted towards more inclusive standards of progression in the first half of the 20th century, later decades witnessed increasing accusations that the relaxation of promotional standards was responsible for the decline in educational quality. These critics claimed that inclusive promotional standards had eroded the 'floor' of achievement, which in turn lowered its 'ceilings' and compromised excellence among the top-performing students.

A frequent argument is that inclusive educational systems lack adequate incentives to elicit effort from students. Following this perspective, school reforms have identified more narrow sets of necessary skills that students must master in each grade or block in order to pass the grade. Accordingly, these new curricula include ad hoc standardised tests to measure the grade-related minimum competency level. This shift corresponds with a renewed emphasis on selection policies, such as GR. However, advocates of GR have challenged the methodological reliability of previous research and underlined its substantial shortcomings. According to Lorence (2002), 'the claim against the effectiveness of grade retention is generally based on a few often-cited studies and reviews of research literature that are interpreted as being conclusive evidence' (p.14).

Also, Lorence (2002) has enumerated essential limitations to the work of Holmes et al. (1984, 1989) and Jimerson (2001). First, Lorence has criticised the robustness of their meta-analyses. When Holmes and Jimerson compared studies, they did not take into account the effect size variations of GR due to the sample size, the degree of similarity between the retained and control groups, measures of academic performance included in the analysis, and the existence and the entity of additional assistance to promoted students (Lorence 2002). Allen et al. (2009) have further remarked that it is a common practice to estimate more than one effect size each study, and the comparison in the case of multiple

(and hence nested) effect sizes represents a violation of the statistical independence assumption between individual observations.

Lorence (2006) has applied stricter criteria to re-analyse the studies of Holmes (1984) and Jimerson (2001) and concluded that ‘there is no overwhelming body of scientifically sound evidence demonstrating that making academically challenged students repeat a grade is ineffective or harmful’ (p.765).

The critique not only targets the methodological reliability of the meta-analyses but also aims to identify a minimum set of GR-related research quality criteria to ensure the inclusion of only empirical studies that meet the standard. The critical issue is the adequacy of pre-GR controls. While randomised experimental designs are unfeasible (at least in the present day), non-experimental settings suffer from the undeniable fact that students' characteristics are associated with both the treatment selection and the academic post-treatment outcomes. To overcome this critical limitation, educational researchers have explored two strategies: statistical control and matching (Allen et al. 2009, Lorence 2006).

Statistical control methods treat GR as the binary independent variable in regression models and consider its covariation with some subsequent outcomes. The minimum set of required controls necessarily includes specific measurements of prior-to-grade repetition\-promotion scholastic performance as well as information about the gender, familial context and migration background of students. *Ceteris paribus*, one should interpret GR's coefficient as the effect or impact of GR on a specific outcome.

However, this approach has significant limitations. First, any interpretation of the effect of GR must be, as noted, *ceteris paribus*, which is problematic to some extent. As Shepard has critically commented, ‘if retained students had entered school white, middle class, higher performing and non-handicapped, they would have come closer to the achievement levels of the non-retained comparison groups’ (in Alexandre et al. 2003, p.271). Second, the sample size limits the set of controls that the model can exploit. Third, the functional form of the relationship between covariates and GR must be correctly specified (i.e. the effect remains unidentified when the analytical approach is non-parametric). Finally, the ‘treated’ units (i.e. students retained at the end of the year) enter into the analysis with all of the non-treated units (i.e. students promoted at the end of the

year). Regression models exploit all of the observations in the data set. They compare the outcomes of repeaters, who are often underachievers, members of ethnic minorities and from low socioeconomic families, with those of drastically different students (i.e. their well-performing, white and wealthy peers).

The alternative approach is to control for pre-grade repetition differences by selecting a comparable group of low achieving but promoted students. A plethora of non-experimental methods for impact evaluation can facilitate this task. The main methods are propensity score estimation, regression discontinuity design, the difference in differences design and the instrumental variable approach. The goal is to account for all of the relevant variables that influence the selection process in the treatment.[2]

Lorence (2006) has noted that only 10 out of 18 of the studies quoted in Jimerson (2001) included matching of past performance. However, studies that match only in previous achievement and cognitive performance are likely to ignore other relevant factors, such as non-cognitive performance, socioeconomic status and critical contextual effects. A certain degree of arbitrariness characterises matching strategies, and scholars contend with challenges that do not always have unambiguous solutions. One of the most prevalent debates concerns the following question: Is it more appropriate to match based on grades or of age? Two main alternative logics are possible: same-age and same-grade (Alexander et al. 2003; Lorence 2006). In same-grade comparison, retained students are older than others in their matched group, but they attend the same grade. In contrast, same-age design estimates GR's effect by comparing retained students with their peers. Thus, it measures the gap between the two groups within the same academic year, but they may enrol in different grades.

[2] Cf. Chapter 4 for the methodological discussion of the propensity score method that is utilised in the present work.

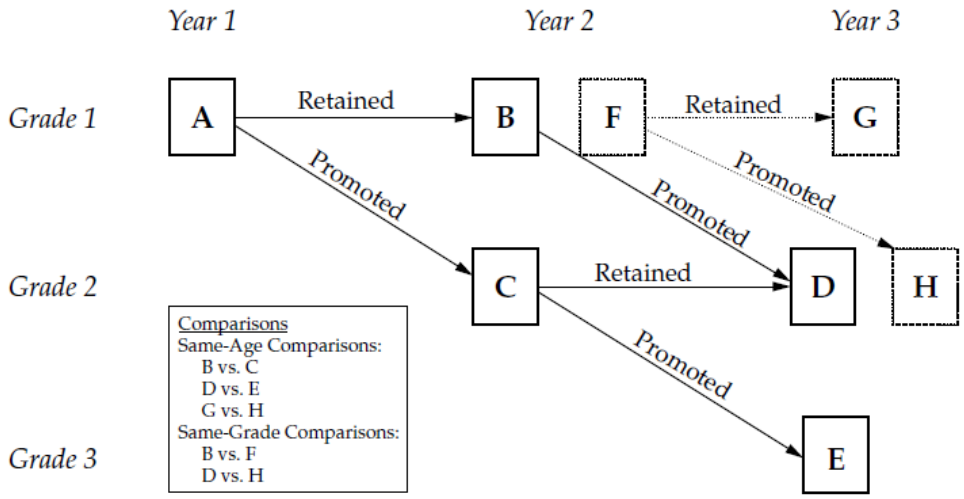


Fig. 2.1: Illustration of same-age and same-grade comparison. Solid lines: members of cohort 1; dotted lines: members of cohort 2. Source: Alexander et al., 2003.

The choice between these two logics affects results (Holmes 1989; Shepard 2004; Alexander et al. 2003; Lorence 2006; Allen et al. 2009). The same-age comparison tends to detect the negative effect of GR. Conversely, same-grade comparison generally favours grade repetition over promotion. In both cases, gaps between groups appear to narrow over time. Thus, the decision should rely upon the context. The same-grade logic focuses on the total effect of an additional academic year; meanwhile, in the same-age comparison, the one-year loss is a crucial implication of GR.[3]

An additional consideration is the scale of measurements for the outcome variables (Lorence 2002). Scales can be grade-specific or grade invariant. Grade-specific measures are problematic if the aim is to compare changes over time or outcomes between two groups in different grades.

Ultimately, Lorence (2002) has underlined that many studies have exploited non-random or too small of samples. A lack of representativeness jeopardises the power to

[3] Many scholars (e.g. Wilson 1990; Karweit 1999; Shepard 2004; Lorence 2006) have maintained that same-grade comparison is a better fit if one assumes that the purpose of GR is to provide the children with the necessary additional time to re-align their competences and scholastic performance. However, non-peer comparison leaves room for numerous alternative explanations (e.g. retained students have taken the same curriculum twice, or retained students are more mature than their classmates).

extend results, especially to sub-groups, such as minorities. Small samples fail to provide enough statistical power to reject differences between retained and promoted students.

Allen et al. (2009) have conducted an additional systematic review of 22 papers on the U.S. context published between 1990 and 2007. They have maintained that 'studies employing poor methodological controls for non-equivalence at pre-grade repetition between retained and promoted children produced more negative effects of grade repetition on achievements' (p.13). Thus, studies able to control for previous characteristics indicated a less harmful impact of GR compared to studies that could not. Nevertheless, Allen et al. (2009) have concluded that evidence does not favour GR intervention.

2.3.1.7 Latest relevant works in the US

The scientific standards that Lorence and Allen et al. identified have stimulated the production of new empirical evidence. For instance, significant positive effects have been found by a population study of low-achieving elementary students in Texas (Lorence et al. 2002). The authors analysed an entire cohort of academically challenged students from 1994 to 1999. They matched promoted low achievers with comparable retained students in the same grade. The results reveal a positive effect of grade repetition during the first year after GR. In particular, retained students were more likely to pass a standardised test examination the subsequent year as compared to their (matched) promoted peers. This effect diminished over time but was still positive after six years.

Nonetheless, a similar portion of comparable promoted students would eventually pass the same examination by the two\three subsequent year. On the long run, the two groups were found quite equally performing after the treatment. In conclusion, the authors have agreed with Alexander et al. (2003) that neither social promotion nor grade repetition alone can allow low achievers to catch up. [4]

Wu et al. (2008) have investigated GR's effect with a propensity score estimation strategy. From an initial sample of low-achieving elementary students in Texas who were at risk of GR, they identified a sub-sample of comparable retained and non-retained

[4] Lorence and colleagues attribute these results to the capacity of the Texas educational system to make schools and teachers accountable. Texan accountability system, in the authors' belief, creates efficient incentives to make educational actors improve students' performance.

children. They successfully matched 97 pairs of students on 72 variables, which included sociodemographic factors, reports from teachers, parents and peers on behavioural problems and hyperactivity of the subject, measurements of personality traits and teacher-student relationship quality. Their results illustrate that in the short term (three years), GR reduced the rate of growth in mathematical skills. The authors detected no significant impact on reading skills but identified some moderators of GR's effect. Children with limited linguistic proficiency suffered more from GR, and the effect of promotion was more beneficial for children with peer- and parent-reported conduct problems.[5]

Hughes et al. (2017) have relied on the same initial sample as Wu et al. (2008) to test the long-term effects of GR. They identified 171 pairs of comparable students whom they matched on 65 covariates through propensity score matching. They found that grade repetition in primary school raised the dropout rate by the age of 16.

Another matching study by Griffith (2010) has investigated the effect of grade repetition. Based on the U.S. National Education Longitudinal Study, the author successfully matched 878 retained students with never-retained peers in terms of reading skills, gender, ethnicity and SES. The analysis compares reading achievements between the two groups. The findings indicate that non-retained students attained scores that were higher than those of their coupled retained students.

When possible, scholars have explored the effect of GR through methods besides propensity score estimation. Eide and Showalter (2001) instruments an exogenous variation in kindergarten entry dates across U.S. states. They assumed that the kindergarten entry period would affect the probability of experiencing grade repetition early in one's career but would not impact subsequent outcomes. The two authors ultimately found a (not statistically significant) positive effect on dropout rates and labour market earnings.

Besides, West and Schwerdt (2012) have exploited the introduction of standardised-test-based promotion in Florida at the end of the third grade. The share of retained students jumped from 2.8% to 13.5%. By a regression discontinuity design, the authors observed short-term gains in both reading and mathematics among retained

[5] The authors claim that "their good behavioural regulation allows them to rise to the occasion" (p.102).

students. Interestingly, grade repetition in third grade sharply reduced the likelihood of being held back in subsequent grades.

Similarly, Ferreira et al. (2018) have taken advantage of a change in GR regulation in the Columbia district to conduct another quasi-experimental study. By comparing achievements before and after the reform, they found that retained students improved in language test scores but demonstrated no improvement in math scores. Meanwhile, Jacob and Lefgren (2009) have examined the introduction of test-based promotion in Chicago Public schools to overcome selection concerns. They have reported that the determinate change in policy introduced a non-linearity between the current achievement and the probability of repeating the year. With this claim, the authors have proposed that GR exerts a significant negative impact on the probability to complete high school among eighth-graders.

These sections have presented a review of the U.S. empirical literature on the impact of GR. Notably, almost a century of public and specialistic debate, wherein passionate scientific enquiries have endorsed both sides, has failed to produce conclusive results. Thus, the question remains: does GR help students?

2.3.2 The debate in France

Because of its widespread use in France, GR has attracted significant attention within the country over the last 20 years. Indeed, compared to other European countries, France has long been renowned for its high rate of grade repetition (EU Commission 2018). In the academic year 1960-61, 22.1% of first-grader pupils were at least one year late; by the seventh grade, more than half of the students were overage (Paul and Troncin 2004). In 2003, more than 40% of 15-year-old students were at least one year late (HCEE 2010).

The debate has often situated researchers and teachers in opposition (Paul 1997). Even though most educational researchers have condemned this practice, belief in the beneficial effect of GR has persisted among educators (CNESCO 2015). French teachers (and parents) often believe that it is necessary to hold back students when they do not accomplish crucial learning goals (Troncin 2004). Moreover, they perceive that students are more exposed to failure when they try to build forthcoming knowledge over feeble

bases. To endorse their conclusion, teachers have often recalled successful cases of retained students who eventually recovered. Leboulanger (1995) has noted that teachers use GR with the definite aim of 'adjusting' the profile of the student, and they believe that low-achieving students deserve an additional chance to recover.

Despite teachers' positive attitude, the number of retained students has diminished in the last 10 years (DEPP 2016), and France is now the country with the steepest decrease in grade repetition rate in contemporary Europe (HCEE 2012). Top-down recommendations and legislative interventions have certainly contributed to this trend. In 2014, a legislative decree (n.1377) stated the exceptional character of GR, to which recur:

'*pour pallier une période importante de rupture des apprentissages scolaires. Il fait l'objet d'une phase de dialogue préalable avec les représentants légaux de l'élève. La décision de redoublement est prise après avis de l'inspecteur de l'éducation nationale chargé de la circonscription du premier degré. En cas de redoublement, un dispositif d'aide est mis en place, qui peut s'inscrire dans un programme personnalisé de réussite éducative*'.[6]

The concept is that:

'*une prévention de l'échec, un accompagnement des élèves en difficulté tout au long de l'année et des possibilités d'aménagement de la scolarité d'une année sur l'autre offrent de meilleures chances de succès*' (HCEE 2012, p.5).[7]

The effect of the legislative shift has been significant in middle schools. Between 2013 and 2016, the GR rate dropped from 2.2% to 0.6% (ibidem). The reduction of this rate has also been evident in high schools, where it fell from 7.8% in 2013 to 4.3% in 2016. However, the decrease in the GR rate significantly varies according to the school's composition. Specifically, it has been substantial in low-SES schools but considerably more limited among wealthy schools (DEPP 2016).

[6] "to compensate for a significant period of disruption in school learning. It is the subject of a preliminary dialogue phase with the student's legal representatives. The decision to repeat a year is taken after consulting the national education inspector in charge of the first-degree district. In the event of repetition of a year, a support system is set up, which can be part of a personalised programme for educational success."

[7] "prevention of failure, support for pupils in difficulty throughout the year and possibilities for adjusting schooling from one year to the next offer better chances of success"

There is a substantial amount of evidence regarding GR in the French context. Seibel and Levasseur (1983) have analysed a sample of French first-grade students. They classified the students into ability groups and compared promoted versus retained children. The findings reflect no significant gains among repeaters.[8] Meanwhile, Grisay (1993) has arrived at a negative evaluation of GR after observing the careers of high school students, and research by Caille (2004) that followed a sample of students after grade repetition has reported that late repeaters suffered less from GR compared to early repeaters.

Also, Cosnefroy and Rocher (2004) have claimed that GR enlarged inequalities in educational achievements and track choices. They have also found that repeaters were less motivated and attached to the school's values. These authors have underlined three significant additional findings. First, even when standardised scores were equal, retained students were under-evaluated by teachers in comparison to their classmates. Second, when the analysis accounted for marks by teachers, retained students exhibited lower educational ambition. Third, on the bases of both marks and ambition, they were less likely to choose a general track (*ibidem*).

Troncin et al. (2005) have carried out an extensive empirical examination of GR's risk factors and impact. The work relies on a sample of 3,000 first graders in central France. The researchers matched retained versus promoted students in the early stages of first grade. They then followed the pupils for the three years following the first grade. They observed that to-be-retained students began to demonstrate worse performance than that of their peers during the scholastic year that ended with grade repetition.

In contrast, to-be-promoted students reflected increasing improvements. After GR, the academic performance of repeaters improved; however, the gains were smaller than the advancements of their matched promoted students. Notably, neither promoted nor retained students managed to close the gap with other peers.[9]

Alet et al. (2013) have investigated GR's impact on achievement by analysing French students' panel data within an instrumental variable (IV) framework. To this end,

[8] The reader should notice that the two authors are not able to control for prior-to-retention performance.

[9] When compared to the entire sample, matched promoted students are much more likely to score among the bottom 10% and to experience future grade repetitions.

they compared retained students' test scores during the first cycle of primary school (first and second grades) with their scores on the same test during the third and sixth grades. To control for endogeneity of selection in the treatment, they employed the students' quarter of birth as IV. The underlying assumption was that the quarter of birth has a direct effect on the probability of GR but no direct effect on post-grade repetition achievements. The study found short-term achievement gains (GR improved third graders' test scores), and the effect was consistent across genders and SES.

Moreover, the primary benefit disappeared in the medium term. The scores of repeaters in sixth grade were lower than the scores they would have attained if they had avoided the repetition.[10]

Majoub (2017) has utilised the same panel to investigate GR's impact in ninth grade when the repetition occurs in sixth grade. He controlled for endogeneity by instrumenting the quarter of birth. The author noted a positive effect of GR on test scores and the likelihood of graduating from junior high school.

Even though French institutions have promoted (and managed to impose) a substantial limitation to GR, the empirical results are not conclusive. On the one hand, the worrying disparities within the selection of GR suggest that the practice is applied mainly to low-SES and migrant students, who are over-represented among low achievers. On the other hand, the impact of GR does not have a clear direction, and it is likely to be mediated by individual and contextual factors.

2.3.3 Relevant impact research in other countries

A selection of empirical material from other countries concludes this literature review on the impact of GR. This collection does not aim to be exhaustive. It specifically collects studies that convincingly controlled for prior-to-grade repetition scholastic performance.

[10] Results are consistent when the authors exploit as an instrument the individual performance ranking in the class. They assume the relative ranking within class influences teachers' decision but has no impact on future achievements.

2.3.3.1 Dutch-speaking Belgium

Baert et al. (2015) have estimated the effect of GR in ninth grade. Although students in Flanders exceed the OECD average in terms of achievements, the link between the latter and social origins also exceeds the OECD average. Grade repetition is standard in high school. To avoid it, students can opt to enrol in a lower track. Based on a representative panel sample, Baert et al. (2015) have developed a dynamic discrete choice model.[11] Through a series of simulations, they observed negative long-term impacts of GR on dropout rates, schooling delays and subsequent downgrading. The findings suggest that forced track changes can improve educational outcomes, though to a lesser extent than is achieved by allowing the students to progress unconditionally.

2.3.3.2 Switzerland

Bonvin et al. (2008) have investigated the effect of GR in the Swiss context by matching retained and non-retained students in panel data. In comparison to students in the same grade, retained students exhibited an improvement in achievement. However, the juxtaposition of their performance with that of their peers reflects a negative impact of GR. The authors have concluded that GR warrants a negative evaluation because it does not provide sufficient evidence of its effectiveness.

2.3.3.3 Portugal

Nunes (2018) has matched students according to standardised test scores in the year of grade repetition. Specifically, the study matched comparable retained and promoted fourth graders on the sixth-grade test score. Among students who matched on both test score and socioeconomic status, the author noted a small, statistically significant positive effect of grade repetition.

2.3.3.4 Brazil

Koppensteiner (2013) has investigated the effect of the introduction of automatic grade progression in Brazilian elementary schools in 1993. The study has identified the effect through a difference-in-differences analysis that exploits variations in the reform's application in terms of time and space. By analysing repeated cross-sectional data on

[11] They claim that, in contrast to regression discontinuity design, this methodology considers potential heterogeneity of the effect.

achievement, the author has estimated an adverse effect of automatic progression on test scores. The study concludes that automatic promotion had a disincentive effect on student performance.

2.3.4 Grade repetition in Europe

Institutional arrangements, diversity of historical traditions and developments in education systems influence the features of grade repetition (Goos et al. 2013, UNESCO 2012). This chapter concludes with a descriptive exploration of such features across European countries.

2.3.4.1 Merit promotion, social promotion and horizontal differentiation

The introduction of the present chapter has addressed the two concepts of merit promotion and social promotion. Merit promotion imposes specific promotional standards that students must meet in order to advance in their educational careers. Schools retain students who are under the required threshold. Meanwhile, with social promotion, students' careers are not conditioned by their scholastic performance. The system stratification mediates the effects of both merit and social promotion. In the realm of educational systems, 'pure' forms do not apply.

However, in the opinion of the author of the present thesis, the educational institutions presumably engage with one of the following four main ideal types of 'promotional regimes' (see Table 2.2 below): pure social promotion, social promotion within differentiated promotional standards, merit promotion within differentiated promotional standards and pure merit promotion.

Tab. 2.2: Promotional regimes and their influence over grade repetition rate

	No differentiation	Differentiation
Social promotion	Low GR	
Merit promotion	High GR	Grade repetition mediated by different promotional standards

In pure social promotion regimes, the rate of GR is close to zero. The same figure should emerge from social promotion regimes within differentiated promotional standards; the only difference is that students are grouped by ability within the same grade. Grade repetition is widespread in merit promotion regimes within differentiated promotional standards. Nevertheless, stricter allocation of students into tracks by an institution corresponds to a milder GR rate. Institutions that foster pure merit promotion should produce a high rate of grade repetition.

Of course, the typology does not apply to the entire educational system. Instead, promotional standards should be heterogeneous within each educational system, depending on the cycle or level. For example, primary schools can be characterised by a social promotional regime, whereas high schools can rely on a differentiated merit promotional regime.

2.3.4.2 European educational systems and GR

The focus on European countries in this section helps to situate Italy, the case study, within a broader, fruitful context of analysis. European educational systems usually fall into four main groups: the Nordic inclusive model, the individual choice model, the early tracking model and the mixed tracking model (Blossfeld et al. 2014; Benavot and Resnik 2006; Vaniscotte 1994).

In Scandinavian countries, a single structure school ('l'école unique') is characteristic. This schooling aims to provide the same education for all students for the most prolonged period possible. Horizontal differentiation occurs late in a student's career (generally at age 16). Normative evaluations, such as teachers' notes (cf. Chapter 3), do not appear until eighth grade. This inclusive model is strongly oriented towards social promotion and prioritises same-age progression.

The individual choice model derives from the Anglo-Saxon tradition. In contrast to the learning group aims that are privileged by the Scandinavian structure, this model promotes the acquisition of students' autonomy and their full development as individuals. On the one hand, it favours continuity between the lower and upper secondary cycles, during which students share a standard core curriculum. On the other hand, starting from age 14, students integrate their curriculum by selecting courses. The score obtained at the end of 10^{th} grade limits their choices in the last three years. Grammar (or modern) schools

are the highly selective and elitist alternative to the polyvalent high school. Institutions regularly measure student performance and evaluate schools according to their students' results.

Continental Europe begins tracking students in middle school. In this context, the goal of educational institutions is to prepare students for labour market entry. The 'dual system' allows students to accumulate early on-the-job skills. In Germany, students generally lay the foundation for their professional future at the age of 10.

In Southern Europe, students follow a unique path until the end of middle school. This mixed tracking model is a compromise between the Scandinavian and Continental models. Students share the same core curriculum until the end of middle school. Then, from the beginning of high school, the comprehensive schooling leaves space for a late tracked system. Table 2.3 displays the typology of promotional regimes and, accordingly, expectations for GR's use.

Tab. 2.3: A typology of promotional regimes

	Differentiation in middle school	Differentiation in high school
Social promotion	-	Inclusive model
Merit promotion	Early tracking model	Individual choice model; mixed tracking model

2.3.4.3 Empirical evidence

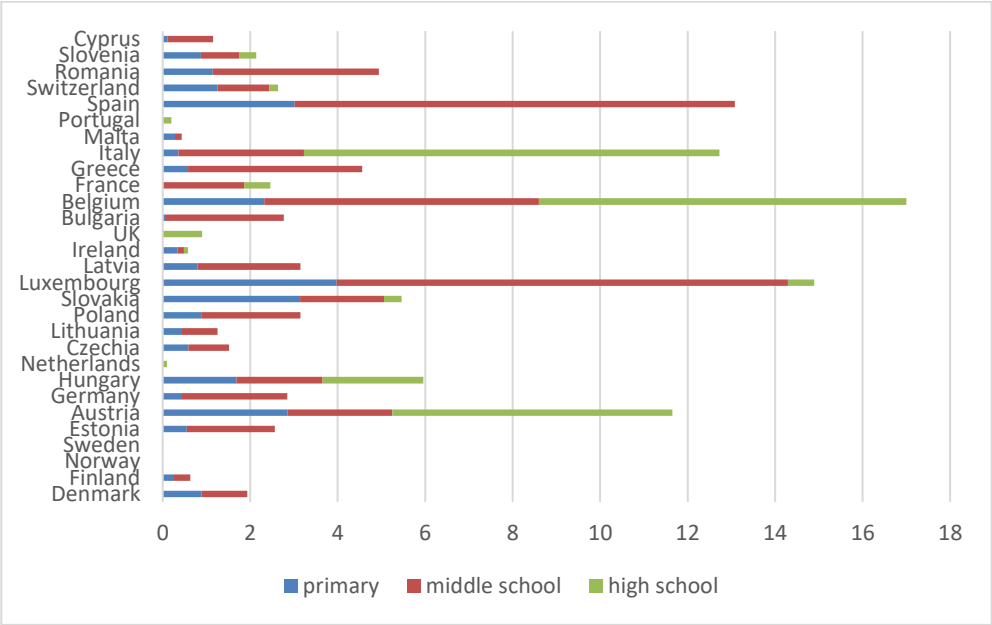
To corroborate the theoretical expectations of the present research, it gathers cross-country data from two sources, namely UNESCO and the OECD. The Education and Literacy section of the UNESCO Institute for Statistics has provided the GR rate for primary and lower secondary schools, which they compute by dividing the number of repeaters in a given grade at time t+1 by the total number of students in the same cohort enrolled in the same grade at year t (UNESCO 2019). Unfortunately, some data are

missing for certain European countries, especially data regarding primary school. Moreover, UNESCO has not provided any data for upper-secondary levels.

The present study uses the 2015 PISA to fill this gap, as it provides the most recent illustration of the occurrence of GR in high schools. The PISA questionnaire asks the following question to students: 'Have you ever repeated a grade?'. Students must also specify the number of times and the level at which the repetition occurred. Unfortunately, PISA's respondents were all 15-year-old students, which presents a severe data limitation given that the aim is to examine GR in high school. In many European countries, the age at which students enter high school is later than 15 years old. In another group of like countries, students consistently enrol in high school at age 15. In the few remaining countries, the transition occurs one year earlier at age 14. Therefore, PISA is likely to underestimate high school grade repetition rate drastically.

Because of these caveats, the present research provides data for each level of school for each country (when possible). In Figure 2.2 below, the blue, orange and grey bars represent the GR rates for elementary, middle and high schools, respectively. The data refer to 2015. Countries are by type of educational structure. Except for Denmark and Estonia, the GR rate is negligible among countries with an inclusive model. No data are available for the high school period, as students enrol in high school at age 16.

Fig. 2.2: Grade repetition in the European countries at 15 years old



However, the situation is more ambiguous in continental Europe, where middle school GR is not negligible and, for countries that provided data, the high school GR rate is not either. Luxembourg is notable for its considerable GR rate in lower secondary school. On their part, Anglo-Saxon countries do not offer much data. It appears that Ireland has few repeaters, and the GR rate in the first high school year is limited. Grade repetition is also limited in UK high schools, which may be partly explained by streaming mechanisms and the critical autonomy at the school level to set promotional standards. Ultimately, when data are not missing, the GR rate is generally higher in the Southern European countries than in the other two groups. Belgium, Spain and Italy are prominent as GR-intensive countries.

The data support some general remarks. First, with a few exceptions, GR in primary school represents a limited portion of the total GR within each country. At this educational level, most of the countries favour same-age progression and reserve GR for peculiar cases. Second, streaming in the UK seems to be an effective alternative to GR in at least the first two years of upper-secondary school. Nevertheless, the developed accountability system prompts the differentiation of promotional standards and is likely to impact the selection at the enrolment. No data are available after the split in 11th grade.

Last, but most importantly, the mediating effect of early tracking is complicated. On the one hand, the GR rate is not negligible in tracked systems, which may suggest that tracking has a mediating effect, but merit promotion is still the preponderant regime within tracks. On the other hand, the GR rate is higher among mixed tracked systems than in continental Europe. Therefore, the tracks have a different influence on merit promotion. Moreover, differences may occur concerning how institutions sort students into tracks. In differentiated lower secondary school systems, student performance substantially constrains individual choice. This sorting mechanism reduces the gap between promotional standards and individual performance; consequently, the GR rate is limited.

Conversely, in Southern European educational systems, students and their families are free to choose the track. This sorting mechanism does not guarantee a match between individual scholastic performance and the schools' promotional standards. In such a context, GR resembles a 'post-orientation' tool for adjusting to this displacement.

2.4 Conclusions

This chapter has engaged with composite research on GR and underlined several relevant issues.

First, GR targets under-achievers; as such, it is pervasive among specific social groups. International reports and empirical works have evidenced that this practice is applied disproportionately to males, migrants and ethnic minorities, and disadvantaged kids when compared to females, native and white students, and privileged peers, respectively.

When researchers have had access to prior-to-grade repetition data, their findings have generally supported the claim that ascribed characteristics exert an autonomous influence on the probability of GR. Independent of their correlation with academic performance, factors such as social class and migration background directly contribute to the occurrence of GR. Similar results have emerged from empirical investigations of decision-making by teachers, which have indicated that teachers' evaluations are likely to embrace factors that extend beyond the actual individual performance.

Although many scholars have adopted a critical position towards grade repetition, the debate over the impact of GR has yielded inconclusive results. This chapter has discussed the impact of GR by inspecting the relevant empirical works that have advanced this debate. Such display clarifies the link between GR and the broader consideration of merit promotion and promotional standards. The results are incredibly mixed in each of the main investigated dimensions (achievement, social adjustment and psychological impact). In the grade repetition research field, the earliest experimental studies are from the 1950s. Scholars have since relied on non-experimental techniques. This review has emphasised that the results are, in many cases, sensitive to the data on which the analysis relies. Data availability and statistical techniques often constrain identification strategies and variations in the composition of controls or the criteria of matching impact results across studies as well.

Researchers also often encounter a critical trade-off: more accurate controls imply a smaller number of compared individuals. Furthermore, the choice of comparison of (the matched) group (in terms of age/grade) influences the outcomes. The results drastically depend on the context, and any cross-country inference is problematic. Educational

practices and beliefs shape the framework in which GR circumstances operate. Grade repetition varies across countries in terms of motivation, timing, target and consequences. Even the political debate decisively influences GR's occurrence (see the debate on merit versus social promotional standards).

Institutional arrangements are a critical factor in the analysis of GR. The mechanisms of differentiation mediate the merit and social promotion regimes. In this respect, the cross-country analysis illustrates a complex role of tracking. Theoretically, tracking should mediate the occurrence of GR by differentiating promotional standards. It should actively influence promotional regimes when the system allocates students according to their performance. When, conversely, families are free to choose, GR consistently intervenes within tracks. Thus, GR is a form of post-tracking orientation that compensates for the offset between individual performance and within-track promotional standards.

2.4.1.1 Strong beliefs, weak evidence

The empirical material collected in this review calls for a negative evaluation of the grade repetition as a tool to tackle academic failure. Nonetheless, results far from being unambiguous has often emerged from many contexts, research designs, and educational outcomes. In sum, Alexander et al. (2003) are probably justified in arguing that GR research is inhabited by 'strong beliefs' that are supported by 'weak evidence'.

Such additional research is of importance for the present case study, Italy. In the context of Italian schools, GR is negligible at the primary level and rather small in middle schools. Grade repetition eventually begins to increase in high school following the orientation into tracks drastically. Although the topic of GR has periodically received attention, the educational practice remains ambiguous in Italy. Apart from a few institutional descriptive reports, no empirical research has tried to infer the socioeconomic risk factor that correlates with grade repetition occurrence. In the same manner, no work has explored its consequences. Hence, the need to fill this gap is substantial.

Chapter 3: The Italian context

3.1 Introduction

Grade repetition remains substantially understudied in Italy. On the one hand, appropriate data is scarce: Italy lacks any specific official report on the topic of GR and no longitudinal study on students' career exists at the national level. On the other hand, as opposed to France and the U.S., the Italian public has never systematically addressed the issue of grade repetition. The present chapter proposes an examination of the relevant institutional features of the Italian educational system concerning GR.

The chapter first looks at the institutional dynamics over time by tracing a brief history of promotional standards and promotional regimes in Italy, as defined in Chapter 2. Promotional standards refer to the specific learning contents set by schools. Promotional standards are usually cycle- and grade-specific. For example, primary schools require lower promotional standards than middle schools insist on a very different curriculum; on the contrary, academic tracks require higher promotional standards than vocational high schools.

In contrast, promotional regimes are the reflection of the dominant logic of students' progression at each stage of the educational systems. Promotional standards emphasise either individual selection or learning. Promotional regimes that favour individual selection would define a logic of merit-promotion. Promotional regimes that favour group learning would configure a logic of social promotion. While social promotion preserves the steady progression of peers, merit-promotion gathers together students by performance (hence, it stops kids who do not reach promotional standards).

Institutional macro factors influencing micro-level decisions are likely to substantially contribute to the dynamics of educational inequalities (Barone 2019). For example, the degree of selectivity of each grade, track or cycle impacts the perceived chances of success. At the institutional level, promotional standards and promotional regimes have been crucial indexes of the social profitability of the educational investments.

Afterwards, the chapter offers a detailed description of the nowadays Italian educational system. In this regard, it also reviews the legislative frame that governs grade repetition in Italian schools.

In conclusion, the chapter presents the (rather scarce) descriptive data on grade repetition at the national level.

3.2 Promotional standards and regimes in the History of Italian educational system

The competition over educational qualifications is a strategic terrain in the struggle between social groups over social mobility (Collins; Boudon) and Italy is in no way an exception (Barbagli 1974).

As Borghi (1951) has argued, two main traditions have struggled over the definition of the educational structure. On the one hand, the authoritarian tradition, which had been the dominant paradigm for the first century of the country's history, has aimed to separate the education of the elites from that of the lower social strata. On the other hand, the democratic tradition, which has gained legitimacy in the post-war period, has ultimately aimed at overcoming the social segregation in schools and providing a comprehensive education for all.

3.2.1 The mechanisms of social selection in Italian schools

The Italian educational system still marked by worrying levels of social selection. Baldacci (2014) has identified four main mechanisms that have hindered the participation of underprivileged kids: selection *via* 'dead-ends', selection *via* under-age dropout, 'indirect' and 'direct' selection.

The first refers to the constant presence of early dead-ends, which had for a long time prevented lower social strata from attaining secondary (and later tertiary) educational titles. Key reforms in the democratic period had indeed increased the level of middle and high school openness and the decreased level of horizontal differentiation.

The second mechanism has been the highly tolerated dramatic share of under-age dropout. Illegal dropout from compulsory education has often jeopardised the impact of institutional reforms. Formally, education or training was compulsory to the age of 14

already in 1923. In facts, at least until the middle-70s, authorities had mostly tolerated high rates of early dropouts (*ibidem*).[1]

The third mechanism has been the direct selection on academic performance, which has occurred through the extensive use of grade repetition.

The fourth and last mechanism is instead an indirect form of selection. A dramatic heterogeneity has indeed characterised Italian schools in promotional standards. The indirect selection has typically occurred at the early stages of education. Promotional standards in rural and in urban elementary schools were by no means comparable in the first half of the last century. Even today, the dramatic social segregation of middle schools has a definite impact on the level of academic preparation at high school enrolment. However, the indirect selection is not limited to lower cycles. For example, the type of secondary degree substantially predicts the odds of dropout and degree.

3.2.2 The authoritarian genesis of the Italian educational system

Table 3.1 summarises the main institutional developments of the Italian school system. Scholars usually trace the origin of the Italian school system back to 1859, when the Kingdom of Sardinia promulgated the law Casati (Schizzerotto and Barone 2006).[2] With the 1861 unification of the Peninsula, the law applied within the monarchy's new borders.[3]

[1] Nowadays, compulsory education lasts for 10 years and covers the age range between 6 and 16 years. The fulfilment of the compulsory education is aimed at obtaining a secondary school diploma or a three-years professional qualification by the 18th year of age. Cf. next section of the present chapter for a detailed overview of the contemporaneous Italian educational system.

[2] Before the nineteen century, formal education concerned a tiny portion of the Italian population. Wealthy families employed private tutors to cultivate their offspring. The economic development, together with the influences of Jacobinism and the 1848 revolutions, brought about reformist ideas in the field of education (Del Passo 2003). In 1809, the Italian educator Vincenzo Cuoco presented to the Bonapartist King of Naples Gioacchino Murat the 'Report for the Organisation of the Public Education'. "Education- Cuoco claimed- must be universal, public, and uniform". Later, the Bonapartist regime of Naples and the 'enlighted' Austrian government in Lombardy acknowledged the demand for universal primary education by establishing ad hoc study commissions. The results of their work would have eventually remained unheard. Indeed, not much of these reformist aims inspired post-unification Italy's educational system.

[3] The law represented a significant effort for the former Kingdom of Sardinia's growing bureaucratic machine. It comprehended 380 articles and legislated over the entire system of education. After primary school, further education fulfilled a key goal for the newborn Kingdom, namely, that of select and shape professionals and executives for specific economic and bureaucratic sectors of the new State. See 'Education and the State', by Green. Under many aspects, the building of the Italian school system answered to the bureaucratic needs of the new State.

On papers, the kingdom established a minimum level of universal basic education through the institution of primary schools. Nonetheless, thousands of children did not even reach the six compulsory years of education, and many dropped school just after the first primary school cycle.[4] Moreover, the dramatic geographical heterogeneity in the provision of education in terms of both structures and professionals further jeopardised the universal education goal.[5] After primary school, students faced an early tracking that guaranteed a strict separation between the popular classes and the elites' education.

Tab. 3.1: Summary of the main interventions in the Italian educational system

Year	Dominant tradition	Law	Description
1859	Authoritarian	Casati	A minimum level of universal primary education, but a strict social separation of educational paths
1923		Gentile	The early dead-end of 'Avviamento al Lavoro' prompted further segregation
1962	Democratic	Gui	Middle school became comprehensive
1969		Codignola	Free access to university
2003		Moratti	The category of Lyceum became broader than before
2010		Gelmini	Toward a revival of merit promotion?

[4] Primary schools had a lower cycle (three years) and an upper cycle (two years). The law raised from six to nine the compulsory years of education only in 1904.

[5] In the '60s, many students still were teaching to younger kids to some extent (Lettera ad una professoressa)

Until the second half of the twentieth century, the structure of the Italian educational system would have remained virtually unchanged. From 1923, the fascist power pushed forward the authoritarian project of the kingdom in the field of education (Del Passo 2003; Baldacci 2014). [6]

By the law Gentile, the regime emphasised the bureaucratic control of schools by putting primary school under the direct state control. It also radicalised the selective character of the Italian schools by splitting the lower secondary level into two main branches. The 'noble' branch channelled students towards upper secondary education (technical or academic).

The other, called *'Avviamento al Lavoro*' (in English, initiation to the job) prepared students for an early labour market entry. After primary school's completion, the *Avviamento al Lavoro* gave the necessary manual skills to initiate the kids to an early labour market entry (Schizzerotto and Barone 2006). In facts, the *Avviamento al Lavoro* established an early dead-end, which pulled out the Italian popular classes.

Moreover, the law Gentile further exacerbated the elitist character of the lyceum, to which students could enrol if they completed the gymnasium.[7] The new structure also discouraged girls from enrolling to the normal lyceum by the institution of a reserved and shorter path, the *Liceo femminile*. A rigorous merit promotion regime characterised students' progression: students faced a final exam at the end of each cycle. Overall, the reform had a severe impact on upper secondary cycles' enrolment.[8] [9]

[6] The fascist Minister of Education Gentile believed in the role of education to separate executive classes from popular classes. Already in 1918, in a public letter to the former Minister of Education Berenini, Gentile was hoping for "the crowd" that was lowering the overall standards in middle schools to clear out (Resto del Carlino, 1918).

[7] Only in 1940, with the Bottai Law, the technical and academic branch of junior high school merged.

[8] The number of students attending secondary schools dropped by about a hundred thousand between 1922 and 1925 (Schizzerotto and Barone 2006)

[9] Later interventions under the regime focused on curricula and propaganda's activities in schools. By the late '30s, the fascist ideology had permeated the Italian school thanks to new programmes and the central role of youth associations in the propaganda (Del Passo 2003).

3.2.3 The democratic turn

With the fall of the fascist regime in 1943 and the advent of the new Republican Constitution of 1948, left parties (and particularly the Communist Party) hoped to impress a substantive democratic turn on the authoritarian character of the Italian school system.

Nevertheless, the deterrent action of the catholic democrats, interested in the containment of any radical political option, managed to postpone any relevant structural reform for almost two decades (Del Passo 2003; Baldacci 2014).

Eventually, the turbulent (and unequal) economic development and the drastic growth in scholastic participation called for a sensible modification of the selective character of the Italian educational system (Schizzerotto and Barone 2006).

In this new cultural and political atmosphere, the printing of the book "Letter to a teacher" (Milani 1967) represented the collective and passionate condemnation of the unequal character of the scholastic selection and denounced the enormous expulsion of low-status children from schools. In the book, many voices are radically critical to grade repetition, which discriminated and pushed out from schools thousands of poor children.

At the institutional level, the democratic turn materialised through seventeen years of educational reforms between 1962 and 1979.

In 1962, the fourth centre-left government of the Prime Minister Fanfani abrogated the *Avviamento al Lavoro* and established a comprehensive middle school. In 1968, early education became (formally) universal by the creation of public kindergarten schools. In 1969, the law Codigola allowed free access to tertiary education, as it established that high school graduates could enrol to university regardless of the type of degree. In the early 70s, elementary schools started to provide to parents full-time education for their children. In the same years, the first courses for adults appeared. In 1977, another intervention banished the remedial exams in primary and middle schools. Eventually, in 1979, the debate over the teaching of Latin in middle schools ended with its elimination from learning programmes.[10]

[10] The teaching of Latin has always represented an actual barrier for the popular classes in Italy (Barone 2006). However, as noted by Cha (1991), foreign languages have started to play the credentialist role previously played by Ancient languages. For further reading on curricula changes: Benavot, Aaron,

By the 1980s, the Italian educational system had drastically transformed. The formal differentiation of promotional standards in primary and middle schools had fallen apart. The strict merit promotion logic had started to loosen up in favour of practices of social promotion.

Nonetheless, the democratic turn had affected mostly the lower cycles of education and, even in the face of the drastic growth of high school enrolments, the reforms had mostly safeguarded the segregated features of the upper secondary cycle.

3.2.4 The last years

The picture has slightly changed in the last years. In 2003, the centre-right Minister Moratti 'promoted' to Lyceum the art, linguistic and teaching schools. This marginally eroded the segregated character of the Italian upper secondary school.

However, the reform did not share the democratic intents of the previous period (Baldacci 2014). Quite the contrary, the reform addressed a rather vague idea of modernisation of the Italian educational system. The idea of school as a mean for social emancipation had left room for the centrality of human capital as a mean for full individual realisation.[11]

In 2008, the new Minister Fioroni restored the remedial exams in high schools (which a 1994 intervention had abrogated) and, between 2008 and 2010, the Berlusconi's Minister Gelmini has reintroduced numeric evaluations (also on behavioural conduct) in primary and middle schools.[12]

Hence, an ambiguous orientation and a lack of any clear project or ideal of school characterise the current phase (Baldacci 2014). Although the democratic push had run out, a coherent paradigm has not yet replaced it. Nevertheless, the growing emphasis on

and Cecilia Braslavsky, eds. 2007. School Knowledge in Comparative and Historical Perspective: Changing Curricula in Primary and Secondary Education. CERC Studies in Comparative Education 18. Dordrecht: Springer.

[11] The concept of 'individualisation' of educational paths and instructions has made at this time its appearance. Berlusconi's 3i can summarise the modernisation narrative of Moratti's reform (in Italian): Impresa, Internet e Inglese (Enterprise, Internet and English).

[12] A scale between 0 and 10 replaces a more qualitative and ordinal method of evaluation.

individual merit in a human capital perspective could give new life to the logic of merit promotion.

3.3 Grade repetition trend

Against the above described institutional dynamics, this section reports grade repetition's trends over the last 80 years. The Italian National Institute of Statistics (ISTAT) provides data on the percentage of repeaters since 1945 in elementary, middle and high schools.[13]

Two different graphs plot the annual percentage of repeaters (lines) along with the number of students (bars) attending each cycle or, for the upper secondary cycle, each kind of high school.[14]

Figure 3.1 plots data for elementary and middle schools in blue and in orange, respectively. As described above, the lines represent the percentage of repeaters for each cycle (right y-axis), while the bars the number (in thousands) of students for each cycle (left y-axis).

The number of primary school students remained relatively constant during the first half of the century, and then it has recently diminished due to the demographic slowdown. A drastic reduction of repeaters rate, which has eventually flattened closed to zero, followed the full participation in primary education. Hence, since the 1980s, social promotion became predominant in primary schools. Full participation and social promotion anticipated the institutional democratic turn.

The link between participation and promotional regimes also emerges in middle schools (Figure 3.2). The number of students (orange bars), which increased in the first half of the century, has then flattened following the last decade's demographic compression.

[13] While grade repetition rate is the number of retained students at the end of each year over the total number of students, the percentage of repeaters refers to the share of enrolled repeaters for each year or grade. This work assumes that the percentage of repeaters is a decent proxy for grade repetition rate (see graph on GR's rate in Appendix N). The *Annuario Statistico* of ISTAT provides for the share of non-admitted students from the academic year 1970-71 on. Please refer to Graph A2 in the Appendix for GR's rate in primary and middle school. Table N reports GR's rate by high school type for two recent cohorts.

[14] For the same period, Graph A1 in the Appendix shows the percentage of completion for each educational level.

The marked downward trend in repeaters' rate has started from the creation of the comprehensive middle school in 1961. After a quite odd discontinuity, attributable to the abrogation of the remedial exam, the share of repeaters has remained relatively stable (around 10%). The marked decrease is contemporary to the democratic turn, although the decrease in grade repetition in elementary schools is likely to have concurred.

Figure 3.2 reports the participation and repetition trends for upper secondary schools. For the sake of clarity, it groups the high school types in two: the lyceums (blue) and the non-lyceums (green). Again, lines and bars represent respectively the share of repeaters (right y-axis) and the number of students (left y-axis).

The number of students in non-academic tracks increased continuously until, in the 90s, it has started to slow down. The drop in the academic year 2003-2004 is attributable to the Moratti's reorganisation, which 'upgraded' some non-lyceum to lyceums. Repeaters rate markedly dropped twice: in the second half of the 1960s and at the beginning of the 1990s.

The much smaller number of students in the academic tracks has slowly increased (again, the jump in the 2000s is attributable to institutional reorganisation). The repeaters rate dropped by five percentage points in the 1960s. Then, it has remained quite stable.

The gap in the share of repeaters between non-lyceums and lyceums has, instead, persisted over the observed time.

Eventually, Figure 3.3 focuses on the most recent dynamics of high school grade repetition. It displays the repeaters rate by tracks from 1971 onwards by the four principal types of school, namely, vocational, technical, traditional lyceum, and the other academic tracks. (Cf. next section for a detailed presentation of the type of Italian high schools). Except for technical high schools, the remaining tracks show relative stability of GR during the last decades.

Fig. 3.1 Share or repeaters in primary and middle schools. My elaboration of ISTAT data- Serie Storiche. Left axis: the number of students is in thousand. Right axis: percentage of repeaters each cycle. **Missing data for middle school repeaters before the academic year 1953-54 and between the academic years 2002-2003 to 2008-9.

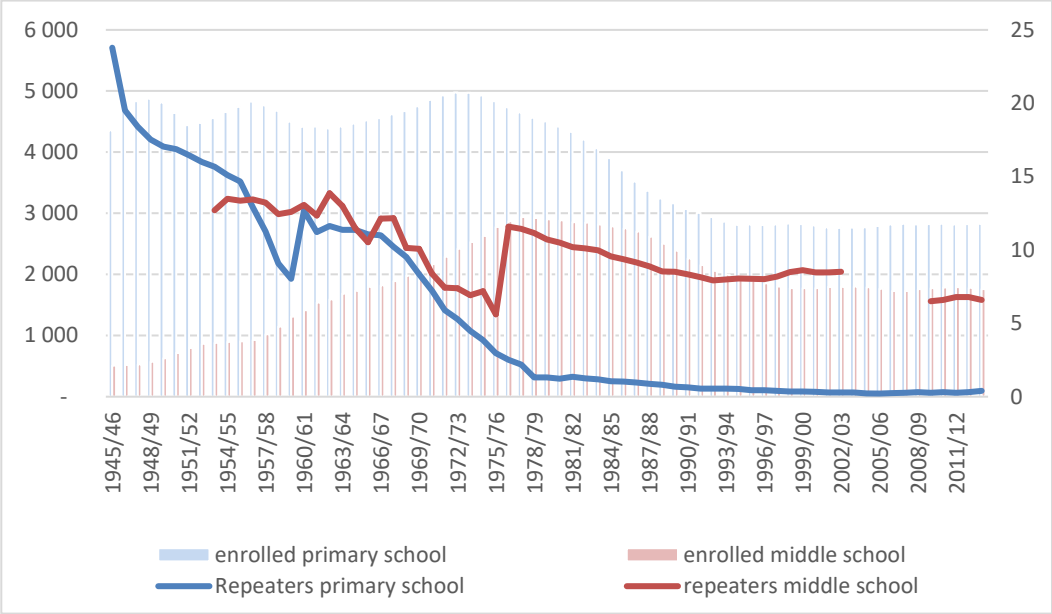


Fig. 3.2 The share of repeaters in upper secondary schools by high school type. My elaboration of ISTAT data- Serie Storiche. Left axis: the number of students is in thousand. Right axis: percentage of repeaters each cycle. Missing data before a.y. 1953-54; data are missing for non-lyceum a.y. 2009-10.

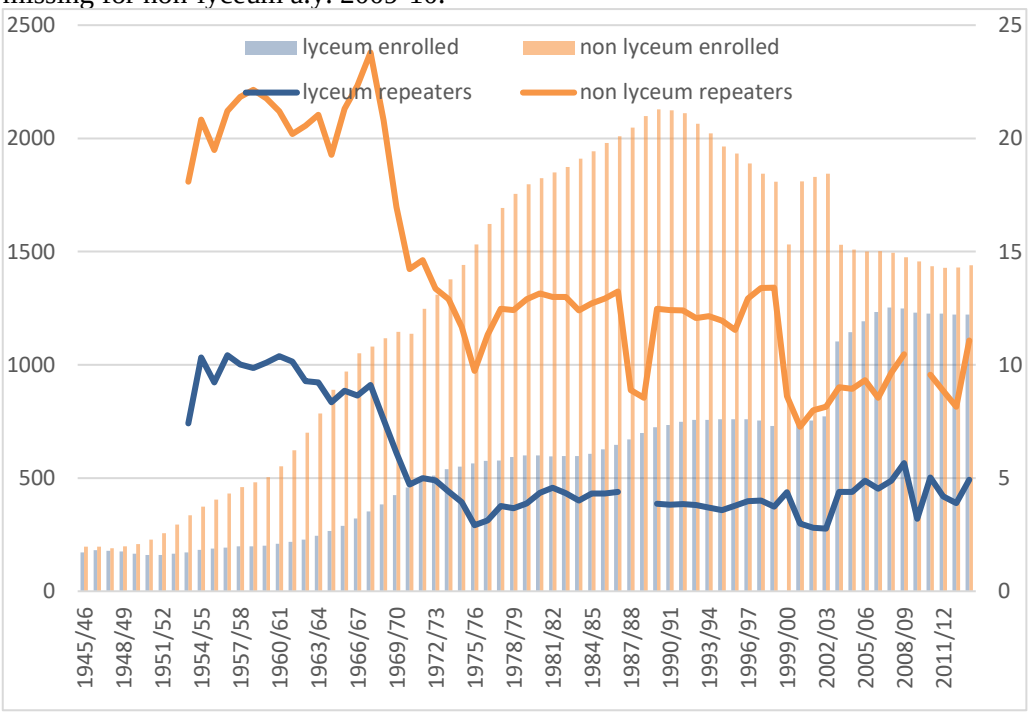
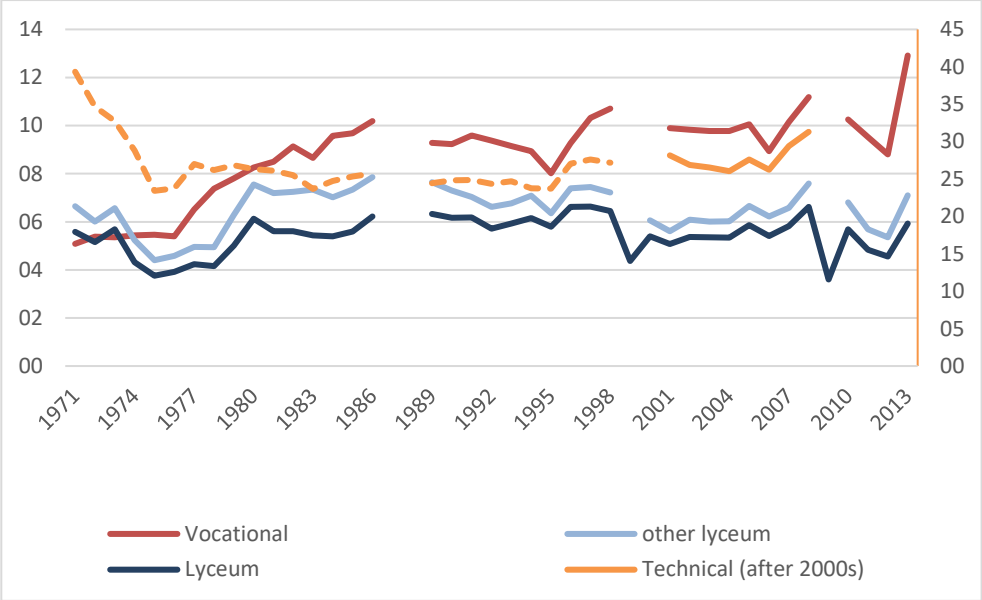


Fig. 3.3: Repeaters rate by tracks from 1971 onwards by the four principal types of school. My elaboration of ISTAT data- Serie Storiche.



Tab. 3.2: Summary table of the results. It summarises the results of this historical overview. From 1859 to the early 1960s, the authoritarian agenda dominated the Italian educational system. It established a strictly selective school and, at all educational levels, made extensive use of individual selection and merit-promotion.

Period	Main interventions	Promotional regime	Promotional standards	Enrolment	GR's practice	Other types of selection
The authoritarian period. From 1859 to the end of the '50s	Casati (1859); Gentile (1923)	Merit promotion at all levels	Tracking from middle school	Increasing at primary schools	Widespread at all levels	Illegal evasion of compulsory education; Dead ends
The democratic period. From 1962 to the end of the '80s	Comprehensive middle school (1962); kindergarten (1968); abrogation of remedial exams (1977); removal of Latin in middle school (1979)	The substantial growth of social promotion in elementary and middle school	Tracking in high school	Increasing at all level (mass education starts here)	Decreasing in primary and middle schools	Illegal evasion of compulsory education
The contemporary period	Moratti (2003); Gelmini (2008)	Non-systematic turns towards a new merit promotion	Lyceum 'upgrade' for artistic, linguistic and teaching schools	Increasing in lyceums (stabilised in the last decade)	Stable in high school; consistent gap between tracks	Indirect selection in primary and middle schools (?)

The Italian educational system tried to limit the push of increasing participation. It hence channelled students in different tracks at early stages and subjected students to severe end-of-cycle and end-of-year examinations, and it tolerated illegal dropouts. In this regard, grade repetition was a widespread practice already in elementary schools.

A substantial democratic turn occurred in the 1960s. The number of students in middle school drastically increased, and a series of policies radically changed the structure of primary and middle schools. Social promotion soon has become the norm within the first two cycles of education, in which grade repetition has drastically decreased, reaching today's levels. Nonetheless, early (illegal) dropout remained substantial and, in high schools, merit promotion persisted, although mitigated by differentiated promotional standards.

Elements of ambiguity characterise the contemporary period. Both the democratic push and the pressure exerted by the enrolment have run out. Nonetheless, the last interventions have aimed at reinvigorating merit promotion through the emphasis on individual merit and human capital. Furthermore, as Baldacci (2014) points out, indirect selection continues to target underprivileged kids in elementary and, particularly, in middle schools.

3.3.1 The contemporary Italian educational system

In Italy, the compulsory age for education is 16. After five years of primary education, at the age of eleven, students enrol in the three-years lower secondary school (from 6th to 8th grade). At the age of 14, pupils are free to choose whether to enrol to high schools or to fulfil compulsory education in the *Formazione Professionale* (vocational training).

The *Formazione Professionale* consists of a plethora of three-years training courses for early integration into the labour market. The local administrations (the Regions) are responsible for the activation and the content of *Formazione Professionale*, so to adapt the supply to the local labour market characteristics. At the end of the course, students obtain a specific diploma (*Qualifica*), which, however, does not entitle them to enrol to tertiary education.

3.3.1.1 The structure of high schools

High school lasts five years (from 9th to 13th grade). There are several types of high schools.[15] Three of them are traditional lyceums. The *Liceo Classico* focuses on the study of humanities and classical culture. The *Liceo Scientifico* promotes the acquisition of knowledge and methods typical of mathematics, physics and natural sciences. The *Liceo Linguistico* addresses the study of different linguistic and cultural systems. Three additional types of lyceums exist. The *Liceo Artistico* focuses on arts and the study of aesthetics. The *Liceo delle Scienze umane* focuses on the study of psychological and sociological theories. Eventually, the *Liceo Musicale e Coreutico* addresses the theory and the practice of music and dance.

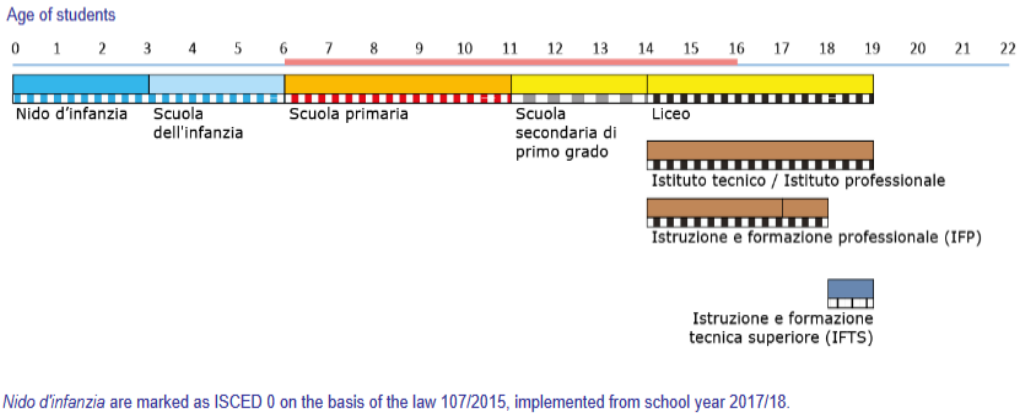
Technical high schools provide both theoretical and practical knowledge for specific economic sectors. Students can enrol in two main sectors: economics and technologic. Like technical high schools, the vocational high schools equip students with both practical and theoretical knowledge. However, the vocational track prepares students for direct entry into the labour market. Students can enrol in two main sectors, namely, services and handicraft. A higher degree of content-individualisation characterises Vocational high schools as compared to Technical tracks.

Unconditionally to the chosen track, high school completion provides students with a degree (the *diploma*). Anyone with a *diploma* can enrol at university.[16]

[15] Descriptions and finalities of each kind of high school retrieved at the following link in May 2019: https://www.miur.gov.it/web/guest/scuola-secondaria-di-secondo-grado

[16] Nevertheless, while the transition rate to university is around 73.8% among who hold an academic degree, it is sensibly lower for those who completed technical (33.1%) and vocational (11.3%) paths (MIUR 2017).

Fig. 3.4: The structure of Italian education. Source: Eurydice



3.3.1.2 Social segregation in nowadays Italian schools

Substantial socioeconomic inequalities still characterise the Italian educational system (Pisati 2002; Triventi 2014; Contini and Scagni 2013). Social background exerts a critical influence already in the early stages of education (Gambetta 1987; Schizzerotto and Barone 2006).

Students belonging to privileged social *milieus* get higher notes at the end of middle school. Higher notes, in turn, raise their probability to enrol in the academic track. Unconditionally to their performance, high-SES kids opt for the lyceum, that is, make more ambitious tracking choices.

Despite the formal easy access of the educational system, the 'decision effect' (cf. Chapter 1) is much more significant in Italy than in other tracked systems around Europe (Jackson 2013) and accounts for more than half of the social background influence in Italy (Contini and Triventi 2016).[17]

During the last century, a consistent reduction of inequalities in access to upper secondary education has occurred. However, long-term trends show that the social background effect on track choices has enlarged (Panichella and Triventi 2014). The gap in lyceum enrolment among social classes is persistent: in recent cohorts, it has only

[17] Horizontal segregation in high school triggers inequalities at the university level. As recalled in Note 18, high school choice is of crucial importance, and both transition rate to tertiary education and degree completion rate drastically vary across tracks. Secondary effects have an essential impact on the overall stratification of educational attainments.

slightly declined (Guetto and Vergolini 2016). The number of high schools graduated from traditional lyceums has progressively slowed down in the last decade.[18]

Educational segregation in enrolment and achievements severely concerns immigrant students as well, whose share has known a rapid growth in the last years (Santagati 2013). As Santagati (2015) points out, the picture of the Italian policies concerning immigrant students is ambiguous. On the one hand, Italy has opted for a fully inclusive pattern of integration of incoming pupils. Schools typically assimilate international students in regular classes, although they might enrol the student in a grade in which she\he is overage to nurse language learning (postponed enrolment). In the last two decades, schools have promoted an 'intercultural approach', in the effort of emphasising the richness of cultural diversity. Against this scenario in the educational realm, Italy is one of the countries with the strictest rules governing citizenship-acquisitions, the latter based mostly upon jus sanguinis. The two elements contribute to the contradictory situation in which immigrant students are typically enrolled in schools but mostly left without any juridical right of citizenship.

According to the official statistics published by the Ministry of Education (MIUR 2019), the share of students without Italian citizenship is slightly below 10%. Choices of enrolment in the secondary schools are significantly stratified, although there is a slight increase of enrolment in the Traditional Lyceums by Italian-born international students. In facts, immigrant students are just 4.2% in the traditional lyceums, while they reach 8.2% and 12.5% in the technical and vocational high schools respectively.

Even if, on average, immigrant students perform below the national mean, they significantly choose less often academic tracks as compared to their equally performing peers with Italian citizenship (see Appendix 2 for additional material, and comparison with the data used in this thesis).

Immigrant students are often older as compared to their classmates. At the age of 14, only 57.4% of the immigrant students attend the 9th grade. At the age of 18, the percentage of immigrant students in 13th is 34%. Overage immigrant students might be such due to postponed enrolment or past grade repetitions. Delay in grade progression

[18] While the share of traditional lyceum graduated has increased continuously in the last decades, from the academic year 2003/04 on variations have been at best of 3 percentage points (our elaboration of ISTAT data).

associates with a considerable dropout risk. In facts, both grade repetition and dropout rates are particularly high among immigrant students (ibidem).

3.3.1.3 The legislative framework of evaluation and grade repetition

This section provides an overview of the legislative framework relative to evaluation and, in particular, grade repetition.

The promotional regime in primary and middle school should be compliant with the Ministerial Decree n. 62/2017. Generally, students in elementary schools automatically advance to the next grade, even in the presence of only partially reached learning levels. When individual achievements are unsatisfactory, schools must activate *ad hoc* recovery strategies. In 'extraordinary cases', teachers can opt to retain a student to the unanimity. More open to grade repetition is the promotional regime in middle schools, in which teachers can decide to retain students if achievements are not satisfying.

The picture changes in upper secondary schools, in which grade repetition should abide by the DPR 122\2009 (*Decreto del Presidente della Repubblica*) and the OM 92/2007 (*Ordinanza Ministeriale*). According to the legislative framework, the evaluation is the autonomous and competent expression of the teacher. The evaluation involves teachers' professional competences as an individual and as part of a collective, i.e., in coordination with the other colleagues. The evaluation concerns the students upon three main aspects: their learning process, their behavioural conduct, and their overall scholastic performance.[19]

Teachers arrange both intermediate and conclusive examinations to test students' scholastic performance. An *ad hoc Piano dell'Offerta Formativa* (Plan for Formative Goals- POF) outlines the learning contents and goals for each year and the three subsequent years. Each school, in line with the established Ministry of Education's standards, draw-up its own POF.

The *Collegio dei Docenti* (teacher's assembly; henceforth CD) adapts the POF to the specific needs of the class. Moreover, the CD defines the criteria and procedures of

[19] Behaviour is an essential part of the evaluation with the expressed aim to teach students that 'their freedom is realised in respect of the others and of the rules governing civil institutions' (OM 92/2007)

the evaluation. The CD also assures a prompt communication of the outcome of the evaluation to families.

According to the law, the evaluation has an educational goal as it concurs to the self-evaluation process, encourages personal responsibility, and improves students' academic performance by stressing their potentialities and weaknesses. As such, the evaluation is a process that accompanies students through their whole formative career.

3.3.2 Teachers' expectations

As mentioned above, the CD and the school's director (or one of her\his delegate) oversee the evaluation. If necessary, CD votes and decides by majority.

In the specific case of the upper secondary schools, the end-of-year final evaluation works as follows. Each teacher recommends a mark for her\his subject. Students are admitted to the next grade (promoted) when they obtain a mark greater than or equal to 6\10 for each subject in the POF, including the behavioural conduct.

In the case of underachievement in one or more subjects, the evaluation is done considering 'the possibilities the student meets the expected formative objectives and contents by the end of the year' (OM 92/2007).

When the chances of recovery are remote, teachers directly retain students, who must hence enrol in the same grade the following year. When the chances of recovery are substantial, the CD may postpone its judgement and arrange an individualised plan of recovery to help the student to catch up. Catch up activities, as essential parts of the POF, are compulsory for underachieving students in any intermediate and the final evaluation.[20]

Before the start of the following academic year (usually at the end of August), the CD must verify through an *ad hoc* examination whether students have cached up. It eventually deliberates students' 'admission' (promotion) or, in case of unsatisfactory results, their 'not admission' (grade repetition).[21]

[20] The recovery activities must last at least 15 hours. By expressed will, parents can exonerate their children.

[21] Between 1994 and 2008, the 'reparation exams' in high schools had been substituted with the mechanism of 'formative debts'. Underachieving students in a specific subject received a 'formative debt' that they had to recover by the end of the first half of the subsequent year.

The law makes express reference to teachers' expectations, which are a vital element in influencing students' performance. Therefore, teachers' decision not only looks at the effort made by failing students before the beginning of the subsequent year but also considers the expectations of academic recovery. As such, grade repetition is likely to vary according to the vast room that the legislative framework leaves to school policy and the professional activities of teachers.

3.3.3 Grade repetition's distribution

Table 3.3 below displays the share of non-admitted students by grade for two recent academic years, which coincide with those of the empirical work presented in the next chapters.

In middle schools, grade repetition rate remains between 2% and 4%. Grade repetition rate is substantially larger in high schools. The trend is slightly toward a reduction of grade repetition share.

Tab. 3.3: The share of non-admitted students by grade for two recent academic years

AY	6th	7th	8th	9th	10th	11th	12th	13th
2013-2014	3.8%	3.2%	2.8%	16.3%	10.5%	9.7%	7.2%	4.2%
2014-2015	3.4%	2.1%	2.8%	14%	8.4%	7.8%	5.3%	4.6%
2015-2016	3.2%	2.7%	2.4%	14%	8.3%	7.8%	5.5%	4%
2016-2017	2.5%	2.1%	2.1%	13.4%	8.3%	7.6%	5.5%	3.8%

Table 3.4 displays the geographical distribution of repeaters by macro-regions (source: ISTAT 2018), and it outlines the precise data for the regions included in the present study, namely, Piedmont, Lombardy and Veneto. With the only exception of the Islands (Sicily and Sardinia), the gap between areas is negligible. Piedmont and Veneto are slightly under the national mean, while Lombardy is slightly above.

Tab. 3.4: The geographical distribution of repeaters in Italy

		Repeaters' rate
North West		6.7
	Piedmont	6.3
North East		6.1
	Lombardy	6.8
	Veneto	5.7
Middle		6.2
South		6.3
Islands		8.5
Italy		**6.6**

3.4 Conclusion

The previous pages have aimed at framing the grade repetition into the Italian context.

The institutional and structural setting of the Italian schools moved from a strict merit-based, highly selective system to more open and social promoting scholastic institutions. Primary education is now comprehensive, and virtually all students regularly advance throughout it. Social promotion regime also characterises middle schools.

The picture is still different in high schools. Students face different promotional standards according to the chosen track and, within these standards, grade repetition remains substantial. The photographic evidence gathered in the present Chapter has suggested that a merit promotion regime is still mainly in place at the high school level.

The present chapter has concluded the literature overview of the thesis. The next chapters will instead guide the reader into the extensive empirical analysis of the present work.

Chapter 4: The dataset

4.1 Introduction

As mentioned in the Introduction, the empirical investigation of this work has the following two research aims. The first one is to inspect the determinants of grade repetition and to estimate the contribution of socioeconomic status and migration background beyond their influence over prior academic performance.

The second aim is to examine what happens next to grade repetition. On this subject, the approach is both descriptive (as it looks and compares academic careers among repeaters) and evaluative (as it tries to estimate the causal impact of grade repetition on subsequent careers).

The empirical analysis relies on an extensive longitudinal dataset of unique students' school careers. A co-signed protocol by the research unit, the Italian Ministry of Education and The National Institute for the Evaluation of the Educational System (INVALSI) has ruled both the data releases and the privacy warranties.

Although each of the three following empirical chapters gives the necessary pieces of information of the exploited data, the present chapter displays a detailed description of the data releases and the database structure.

The intense work of linkage between administrative and socioeconomic individual records successfully retraced students' whole academic careers. The resulting database is an authoritative source of insights in a deeper understanding of the determinants and the consequences of grade repetition.

The thesis mainly focuses on two students' cohorts. The analysis makes use of Cohort A to estimate grade repetition determinants in Chapter 5. Chapter 7 exploits Cohort B to analyse after-grade-repetition consequent careers.

4.2 The sources of data

As anticipated, the empirical investigation relies on an *ad hoc* built a longitudinal database of students' careers. The database is the product of incorporation between two sources of information: The National Register of Students (*Anagrafe Nazionale degli*

Studenti- ANS) and The National Institute for the Evaluation of the Educational System (*Istituto Nazionale per la Valutazione del Sistema di Istruzione e formazione*- INVALSI).

4.2.1 The National Register of Students (ANS)

ANS collects administrative data from Italian schools. The centralisation of students' data is a new practice in Italy. In 2005, the legislative decree n.76 (D.Lgs 76/2005, art.3) established the creation of a 'national system of registers of students' for the primary and secondary educational cycles. However, only in 2010 a decree by the Ministry of Education (DM 74/2010) has enabled the schools to start the collection of data by setting up an actual register.

The register can work because ANS assigns to students a unique individual identification number that follows them throughout their entire educational careers (from primary school to tertiary education). ANS refers to this code as *'Codice SIDI*' ('Sistema Informativo Dell' Istruzione'- Schooling Information System, henceforth SIDI).

The ANS obtains students' information directly from schools. In facts, the systematic collection stills an ongoing process, and the 2010 collection was dramatically partial. Moreover, up to the present days, substantial portions of the country have left uncovered due to limited or null data transmission.

As shown by Table 4.1, ANS stores data following three main criteria: the scholastic year, the educational level and the type of incoming data.

Tab. 4.1: ANS releases

Scholastic years	Educational level	Type of incoming data
• 2013-2014 • 2014-2015 • 2015-2016 • 2016-2017	• Middle school (from 6th to 8th grade) • High school (from 9th to 13th grade)	• Personal information and status of enrolment • Evaluations

The data release regards the secondary educational level, which in Italy comprehends a lower cycle of three years and an upper cycle of five years (cf. Chapter 3). The research team involved in the project received the administrative records of four

academic years (2013-14, 2014-15, 2015-16, and 2016-17). As mentioned above, data regarding previous spells are severely incomplete.

The reader should also notice that, as mentioned earlier, the geographical unbalances in data collection circumscribed the data-release to the population of students enrolled in three Northern Italian regions of Lombardia, Piemonte and Veneto.

For each academic year, ANS keeps two separate archives on students' career. The first one collects their biographical and enrolment information. Variables on enrolment comprehend the class and the school a student is enrolled in, and the status of enrolment.

From ANS, the research team obtained the total of sixteen separate data releases (that is equal to the number of academic year times the number of educational levels times the number of archives). For example, a data release could cover the biographical information of all the students enrolled in middle schools in the scholastic year 2014-2015.

In practice, ANS set an encryption algorithm of SIDI codes to safeguard students' privacy. While the research unit did not have access to the encryption algorithm, ANS shared it with INVALSI, which could hence proceed to the data linkage.

4.2.2 The National Institute for the Evaluation of the Educational System (INVALSI)

INVALSI is an autonomous institute under the supervision of the Ministry of Education (D.Lgs 258/1999), established in 1999. Its key-goal is to evaluate the efficiency and the efficacy of the Italian educational system by framing it in the international context. From 2007 on, INVALSI has subjected students to standardised tests in Italian, Mathematics and English at different points of their academic carrier. The standardised test derives from the Programme for International Student Assessment (PISA) test on literacy and numeracy.[1]

[1] PISA test has gained growing legitimacy as the worldwide standard for the evaluation of the performance of students and, consequently, of educational systems. The fact has risen many concerns around a supposed ongoing 'PISA shock' experienced by many educational systems. In 2014, an open letter to the PISA's Director Andreas Schleicher addressed by many academics around the world has tried to

INVALSI also asks students to fill a specific survey, a Student Survey, which collects critical information on their socioeconomic background and allows to compute the PISA's Economic, Social and Cultural Status (ESCS). The ESCS index combines the parental occupational status, the highest level of parental educational attainment, and the information on the wealth and the educational resources in the household.

As Table 4.2 shows, for each year of the collection, INVALSI has allowed access to the results of the standardised performance test. When the collection is in fifth and 10th grade, INVALSI also released socio-economic data collected within the Student Survey. To sum up, INVALSI released a total of 22 single data collections.

Tab. 4.2: Data releases by INVALSI

Grade in the scholastic year 2013-2014	Data collections (points in time)
6th	5th (2012-13); 8th (2015-16)
7th	5th (2011-12); 8th (2014-15)
8th	5th (2010-11); 8th (2013-14); 10th (2015-16, 2016-17)
9th	8th (2012-13); 10th (2014-15, 2015-16)
10th	8th (2011-12); 10th (2013-14)
11th	10th (2013-14)

4.3 The dataset

This section pictures out the final database, that is, after the intense work of elaboration and linkage described in Section 4.4 of the present chapter.

4.3.1 Variables

Table 4.3 displays the variables in the present analysis. Each student is identified by a unique code, which is the encrypted version of the SIDI code (see above). Besides, to each student, I assigned an increasing cohort number that corresponds to the first-observed academic year-grade combination.

summarise these concerns ("OECD and PISA Tests Are Damaging Education Worldwide - Academics | Education | The Guardian." n.d. Accessed July 28, 2019. www.theguardian.com). They claim that PISA has contributed to a substantive escalation in the exclusive reliance of quantitative measures of academic performance and consequently has 'harmed our children and impoverished our classrooms'.

Tab 4.3: the list of the variables (individual level only)

Variable	Description	Source
Student id. number	An encrypted version of the SIDI unique code, which identifies a student (see in the text)	ANS
Cohort id number	A progressive number which identifies students by the first academic year and the first grade they appear in the dataset	Original elaboration from ANS
Gender	As Female or Male	ANS
Over-age	Whether the student is first-time observed in a grade in which is older than the modal age. Time-invariant.	Original elaboration from ANS
Migration status	Derived from the first recorded citizenship and the province of birth (see in the text).	Original elaboration from ANS
Socioeconomic status index (ESCS)	PISA's Economic, Social and Cultural Status (see above)	INVALSI
Parental education	Parents' highest education in three cuts: no upper secondary diploma; upper secondary diploma; tertiary degree or higher.	INVALSI
Mark in Italian	End of year mark in Italian assigned by teachers. Range: 0-10, with 6 as passing mark.	ANS
Mark in Mathematics	End of year mark in Maths assigned by teachers. Range: 0-10, with 6 as passing mark.	ANS
Mark in Behaviour	End of year mark in Behavioural conduct assigned by teachers. Range: 0-10, with 6 as passing mark.	ANS
Invalsi Italian test score	Standardised to the Italian population mean. Range: ±5	INVALSI
INVALSI Mathematics test score	Standardised to the Italian population mean. Range: ±5	INVALSI
Educational outcome: grade repetition *versus* promotion	End of year teachers' determination about a student's progression (see in the text).	ANS
School id.	Alphanumeric number of schools	ANS
Class id.	Progressive identification number of the individual classes	Original elaboration from ANS
Track id. and type	At the class level. Reduced to four categories: traditional lyceums, other lyceums, technical and vocational high schools (see in the text)	Original elaboration from ANS

In addition to gender, two main pieces of biographical information exist. First, whether a student first observed over-age as compared to the modal age of her classmates (in the first year of observation). This information is time-invariant. Second, each student displays a migration status derived from the intersection of two ANS' pieces of information: (i) country of first citizenship and (ii) province of birth. As shown in Table 4.4, the dataset identifies three categories of migration status: Italian for with Italian citizenship and born in Italy; second-generation immigrant for those born in Italy without

the Italian citizenship; and first-generation for those born abroad and without the Italian citizenship.

Tab. 4.4: Migration status

		Italian citizenship	Foreign citizenship
Province of birth:	Italian	Italian	Second-generation Immigrant
	Missing	-	First-generation Immigrant

Besides the ESCS index (see above), the dataset includes the highest parental educational attainment that allows identifying the cultural endowments of the students. The variable displays three categories: no high school diploma; high school diploma; tertiary degree or higher.

As already outlined, the information on the academic performance of the students is quite informative. On the one hand, it embraces end-of-year marks in the two main subjects (*i.e.* Italian and Mathematics) and the behavioural conduct, the latter contributing to the final evaluation of the students by teachers (see Chapter 3). On the other hand, in 5th, 8th, and 10th grade (see Figure 4.1 below), the dataset stores the scores of the Invalsi standardised test in Italian and Mathematics.

Naturally, the data contains the final educational outcome, which is the crucial variable in the present empirical investigation. At the end of each academic year, teachers decide whether the student passes the grade (promotion) or she\he must repeat the year (grade repetition).[2]

Besides the individual information, the data also collects a rich set of contextual information. For each student, it is possible to know the class and the school, as the type of track (Traditional Lyceum, Other Lyceum, Technical and Vocational high school; see Chapter 3).

In order to shed some light on the school-context, the analysis includes the school-average values of the variables at the individual level (not shown in Table 4.3): percentage

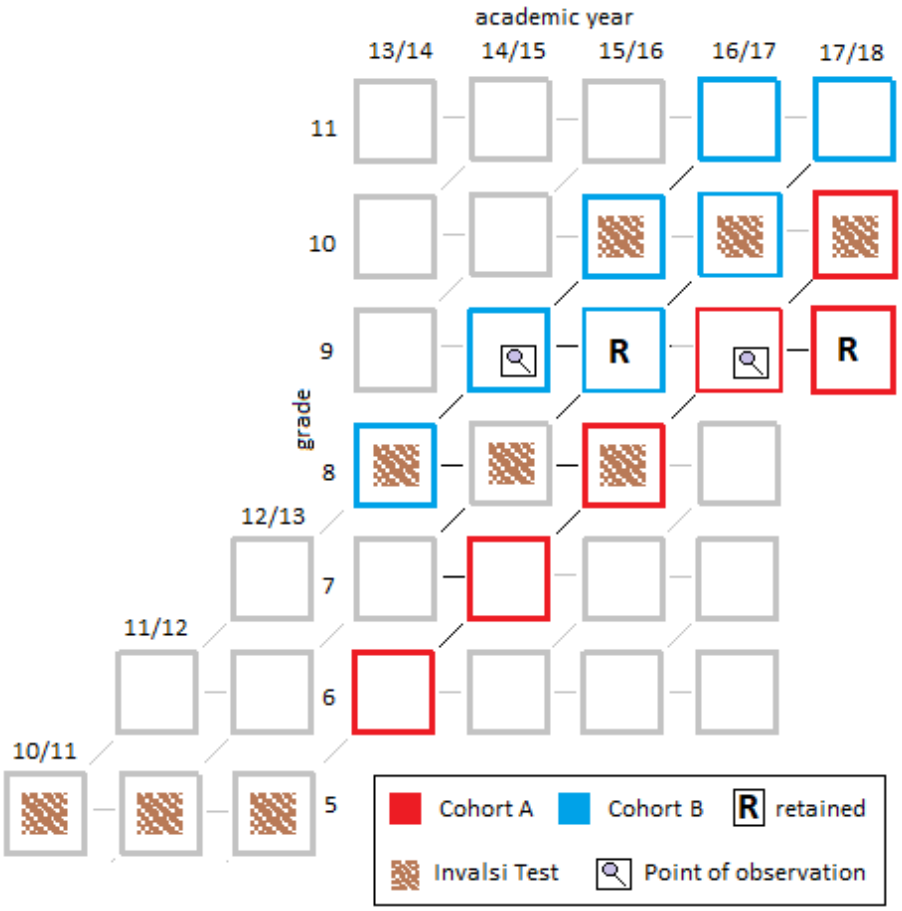
[2] In fact, ANS registers the status of the students at the end of the academic year: admitted to the next grade; not admitted to the next grade due to too many days off; not admitted to the next grade due to grade retention.

of female, over-age and immigrant students; average ESCS, or proportion of parents with a tertiary degree within a school; average obtained marks and test scores.[3]

4.3.2 The student cohorts

The empirical investigation of the present chapter the list of individual students that belong to three specific cohorts of students. The scheme in Figure 4.1 summaries the panel structure of the data.

Fig. 4.1: the cohorts of students in the empirical analysis



[3] Marks are likely to be anchored to the average of the school. Hence, the mark is considered as the individual deviance from the school mean. End-of-year marks are endogenous if one is interested in the educational outcome of the same year. Hence, I take the previous year individual deviance from the school-average mark. Refer to Chapter 5 of the present thesis for a systematic discussion on this issue.

4.3.2.1 Cohort A

Cohort A embraces the careers of those students enrolled in 6th grade in the scholastic year 2013-14. The empirical analysis exploits this cohort to measure the impact of socioeconomic and migration background on the probability to get retained at the end of 9th grade.

Figure 4.2 shows the students' flow relative to cohort A. The dataset contains 183,903 unique students attending the first year of middle school in 2013-14. In Figure 4.2, the percentages are relative to each transition.

For each year, the share of grade repetition is in red. There is, moreover, a portion of non-retained leavers that does not appear the subsequent year in the administrative Register (in black). Although limited in 2014-15 and 2015-16, attrition is consistent in the transition toward high school. Indeed, 14% of the 8th graders do not enrol in any of the schools of the dataset. In Appendix 1, the reader can find an analysis of the composition of the observed early dropout.

The possible reasons of early dropout from the database could be: i) the student moved to a school outside the three Regions of observation; ii) the student illegally dropout from school; iii) the student enrolled to vocational training. Enrolment in vocational training is possible only after middle school completion and, hence, it is likely to explain a consistent share of dropouts in the transition toward high schools. The Ministry of Education has access to vocational training's online enrolments at the end of middle school, although this piece of information eventually was not incorporated in this work.

The significant share of students split into the different high school kids at the end of eight grade, namely, traditional lyceums, other lyceums, technical and vocational high schools. The analysis in Chapter 5 looks at the grade repetition share at the end of the first year of high school.

Table 4.5 displays Cohort A's descriptive statistics for the 9th graders in the academic year 2014-15. The average grade repetition rate is 12.7%, but there is a considerable variation within high school types: it is 7.7% among traditional lyceums while higher than 16% in technical and vocational high schools.

4.3.2.2 Cohort B

The second cohort of observation, Cohort B, includes the careers of those students enrolled in 8th grade in the scholastic year of 2013-14. The analysis employs cohort B for two purposes. On the one hand, it considers retained students at the end of 9th grade in 2014-15 intending to study the influence of socioeconomic and migration background on the after-grade repetition careers (cf. Chapter 6). Second, the analysis selects similar retained versus non-retained students at the end of 9th grade in 2014-15 and compares them to estimate the impact of grade repetition on specific educational outcomes (cf. Chapter 7).

Table 4.6 shows the distribution of students in Cohort B over the variables in the analysis. Instead of the ESCS index, as a proxy of social position, Cohort B uses three cuts of the highest parental education attainment (no secondary degree, secondary degree, and tertiary degree).

Figure 4.3 illustrates the aggregate trajectories of Cohort B's retained students at the end of 9th grade in 2013-14. Differently to Figure 4.2, the percentages in Figure 4.3 refer to the total number of retained students, that is, around 25 thousand of students.

The chart eloquently shows that more than two-thirds of after-grade-repetition careers are problematic. In respect of the total number of retained students in 2013-14, the share of students who immediately drop out (32.5%) roughly equals that of students who reach the 11th grade (33.2%) in 2016-17, and dropout reaches 64.1% within the subsequent three years.

Tab. 4.5: Cohort A descriptive statistics

	Trad. lyceums	Other lyceums	Technical	Vocational	Total
Grade repetition	7.6%	10.4%	16.4%	16.6%	12.7%
Female	45.9%	77.0%	34.6%	49.2%	50.1%
Over age	1.4%	3.1%	5.4%	14.8%	5.1%
1st gen.	2.3%	3.1%	5.8%	9.1%	4.7%
2nd gen.	3.7%	4.0%	7.3%	9.5%	5.9%
ESCS	0.44	0.08	-0.22	-0.42	0.00
mis. ESCS	6.1%	9.6%	11.9%	29.7%	12.4%
Italian	8.1	7.6	7.0	6.5	7.3
Maths	8.3	7.2	7.0	6.4	7.3
Behaviour	9.2	9.0	8.6	8.5	8.8
Italian score	0.76	0.31	-0.07	-0.58	0.18
Maths score	0.97	0.04	0.01	-0.66	0.19

Tab. 4.6: Cohort B descriptive statistics

	Trad. lyceums	Other Lyceums	Technical	Vocational	Total
Grade repetition	8.3%	12.0%	19.1%	18.9%	14.7%
Female	45.3%	78.5%	35.7%	46.0%	49.3%
Over age	1.9%	5.0%	9.1%	22.8%	8.7%
1st gen	3.1%	4.4%	8.1%	12.3%	6.7%
2nd gen	2.7%	3.1%	5.5%	6.5%	4.4%
no diploma	13.3%	24.7%	36.9%	45.0%	29.6%
diploma	40.1%	44.9%	41.4%	25.8%	39.2%
degree	37.7%	18.6%	8.6%	3.9%	17.4%
mis.	8.8%	11.8%	13.1%	25.2%	13.8%
Italian	8.2	7.2	6.9	6.4	7.2
Maths	8.0	7.5	6.9	6.5	7.2
Behaviour	9.2	9.0	8.6	8.3	8.8
Italian score	0.73	0.27	-0.08	-0.56	0.15
Maths score	0.92	0.06	0.05	-0.54	0.20

Fig. 4.2: Sankey chart for Cohort A

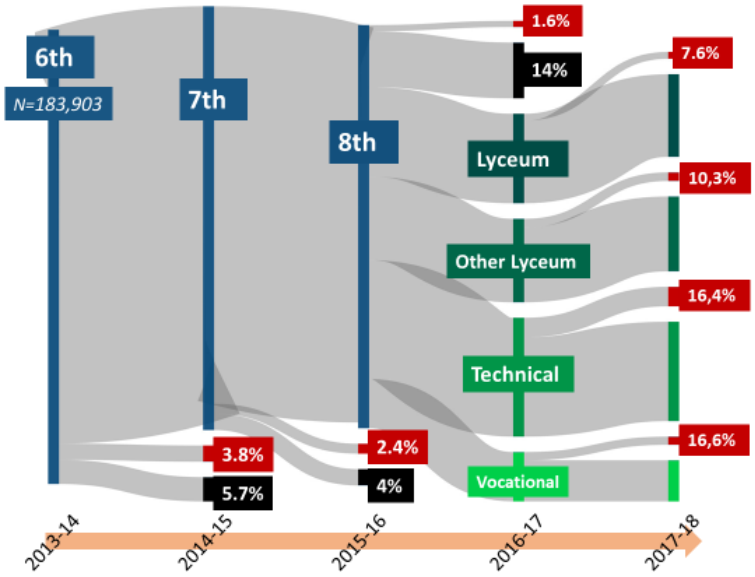
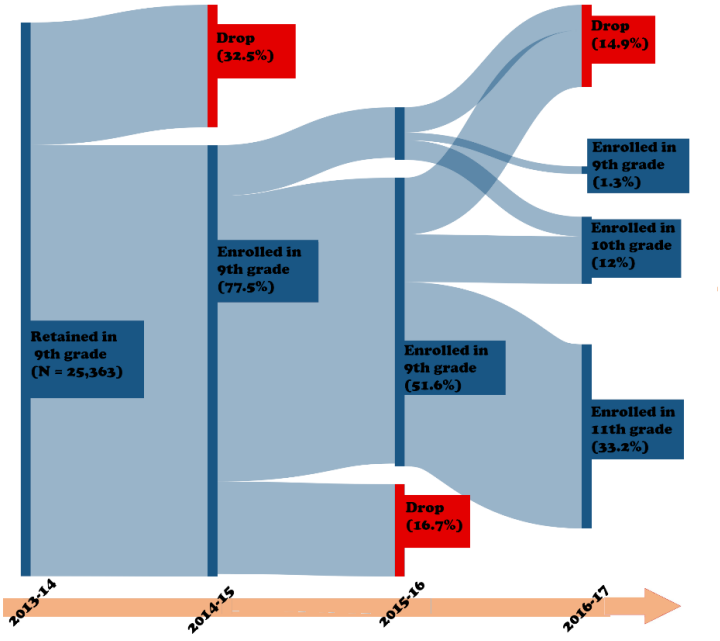


Fig. 4.3: Sankey chart for retained in Cohort B



FOR FIGURE 4.2 AND 4.3**: Red boxes: GR share. Black boxes: dropout share. The higher share after the eighth grade is due to the exit to vocational training programmes, which do not classify as high schools and do not allow enrolment in tertiary education — **NB: percentages by academic year for Figure 2; percentages refer to the total number of retained students in 2013-14 in Figure 4.3.

4.4 The database construction and its inner structure

The large amount of data in this work has been harmonised in a relational database. Before moving forward in the data description, this section briefly introduces some basic knowledge of relational database theory.

4.4.1 The relational database (brief notes)

As Elmasri and Navathe (2015) point out, relational databases are a collection of relations. Each relation represents a real-world relation or a state of reality. For example, a student with her\his characteristics. A relational schema is a type of assertion on reality. It comprehends the relation name and a list of attributes:

$$STUDENT\ REGISTER(SIDI\ code, gender, age, \dots, citizenship) \qquad (1)$$

while the relation itself is a collection of *n*-tuples (the rows) and *n*-attributes (the columns) in a simple flat table, as Table 4.7 below exemplifies.

Tab. 4.7: An example of a flat table

Student Code	*Gender*	*Age*	*…*	*Citizenship*
102310391	F	12	…	Italian
102310392	M	15	…	Foreign
102310393	M	13	…	Italian
…	…	…	…	…

Each tuple (row) is a fact of the assertion represented by the relational schema. Importantly, each relation displays at least one key attribute that must be unique, not null and time-invariant. Indeed, 'two distinct tuples in any state of the relation cannot have identical values for (all) the attributes in the key' (ibidem p.159).

A primary key identifies a tuple in a relation uniquely. Following the example of relation number 1 above, the SIDI (as the key) uniquely links each student to some characteristics, namely, gender, age and citizenship. The relation state at a given time, *i.e.*

the current relation state, reflects only the valid tuples that describe a specific state of the real world. Each value in a tuple is an atomic value (*ibidem*), for it is not divisible in the real world.

Relations connect through foreign keys. A foreign key is an integrity constraint between one or more tables; that is, it identifies an attribute in a relationship that is present in another relation and allows to connect the two relations. For example, the relation:

$$CAREER\ TABLE(observation\ id, SIDI\ code, \dots, educational\ outcomes) \quad (2)$$

identifies the academic situation of students in a single academic year, including the final educational outcomes. In Relation 2, the $SIDI\ code$ links the relation $STUDENT\ REGISTER$ to that of $CAREER\ TABLE$; that is, the $SIDI\ code$ is a key in $STUDENT$ and a foreign key in $CAREER\ TABLE$. Differently to what happens in Relation 1, in the $CAREER\ TABLE$, each $SIDI\ code$ may (and it is likely to) appear more than once. In contrast, $the\ observation\ id$ is the primary key in the $CAREER\ TABLE$ and is unique for each tuple.

Eventually, a single relational database would typically consist of many inter-connected relations. A database schema is a set of relations:

$$DATABASE(STUDENT\ REGISTER, \dots, CAREER\ TABLE) \quad (3)$$

4.4.2 Database building and tables description

Figure 4.4 displays the database schema. Each rectangle is a relation, that is, a flat table of tuples and attributes. Attributes that are primary keys are in bold, while foreign-keys are in italics. (When both bold and italics, an attribute is both a primary and a foreign key).

Fig. 4.4: The relational structure of the database

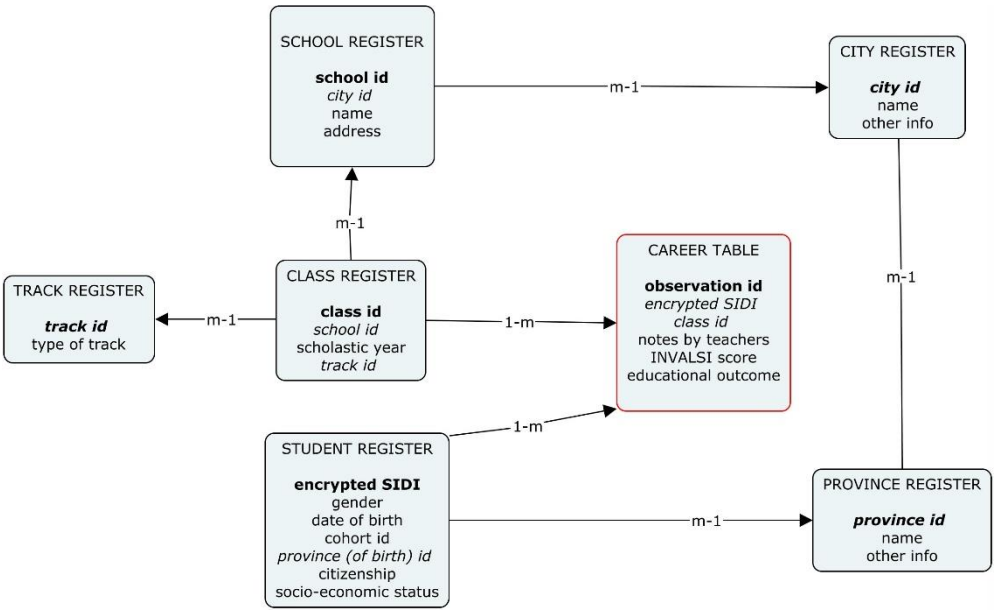


Figure 4.4 represents the connections among relations with arrows. The ‘m' and the '1' shows which kind of connection links the two relations: when 'm', the foreign key in the incoming\outcoming relation corresponds to multiple tuples (rows); when '1', a foreign key in the incoming\outcoming relation corresponds to a unique tuple (row). For example, 'm-1' corresponds to a tie in which many tuples of one relation link to a unique tuple in another relation.

What follows is a description of each table and the variables (tuples) that collect.

- *School register* table collects the information on each school present in the database
    - *School id,* a unique code for each school
    - *Name and address (with the city id)*
- *Class register* table collects the information on each class present in the database.
    - *Class id,* a unique code for each class at each academic year
    - *School id*
    - *Academic year*
    - *Track id.*

- *Track register*
    - *Track id,* a unique code for each track;
    - *Type of track*
- *Student register*
    - *SIDI code,* a unique code for each student;
    - *gender*
    - *date of birth*
    - *cohort id,* which identifies students according to the academic year and the grade of the first observation
    - *province of birth, if first recorded citizenship is Italian*
    - *first recorded citizenship,* which is Italian or non-Italian
    - *socioeconomic status.* Socioeconomic status is recorded as ESCS index (see above) or as the highest parental educational attainment between three levels (no secondary degree or no diploma; secondary degree or diploma; tertiary degree)
- *Career table*
    - *Observation id, the unique numeric code assigned to each administrative record in the data set*
    - *SIDI code (see above)*
    - *Class id*
    - *Marks by teachers.* Marks are end-of-year marks in Italian, Maths and Behaviour. The mark is numeric and goes from 0 to 10, with 6 as the satisfactory threshold.
    - *INVALSI standardised scores.* INVALSI provides a Rasch-based score on both Italian and Mathematics.
    - *Educational Outcome.* This variable registers the status of the students at the end of the academic year: admitted to the next grade; not admitted to the next grade due to too many days off; not admitted to the next grade due to grade repetition.

The two additional tables of cities and administrative provinces are linked only as a reference to the structure of the database. Nonetheless, the relational form could comfortably accommodate new relevant contextual information.

Chapter 5: Equally Performing, Unfairly Evaluated

"What's your rank? What's your number? Are you descending or are you climbing?
The ranking is nothing more than a shooting and you know it! You can't claim to have won the game.
I will be flunked to life. You're gonna waste your time looking for a result. Always flunked I will be."
Punkreas- Sotto esame (2000)

5.1 Introduction

The chapter represents the empirical core of the thesis, and it models the probability of repeating an academic year. Although grade repetition is a contested educational practice, scholars have understudied it in the Italian context. Because of this, the present analysis intends to clarify the influence of a comprehensive set of risk factors and, in particular, it emphasises the influence of socioeconomic and migration backgrounds on the probability to incur in a grade repetition at the end of 9th grade.

As such, the present analysis is intimately linked to the research field of educational inequalities, as it seeks to examine the influence of such backgrounds of students on their probability of grade repetition during the first year of high school. The analysis herein relies on a rich array of academic performance measurements, which allow for estimating the residual influence of parental resources on that probability. To the author's knowledge, this study is the first attempt to address the issue with such comprehensive data in the Italian context.

In general, grade repetition, the practice of denying a student progression in his or her educational career for an additional year, might occur when teachers believe that a student has not met the learning goals for a given grade. The primary purpose is to give the failing student another chance to assimilate the learning contents and catch up to his or her peers.

Opponents of grade repetition have maintained that the practice potentially harms the self-esteem of repeaters and eventually incites their definitive disaffection towards school (for an extensive review of these themes, cf. Chapter 2). Grade repetition is an intervention that targets the bottom tail of the academic performance distribution; thus, it

is significantly likely to affect disadvantaged students (*i.e.* those in low socioeconomic strata or\and with a migration background). International reports broadly attest that the odds of having repeated a grade are consistently higher for disadvantaged students.[1]

The following inequality represents grade repetition's selection:

$$P(GR|\boldsymbol{B}_{dis}) > P(GR|\boldsymbol{B}_{adv}) \tag{1}$$

Where $\boldsymbol{B}_{dis}$ and $\boldsymbol{B}_{adv}$ are disadvantaged and advantaged (socioeconomic and migration) backgrounds, respectively. Hence, the probability of experiencing a grade repetition (GR) given a disadvantaged social background is higher than the probability to experience a grade repetition given an advantaged social background.

Undoubtedly, inequality (1) comes not unexpected. The overall distribution of educational attainments partially results from differentials in scholastic performance, which varies across social classes *via* parental resources and influence (cf. Chapter 1). On average, students with disadvantaged (advantaged) background perform worse (better) than their fellow advantaged (disadvantaged) peers. Boudon (1973), among others, has identified the socially driven achievement gap as the 'primary' or 'performance' effect of social background. Less privileged students are likely to be over-represented among repeaters simply because they are likely to be over-represented among low-performing students.

The main goal here is to measure the residual contribution of a student's background once the analysis accounts for the performance effect. Following on inequality (1), the present chapter engages in the estimation of the gap:

$$P(GR|\boldsymbol{B}_{dis}, AP_{low}) > P(GR|\boldsymbol{B}_{adv}, AP_{low}) \tag{2}$$

Where AP_{low} indicates a low academic performance. The goal is to measure the gap in grade repetition probabilities between disadvantaged and advantaged students who are comparable the most in terms of academic performance. If inequality (2) is accurate,

[1] The 2017 report on scholastic dropout by the Italian Ministry of Education states that 'the percentage of pupils who have dropped out of the school system is 5.1% for repeaters pupils, and 0.4% for pupils in good standing' (MIUR 2017, p.23). However, this knowledge is of little help to our investigation, as the report does not control for any index of previous academic performance.

then disadvantaged students experience an above sharper grade repetition probability even when they display the same (low) level of academic performance.

Given this, two research questions are at the core of the present investigation:

(Q1) *Do socioeconomic and migration background have a residual effect on the probability of being retained, given low performance? (Q1a: By which possible mechanisms?)*

(Q2) *Does the schooling environment influence GR's occurrence?*

The outline of the chapter is as follows. The next section presents a detailed data description. Afterwards, the chapter models grade repetition probabilities for a cohort of northern Italian-based sixth-grade students in 2013/14 who successfully transitioned to the 9th grade (high school) in 2016/14. It then observes the outcome of the ninth grade for promoted (who pass the grade) versus retained (who need to repeat the grade) students and estimates the contribution of socioeconomic and migration background. The chapter also describes the methodological strategy of this study in the respective section. The primary objective is to estimate the contribution of the individual background on the probability of grade repetition. Ultimately, it also explores the contribution of the school context in terms of how the composition of the student body influences individual grade repetition's probabilities. The last part of the chapter proposes some mechanisms that concur to explain the present findings.

5.2 Data and methodology

In the Italian school system, students ordinarily complete middle school in by the age of 14. Subsequently, they must enrol in high school or opt for vocational training (cf. Chapter 3). When enrolling in high school, students must choose from four main tracks: traditional lyceums, other lyceums, technical and vocational. Although university enrolment is open to any person who holds a secondary degree, academic-oriented and

generalist tracks (traditional lyceums and other lyceums) specifically prepare students for university enrolment. Nonetheless, a drastic social segregation flaws the high school tracks and, while traditional lyceums are highly selective, vocational high schools tend to attract socially disadvantaged students.

5.2.1 Cohort A

The data include the population of northern Italian students in the first year of middle school (cf. Chapter 4). The present empirical investigation merged the students' administrative records with the data from an autonomous evaluation institute, namely The Italian National Institute for Evaluation (INVALSI). This institute cyclically subjects students to standardised performance tests. Moreover, it collects data about their socioeconomic situations. The administrative data of students correspond to four academic years (from 2013\14 to 2016\17). Students who complete middle school and enrol in one of the four high school tracks in the following year constitute the object of the present investigation. Figure 4.2 (p.108) displays the aggregate flow.

5.2.2 Variables

Our dependent dichotomous variable is the event of grade repetition at the end of ninth grade. Its value takes zero if students pass the grade (GR = 0). When they must repeat the year, it is equal to one (GR = 1).

The two key covariates are the socioeconomic status (SES) and migration background (MIGR). The present empirical analysis operationalised SES by using the PISA Index of Economic, Social and Cultural Status (ESCS) (range ±2), which INVALSI computes by factorial analysis of parental occupation, parental educational attainments and household cultural resources.[2] Scholars have broadly utilised the ESCS index for

[2] INVALSI is an autonomous institute under the supervision of the Ministry of Education (D.Lgs 258/1999). The INVALSI institute was founded in 1999 with the key goal of evaluating the efficiency and efficacy of the Italian educational system by framing it in the international context. From 2007 onwards, INVALSI has subjected students to standardised tests in Italian, mathematics and English at different points in their careers. INVALSI designed its test according to (and hence comparably to) the PISA international test.

international research on educational inequalities. The ESCS index is a mix of parental occupation, parental educational attainment and home environmental resources.[3]

Migration background is the combination of the first recorded citizenship and the province of birth; both recorded in ANS (cf. Table 4.4. in Chapter 4). Migration background takes three categories: Italian-born students with Italian citizenship; Italian-born students without Italian citizenship, *i.e.* second-generation immigrant students; Foreign-born students without Italian citizenship, *i.e.* first-generation immigrant students.

Our control variables include gender, whether or not the student is in the 6th grade observed over-age as compared to the modal age of the cohort, and prior academic performance (8th grade).

As described in Chapter 4, the data in this thesis cover a rich array of controls for past performance. The insights regarding past performance derive from two sources. For each observed academic year from sixth to ninth grade, ANS receives from schools the end-of-year marks from teachers in the two main curricular subjects, namely Italian and mathematics. ANS also collects data on behavioural conduct, which formally supplements the academic evaluation. The three marks are on a 10-point scale in which 6 out of 10 reflects sufficiency. The data also integrate INVALSI's standardised test scores in Italian and mathematics at the end of eighth grade. As already highlighted in Chapter 4, the INVALSI and PISA tests are highly comparable.

There are substantial gains from jointly utilising the two sets of performance measurements. The overlap in content between teachers' marks and standardised tests is somewhat limited, as they mirror the various intertwined and complex dimensions of the academic performance. Teachers' marks mainly cover curricular matters. Marks also incorporate the evaluation of crucial non-cognitive skills, such as aptitude towards school and academic and job aspirations. Teachers must oversee broader learning achievement, which might be registered by standardised tests. At the same time, the use of the INVALSI

[3] ESCS index 'was created on the basis of the following variables: the International Socio-Economic Index of Occupational Status (ISEI); the highest level of education of the student's parents, converted into years of schooling; the PISA index of family wealth; the PISA index of home educational resources; and the PISA index of possessions related to "classical" culture in the family home' (https://stats.oecd.org).

scores offers a performance indicator that is consistent across learning environments and thus provide for a between-schools comparable measure.

The models account for both high school and middle school contextual effects (see the methodological section). High school contextual effects include the schools' means for the individual-level variables (e.g. the share of migrant students or the average eighth-grade score in mathematics). These effects are an object of study (see Q2, which investigates the contextual schooling effect).

Meanwhile, middle school contextual effects allow us to control for the heterogeneity of grading in eighth grade. Teachers are likely to mark students concerning the distribution of academic performance within schools. When students cluster by social class in middle school, teachers' marks are no longer comparable between schools because teachers assign the same mark to different levels of academic performance. The model overcomes this significant limitation by controlling for the middle school context.[4]

5.2.3 Descriptive stats

Table 5.1 compares promoted and retained students. Retained students comprise 12.8% of the population. Among retained students, there are fewer girls (39% versus 52%). Moreover, among retained, both first- and second-generation immigrant students are present at a rate higher than 10%, which is almost double their population average. The ESCS index is -0.125 among retained students, while the population average is 0.16, and the admitted students' average was 0.20.

The two groups display substantial differentials in marks and test scores, and, as expected, promoted students outperform retained students. Except for the behavioural conduct, average teachers' marks and standardised test scores were significantly lower among retained students than among admitted students. The group of retained students contains fewer girls, a larger share of migrants (12% versus 6% in the whole population) and a far lower SES (-0.125 compared to 0.20 among promoted students).

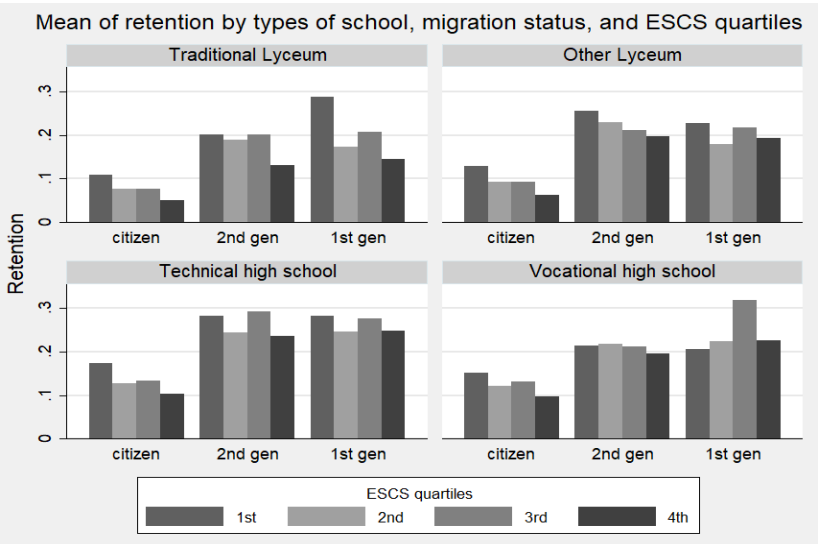
[4] I centre the individual-level variables on the mean of the population. Thus, when individuals take zero, they are at our data mean on that specific variable. The aim is to give a more explicit interpretation of our results.

Tab. 5.1: Descriptive statistical comparison between admitted and retained students

	Retained (12.8%)	Non retained (87.2%)	Total (100%)
Female	51.7%	39.3%	50.2%
Age old	3.8%	12.7%	4.9%
2nd gen. immigrant	3.8%	10.2%	4.6%
1st gen. immigrant	4.9%	11.5%	5.7%
ESCS index	0.2	-0.13	0.16
Italian mark (8th grade)	6.46	7.48	7.35
Mathematics mark (8th grade)	6.31	7.45	7.31
Behaviour mark (8th grade)	8.15	8.96	8.85
Italian test score	-0.49	0.29	0.19
Mathematics test score	-0.6	0.31	0.19

Moreover, the share of retained students drastically varies between high school types. While it is 7.68% in traditional lyceums and 10.39% in other lyceums, it is twice that rate within technical (16.55%) and vocational (16.84%) high schools. Figure 5.1 (below) focuses on grade repetition distribution across the three key dimensions of the track, socioeconomic background and migration background. For each high school type, the graphs plot the observed share of grade repetition (vertical axis) for migration and socioeconomic background. The four horizontal bars represent the SES quartiles ranging from the bottom (0-25th) to the top (75th- 100th) of the horizontal axis.

Fig. 5.1: The distribution of retained students across ESCS quartiles, types of high school and migration background



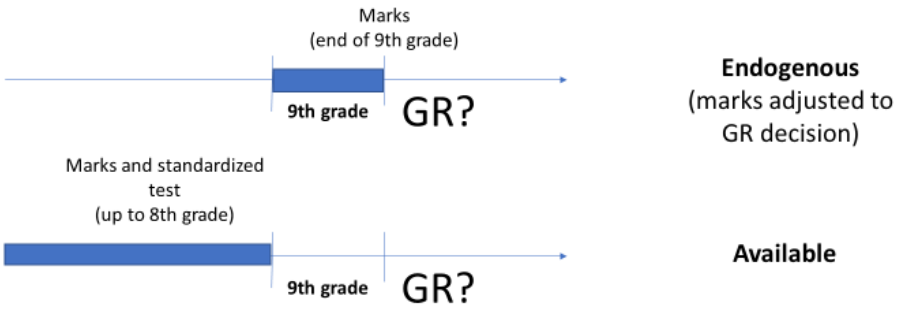
Unconditionally to high school type, high-SES Italian citizen students, the observed grade repetition is lower as compared to low-SES Italian citizen students. The picture is similar as compared to that of second- and first-generation immigrant students, although differences are less pronounced ambiguous. In facts, especially in non-academic tracks, differentials in grade repetition distribution are much smaller, and the trend is not always distinct.

5.2.4 The models

As noted, the goal of this empirical investigation is to model grade repetition at the end of ninth grade. The analysis also aims to estimate the contribution of socioeconomic and migration background to the selection into grade repetition, after controlling for past academic performance. The identification of the model relies on a rich vector of past academic performance, which embraces both subjective (teachers' marks) and standardised (INVALSI test scores) measurements.

The optimal estimation of the socioeconomic and migration's residual contribution should rely on the performance information the closest to the grade repetition. Unfortunately, I am compelled to discard the ninth-grade marks, as it is highly likely that teachers adjusted their marks according to their decision. As such, ninth-grade marks are expected to be endogenous concerning the GR decision. Under this assumption, if a teacher wants to promote a student, they will slightly increase the student's mark to raise them above the threshold. Conversely, if they wish for a student to repeat the year, they will adjust the student's marks downwards to place them under the threshold. Given these considerations, the information about academic performance is exogenous up to the end of the eighth grade (Figure 5.2).

Fig. 5.2 Ninth-grade marks are endogenous; analysis used information about past academic career up to the end of the eighth grade



5.2.4.1 Model Zero

First, the analysis fits a simple set of logistic regressions to offer a usable graphical indication of the GR social gap. In Model Zero, only the bottom and the top quartiles of the SES distribution contributes to the estimation; this enhances the clarity of the results. Each logistic regression fits the data for each of the two groups and each of the four high school tracks (eight in total). At this stage, the only explanatory variable is an *ad-hoc* summary index of the five performance measures: marks on Italian, mathematics and behaviour and test scores in Italian and mathematics.[5]

The text refers to the following specification as Model 0a:

$$P(GR = 1) = \beta_0 + \beta_1 X \; for \; each \; combination \; of \; T \; and \; SES_q \qquad (3)^6$$

Where T is high school track, SES_q is the SES quartile (top or bottom 25%) and X is the performance index. Model 0b runs the same specification while stratifying for migration background:

$$P(GR = 1) = \beta_0 + \beta_1 X \; for \; each \; combination \; of \; T \; and \; MIGR \qquad (4)$$

[5] The index is the first component of a factor analysis, which includes all performance measures. Eigenvalue = 2.41. Factor loading: Marks: Italian: 0.77; Maths: 0.75; Invalsi: Maths: 0.80; Italian: 0.78.

[6] For the sake of clarity, I present the equations in their linear form.

5.2.4.2 Multilevel specifications

Next, the aim is to accurately measure the influence of socioeconomic and migration background on GR probability. Hence, one concern is to control for school-level unobservable heterogeneity, which would yield biased estimates (cf. related Appendix 4). In avoiding such risk, the analysis performs a school-level random intercept logistic regression. The model focuses on two main covariates: the socioeconomic index (SES) and the migration background ($MIGR$). As noted, it controls for a rich vector of variables, including previous academic performance. Furthermore, the model accounts for the effect of the schooling environment by considering the contribution of high school contextual effects. In the context of this empirical investigation, the high school-level variables indicate the contextual effect; that is, the additional influence of school composition on GR probability that is not accounted for at the individual level) (cf. *ibidem*).

Random effect models consider the contribution of the hierarchical levels in clustered data. The correlation of unobserved factors among same-cluster observations would otherwise produce biased standard errors (Grilli and Rampichini 2016). Another advantage of the random intercept model is that it enables investigation of the school effect by regressing over the school's means of the individual-level variables (see Chapter 4 and related Appendix 1 for a discussion on centring decisions).[7]

The model also includes the middle school averages of variables. Teachers evaluate students concerning each school context. A student who attends a low-demand middle school would display higher marks than a similar student enrolled in a high-demanding academic environment. Since low-demand schools are likely to be attended by disadvantaged students, the (non-standardised) teachers' marks might be problematic to compare. For example, high-SES peers are likely to earn lower grades in middle school simply because they attend more highly demanding middle schools. Hence, the bias can

[7] The main advantages of fixed effects (Fes) are that the analyst avoids any specification of the cluster effects (and their correlation with the covariates) and that FE is a valid alternative when the number of clusters is minimal. However, fixed effects represent a loss in efficiency (high number of parameters) and preclude the measurement of cluster-level effects. Fixed effects may produce inconsistent estimations in the non-linear model: fixed effects inference on the observed clusters, whereas REs inference on a population of clusters. Notably, our results are consistent with FE specifications (namely high school FE, middle school FE, and high school × middle school FEs). The reader can find the detailed results in Appendix 4.

lead to an overestimation of the real GR gap. Nevertheless, middle school contextual effects in the model mitigate the heterogeneity in the assignment of middle school marks.

The analysis comprehends three specifications of a random intercept logit model. The first returns to Q1 (Do socioeconomic, and migration background have a residual effect on the probability of being retained, given low performance?), and focuses on the individual level by including the interaction between the socioeconomic index (SES) and the first citizenship ($MIGR$):

$$P(GR = 1) = \beta_0 + \beta_1 SES_{ij} + \beta_2 MIGR_{ij} + \beta_3 (SES_{ij} \times MIGR_{ij}) + \beta_4 C_{ij} + \beta_5 \bar{X}_j + \beta_5 \bar{X}_Z + \varepsilon_{ij} + \mu_j \quad (5)$$

Where i, j and z are the subscripts for the individual, the high school and the middle school levels, respectively, C is the vector of the control variables in the model (gender, age and the array of academic performance) and X is the vector of all variables (control variables plus SES and $MIGR$) to account for the contextual effect both in high school and middle school.

Models 2 and 3 explore two cross-level interactions. The goal is to investigate whether the school context limits or favours the importance of a student's background (in terms of socioeconomic index and migration). Therefore, the model jointly interacts individual-level background and the respective school averages.

Model 2 interacts the individual SES (SES) with the average high school SES ($\overline{SES_j}$):

$$P(GR = 1) = \beta_0 + \beta_1 SES_{ij} + \beta_2 CIT_{ij} + \beta_3 (SES_{ij} \times \overline{SES_j}) + \beta_4 C_{ij} + \beta_5 \bar{X}_j + \beta_5 \bar{X}_Z + \varepsilon_{ij} + \mu_j \quad (6)$$

This model aims to clarify the school contribution in terms of the context. The purpose is to explain how the composition of the school affects the individual outcome above a student's SES.

Meanwhile, Model 3 investigates the interaction between the individual migration background and the share of migrant students at the high school level ($\overline{MIGR_j}$)

$$P(GR = 1) = \beta_0 + \beta_1 SES_{ij} + \beta_2 CIT_{ij} + \beta_3 (MIGR_{ij} \times \overline{MIGR_j}) + \beta_4 C_{ij} + \beta_5 \bar{X}_j + \beta_5 \bar{X}_Z + \varepsilon_{ij} + \mu_j \quad (7)$$

Figure 5.3: Gap in grade repetition probabilities for two groups of students on ESCS. Blue lines: top 25%. Red lines: bottom 25%.

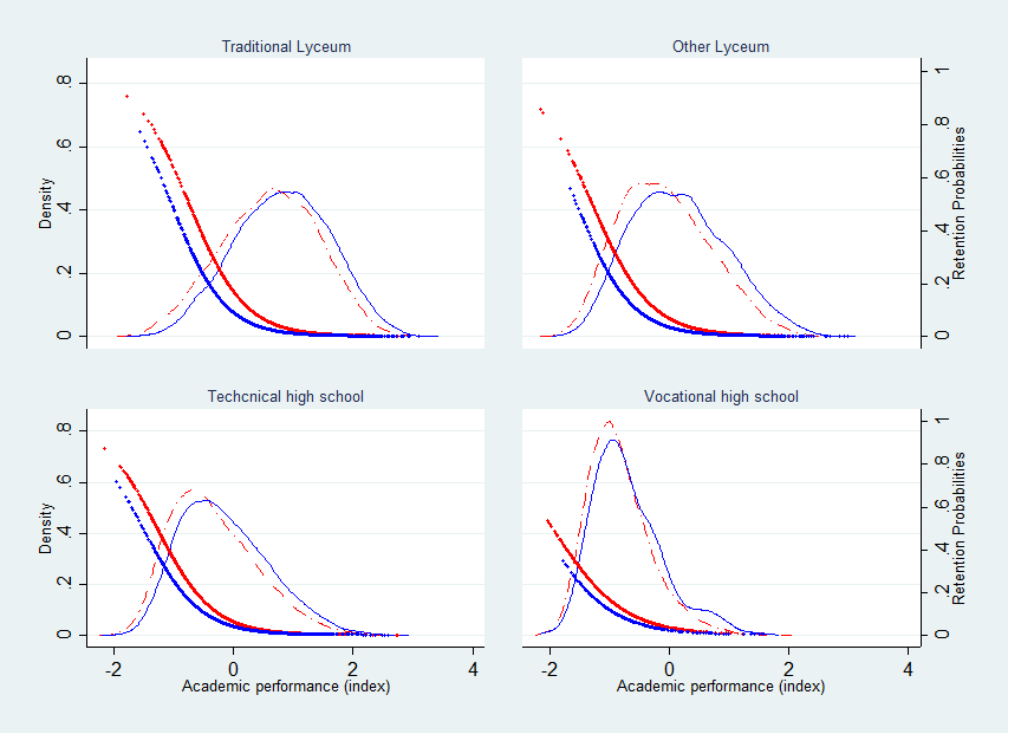
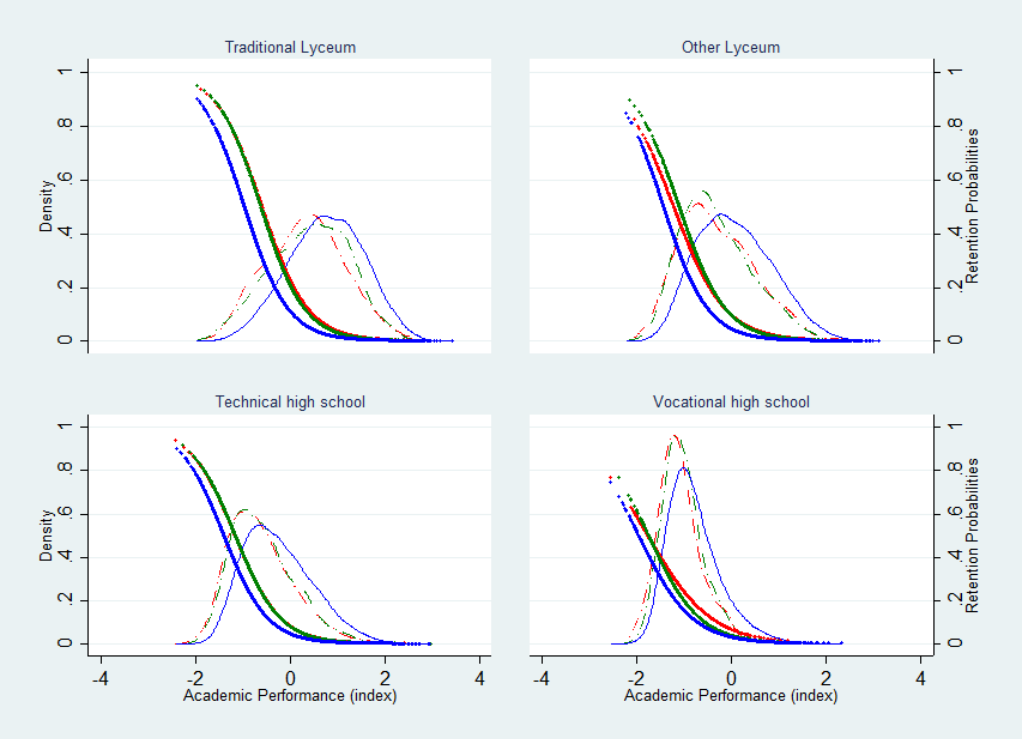


Fig. 5.4: Gap in grade repetition probabilities for migration statuses. Blue lines: citizen students. Green lines: 2nd generation immigrant students. Red lines: 1st generation immigrant students.



5.2.5 Results

Figures 5.3 and 5.4 (above) display the graphical outcomes for Models Zero-a and Zero-b. Each graph displays a high school track. In Figure 5.3 (Model 0a), the red and blue lines correspond to the bottom and top 25% of SES distribution, respectively. In Figure 5.4 (Model 0b), the red, green and blue lines represent first-generation immigrants, second-generation immigrants, and citizen, respectively. At each point in the academic performance index (horizontal axis, centred on the population average at 0), the graphs plot its group distributions (the bell curves in the background) and the estimated GR probabilities (the logistic curves).[8]

As expected, the graphs reveal a systematic achievement gap between the groups. The mean of the curves also differs from one track to another, with the traditional lyceum at the top-right and the vocational high school at the bottom-left. The graphs likewise illustrate that GR probabilities dramatically increased among low-performing students. One point below the population average, one out of five students is likely to experience a grade repetition. Two points below, the predicted GR probabilities are around 40%.

However, and most importantly, the 25% and 75% percentiles of SES distribution display substantial differentials in GR probabilities at any given point of the academic performance index. The same applies when Model 0b stratifies for migration background. There are marked differentials between migrant and citizen students in terms of GR probabilities, at equal levels of academic performance. The gap is accentuated in traditional lyceums and at low-performance levels.

Although controlling only for school track and past performance, Models 0a and 0b have the advantage of graphically illustrating a consistent gap in GR probabilities between socioeconomic and migration backgrounds. At any given point of the academic performance distribution, low-SES and immigrant students are well above in terms of GR risk as compared to their socially more advantaged peers.

[8] The graphs derive inspiration from Erikson and Jonsson theoretical work on primary and secondary effects (1996). Their intent was to graphically conceptualise gap in educational choices while holding ability constant.

To confirm these descriptive findings – and, more importantly, to accurately measure the GR gap – the empirical strategy gets into more complex specifications.

5.2.5.1 The social gap in grade repetition probabilities

Consistent with Models Zero-a and Zero-b, the results for Model 1 suggest the existence of a dramatic social gap in GR probabilities. On the y-axis, Figure 5.5 displays GR predicted probabilities; on the x-axis, it reports the selected SES percentiles. The graph plots three academic performance levels. The blue line corresponds to the average-performing students on all the performance measures. Meanwhile, students on the green line are one half standard deviation under the mean, and the red line marks one full standard deviation under the population mean.[9]

Figure 5.6, to acknowledge the interaction effect between SES and migration background, splits the predicted probabilities into citizen students, second-generation immigrant students, and first-generation immigrant students (only for the worst-performing group at -1 s.d. from the population mean).

The outcomes confirm that a dramatically higher GR risk accompanies low-SES and migration background. For these profiles, the GR risk is consistent across the three performance groups. While the absolute difference expands for low performers, the relative risk within the academic performance levels is rather constant. Ceteris paribus, a child in the first percentile, faces around twice the risk of being retained compared to a similar peer in the 99th percentile. For example, at a half standard deviation below the mean academic performance, a student in the first percentile has a 25% probability of being retained at the end of ninth grade, while a comparable peer in the 99th percentile experiences GR only 11 out of 100 times. On a shorter distance, a student in the 25th percentile is more exposed by 1.23 times to GR as compared to a peer in the 70th percentile of ESCS distribution (which translates in 1.3, 3.2 and 5.8 percentage points difference in the three performance groups).

[9] The shortcoming of this choice is it produces predictions for severely under-performing students and hence the resulting large GR probabilities in absolute terms would not reflect the descriptive statistics. However, I am interested in the relative differentials and gaps across SES and between migration backgrounds.

Fig. 5.5: Predicted GR probabilities by SES percentile and academic performance

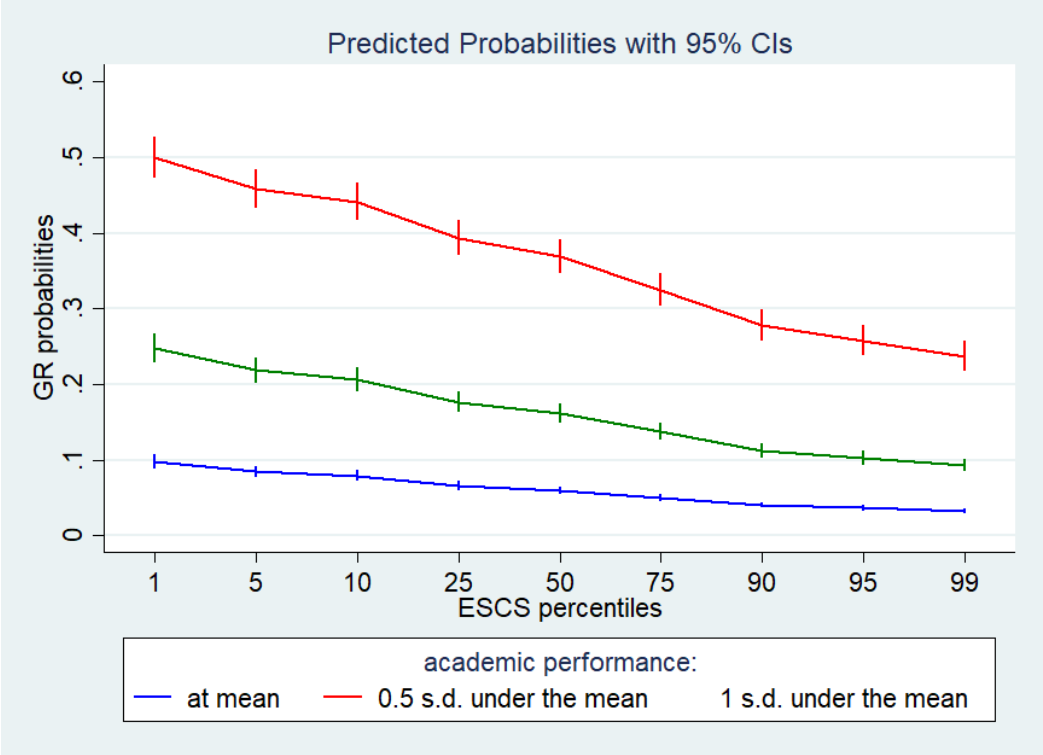
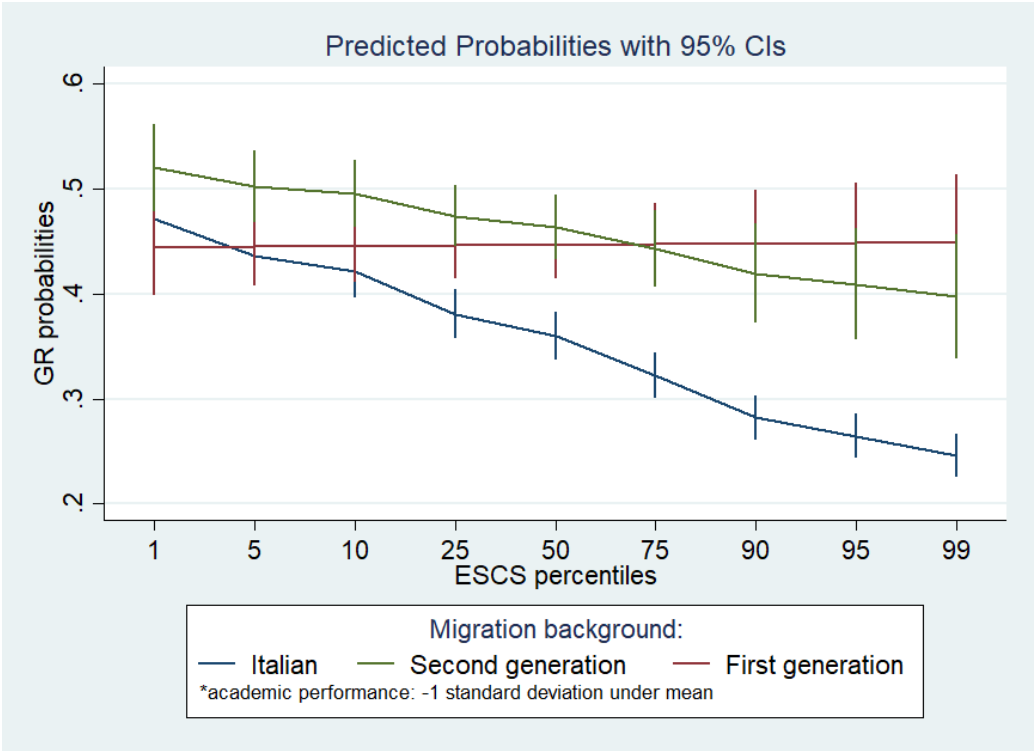


Fig. 5.6: Predicted GR probabilities by SES percentile and migration background for those students at negative one standard deviation in all performance measures



Crucially, these substantial differences hold despite controlling for a diverse array of academic performance measures and using other relevant controls and contextual variables (i.e. high school and middle school contextual variables). Therefore, the socioeconomic background exerts a dramatic influence on GR probability when academic performance is held constant.

The analysis also highlights the additional deprivation of migrant children. Figure 5.6 shows in red, green and blue lines first-generation immigrants, second-generation immigrants and citizen students, respectively.

Second-generation immigrant students significantly diverge from the blue line, and the gap gets wider as they move through the SES percentiles. For example, odds in the 1st percentile are 52% (2nd generation) versus 47.2% (citizen); the gap grows when approaching the median value of the ESCS distribution (46.3% against 35.9%), and at the 99th is 39.7% (2nd generation) versus 24.6% (citizen). Hence, high-SES migrant children are subject to GR significantly more often compared to their high-SES native peers, ceteris paribus.

The interaction between socioeconomic and migration background appears to be instead null on GR risk for first-generation immigrant students, as the flat red line in the graph. The GR risk for this group is always comparable to that of low-SES citizen students, regardless of parental resources.

Therefore, the estimates from Model 1 support the hypothesis of a social gap in GR probabilities along the lines of socioeconomic and migration background. At this point, the focus shifts to the second research question, which concerns school contextual effects.

Figure 5.7 presents the individual SES average marginal effects (y-axis) when interacted with the average SES at the high school level (x-axis). For each school track, the graph displays the marginal effect of SES for the school at the 25^{th}, 50^{th} and 75^{th} percentiles. From left to right, the figure displays the median-SES vocational high school at -0.43 (blue circle), the median technical track at -0.21 (yellow circle), the other academic track at 0.10 (grey circle) and the traditional lyceums at 0.36 (green circle), respectively.

Higher school SES average corresponded to a more significant negative impact of the individual SES. For example, one point of SES increment in a median vocational high school (i.e. at -0.43 SES index points under the SES population average) would affect the GR probability by -1.5 percentage points, whereas the same increase would have an impact of -3.3 points in a median traditional lyceum. In other words, the impact of the individual SES is more substantial in schools that are more affluent in terms of the composition of their student body. It is imperative to consider the implications of this finding for the highly segregated features of the Italian high school system that are apparent in the data. Notably, the tracks naturally form a line along the x-axis, with traditional lyceums positioned at the top-right of school SES distribution.

The same logic works in Figure 5.8, which consists of two graphs that plot the outcomes from Model 3. As the data and methods section has described, Model 3 focuses on the interaction between the migration background and the share of migrant students at school. On the horizontal axes, the graphs display the total migrant share at the school level. On the vertical axes, the graphs plot the average marginal effect (AME) for a second-generation immigrant background (upper graph) and a first-generation immigrant background (lower graph) on GR probability. For each track, the graph presents the AME for the 25^{th} (triangle), 50^{th} (circle) and 75^{th} (rectangle) percentiles in terms of total migrant school share. As noted in Chapter 4, school-types display significant differences in their average migrant share, which is only 3% in traditional lyceums, 4% in other academics, 6% in technical high schools and over 10% in vocational high schools.

The interaction between first-generation immigrant student and total migrant school share seems to be absent (upper graph in Figure 5.8). AMEs are statistically meaningful and positive in traditional lyceums. However, the effect seems to show a generalised higher GR risk for first-generation as compared to Italian students,

For the first-generation immigrants (lower graph, Figure 5.8), the AME's effects are small but not negligible. They range between 1 and 5 percentage points. As concerns the cross-level interaction, the higher the migrant students are in number, the higher the impact of being a first-generation immigrant. However, the slightly positive link between school and individual level is rather weak, and differences among schools are not statistically significant.

Therefore, the analysis highlights a dramatic social gap in GR probabilities. Socioeconomic and migration status critically contribute to the probability of experiencing grade repetition at the end of the ninth grade. Once again, the estimation controlled for past performance and relevant contextual effects.

The results of the present empirical analysis should raise severe concerns upon the GR selection process. Low-SES and immigrant students have systematically higher GR probabilities. In a single academic year, children who displayed the same performance were highly likely to experience opposite outcomes: grade repetition for the disadvantaged pupils and promotion for their advantaged peers. Besides, the school context is significant. The average SES in schools interacts with the individual SES. As a result, disadvantaged students in traditional lyceums contend with higher GR probabilities.

Fig. 5.7: Average marginal effects of individual SES by high school SES and type of school

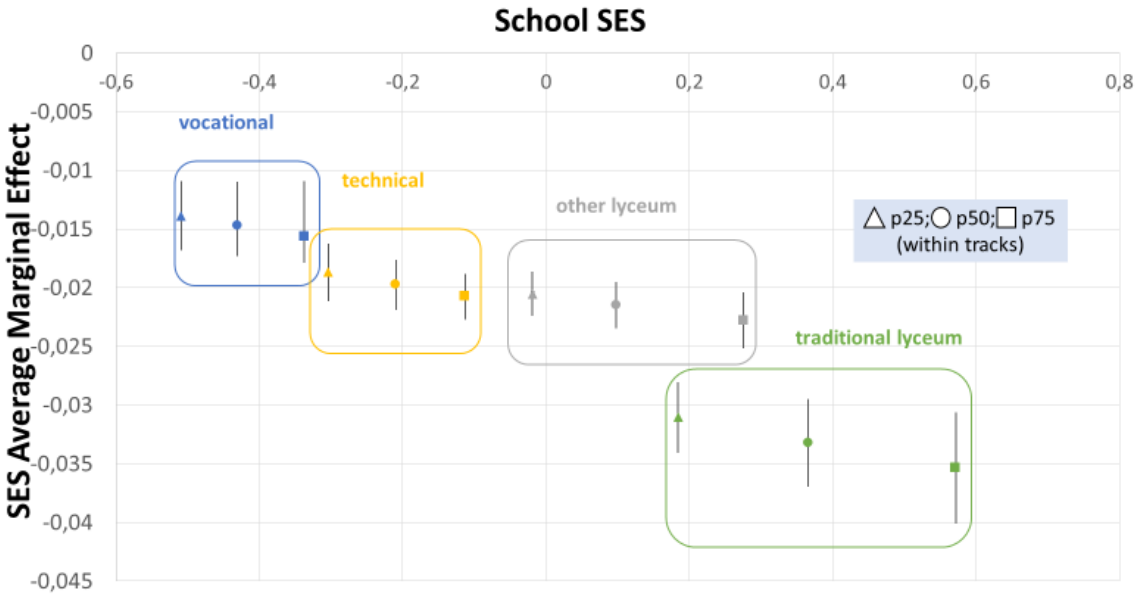
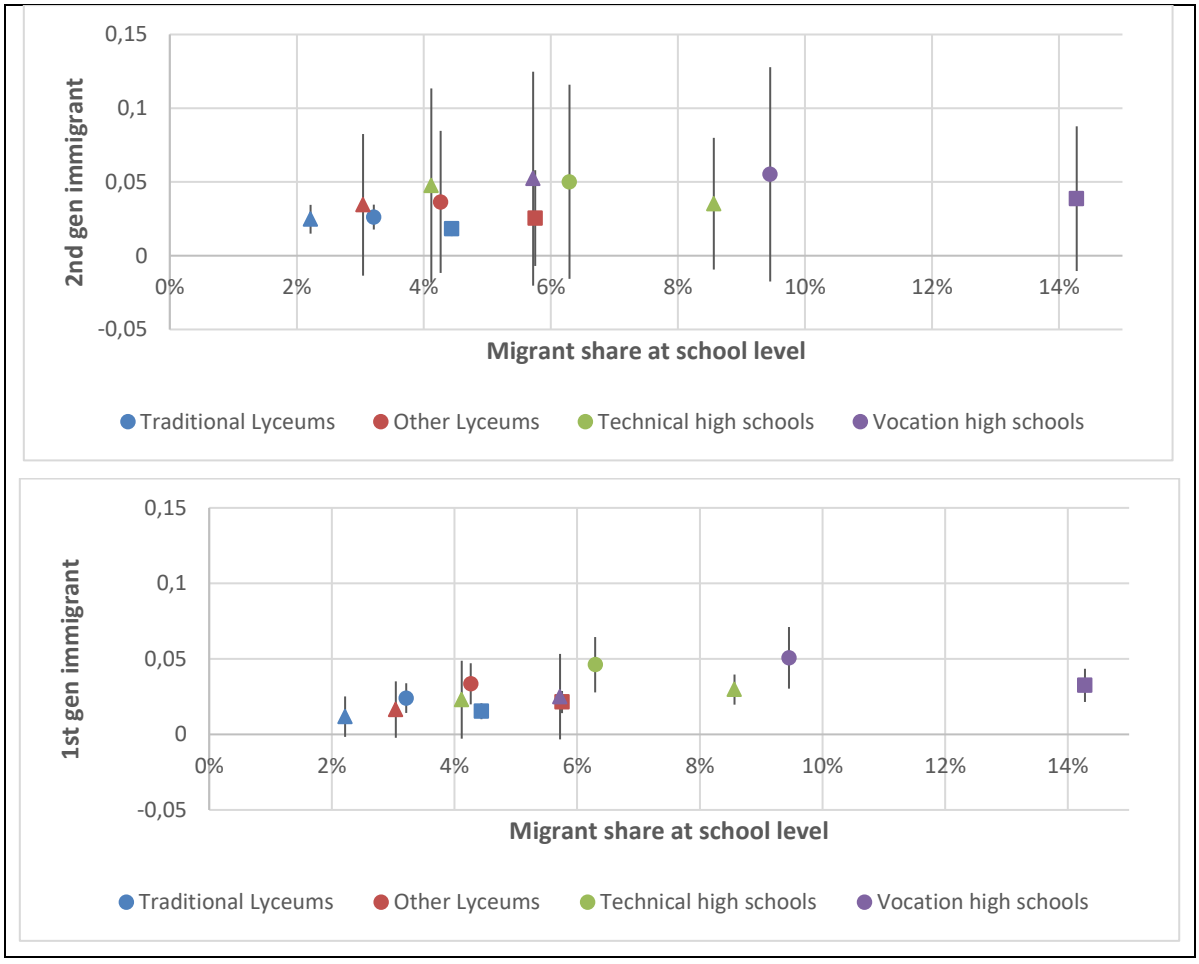


Fig. 5.8: Average marginal effects of migration background by type of school and within types of school percentiles of school migrant share (triangle p25; circle p50; square p75)



5.3 Discussion and conclusions

Disadvantaged students predominate the ranks of low performers, who are more likely to miss the learning thresholds that schools establish. Nevertheless, gaps in performance are excessively broad to explain the estimated differences in grade probabilities. Thus, the analysis suggests a substantial gap in grade probabilities that is unrelated to past performance.

The empirical results in this chapter unambiguously illustrate that the 'GR social gap' is broad and consistent across learning contexts. Even after accounting for academic performance, disadvantaged kids (i.e. with low SES and a migration background) had a drastically higher risk of GR at the end of ninth grade compared to their advantaged peers (i.e. those with high SES and a native background). Socioeconomic and migration backgrounds contribute dramatically to GR probabilities. Moreover, the GR social gap is even more severe in privileged scholastic contexts; accordingly, higher SES of the student body corresponded to a stronger impact of individual parental resources.

5.3.1 Potential mechanisms

Because of the empirical results, and consistent to the theoretical background detailed in the previous chapters, this section discusses how parental socioeconomic and migration background could exert an effect through two main channels: parental support during ninth grade and teachers' expectations for future academic prospects.

5.3.1.1 Parental support

Regarding the parental support hypothesis, there is no reliable information about academic performance during the ninth grade, so the analysis was not able to exclude that achievement gaps might emerge after the transition to high school due to differentials in parental support. Parental support is a function of parental socioeconomic and migration background. Families experience various constraints to the amount and type of resources that they can mobilise to promote their offspring's academic performance.

Fig. 5.9: Two equally low-performing students at the end of eighth grade displaying different academic performance after one year (i.e. at the end of ninth grade).

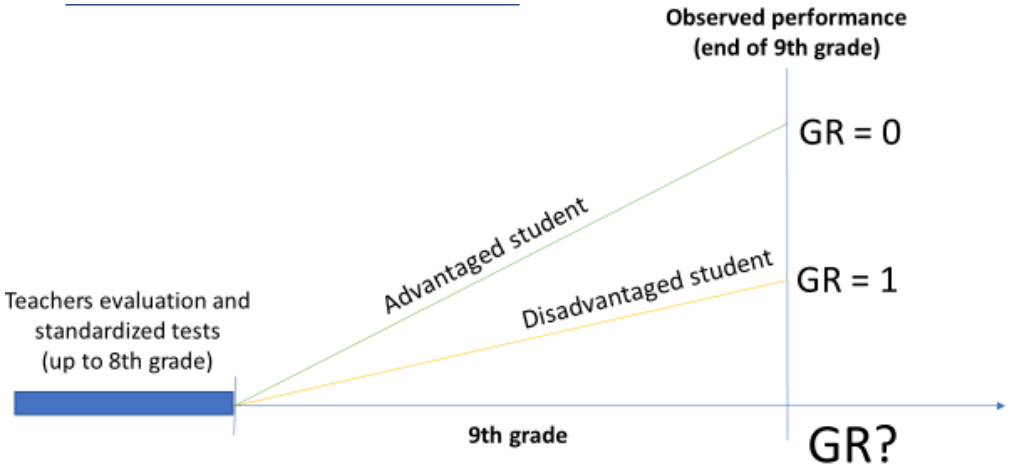


Fig. 5.10: Two equally low-performing students at the end of ninth grade are evaluated differently due to expected performance during the subsequent year.

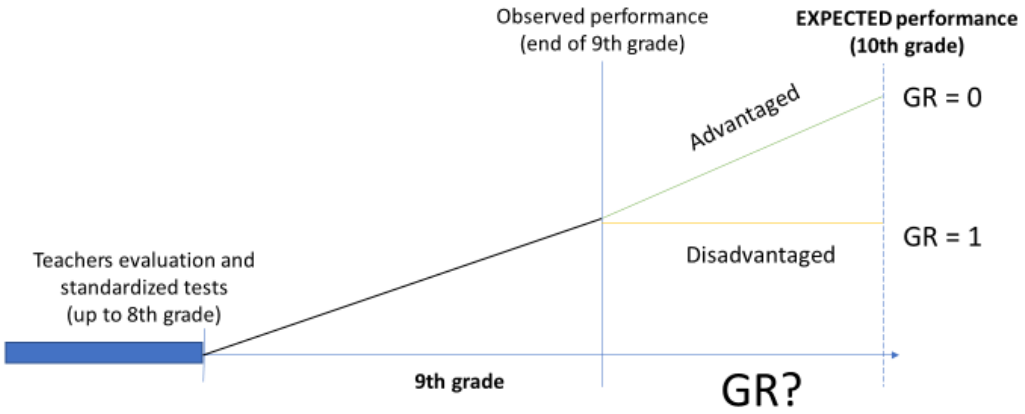


Figure 5.9 (above) exemplifies the scenario of two eighth-graders. The two lines indicate that the two students belonged to different social groups. The two were similar in marks and test scores until the end of eighth grade; thus, in terms of academic performance, they had an identical likelihood of failing ninth grade.

It is again notable that the available array of academic performance controls is rich. Comparable students in terms of performance are hence easily identifiable; i.e. same low marks in Italian, mathematics and behaviour (throughout middle school), and equally low standardised test scores in Italian and mathematics.

Imagine the scenario that, in just one year, the differentials in parental resources resulted in a gap in performance at the end of the year. Investment constraints limit action by families. Hence, it would appear that advantaged children can overcome the learning threshold that schools set, but their disadvantaged peers do not meet it. If parental support is the driver, then the gap in GR probabilities will emerge from the disparities in constraints that parents experience during their child's ninth-grade, which will lead to a substantial achievement gap at the end of the scholastic year.

The gap must be wide enough to provoke different outcomes (i.e. the advantaged child passes the grade, whereas the disadvantaged one gets retained). Given two equally low-performing students at the end of eighth grade, the one who can depend on superior cultural, economic and even social parental resources may be more likely to recoup. There are several types of parental involvement. For instance, parents can mobilise their economic capital to buy private coaching and tutoring for their children that could address their ability gap. Parents can also have a direct role in catching up by supporting their children through their cultural knowledge and skills. More highly educated parents are more likely to be able to help their children with homework and home studies.

Moreover, social skills are probably significant; some parents might appear to be more engaged with school life than others to elicit higher marks from teachers.[10]

[10] The crucial role of parental investments is even more prominent given that I analyse the year after the peculiar transition from middle to high school, which is likely to require parents to provide more strategic and focused educational investments.

5.3.1.2 Biases in expectations on the side of teachers

Notably, in the hypothesis regarding parental support, teachers merely 'register' the noticeable achievement gap and decide consequently. Thus, students who meet or exceed the set learning threshold are secure, while those who underperform get retained. No biases concur in their decision.

However, teachers might possess biased expectations for students' academic prospects. Teachers oversee the academic performance of students and, ultimately, decide to promote or to retain them. Teachers check whether a student's performance meets the learning threshold. When they evaluate that the student has not fulfilled the expected learning goals, they must form expectations about the student's likelihood of catching up. Thereby, teachers estimate the anticipated performance of failing students should not be retained. Literature has consistently emphasised the likelihood that teachers have biased academic expectations towards certain social groups (Becker 2012; Ferguson 2003; Gershenson 2003; Gregory and Huang 2013). Teachers' expectations are at the core of the GR decision, and those teachers tend to underestimate the academic prospects of low-SES students.

When expectations are biased, one might not need an actual achievement gap to explain the differences in GR probabilities. Figure 5.8 illustrates a scenario in which teachers have biased expectations. In this case, the two exemplary equally low-performing and socially distant children exhibited identical performance trajectories until the end of ninth grade. They both ended up below the learning threshold of the school. However, at the end of ninth grade, the more advantaged one elicited higher expectation from teachers and passed the grade.

In facts, teachers estimate that the performance of wealthy students will meet the learning goals of the next grade without the need for GR. Meanwhile, they anticipate a different performance trajectory for disadvantaged students. The result is the observed GR gap. The claim here is that the gap in expectations of recovery may arise from the intrinsic uncertainty which characterises GR decisions (and in fact processes of evaluation in general; see Marcoux and Crahay 2008). Grade repetition implies that teachers must maintain certain beliefs about student academic prospects. When information is not complete, salient social characteristics might influence the decision

about GR; thereby, teachers might infer the academic prospects of a student by relying on the average value observed within the social group to which that student belongs.

In the school realm, socioeconomic and migration status could function as informal academic standards of evaluation. In their relationship with teachers, students equipped with extensive cultural, social and economic resources demonstrated more suitable behavioural traits or could signal higher academic or occupational aspirations to others. In contrast, the disadvantaged students may convey a lack of posture or effort in the school environment that could suggest that they need to be punished or protected from failure. For example, teachers might believe that the average performance prospects for a child with non-tertiary-educated parents are lower than that of a peer whose parents possess a tertiary degree. Consequently, teachers might rely on such a belief when making decisions about GR.[11]

Moreover, there is an additional reason to devote extra care to account for teachers' expectations, and it somehow lies between the two mechanisms discussed above. The academic expectations that teachers express are likely to directly influence the performance of students to the extent that they act on students' self-perceived ability.

If teachers have biased expectations, then disadvantaged students comply with low expectations and thus start to undervalue their actual academic performance, including their future academic prospects. They even rationally conclude that any effort to improve their performance would be useless because their chances of success are limited. In this way, disadvantaged students convince themselves that they are not capable

[11] In a way, statistical discrimination is politically unproblematic because it does not imply any unfair role of schools and resolve in information bias. It indeed emerges as the optimal result of a systematic pattern of decision when a high level of uncertainty is involved. I know that GR decisions are founded on teachers' expectations, which, by definition, are estimations that are subject to variable margins of error. Statistical discrimination emerges from a lack of information on the part of teachers; the integration of further information should also mitigate the statistical discrimination. Hence, I conclude that teachers who spend more time with students are more capable of collecting superior information about students' individuality and peculiar characteristics and, hence, of substituting inferred information with actual knowledge. The radical alternative is stereotype discrimination. Although many scholars have widely criticised this view (see Arrow 1971), vast literature on the topic exists in sociology. Biases in expectations might also emerge from deliberate discrimination. According to Goffman (1963), stigma is the relationship between an attribute and a stereotype, which leads to a significant status loss in the stigmatised subject (Link and Phelan 2001). While in the case of statistical discrimination, some salient social attributes (e.g. citizenship and social class) are employed as drivers of an 'efficient' decision under uncertainty, the same labels are likely to be linked with stereotypes here, which provokes a degradation of expectations. If the academic or class environment devalues a migrant identity or background from a disadvantaged socioeconomic milieu, teachers might evaluate children with those social identities more harshly, even when they are fully aware of their actual ability. See Chapter 1 for a critique.

of reaching the required promotional standard. Meanwhile, advantaged students comply with high expectations and consequently tend to overvalue their actual academic performance as well as their future academic prospects.

The influence of teachers on students' self-perceived ability should cast severe doubt on the efficacy of GR as a motivational scheme. From the teacher perspective, GR operates as an incentive to elicit effort from low-performing students. The threat of GR (i.e. the expected psychological, social and economic costs of repeating an entire academic year) should encourage students below the performance threshold to work more diligently.

However, the GR incentive provokes effort among low-performing students only when the cost of repeating a year exceeds that of investing in academic performance. If disadvantaged students perceive that their likelihood of catching up is minute, then they might lose motivation and accept GR rather than exerting more effort to improve their performance. The incentive of GR fails when disadvantaged students comply with low expectations and, in turn, experience a discouraging effect.

Evaluation biases have one last crucial implication. If one admits biases on the part of teachers, then the empirical analysis conducted in the present chapter is conservative (i.e. it identifies only the lower bound of the GR social gap). In the literature (Autin et al. 2018; Barg 2013; Pit-ten Cate et al. 2016), there is a clear tendency for teachers to under-mark students who come from disadvantaged backgrounds. For two equally performing students, the one who derives from a disadvantaged group is more likely to receive lower grades. Accordingly, concerning the occurrence of GR, disadvantaged retained students are likely to possess a higher 'true' ability compared to advantaged retained students. Ultimately, such finding implies that the estimated gap is likely to constitute only a lower bound of the real gap.

According to Roemer (1998; 2000), there are two pivotal dimensions of inequalities of opportunities within contemporary Western societies. The first dimension concerns the levelling of the starting points of people in the competition for jobs and social positions. Those who have endorsed such claim have argued that society should compensate for circumstances of inequality that are beyond one’s control. Therefore, social competition should be based only on the free will and effort of each person. The

second conception frames equality as 'non-discrimination' or the merit principle. The merit principle prescribes that only relevant attributes for the performance of a social position or job should legitimately enter in the social competition; since gender, ethnicity and social class are irrelevant to the performance of a job, they should not count as eligibility criteria.

On the one hand, the analysis of the factors that influence GR underlines the need for a levelling of the playing field. The empirical evidence presented in this chapter illuminates the ineffectiveness of high schools in supporting the academic performance of weak students. The gap in GR expands as one moves along the left tail of the ability distribution. Hence, much of the efforts to help low-performing students recover must derive from the students' families. Moreover, inequalities exacerbate in wealthy schools (*i.e.* most of the traditional lyceums). The school system should implement compensatory policies that aim to fill the achievement gap. Advantages accumulate rapidly, and the first year following the transition into secondary education should represent a delicate turning point to policymakers who are interested in equalising educational opportunities.[12]

On the other hand, the analysis also advocates for an improved understanding of potentially discriminatory teaching practices in terms of whether they follow the 'merit principle' in decision-making about students. Teachers who have biased expectations tend to underestimate or overestimate students' academic prospects. Teachers' expectations are determinant in GR decisions and also affect students' self-perceived performance. Socially driven biases in academic expectations represent a de facto violation of the merit principle.

As a final remark, it is worth noting that the substantial level of inequality is likely to intensify if GR imposes negative consequences on subsequent educational outcomes. Any potentially harmful academic effect of GR will disproportionally target disadvantaged students. Consequently, educational inequalities will substantially increase.

[12] Cf. Chapter 3 on the Italian school system. High school completion is of fundamental relevance to the academic and professional future of students. The diploma is a crucial objective for almost any outcome in the labour market. A traditional lyceum diploma actively fosters students' chances of reaching a subsequent tertiary title.

Chapter 6: Is there life after GR? The trajectories of the repeaters

6.1 Introduction

Chapter 5 has shown how the role of social and migration background dramatically contributes to the social gap in grade repetition probabilities. The empirical results have pointed out that parental resources explain much of the variability in grade repetition risks among equally (low) performing kids.

The present chapter aims at investigating what happens to repeater students after grade repetition. On the one hand, in the teachers' intentions, repeaters should beneficiate from the additional time, and eventually catch-up with the necessary competencies and knowledge to progress in their academic career.

On the other hand, students who experienced grade repetition might downward adjust their academic prospects. Perceived (subjective) chances of success are one pivotal driver in educational choices (cf. Chapter 1). Repeaters are likely to make less ambitious academic investments.

Past GR might likewise carry stigmatisation in the classroom; that is, both teachers and peers are inclined to hold lower academic expectations towards repeaters. Under this light, a past grade repetition can represent for a student a negative lock-up situation that leads towards a cumulation of disadvantages (DiPrete and Eirich 2006; Bernardi 2014).

Social structures might disproportionally reward successful individuals while undermining the others according to a pre-existent distribution of merit. As Merton (1968; 1988) pointed out discussing the Matthew effect, the primary advantage of an individual (or of a group of individuals) 'makes for successive increments of advantages such that the gaps between the haves and the have-nots […] widen until dampened by countervailing processes' (1988, p.606).[1]

[1] Merton takes the cue from Saint Matthew (25:29): "For to everyone who has will more be given, and he will have abundance; but from him who has not, even what he has will be taken away". Merton applies the Matthew effect to the field of scientific work. He claims that "the social structure of science provides the context for this inquiry into a complex psychosocial process that affects both the reward system

Therefore, early productivity and merits produce high levels of reputation, which in turn enhances a virtuous cycle of productivity, reward and notoriousness. Alike, early failures might considerably jeopardise future successes.[2]

Halo effects propagate disadvantages (Thorndike 1920). The psychological reactions produce halo effects to a (negative) trait, which interferes with the evaluation of other traits. For example, a student with brilliant notes in one subject stimulates more favourable judgements in other subjects. In the same way, a past-GR harms subsequent brilliant notes.

Mechanisms working 'beyond the shoulders' of actors are also likely to influence the cumulation of disadvantages. Unconscious cognitive involvements, such as attitudes, self-projections and stereotypes, are pervasive and exert a significant interference in deliberate judgements (Greenwald and Banaji 1995). For both the student and teachers, a past academic failure can constitute an (unconscious) interference in the judgement of the actual performance.

However, the accumulation of disadvantages varies across social groups. Families experience different material and cultural constraints, which limit the amount and quality of parental support. The expectation is that students who can rely on a favourable amount of resources have more chances to get back on tracks.

Additionally, grade repetition is likely to influence students' preferences to different degrees. As Lucas (2009) has claimed, signals on academic ability and chances of success influence to a much lesser extent on high-status students and parents.

Educational attainment, as a critical positional good, is particularly valuable to high-status families. They face a considerable risk of social demotion, which pushes them

---

and the communication system of science" (p.57). The reward system damages unknown scientists, for it awards known scientists much more than young or unknown scholars. Moreover, the discoveries of prominent scientists get much more attention to the scientific public. Merton argues that there is "a correlation between the redundancy function of multiple discoveries and the focalizing function of eminent men of science" (p.62). See also: Merton, Robert K. The sociology of science: Theoretical and empirical investigations. University of Chicago Press, 1973; and: Merton, Robert K. 1988. "The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property." Isis 79 (4): 606–23.

[2] Another way to think at cumulative disadvantages is as the sum of the dimensions of the disadvantages experienced by individuals who belong to multiple oppressed social groups. For example, in the classic Blau and Duncan's investigation, the authors observe that Afro-American workers experience the double disadvantage of being poor and belonging to an ethnic minority (Blau and Duncan 1976).

to invest in education despite the signals they get from teachers. Academic achievement for high-status children is also less reliant on past achievements (Bernardi 2014).

6.1.1 Aims of the chapter

Against this background, the present chapter follows a cohort of repeaters for the subsequent two years after grade repetition and models their academic trajectories. Moreover, the analysis intends to investigate whether there is a further increase in educational inequalities among repeaters.

After the description of data and methodology, the chapter provides detailed descriptive figures for both retained and promoted students. It follows the presentation of the results of the regressions and the discussion.

The present text speaks of repeaters when referring to students who have experienced a grade repetition, whereas it speaks of promoted in regards to students who make to the next grade.

6.2 Data and Methodology

As detailed in Chapter 4, the data of Cohort B refer to the population of northern Italian students who enrolled in the first year of high school in the academic year of 2013-2014. The integration of the administrative records with the socioeconomic and performance information coming from the Italian National Institute for Evaluation (henceforth, INVALSI) allows for a detailed reconstruction of the academic careers of these students.

Out of the 146208 individual observations, 22304 (14.69%) experience a GR at the end of 9th grade in 2014. Among the retained students, more than 7 thousand of students (the 32.5%) had left school by the end of the subsequent year. Of the 77.5% enrolled in 9th grade for the second time, only the 51.6% and the 33.2% of the retained in 2014 made to, respectively, the 10th and the 11th grade. By the end of the fourth year of observation, 64.1% had dropped school.

6.2.1 Variables

The outcome of interest is the trajectories of repeaters. The analysis deductively clusters the academic careers of students into five topic sequences.

As detailed in Table 6.1, for each student in Cohort B, four turning points are available for observation. Soon after grade repetition, students must decide their 2015 enrolment. There are three separate alternatives, namely, dropout, enrolment in the same kind of school, and downward enrolment. School changes can be either horizontal (when a student enrols in a different school without choosing a less demanding type of school) or downward (when a student enrols in a less demanding track).[3]

The other point of observation is at the end of the same year and register the educational outcome, which can be a promotion, a grade repetition, or a non-admission. Non-admission occurs when students do not reach the minimum amount of school hours and hence represents a de facto dropout. Both the enrolment decision and the academic outcome were observed again in 2016.

Tab. 6.1: summary of the observation points and outcomes

Point of observation (2015 and 2016)	Outcome
Enrolment decisions	• No enrolment (dropout) • Enrolment in the same school • Horizontal school change • Downward school change
Academic outcome	• Promotion • Grade repetition • Non-admission (dropout by the end of the year)

The four observation points combine in 94 single student's academic sequences, which fall into five main clusters or patterns of academic careers:

[3] The order from the most to the least demanding is the following: Traditional Lyceums, Other Lyceums, Technical high schools and Vocational high schools (cf. Chapter 3). Due to the minimal number of retained students choosing an upward change, I collapse the horizontal and the upward change in a unique category.

- *Linear*. After the grade repetition in 2014, a student gets promoted two times (in 2015 and 2016) without experiencing any downward change
- *Downward change.* After the grade repetition in 2014, a student gets promoted two times (in 2015 and 2016). However, she\he opts for a downward change at least once
- *Grade repetition.* A student gets retained (again) at least once in 2015 or 2016.
- *Dropout.* After the grade repetition in 2014, a student appears as dropout by the end of 2016.
- *Return.* After the grade repetition in 2014, a student appears as a dropout at least once; nonetheless she\he re-enrols by the end of 2016.

The five clusters compose the categories of the dependent variable in the present empirical analysis.

Table 6.2 displays the covariates that aim at explaining the variation in the academic careers of the repeaters. The model focuses on two covariates, namely, the migration background and the social background. As in Chapter 5, students fall into three categories: Italian citizens, second-generation immigrant students, and first-generation immigrant students.

The analysis takes into account social background by including the highest parental educational attainment. According to the dominance criterium, parents groups in: no high school degree (No diploma), High school degree (Diploma), and Tertiary or higher degree (Degree).

The model also includes gender, whether the student is over-age, and prior academic performance. On this matter, the analysis accounts for the end-of-2014 marks from teachers in the two main curricular subjects, namely, Italian and Mathematics, and for the mark on behavioural conduct. The three notes are on a ten points scale, where the

6/10 is the sufficiency. The model also controls for the INVALSI's standardised scores in Italian and Mathematics.[4] [5]

6.2.2 Descriptive statistics

Table 6.2 displays the distribution of the covariates between promoted and repeaters in 2014-2015 in Cohort B.

Tab. 6.2: Descriptive statistics: retained and promoted students in Cohort B.

		Retained N= 22304 (14.69%)	Promoted N=129514 (85.31%)
	Female	39.1%	51.2%
	Older students	20.9%	6.6%
	2nd gen. immigrant	8.3%	3.7%
	1st gen. immigrant	13.4%	5.5%
Parental education	No diploma	40.0%	27.7%
	Diploma	31.3%	40.5%
	Higher Education	9.0%	19.0%
	(missing)	19.6%	12.8%
Marks (8th grade)	Italian	-0.69	0.30
	Maths	-0.76	0.33
	Behaviour	-0.59	0.23
INVALSI (8th grade)	Italian	-0.56	0.20
	Maths	-0.62	0.21
Percentages across types of high school	Traditional Lyceums	8.31%	
	Other Lyceums	12%	
	Technical	19.08%	
	Vocational	18.89%	

Unsurprisingly, repeaters underperform promoted students and display below-the-mean performance marks and test scores. Among repeaters, there is a predominance of males and over-age students.

[4] There are substantial gains in using the two sets of performance measurements jointly (cf. Chapter 4). The content's overlap between teachers' notes and standardised tests is limited. They are indexes of the various and intertwined dimensions of the academic performance. See Chapter 4 and 5 on this subject.

[5] I reparametrize the individual values on the population means; that is, when a student takes zero, she\he is at our data mean on a specific variable. The aim is to give an unambiguous interpretation of our results.

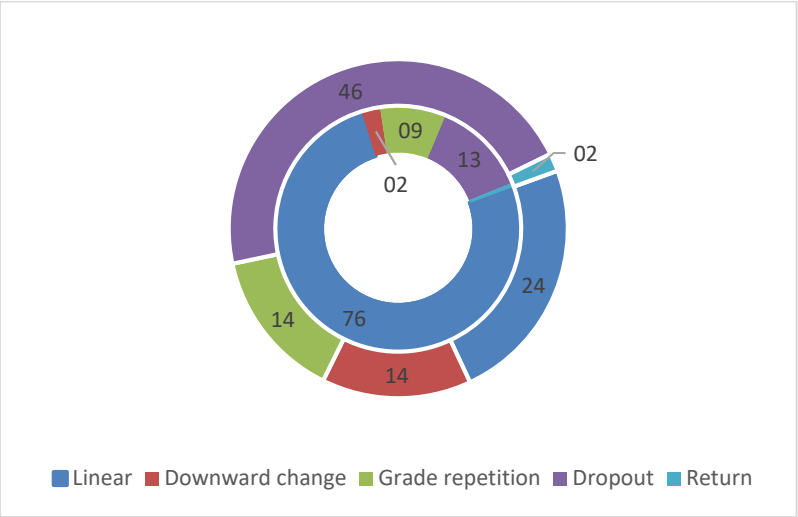
Students with a migration background are over-represented (8.3% and 13.4% of second- and first-generation immigrant students, respectively). The same unbalance exists when one looks at the distribution of parental educational attainment.

Repeaters' parents without a diploma are 40%, against 28% among promoted students. Parents of repeaters who hold a high-school diploma are only 31.3% (versus 40.5% among promoted). Only 9% of repeaters can count on a tertiary-educated parent, whereas this is the case for almost 20% of the promoted students.

Repeaters disproportionately come from less demanding types of high schools. In traditional lyceums, the share is 8.31 against 19% in the non-academic (technical and vocational) high schools.

The pie chart in Figure 6.1 displays the distribution of clusters between the repeaters (internal ring) and, as a reference, the promoted (external ring).

Fig. 6.1: Distribution of clusters between repeater (external ring) and promoted (internal ring)



The share of linear careers among repeaters is rather small (23.6%). The modal category is dropout with 46%. The 14% of repeaters opts for a downward change, and another 14 faces a new grade repetition.

The additional focus of the present chapter is the comparison among repeaters. Figure 6.2 and 6.3 display the distribution of academic careers across parental educational attainment and migration background, respectively.

Fig. 6.2: The distribution of academic careers across parental educational attainment

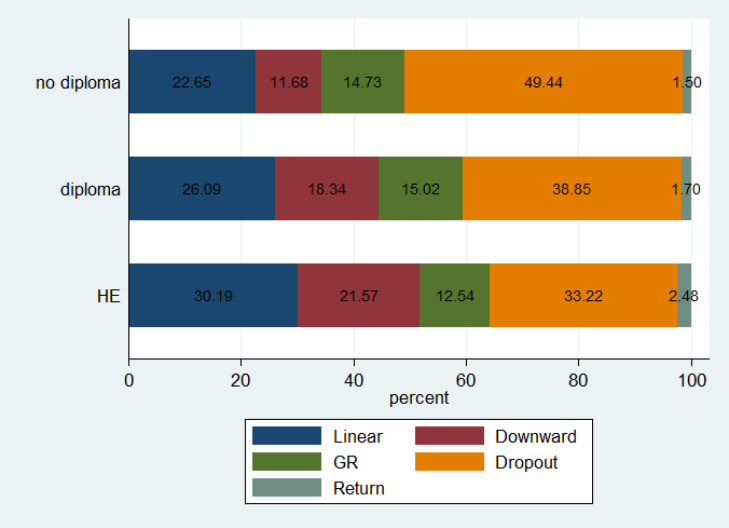
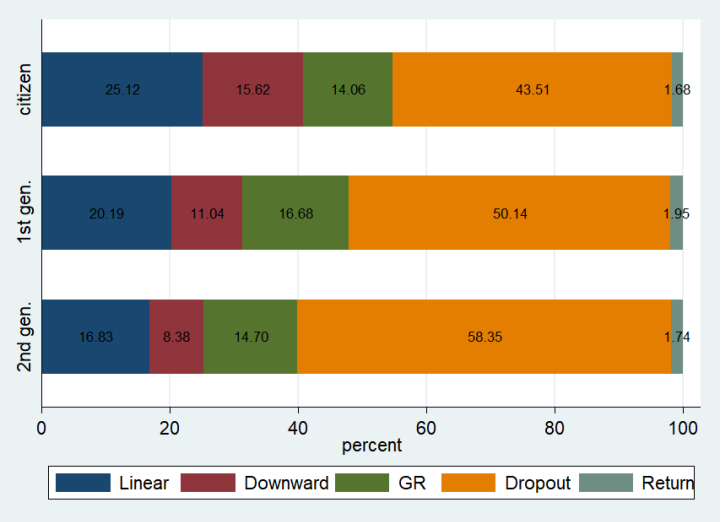


Fig. 6.3: The distribution of academic careers across migration background



A gap in the frequency of linear careers strongly emerges. Approximately one-third of the repeaters that can count on a tertiary-educated parent experiences a subsequent linear career and another third of them drop out of school.

The picture overturns among culturally disadvantaged kids. Less than one-quarter will get through the subsequent two years without a further adverse event. Almost half will eventually drop out of school.

Culturally advantaged kids also frequently change track (21.57%) and come back to school more often (2.48%). One hypothesis is that the holes in their careers are temporary enrolments in (often less demanding) two-years-in-one private schools.

Figure 6.3 plots the distribution of academic sequences for migration background. Non-migrant students are more linear in their academic careers (25% *vs* 17%). Conversely, migrant students drop out more often (58% *vs* 44%), while they are less likely to make a downward change (15% *vs* 21%).

Overall, the evidence suggests that repeaters' achievements are not bright. Moreover, there is a clear-cut divergence between social groups in after-GR careers.

The analysis now aims at modelling the contribution of the covariates on the probability to fall in one of the five clusters of academic careers.

6.2.3 The model

Gender, age, parental educational attainment and migration background are here regressed on the probability to fall in one of the five academic careers (as stated above: linear, change down, new GR, dropout or return) by a multinomial regression.

The specification includes two sets of controls. The first set aims to control for prior performance differentials. It includes both teachers 8^{th}-grade marks in Italian, Mathematics and Behaviour, and INVALSI scores in Italian and Mathematics (also at the end of 8^{th} grade).

The other set of controls takes care of the contextual effect at the time of grade repetition (9^{th} grade). The contextual variables are the within-school average values of

every covariate in the model. The intention is to control for unobserved heterogeneity in grade repetition's decisions due to differentials in promotional standards at the school level.[6]

The equation (in its linear form) looks like the following:

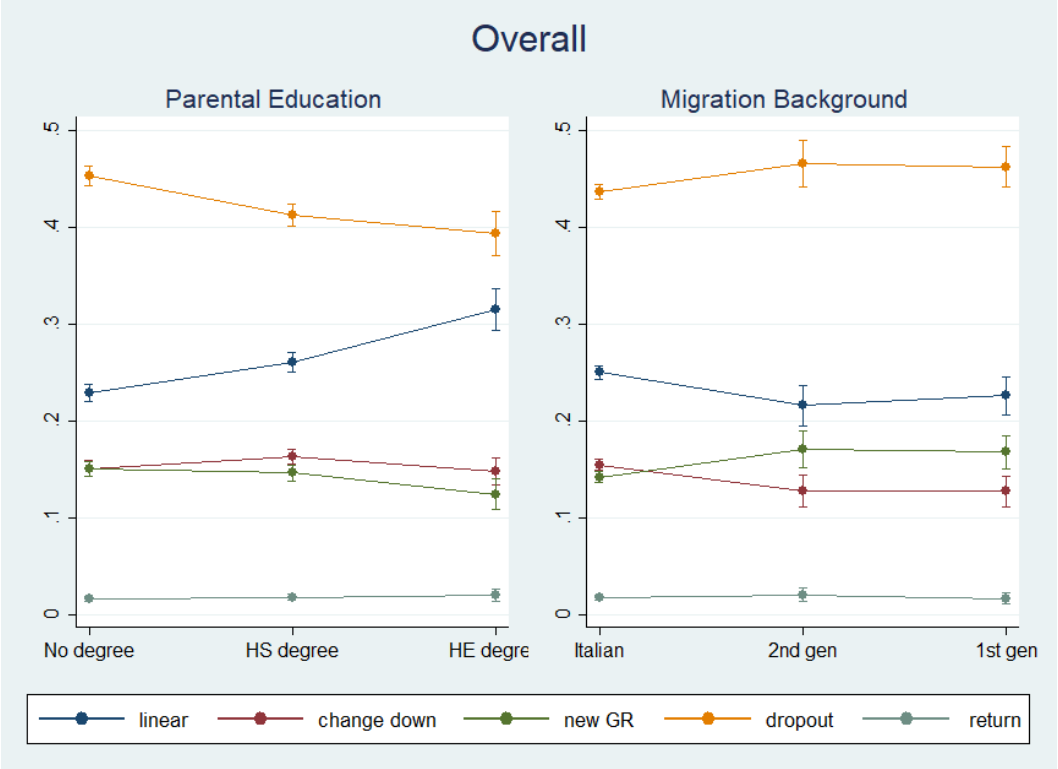
$$P(SEQ = 1|GR_{2014} = 1) = \beta_0 + \beta_1 paredu_i + \beta_2 migr_i + \beta_3 x_i + \beta_4 \bar{x_j} + \varepsilon_i \quad (1)$$

where $SEQ = 1$ is one of the five academic careers and $paredu$ and cit are, respectively, the parental educational attainment and whether the first citizenship is Italian or otherwise. x_i is the vector of the other covariates and the set of individual-level performance controls; $\bar{x_j}$ is the vector for the contextual variables at the *j*-th school.

The results reported below are estimated first on the whole population of retained students and, then, separately for each type of school.

Fig. 6.4: Predicted probabilities for repeaters

[6] Our results are consistent with the ones from two alternative specifications, namely, a set of logistic regressions with fixed effect at the school level; a set of mixed-effect logistic regression with random intercepts at the school level plus contextual effects (cf. the additional material in the Appendix of the present chapter)



6.3 Results

Figure 6.4 (above) plots the model's predicted probabilities for the whole population by parental educational attainment (vertical axis of the left-hand graph) and migration background (right-hand graph).

Overall, the most frequent outcome among repeaters is a dropout. Regardless of the parental cultural resources, the dropout predicted probabilities are always above 40%. Although to a much more limited extent, also the repeaters' risk to downward change and get grade retained is not negligible.

Within this critical landscape, the gap in linear careers between groups is statistically significant and substantial. Students with tertiary degree parents have +5.4 and +8.6 percentage points of experiencing a linear career after GR as compared to students with a diploma and no-diploma parental background, respectively. Culturally advantaged students are considerably less likely to leave school (around +2 and +6 percentage points in respect of the other two groups) and, to a lesser extent, to incur in another grade repetition (around +2 points).

When looking at the migration background, the gap in linear careers is smaller. Italian citizens have +3.4 and +2.4 percentage points as compared to second- and first-

generation immigrants. Immigrant repeaters have higher downward change (around +3 points), (new) grade repetition (+3 points), and dropout probabilities (+3 points) as compared to citizen students. The analysis shows no statistical evidence of differentials between the two groups of immigrants.

As mentioned above, the model separately goes for high school types. The right-hand graph in Figure 6.5 displays the predicted academic careers for traditional lyceum repeaters. Irrespectively to parental educational attainment, the most probable outcome is a downward change; that is, students in traditional lyceums are likely to move to 'less demanding' track when retained. Differences between groups are not statistically meaningful. Nonetheless, the advantage for the culturally endowed pupils emerges when looking at the dropout share. Students with tertiary-educated parents have a 19% probability of dropping out of school after grade repetition, while is 22% and 26% among children with secondary educated and non-educated parents. The return career, while still a marginal phenomenon, is slightly more often chosen by highly educated families.

The right-hand graph presents the predicted careers over migration background. Provided that the overall academic future of repeaters is overall rather mediocre, there is a definite advantage for Italian citizens. Italian citizen students downward change more (+14 over second- and +6 over first-generation) and drop school in fewer cases (+10 over second gen. and +3 over first-generation) as compared to (both group of) immigrant students. The gap is particularly evident when comparing Italians and second-generation immigrants.

Repeaters in other lyceums reflect the overall pattern (Figure 6.6). Dropout is again the prominent outcome for post-GR academic careers. Highly educated parents are much more likely than the others to have children who experience a linear career and do not incur in grade repetition once more. Second-generation immigrants are more likely to leave school: +10 percentage points as compared to Italian citizen students. On the contrary, both groups of immigrant students are the least likely to change towards a less demanding school, with -7 and -8 percentage points respectively as compared to Italian citizenship.

In technical high schools, repeaters are, once again, very likely to drop out (Figure 6.7). Nonetheless, the probability of dropping out is limited among highly educated

families, which instead are consistently more likely to no further encounter obstacles in their academic careers. The gap between migration backgrounds is small — almost the same works for vocational high schools in Figure 6.8. Worthy of notice, the model assumes that students enrolled in Vocational do not have the downward option; that is, the analysis collapse dropout with any potential enrolment in the Vocational and Training system. Differences between groups are limited, although there is the usual gap between parental educational levels and migration background in dropouts.

Fig. 6.5: Predicted probabilities for repeaters in traditional lyceums

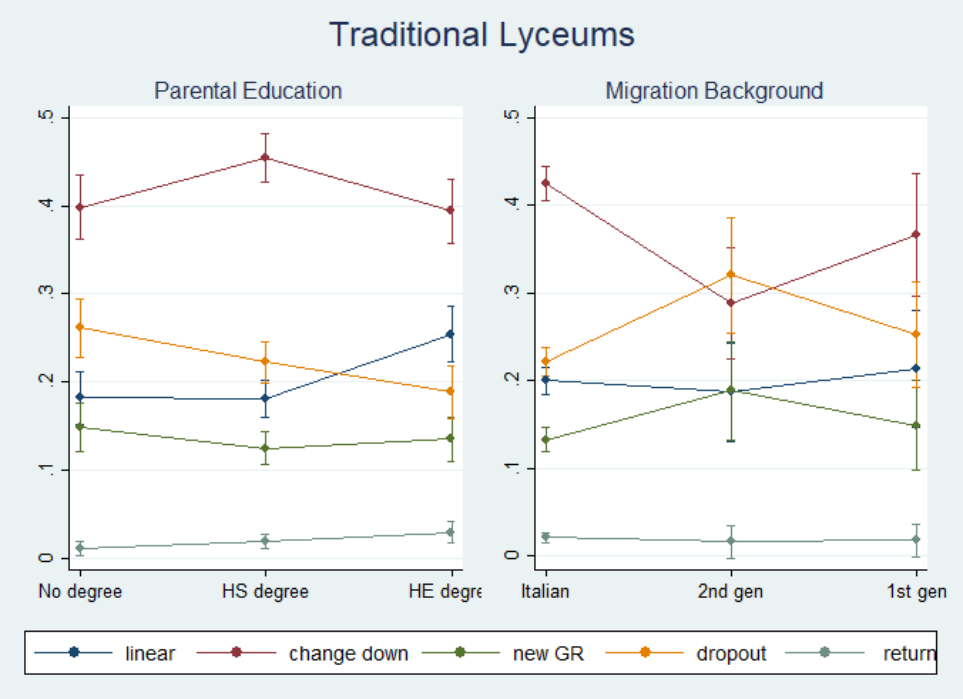


Fig. 6.6: Predicted probabilities for repeaters in other lyceums

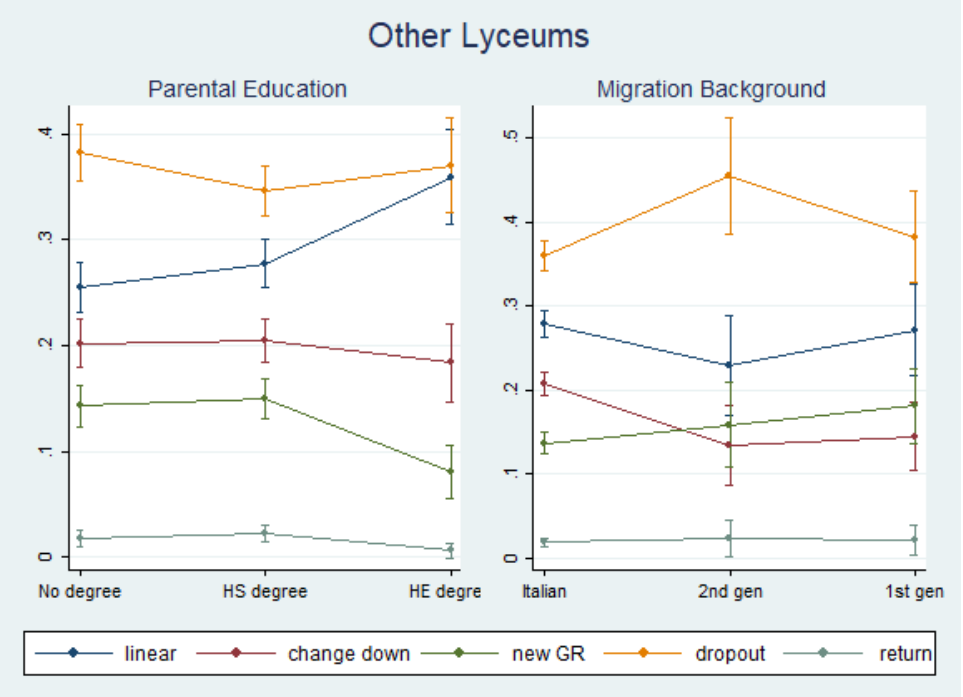


Fig. 6.7: Predicted probabilities for repeaters in technical high school

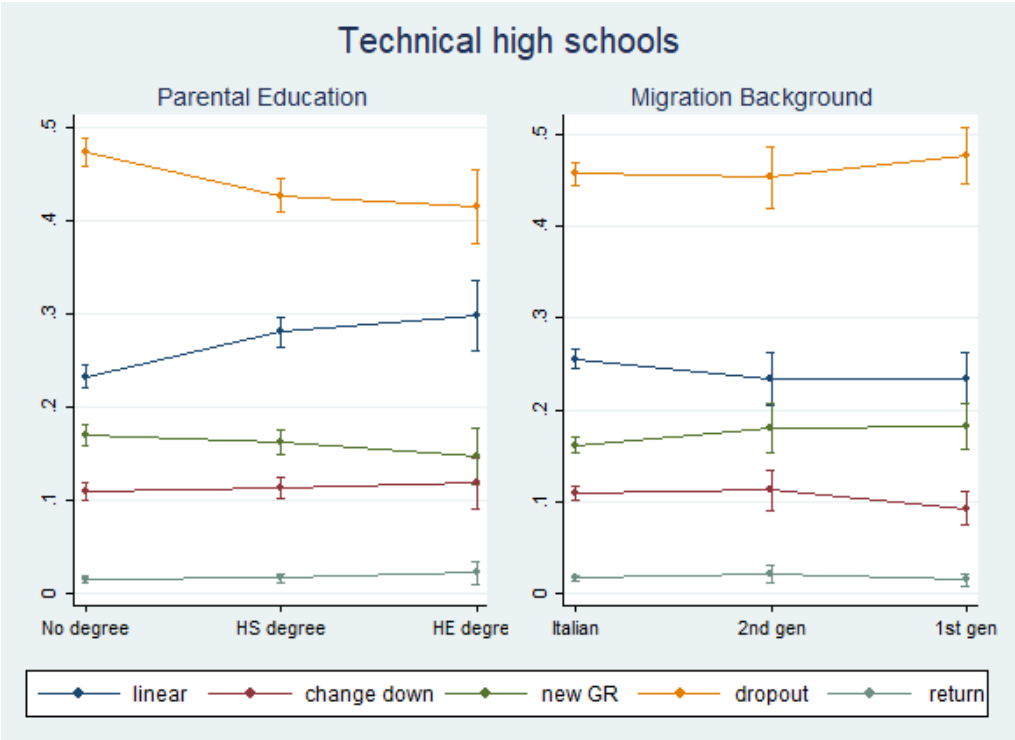
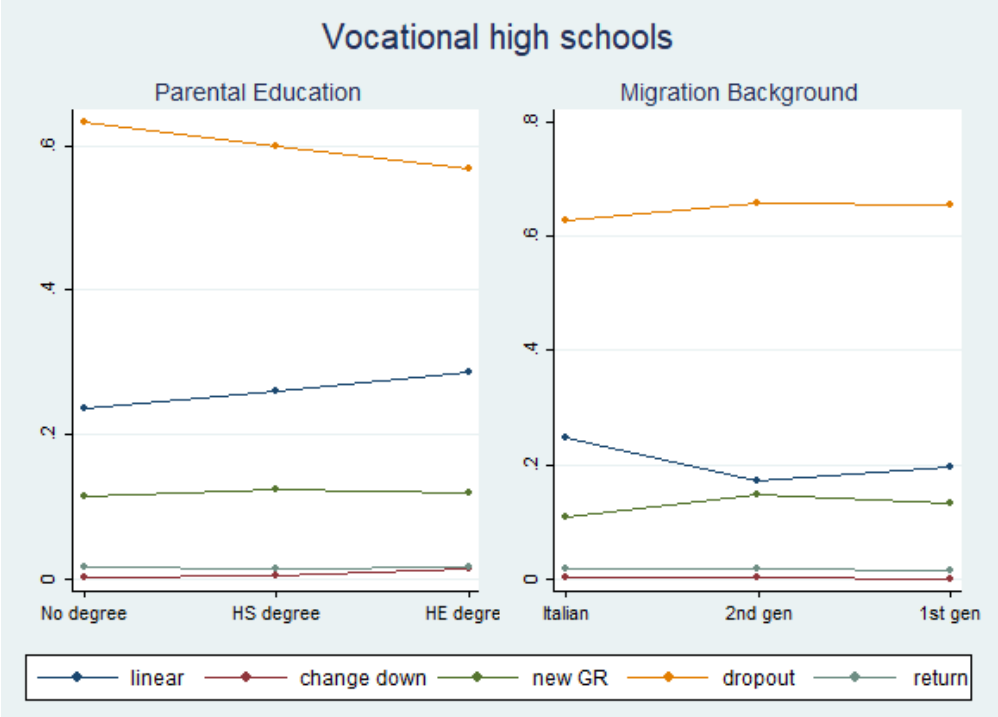


Fig. 6.8: predicted probabilities for repeaters in vocational high school



6.4 Conclusions

This chapter has investigated the subsequent academic careers of repeaters. The results have presented vast evidence of the prevalence of the non-linearity of academic careers among retained students. Regardless of their characteristics, the chances of a linear career are somewhat limited for all repeaters, who instead face dramatic probabilities of dropout.

Many, as the main alternative to dropout, change the type of school toward a (supposedly) less demanding type.

The analysis also has shown that after-GR careers significantly differ according to social and (to a smaller extent) migration background. Students from highly educated families are more likely to experience linear careers and to change school, while are less likely to experience a new grade repetition or to drop out of high school. Although to a smaller extent, the same works for the comparison of migration background.

The gap is significant across all school types. Traditional lyceum repeaters are the least likely to drop out and the most likely to change downward.

To sum up, even though the observed gap in terms of percentage points is not enormous, social groups significantly differ in terms of educational chances after grade repetition.

Chapter 7: The Policy of Grade Repetition

7.1 Introduction

Chapter 5 has underlined a dramatic social gap in GR selection. The wealthier the kids, the smaller their probabilities of repeating the year at school, poor prior academic performance held constant. Then, Chapter 6 has examined after-GR academic trajectories. The analysis showed that no significant catch-up occurred among repeaters. Overall, repeaters usually experience non-linear careers and are particularly exposed to dropouts. The lower the cultural resources of the student, the more troubled the subsequent academic careers were.

The present chapter aims at estimating the impact of grade repetition (also GR henceforth). As such, the next pages ideally conclude the present empirical investigation of this contested educational policy.

The debate over grade repetition's educational impact has been living for decades now (cf. Chapter 2). Alexander et al. (2003) have argued that grade repetition's research is inhabited by 'strong beliefs' supported by 'weak evidence'. While, for almost three decades, many scholars had judged grade repetition as detrimental, a more dubitative (if not positive) attitude towards grade repetition has recently gained legitimacy. Besides, the (perceived) decrease in academic standards has renewed the emphasis on schooling selection mechanisms.

Although the issue of grade repetition has periodically reached the spotlight in the Italian public debate, this educational practice remains dramatically understudied in this context. Except for a few institutional descriptive data (cf. Chapter 3), in the author's knowledge, there is no empirical study that has yet tried to infer grade repetition's causal consequences.

The analysis in this chapter makes use of matching techniques to estimate the Average Treatment effect for the Treated.

The goal is to match comparable retained and promoted students. The general idea behind matching is to mimic an experiment by the post-treatment identification of a

suitable control group to which pair the treated subjects, namely, the repeaters. Within this framework, the selection of the most appropriate control group is crucial.

The matching procedure exploited in this work relies on a rich array of information covering socioeconomic and migration background, marks and standardised scores on past academic performance, and contextual (school-level) variables.

As noted in Chapter 2, scholars generally implemented two matching logics in this field of research, namely, the same age comparison and the same grade comparison (Alexander et al. 2003; Lorence 2006).

In the same-age design, the estimation of the effect of a grade repetition emerges from the comparison with repeaters' same-age peers. As such, the measurement of the outcomes of interest occurs while the two groups of students attend different grades.

Instead, the same-grade design matches repeaters with their after-GR same graders. In this case, repeaters students are usually older than the control group, but both attend the same grade.

Of course, the choice between these two logics affects the results (Holmes 1989; Shepard 2004; Alexander et al. 2003; Lorence 2006; Allen et al. 2009).

Results from the same-age comparison tend to show that GR has adverse effects on subsequent academic careers. Conversely, same-grade comparison generally highlights positive effects of GR. In both cases, gaps between groups appear to narrow over time.

The same-grade logic leaves space for an additional academic year before measuring grade repetition's effect. The argument is that the one-year loss is an intended (and desirable) implication of grade repetition. The contrary is true in the same-age logic, which measures repeaters' performance straight away the grade repetition.

On this subject, many scholars (Wilson 1990; Karweit 1999; Shepard 2004; Lorence 2006) have maintained that the same-grade comparison measurement is closer to the purpose of GR, which is to provide the children with the necessary additional time to re-align their competences and scholastic performance. However, the same-grade matching incorporates as grade repetition's effect a series of factors (e.g. repeaters may

overperform because they have taken the same curriculum twice, or they are more mature than their classmates), both may not relate to any grade repetition's effect. Therefore, the decision should rely upon the context.

The present preliminary analysis favours the same-age design and, as such, estimates GR effect by matching students who got retained in 2014-15 in 9th grade within their cohort. This choice allows for the estimation of the grade-repetition effect also on the probability of school enrolment the year subsequent of grade repetition, which could not be possible when considering same-grade design. Same-grade logic excludes from the ATT's computation those repeaters who do not re-enrol the subsequent year.

7.2 Data and Methodology

The present analysis estimates the ATT on the students who attended the 9th grade in the academic year 2014-15 and got retained. The same-age design looks for matches within the same cohort of students.

Table 7.1 exemplifies the matching logic. The bottom rows are, respectively, the outcomes (t) and the academic years (A.Y.) of the observed outcomes. When time is equal to *t* (and the A.Y. to 2014-15), the students receive the 'treatment' of grade repetition. Please note that each couple *t+1* and *t+2*, and *t+3* and *t+4* correspond to a single academic year: 2015-16 and 2016-17, respectively. Each academic year consists of two outcomes: the first relates to the decision over enrolment, whereas the second the educational outcome.

Tab. 7.1 Summary of the matching logic. Students are matched over prior performance (in green) and high school context (in yellow). In the same-age comparison, there are four available outcomes.

Same-age	8	9	enrolment	outcome	Enrolment	outcome
Treated	8	9				
Time	t-1	t	t+1	t+2	t+3	t+4
A.Y. 20…	13-14	14-15	15-16		16-17	

As in Chapter 6, decisions over enrolment are:

(a) To re-enrol in the same school
(b) To enrol in a different institution, although without changing the type of high school
(c) To enrol in a less demanding type of school (downward change), assuming the following descending ranking: traditional lyceums, other lyceums, technical high schools, and vocational high schools
(d) Eventually, to not enrol, that is, to drop out of high school

The educational outcome can be:

(a) Promotion
(b) Grade repetition
(c) Dropout by the end of the year.

Let us look at the upper part of Table 7.1 (the two lines above the timeline and A.Y. rows), which explicit the same-age comparison. At time *t-1* and *t*, both treated and control group are enrolled in the same grade. The design matches the students upon their academic performance in 8^{th} grade, and over high school track and school characteristics in 9^{th} grade.

Then, matched students are compared over four different outcomes: enrolment in *t+1*, the educational outcome in *t+2*, enrolment in *t+3*, and educational outcome in *t+4*.

Given the observed outcomes, the outcome over which ATT is computed is the type of sequence observed in the students' careers (already identified in Chapter 7):

(a) *Linear career*. Students experience no downward change of school, neither grade repetition, instead they advance linearly throughout the observe time
(b) *Downward change.* Students change high school toward a less demanding type, *i.e.* from traditional to other lyceums, from other lyceums to technical high school, and from technical to vocational high school
(c) *Grade repetition.* Students experience a grade repetition in *t>1*
(d) *Dropout.* Students disappear from the database
(e) *Return.* Students do not enrol for one year, while come back by the end of the second academic year.

7.2.1 Identification strategies

Let us define two potential outcomes Y_0 and Y_1. The Average Treatment effect on Treated (ATT) is defined:

$$ATT(x) = E(Y_1 - Y_0|D = 1) \quad (1)$$

In observational contexts, scholars do not have control over the selection into treatment, which is likely not to be random. Hence, ATT holds if the observable variables capture any relevant difference between treated and non-treated, *i.e.* retained and promoted pupils (Conditional Independence Assumption- CIA):

$$Y_1 \perp D| \quad (2)$$

In words, it means that in order to estimate the average effect of a treatment, the analysis needs to control for the variables that affect both the treatment assignment and the treatment outcome. For example, I expect that prior low performance affects both the probability to get retained and the probability to experience a dropout. Before being able to estimate the effect of the grade repetition, it is necessary to control for the prior performance's influence on dropping out.

Moreover, for each value of X it must be possible to find both treated and non-treated units (Common Support- CS):

$$P(D = 1|X) < 1 \quad (3)$$

Hence, the observables over which the estimated conditioned probability must not wholly define the probability to get the treatment itself.

Therefore, the propensity score should, on the one hand, includes all the relevant variable that influences both D and Y. At the same time, on the other hand, the propensity score should leave some unexplained portion of the variance that can be possibly assumed as independent with the outcomes and can hence be used to determine the matching with control units.

So, the propensity score is a balancing score only when:

$$p(x) = P(D = 1|X = x) = E(D|X = x) \quad (4)$$

$$X \perp D|p(x) \quad (5)$$

such that:

$$ATT[p(x) = \eta] = E[Y_1 - Y_0|D = 1, p(x) = \eta] \quad (6)$$

Under this light, the ATT is the average value of the propensity score's distribution in the identified subpopulation $[p(x) = \eta]$.

The present analysis implements two distinct same-age designs (Strategy A and B) to estimate the impact of grade repetition on subsequent academic careers.

Strategy A performs a within-school one-to-one matching with replacement. Within each school, Strategy A estimates the probability of incurring in grade repetition using all the students in Cohort B:

$$P(GR = 1|\mu_j) = \beta_0 + \beta x_{ij} + \varepsilon_{ij}$$

where x is the vector of individual characteristics, and μ_j is the school of the i-subject. As mentioned above, the model includes controls for both social and biographical controls and academic performance. The analysis accounts for past academic performance (in 8th grade), *i.e.* marks in Italian, Mathematics and Behaviour, and standardised scores in Italian and Mathematics. Jointly with this rich array of ability measures, gender, age (as an over-age student), parental educational (no high school diploma, high school diploma, tertiary degree or higher), and migration background (Italian citizen, first-generation immigrant, second-generation immigrant) enter in the within-school propensity score estimation model.

Following the estimated coefficient, the model predicts the individual probability to get the treatment (the grade repetition) for all the students in Cohort B.

The computed probability (ranging from 0 to 1) operates as a score to find a match for repeaters. The goal is to find, for as many repeaters as possible, fitting control students who share the most similar probability to get retained. The analysis makes use of

propensity score matching to reach an optimal match and estimate the ATT between promoted and retained.

As noted, Strategy A relies on a one-to-one propensity score matching with replacement within strata. In practice, Strategy A assigns a propensity score to students in each stratum, *i.e.* schools. Then, the most similar control within a given propensity score distance (a caliper of .20 propensity score's standard deviations) is selected. The same control might pair with more than one treated student.

Then, the analysis looks also at potential heterogeneities along with high school type, parental educational attainment and migration background.

Strategy A has the significant advantage of ruling out any school-level unobserved heterogeneity linked to the process of selection into grade repetition. Nonetheless, the very same heterogeneity might, on the other hand, undermine the identification of the causal effect when it is related to the outcomes under study. For example, an individual school policy that favours grade repetition as a useful remedial tool increases GR chances, while exerting no effect GR's subsequent outcomes. On the contrary, a strict administrator might, while boosting GR probabilities, also affect dropout probabilities.

An alternative is to spot perfect matches over crucial characteristics. In this sense, Strategy B looks for perfect correspondence on gender, year of birth of the student, migration background and parental educational attainment. Within each group of perfectly matched students, the procedure estimates the propensity score over the past performance measures of Italian, Mathematics and behaviour 8th-grade marks, and Italian and Mathematics INVALSI's standardised scores.

Strategy B is an attractive option with the large-N in the present data. Strategy B has the main advantage of avoiding school-level unobserved heterogeneity. Moreover, it looks for identical individuals over crucial characteristics and hence avoids 'odd' or even unrealistic matches due to propensity score's one-dimensionality.

Tab. 7.2: Matching results

	Strategy A	Strategy B	Total
Repeaters (treated)	15,289 68.5%	18,142 81.3%	22,304
Promoted (controls)	9,099 7.0%	11,784 9.1%	129,514

As mentioned above, a good propensity score should function as a balancing score. Hence, the distribution of the observable characteristics between matched treated and control should not be very dissimilar. The literature (Ho et al. 2007) suggests that the standardised bias, defined as the average difference between the two groups over the standard deviation, should not exceed 10% on each observable (although others are less strict and allows event for .20). The performed robustness checks (provided in the related Appendix) shows that the bias for both the strategies is comfortably under 10%.

Moreover, the same Appendix provides also two graphs (corresponding to each of the strategies) in which I plot the propensity score distribution between treated and control units before and after the matching, showing that the fit after matching can cut-off the difference among treated and control units on support.

7.3 Results

Table 7.3 and 7.4 present the estimated average treatment effects for the retained students as estimated by Strategy A and B, respectively. The overall ATTs are in the first row of each table; then, the tables display the results by school type, highest parental education, and migration status. The tables adjust the values by the number of students who contributes to the ATT's estimation in each stratum (schools in Strategy A; perfectly matched groups in Strategy B). Worthy of notice, results are consistent between Strategy A and B, this denoting a very promising clue that boosts the credibility of these preliminary results.

Tab. 7.3: Average Treatment Effects. Within school propensity score estimation and matching

	Linear	Downward	Subsequent GR	Dropout	Return
Overall	**-0.27**	**0.13**	**-0.06**	**0.19**	**0.01**
Traditional Lyceum	-0.33	0.31	-0.08	0.08	0.02
Other Lyceum	-0.31	0.20	-0.05	0.15	0.01
Technical High school	-0.23	0.10	-0.08	0.20	0.01
Vocational High school	-0.25	0.00	-0.02	0.27	-0.01
No high school diploma	-0.26	0.11	-0.05	0.20	0.00
High school diploma	-0.27	0.15	-0.06	0.17	0.01
Tertiary degree	-0.28	0.20	-0.08	0.14	0.01
Italian citizen	-0.27	0.14	-0.06	0.18	0.01
1st gen. immigrant	-0.25	0.11	-0.06	0.19	0.01
2nd gen. immigrant	-0.25	0.10	-0.05	0.20	0.00

Tab. 7.4: Average Treatment Effects. Perfect matching over gender, overage, migration background and parental education combined with a p-score estimation on academic performance measurements

	Linear	Downward	Subsequent GR	Dropout	Return
Overall	**-0.29**	**0.13**	**-0.04**	**0.20**	**0.01**
Traditional Lyceum	-0.37	0.34	-0.05	0.07	0.01
Other Lyceum	-0.32	0.17	-0.02	0.17	0.01
Technical High school	-0.27	0.10	-0.06	0.22	0.01
Vocational High school	-0.25	0.00	-0.02	0.27	-0.01
No high school diploma	-0.28	0.10	-0.05	0.22	0.00
High school diploma	-0.31	0.16	-0.04	0.18	0.01
Tertiary degree	-0.30	0.20	-0.05	0.15	0.01
Italian citizen	-0.30	0.14	-0.04	0.20	0.01
1st gen. immigrant	-0.23	0.09	-0.07	0.20	0.01
2nd gen. immigrant	-0.25	0.07	-0.01	0.19	0.00

The overall impact of grade repetition on the subsequent academic careers of the repeaters does not seem to be constructive. Among the treated, grade repetition exerts a negative impact on the probability to experience a linear career (-27 and -29 percentage points for Strategy A and B, respectively). Although it reduces subsequent GR of -6/-4 points, grade repetition negatively impacts on two outcomes, namely, downward change and dropout. In facts, repeaters show +13 percentage points on the probability to change to a less demanding school as compared to their matched peers. Alike, treated students are severely more likely (+19/+20 percentage points) to drop out of school as against their promoted classmates.

The GR effect appears to be even more detrimental among disadvantaged students, as shown in Tables 7.3 and 7.4. The effect is the largest among those who had been retained in traditional lyceums (-35.9 percentage points), while it is around -25 percentage points among vocational students.

As it has emerged, grade repetition pushes students to downward change. Nonetheless, the effect is the largest among those who had been retained in traditional lyceums (+31/+34 percentage points) in contrast to the other types of schools in which the downward option is feasible (+20/+17 in other lyceums, and +10 in technical high schools).

GR effect on dropout probabilities also dramatically varies across types of school: it is limited in traditional lyceums (+8/+7), while intensifies in other lyceums (+15/+17), being quite dramatic in non-academic high schools (+20/+22 in technical and +27 in vocational high schools).

Variations across parental education are even similarly striking. Culturally advantaged kids often chose the downward option (+20 percentage points), while the effect is more limited in families with a secondary diploma (+15/+16), and with no-diploma (+11/+10). Instead, the opposite trend is actual regarding the dropout option. GR raises of ++20/+22 percentage points the chances of dropping out from school for repeaters with non-educated parents, while the same values are +17/+18 for the kids with secondary graduated parents, and of 'just' +14/+15 when parents hold a degree.

Unequal heterogeneities of GR effect on the subsequent academic careers are less marked between Italian and immigrant students. Among Italian students, GR has a

tangible impact on raising downward chances (+14 percentage points). The same effect is more limited among first-generation (+11/+9) and second-generation (+10/+7) immigrants. The gap is not relevant when it comes to the GR impact over dropout probabilities, where the three groups show very similar behaviours (around +20 percentage points).

Hence, regardless of the chosen logic for the matching, results point at an adverse GR effect among the treated on the probability to conduct linear academic careers after a GR. In particular, GR induces downward decisions, and especially for traditional lyceums and culturally advantaged kids. Moreover, GR affects dropout probabilities, which are especially severe in non-traditional lyceums, and for kids with non-educated parents.

7.3.1 Discussion and Conclusions

The analysis matched repeaters with comparable promoted following two different logics. The preliminary results call in question the usefulness of grade repetition in the observed cohorts of students.

Overall, the matching underlined that grade repetition (a) severely narrows down the probabilities to experience a linear academic career in the future; (b) generally pushes students toward a downward change of school; (c) slightly lowers students' probabilities of a subsequent grade repetition as compared to non-repeaters; (d) unequivocally increases students' chances to drop out from school; and, (e) it increases (very slightly) the probability to drop out and then return into school. An additional concern relates to the variations of GR effects among social groups: (f) students in academic tracks and culturally advantaged families tend to react to GR by changing school, while (g) students in non-academic tracks and with no- or low-educated parents tend to dropout.

Therefore, the results do not speak in favour of grade repetition as an effective catch up educational policy.

7.3.1.1 Major limitations

It is necessary to underline that the results produced in this chapter still preliminary. The one-to-one matching within calipers is only one of the many available

matching procedures. Moreover, causal identification might rely on methods other than propensity score matching (such as IV).

More importantly, the analysis lacks any detailed investigation of the characteristics of the students that provide support for the ATT in the two logics. Matching within the same school might defile the results being both grade repetition and subsequent outcomes influenced by school policies. Nonetheless, the school context is a crucial source of variation in GR selection, and any further investigation must carefully account for its influence. Perhaps, further analyses should try to operationalise school policies (such as strictness, or trust in grade repetition as a catch-up tool) to exploit between-school variations in order to identify GR impact.

Last but not less important, the present analysis should be able to intervene in the debate same-age versus same-grade comparison more concretely. The theoretical discussion must reach a further understanding of the same-grade comparison in the Italian context and, if feasible, the GR impact must also be estimated following this logic. Consistent results would be an additional clue about the detrimental effect of grade repetition.

7.3.1.2 The detrimental effect of grade repetition

Without denying such major limitations, the overall indication of a detrimental effect of grade repetition is consistent across the different lines of stratification of school type, parental educational attainment, and migration background. Moreover, the analysis consistently detects a dramatic heterogeneity of the grade repetition effect.

Privileged students (*e.g.* enrolled to traditional lyceums, or with highly educated parents, or with Italian first-recorded citizenship) react very differently to grade repetition. Although privileged students are those who react the most to a grade repetition, that is, are those whose careers are the least linear. However, this fact does not necessarily represent a loss when compared to the other repeaters. Their typical response if retained is to opt for a downward reallocation of their enrolment.

The downward change of type of school is the least costly reaction to grade repetition (when you set aside the linear career), for it leaves out any other waste of time and allows the student to continue in its academic career. The analysis shows that

privileged students opt more frequently for a return-type-of-career, which probably implies an enrolment in the private school system that may enhance a catch-up.

Grade repetition exerts more severe effects on underprivileged students (*e.g.* enrolled to other lyceums or non-academic tracks, or with low educated parents). While it is true that they are more likely to 'hold on' after a grade repetition, they also are pushed toward much more costly responses to grade repetition, that is, dropping out of school.

The present chapter, once more, calls in question the usefulness of this device is tackling students' poor performance. Grade repetition not only produces disappointing outcomes for all, but primarily acts on the well-known lines of social stratification within schools and, eventually, substantially contributes to enlarge the social gap in academic performance.

Conclusions

"You say you flunked the lazy and dumb. Then you say that God gives birth to idiots and listless people in the houses of the poor. But God does not make these pranks on the poor. It is easier for the spiteful to be you..."
Don Milani- Lettera ad una professoressa, 1967

"Pampero came back to his land with new eyes and reason. He the pastor who became a great writer of the time With new sounds and words his pages shone in the sun A dedication to the Beautiful Creole vibrated in the wind - Thank you for all your love, our Lady Education."
Murubutu – La bella creola (2016)

An enduring debate has surrounded grade repetition. This thesis has investigated grade repetition as a selection and recovery device. By reconstructing its social determinants and academic consequences, this work has particularly emphasised the practice's implications in terms of the broader mechanisms of educational inequalities and social stratification. The following paragraphs summarise the key contributions of this dissertation.

Chapter 1 has systematised social science research on the determinants of underachievement in schools. An underachieving child is a student who fails to meet the learning demands that have been set by schools for a specific grade. Therefore, the chapter has underpinned the channels through which parents transmit social, cultural and economic resources to their offspring.

The literature review has addressed three main themes, the first of which is the unequal distribution of academic performance across social groups. The concept of ability has evolved over the past decades to include both cognitive and non-cognitive aspects. The review has also described the genetic, economic, cultural and social mechanisms that are involved in the process of intergenerational transmission of academic ability.

However, social groups differ not only in terms of academic ability but also in their academic preferences. Two equally performing kids from different social groups are likely to display different academic behaviours. In the past, underprivileged children have had to blame themselves for their academic failures. The supposition that the 'poor are

dumb' imposes the burden of social stratification on the individual and frames inequalities as an inescapable result of cultural gaps.

In recent decades, the scientific community at large has recognised that the differentials in academic preferences are socially driven. Nevertheless, there is still a fervent dispute over the definition of the pivotal component of the individual educational decisions: the socioeconomic profitability of educational investments. Some scholars have claimed that people primarily view educational qualification as critical insurance against social demotion. Meanwhile, others have argued that people afford primacy to cultural motives, which biases the perceived probability of success.

To date, no empirical study has yielded any conclusive evidence. Much of the parental influence on educational achievement and attainment has yet to be explained. Somehow to get out of the corner, the last decade witnessed renewed research efforts to explore the roles of schools and teachers in perpetuating inequalities. Chapter 1 has addressed empirical evidence of the decisive influence of schools and teachers on both academic performance and educational choices.

Grade repetition targets underachievers. Chapter 2 has systematised the literature on this educational practice with the primary goal of gathering and systematically illustrating the available findings on the benefits and disadvantages of grade repetition. Although many scholars have concluded negative evaluations of grade repetition concerning its pedagogical, psychological and educational impacts, others have suggested pragmatism given the hardly ubiquitous data constraints that such research has encountered as well as the related methodological shortcomings.

Beyond the posture of the scientific community, teachers in many countries still widely adopt grade repetition to mitigate underachievement in schools. Chapter 2 has critically examined evidence of potential biases and inequalities within the process of selection. The review of the research over the impact of grade repetition comprehended two ad hoc sections for the U.S. and French discourses on the subject. In both contexts, the empirical evidence is not conclusive, and the debate over the impact of grade repetition remains open.

Chapter 2 has also investigated the influence of the institutional level on the frequency and dynamics of grade repetition by offering a descriptive analysis of its

application across various European educational systems. In this regard, educational systems are likely to engage in two main logics: social promotion and merit promotion. Social promotion favours group learning over individual learning, and it eases the linear progression of student cohorts across grades.

In contrast, merit promotion prioritises the ability of the individual to meet grade- and track-specific learning demands. Consequently, merit promotion clusters students into grades according to their achievements (or 'merit'). In countries and educational levels that embrace merit promotion, one should expect a considerable share of grade repeaters.

Chapter 3 has delivered a detailed description of the Italian institutional setting to contextualise the subsequent empirical analysis. Although widely accepted and diffused, scholars and teachers have persistently understudied grade repetition in the Peninsula. The lack of appropriate data has undoubtedly contributed. Italy has produced neither a specific official report on the topic nor a longitudinal study of student careers at the national or local level. In this context, the chapter has investigated the relevant institutional and jurisprudential features of the Italian educational system and their developments in terms of grade repetition.

Italian schools have shifted from a strictly merit-based and highly selective system to more open scholastic institutions that significantly adopt social promotion. Lower-secondary education became comprehensive by the 1960s, and the social promotion regime has since characterised both middle and primary schools. However, in high schools, students contend with different promotional standards depending on their chosen track. With these standards, grade repetition has remained substantial, and the merit promotion regime supplies the primary logic for the progression of students through grade levels.

The empirical analysis has relied on an *ad hoc* extensive longitudinal database of academic careers of secondary level Northern Italian students. The available information includes various measures of academic achievement, and socio-economic and migration background (see Chapter 4). The richness of the database makes this work an innovative empirical analysis in the Italian context.

Chapter 5 has examined the determinants of grade repetition by analysing the probability of being subject to grade repetition at the end of ninth grade, which is the first year of upper-secondary education in Italy. At this stage, the analysis clarified the influence of a comprehensive set of risk factors while emphasising the importance of socioeconomic and migration backgrounds.

As mentioned, the rich array of measurements for academic performance before grade repetition has ensured a credible assessment of the enduring influence of parental resources on the occurrence of grade repetition. To the author's knowledge, this study is the first to address the issue with such comprehensive data in the Italian context.

Disadvantaged students are more likely to be underperformers. Nonetheless, past performance cannot fully explain the estimated gap in grade repetition probabilities between social groups. Instead, the 'grade repetition social gap' (as referenced in Chapter 5) is broad and consistent across all learning contexts. Even after controlling for middle school performance, disadvantaged kids (i.e. those with low socioeconomic status and a migration background) in ninth grade are substantially more likely to repeat a grade compared to their advantaged peers (i.e. those with high socioeconomic status and a native background).

Chapter 6 has centred on an in-depth analysis of academic careers after grade repetition. The analysis examines ninth grade-retained students with distinct socioeconomic and migration characteristics. Regardless of their social background, their chance of achieving a linear academic career was limited. Moreover, careers after grade repetition significantly diverged according to social and, to a lesser extent, migration backgrounds.

The empirical analysis of this dissertation has concluded with a preliminary assessment of the causal impact of grade repetition. This assessment employed ex-post matching strategies to compare the educational outcomes of similar retained and non-retained students.

The impact analysis in Chapter 7 has provided no evidence of any positive effect on grade repetition. On the contrary, grade repetition seems to entrench educational inequalities and is likely to widen the gap in school performance.

The present work has encountered several limitations that highlight directions for further research. The first was a lack of external validity in the Italian context. The empirical basis of the analysis encompasses the three most populous Northern Italian regions. In a country such as Italy, which exhibits a marked socioeconomic division between the North and the South, the analysis of this work should be replicated with data from other regions. A nationally representative dataset would allow for deriving general coordinates (and policy orientations). Still, it is notable that the present work is the first to consider this topic in the Italian context, and it relies on a large number of individual observations.

A second limitation was the insufficiency of information about the trajectories of students who disappeared from the national register of students during the observed period. In this regard, it would be advantageous for vocational training to distinguish between dropout and enrolment.

As discussed in Chapter 7, a third limitation derives from the preliminary stage of the causal analysis. Future efforts in this direction should focus on more systematic choices of both the model of selection and the matching method. A more accurate strategy of identification could generate significant knowledge of the phenomenon in Italy and is thus of extreme interest to policymakers and workers in the field of education.

The following paragraphs discuss three additional research perspectives that are crucial for consolidating the overarching interpretation of grade repetition advocated in this work.

The first research perspective aims to investigate the motives of teachers in using grade repetition. The practice exerts both selection and recovery (or catch-up) functions. On the one hand, grade repetition seeks to be an educative and formative tool for underperforming students as well as, undeniably, a recovery device.

The belief that the student would profit from repeating the year informs the teacher's decision. From this perspective, grade repetition could present an opportunity for the retained student to catch up.

On the other hand, grade repetition has a clear selection goal in that it seeks to distinguish students who underperform from those who meet or surpass the threshold.

Students who fail to reach the required learning threshold for a given grade must repeat that year. Grade repetition resolutions are presumably a mix of these two factors (education and selection).

The analysis demonstrates that retained students continue to perform poorly, and they are likely to experience further academic misfortune. The largest share of repeaters has non-linear academic careers that are mainly characterised by a dramatic dropout rate.

Concerning the selection aim, the observed (and unobserved) criteria that teachers apply in their decision accurately describe the future academic performance of the repeaters, which indeed continues to be very modest afterwards. For the same reason, the empirical observation does not favour grade repetition as a recovery (and educative) device because there appears to be no generalised post-grade repetition recovery in place. Although preliminary, results concerning the impact of grade repetition suggest that the effect is consistently adverse.

Teachers presumably have some opportunities to collect at least some information about the subsequent academic careers of repeaters. For example, teachers could instruct classes of repeaters or receive insights from colleagues. If this assumption is accurate, then teachers should recognise that repeaters perform poorly on average. Furthermore, from a comparison of academic careers before and after the grade repetition, teachers may notice that underprivileged children are typically less successful. Thus, further empirical inquiries include why teachers continue to rely on grade repetition despite its poor results in helping underachieving students recover as well as why beliefs have not updated about the efficacy of educational motives behind grade repetition.

This work hypothesises that a marked unbalance exists between the selection for grade repetition and educational motives. Teachers consider grade repetition to be primarily a selection tool; therefore, they do not stress the effectiveness of grade repetition as a catch-up function. At most, they require only a few successful cases of reformed repeaters to be satisfied with the educational effectiveness of grade repetition (see Shepard 1989; Marxoux and Crahay 2008). The core of the grade repetition is precisely the act of imposing grade repetition itself, which is an act of stopping students. Moreover, the emphasis on selection is most harmful to children from disadvantaged backgrounds, as they have the most mediocre prospects.

In this context, it is crucial to explore beliefs that justify grade repetition for teachers. One option is to design a field experiment in which the experimenter presents a treated group of teachers with unbiased figures about the repeaters' subsequent careers. The conveyed information should explicitly incorporate diverse profiles in terms of social class and high school track. The effect of interest is the degree of sensitivity to the new information for each profile, which would presumably be smaller when the teacher assigns importance to the selection (rather than educational) aim of grade repetition.

The second line of research faces the relevance of grade repetition within the broader relationship between performance and evaluation in the school setting. Feedback from teachers has critical implications for grade repetition selection, as it is likely to affect students' self-perceptions. As Chapter 1 has highlighted, the effectiveness of feedback and signals is heterogeneous across social groups. Specifically, disadvantaged students are more likely to devalue their chance of success and disinvest from their academic efforts. Meanwhile, compensatory mechanisms that result from anelastic academic preferences allow advantaged kids to counteract the adverse effects of grade repetition and, in practice, almost ignore the signal.

As a result, among retained students, disadvantaged students have the worst performance after grade repetition. Such poor performance affirms the belief of teachers that such students have the lowest likelihood of success (for interested readers, Appendix N presents a formalisation of this self-feeding loop).

It is worth acknowledging that the influence of teachers is not limited to grade repetition. Upon closer examination, grade repetition is only the apex of a feedback process that extends throughout the entire academic year. From the first day, students demonstrate their performance to teachers, who in turn provide feedback. This feedback might assume the form of official notes or informal oral recommendations; alternatively, it might consist of a punishment or reward. In the case of weak performance, the final point of feedback is the teacher's decision about grade repetition.

In this respect, it would be helpful to clarify the relationship between the performance and the evaluation as endogenously emerging from the school setting. A theoretical model should formalise the interaction, and an experimental setting should verify its implications. The core hypothesis is that specific school or class settings may

enhance or discourage individual academic performance depending on the social class and cultural resources of the student.

Finally, a third line of research wants to exploit social-class variations in behavioural responses to academic failure to dialogue with the literature on decision effects. The very influential relative risk aversion hypothesis formalised by Breen and Goldthorpe (1997) predicts that students would react differently according to the amount of socioeconomic risk implied in the choices, and to the social mobility preferences. Grade repetition here offers the precious chance to study the heterogeneity of the impact on subsequent educational decisions of a downward exogenous shock in the perceived academic prospects.

Therefore, the plan is to integrate further the existing database with indexes able to measure school-specific socioeconomic risk. One way to do it is to rank curricula over their vocational and generalist contents. Vocational-preeminent curricula would be more attractive when students want to minimise the risk of ending up in unskilled jobs as opposed to generalist-preeminent, which instead offer better chances to access to prestigious social positions. Another way is to look at the school-specific labour market and university outcomes of the high school graduates. After a failure, the re-enrolment in schools with minimal labour market entry could imply a higher risk for working-class pupils if compared to their more privileged peers who might be resilient reproducing their social class.

Besides the limitations mentioned above and the advocated perspectives, two statements are hardly contestable, once one looks at the extensive analysis in this work:

(1) grade repetition appears to disproportionally select disadvantaged students;
(2) there is no convincing empirical evidence of the educative or recovery consequences of grade repetition.

Eventually, the author of this thesis shares Labaree's (2010, p.170) commitment to pedagogical progressivism, which strives for 'developing a new process of teaching and learning in the classroom' and, accordingly, for supporting the emancipative role of schools in the transformation of social structures, instead of their mere reproduction.

The more the school delegates the recovery process to families, the more accurate teachers would be in accounting home resources as necessary elements to estimate academic prospects.

Policymakers should hence afford more material and cultural resources to schools, thereby providing deans, teachers and other educational workers with a broader range of tools to radically counterbalance the performance disparities that derive from the exposure of children to different family environments.

APPENDIX 1: External validity

Early leavers

The 18.8% of Cohort A students do not enrol in high school in 2015-16. The table below describes the composition of the early leavers and compares it with the non-early leavers and the entire cohort population.

	Early Leaver (18.8%):		
	No	Yes	Pop
Female	50.1%	39.0%	48.0%
Over age	5.0%	19.9%	7.8%
1st gen Immigrant	4.7%	12.6%	6.1%
2nd gen Immigrant	5.8%	11.5%	6.9%
ESCS index	0.18	-0.50	0.08
Italian Test score	0.18	-0.69	0.09
Maths Test score	0.19	-0.74	0.09
Italian mark	7.3	6.3	7.2
Maths mark	7.3	6.2	7.1
Behaviour Mark	8.8	8.1	8.7

The present thesis presumes that the largest share of early leavers has enrolled in VET institutions. The presumption seems to be consistent with the public data of the Italian Institute of Statistics (ISTAT) on the subject. The next table compares the number of students enrolled in the first year of VET in 2015-16 (source: Table 7.4 in the 'Annuario Statistica ISTAT 2018') against the early leavers in my data.

	Istat data	My data
Piemonte	5372	6,607
Lombardia	15842	18,524
Veneto	7319	8503
Total	**28533**	**33,634**

The unexplained portion of early leavers might be enrolments in schools or VET institution outside of the three regions in the data, or illegal dropouts (if under 16 years old).

APPENDIX 2: Additional material for Chapter 6

Clusters of academic careers in details

In the table below, I display the five-clusters classification of the 94 individual careers. The first and the second number of the individual sequences indicate enrolment decisions in 2015 and 2016, respectively. They take 1 if no change, 2 if horizontal change, 3 if downward change and zero if no enrolment. The third and the fourth number correspond to the academic outcome of 2015 and 2016, respectively: zero is a dropout, 1 is grade repetition, and 2 is promotion.

Cluster	**Description**	**Sequences falling in the cluster**
Linear	After the grade repetition in 2014, a student gets promoted two times (in 2015 and 2016) without experiencing any downward change	1-2-1-2; 2-2-1-2; 1-2-2-2; 2-2-2-2
Change down	After the grade repetition in 2014, a student gets promoted two times (in 2015 and 2016). However, she\he opts for a downward change at least once	3-2-1-2; 1-2-1-1; 1-2-3-2; 2-2-3-2; 3-2-3-2
New Grade Repetition	After the grade repetition in 2014, a student gets retained at least once in 2015 or in 2016.	1-1-3-2; 3-2-1-1; 2-2-1-1; 1-1-2-2; 3-2-2-2; 1-2-2-1; 2-1-1-2; 1-1-3-1; 3-1-1-2; 1-1-1-2; 2-1-1-1; 1-1-2-1; 2-1-3-2; 1-1-1-1; 3-1-1-1; 3-1-2-2; 1-2-3-1; 3-1-3-2; 3-2-2-1; 2-1-2-2; 2-2-2-1; 2-1-3-1; 2-2-3-1; 3-1-2-1; 2-1-2-1; 3-1-3-1; 3-2-3-1
Dropout	After the grade repetition in 2014, a student appears as dropout by the end of 2016.	0-0-0-0; 1-1-0-0; 1-2-0-0; 1-2-1-0; 1-0-0-0; 1-0-1-0; 3-0-1-0; 3-2-1-0; 1-1-1-0; 2-2-1-0; 2-1-0-0; 1-1-3-0; 3-1-0-0; 2-2-0-0; 2-0-0-0; 2-0-1-0; 3-2-0-0; 2-1-1-0; 1-1-2-0; 3-1-1-0; 3-0-0-0; 0-0-2-0; 1-0-2-0; 1-0-3-0; 1-2-2-0; 1-2-3-0; 2-1-3-0; 3-1-2-0; 3-0-2-0; 3-2-2-0; 2-1-2-0; 2-0-3-0; 2-0-2-0; 2-2-2-0; 3-1-3-0; 2-2-3-0; 3-0-3-0; 3-2-3-0
Return	After the grade repetition, a student appears as dropout at least in one out of the 4 observation points. However, s\he re-enrols by 2016	0-0-2-2; 1-0-1-1; 3-0-1-1; 0-0-2-1; 1-0-3-2; 1-0-2-1; 1-0-2-2; 2-0-1-1; 2-0-1-2; 3-0-1-2; 1-0-3-1; 1-0-1-2; 2-0-2-2; 3-0-2-1; 2-0-3-1; 3-0-2-2; 2-0-2-1; 2-0-3-2; 3-0-3-2

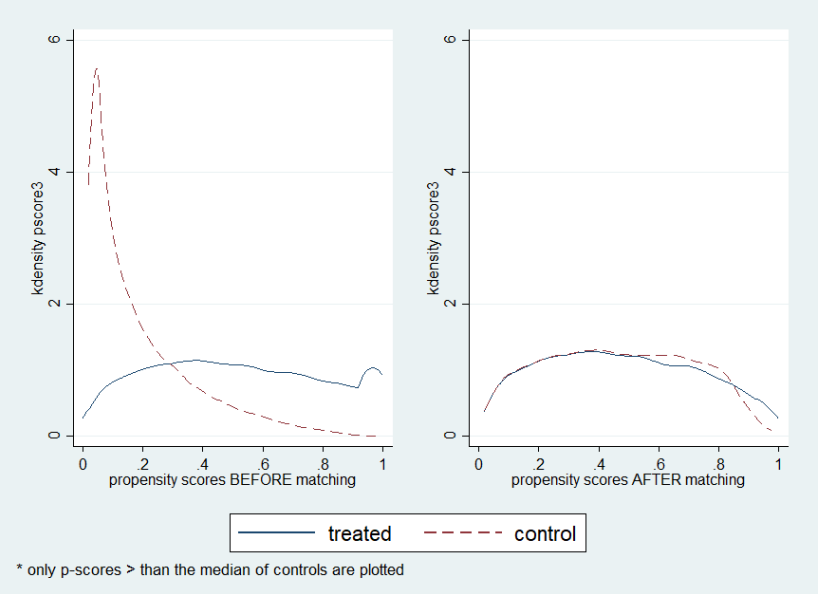
APPENDIX 3: Additional material to Chapter 7

Standardised bias and plots are presented for both strategies. Standardised bias is the weighted difference between the mean of matched treated and controls, divided by the standard deviation of treated (Ho et al. 2007).

Strategy A

	Matched controls		Matched treated		
	Mean	S.D.	Mean	S.D.	Standardised bias
Female	0.399045	0.489729	0.397868	0.489474	0.002404
Over age	0.159919	0.366551	0.170711	0.376269	-0.02944
1st gen. immigrant	0.112761	0.316318	0.119236	0.324077	-0.02047
2nd gen. immigrant	0.070181	0.255466	0.073517	0.260992	-0.01306
No high school diploma	0.356073	0.478864	0.359082	0.479747	-0.00628
High school diploma	0.1038	0.305018	0.100203	0.30028	0.011794
Tertiary degree	0.093466	0.2911	0.098437	0.297914	-0.01708
Italian mark	-0.7204	0.65289	-0.72851	0.650727	0.012422
Maths mark	-0.62563	0.66391	-0.64257	0.671145	0.025516
Behaviour mark	-0.53065	0.875509	-0.54333	0.865458	0.014493
Maths score	-0.43691	0.814004	-0.46848	0.841604	0.038786
Italian score	-0.43411	0.723432	-0.45569	0.74516	0.029823

The graph plots the distribution of the propensity score, before and after matching.

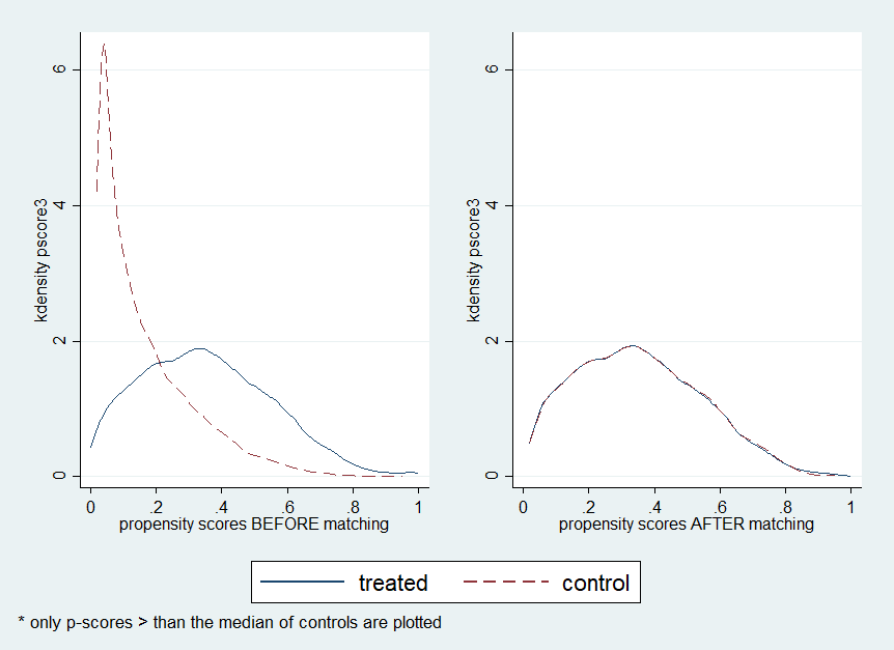


Strategy B

Standardised biases are zero in the categorical variables due to the perfect matching in Strategy B.

	Matched controls		Matched treated		
	Mean	S.D.	Mean	S.D.	Standardised bias
Female	0.397679	0.489439	0.397679	0.489432	0
Over age	0.19	0.39	0.19	0.39	0
1st gen. immigrant	0.128837	0.335034	0.128837	0.335029	0
2nd gen. immigrant	0.078887	0.269573	0.078887	0.269569	0
No high school diploma	0.448509	0.497363	0.448509	0.497355	0
High school diploma	0.349874	0.47695	0.349874	0.476943	0
Tertiary degree	0.097811	0.297071	0.097811	0.297067	0
Italian mark	-0.75548	0.613176	-0.74954	0.63098	-0.00969
Maths mark	-0.67706	0.640616	-0.66666	0.659638	-0.01623
Behaviour mark	-0.56226	0.871613	-0.58145	0.878705	0.022027
Maths score	-0.50753	0.818257	-0.50834	0.847947	0.00099
Italian score	-0.48605	0.730963	-0.49503	0.755447	0.012274

The graph plots the distribution of the propensity score, before and after matching, for Strategy B.



APPENDIX 4: Multilevel Models for Clustered Data and Centering Decisions

Introduction

Multilevel models are designed to deal with data with hierarchically nested systems with more than levels (Snijders and Bosker 2010).

On the one hand, the use of *ad hoc* models when dealing with clustered data allows avoiding the nuisance produced by the dependence among unities in the same group, which violates the independence assumptions of the standard statistical models.

Moreover, within-group dependence might be of scientific interest in many cases. The accurate estimation of higher levels and intra-levels coefficients is of substantive interest for many scholars in social sciences.

Variance

Some relevant preliminaries must be introduced. In a two levels data structure, the random effect ANOVA model allows for the variance decomposition. Any individual outcome y_{ij} of the i-th unit in the j-th cluster would be:

$$y_{ij} = \mu + u_j + e_{ij}$$

With μ as the population mean, u_j as the specific deviation from the mean by the *j-th* cluster, and e_{ij} as the residual random error.

In the random effect ANOVA model, the total variance is equal to $\sigma_u^2 + \sigma_e^2$, and the covariance:

$$Cov(y_{ij}, y_{i'j'}) \begin{cases} 0 \; if \; j \neq j' \\ \sigma_u^2 \; if \; j = j \; and \; i \neq i' \end{cases}$$

On this basis, an important parameter of multilevel models is the Intraclass Correlation Coefficient (ICC), usually denoted with ρ, which shows the degree of similarity between same-group units:

$$\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

With σ_u^2 as the cluster variance and σ_e^2 as the residual variance. In other words, the ICC computes the share of cluster variance in respect of the total variance in the data.

Multilevel models compliant to the following specific notation:

i indexes individuals within the groups ($i = 1, \dots, n_i$)

j indexes the groups ($j = 1, \dots, N$)

Y_{ij} is the dependent variable

x_{ij} is the explanatory variable at the individual level

z_j is the explanatory variable at the group level

The Fixed Effects Model

The simplest way to model a hierarchical data structure is the fixed effect model:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 z_j + u_j + e_{ij}$$

where the beta zero is the value of the intercept, beta one is the contribution of the individual-level variable, and beta two is the coefficient for the group variable Z; ε is the error term. The model also contains the Us, which are cluster-specific parameters. In other words, u_j is the cluster-specific deviation from the mean common to all the units in the j-th cluster.

In the practice, the betas in the fixed effects models represent the within-cluster effect. Betas do not embrace the total effect because the whole between-cluster variation is absorbed by the fixed-parameters U. Conditional to the X, the expected y_{ij} value deviates from the mean by the value of its U_{0j}.

As such, the fixed-effect model needs no assumptions over the distribution of the cluster effects, and hence any possible endogeneity is ruled out.

The Random Intercept Model

U_{0j} can also be interpreted as unexplained group effects, that is, residuals sorted from a population with zero mean and unknown variance, which would define the random intercept model.

The random intercept model can also be derived as a set of hierarchical equations, which has at level-one:

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + e_{ij}$$

and at level-two:

$$\begin{cases} \beta_{0j} = \gamma_{00} + u_{0j} \\ \beta_{1j} = \gamma_{10} + u_{1j} \end{cases}$$

where γ_{00} and γ_{10} are respectively the mean intercept and the mean slope.

When combined, the equations result in the following:

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + e_{ij} =$$

$$= \gamma_{00} + u_{0j} + (\gamma_{10} + u_{1j})x_{ij} + e_{ij}$$

$$= \gamma_{00} + \gamma_{10}x_{ij} + u_{0j} + u_{1j}x_{ij} + e_{ij}$$

in which the last three terms are the random portion of the model.

Researchers should favour random intercept models in respect of fixed-effect models when they hold a substantive interest in the higher-level influence on the individual-level outcome. Fixed effects models leave no unexplained variability between groups, which impede to look for any specific cluster characteristic's contribution to the outcome.

Moreover, the random intercept model offers significative advantages when dealing with small groups. In fixed-effects models, the number of cluster-level variables equals the number of clusters. In the case of many small groups, this is likely to overfit the data while adding little information.

In random intercept models, instead, this problem is avoided. Since the random intercept model assumes that group-effects are randomly drawn from the same population, each of the group is exchangeable, and the mechanism explaining the between-group variation is identical between groups.

Therefore, with many low-populous clusters, you have a significant loss of efficiency (many parameters) and even inefficient estimation of the cluster effect (for it is estimated on a very few cases).

Parameters estimation for the Random Intercept Model

In the random intercept model, I need to estimate both the parameters and the variance components. The random variables are not directly estimated, for they are non-directly observable, and are considered latent variables.

Note that, although the above equation includes the random effects, they are not directly observable. Therefore, they need to be integrated out.

The Full Information Maximum Likelihood (FIML) is a common method of estimation. It maximises the following equation:

$$(\theta) = \int (y|u,\theta) \cdot (u|\theta) \cdot u$$

where $y|u,\theta$ is the distribution of responses conditional on random effects and parameters, and $u|\theta$ is the distribution of random effects conditional on parameters.[1]

There is another common method of estimation, namely, the REstricted Maximum Likelihood (REML), the main difference being that FIML jointly estimates the fixed and the random parameters, while REML keeps this estimation separate. When the number of clusters is much higher than the number of two-level variables the difference between FIML and REML is negligible (Snijders and Bosker 2012).

[1] Although out of the scope of this work, I shall mention that random effects can be eventually predicted as posterior means produced by the empirical Bayes estimation. The empirical Bayes estimation combines the information coming from the population (e.g. the data from each group) and that assumed by the model (e.g. random effects are drawn from a unique population). The empirical Bayes estimation is the weighted average between these two sources of information.

Multilevel Logistic Regression

The equation for the usual linear predictor:

$$y_{ij} = \beta_{0j} + \beta_1 x_{ij} + u_j + e_{ij}$$

is translated into a probability following a Bernoulli distribution thanks to the logit link $logit(\pi_{ij}) = y_{ij}$.

The betas are the conditional effects of the covariates given the random effects:

$$\pi_{ij} = P(Y_{ij} = 1 | u_j, x_{ij})$$

which produces cluster-specific effects, against the population-averaged coefficient produced by a standard logistic regression. While in linear models there is no difference, in non-linear model $P(Y_{ij} = 1 | u_j, x_{ij}) \neq P(Y_{ij} = 1 | x_{ij})$.[2]

The interpretation is that they identify the effect with the random part held constant and, as such, cluster-specific odds ratios compare the outcome within a cluster.

Centering decisions

Suppose that I am interested in including group means $\overline{x_j}$ as covariates. In the multilevel framework, there are different centering alternatives, which have an impact on the interpretation of the coefficients, namely:

- Raw centering: the intercept is the expected value of Y when Xs are zero:

$$y_{ij} = \beta_{0j} + \beta_1 x_{ij} + \beta_2 \overline{x_j} + u_j + e_{ij}$$

- Centering around the Grand Mean (CGM): the intercept is the expected value of Y for an average individual in the sample:

$$y_{ij} = \beta_{0j} + \beta_1 (x_{ij} - \overline{X}) + \beta_2 \overline{x_j} + u_j + e_{ij}$$

[2] Parameter estimation for multilevel non-linear model is more complex. STATA (the software used in this work) estimates the parameters of the multilevel logistic regression using a method of adaptive quadrature, which appears to be unbiased in large samples (see: Rabe-Hesketh et al. 2005).

- Centering Within Contexts (CWC): the intercept is the expected value of Y for an average individual in the group:

$$y_{ij} = \beta_{0j} + \beta_1(x_{ij} - \overline{x_j}) + \beta_2\overline{x_j} + u_j + e_{ij}$$

In GMC, the variance of the intercept represents the between-group variance in the outcome variable adjusted for the level-1 variables.

In CWG, the intercept variance represents the between-group variance in the outcome variable. CWC produces an estimate of the within-group regression that is not confounded by the presence of between-group variance in the predictor variable, whereas CGM and RAS do not remove this variation.

CGM or RAS keep the correlation between individual value and group mean. The interpretation of group means is the effect of level-two composition once the level-one variable has been accounted for.

Conversely, with CWG, the effect of group means is independent of the effect of individual value because CWG rules out the correlation between the levels. Here I want to determine whether the relationship between level-two and the outcome differs from zero.

The interpretation of the coefficients is a within-cluster regression and hence suggests a theory: an individual's relative standing within his group and the effects are dependent on the context.

Kreft et al. (1995) notice that even if RAW and CGM are equivalent, CGM provides a computational advantage for it reduces the correlation between intercept and slope estimated across groups. CGM might produce biased estimates when the model randomly assigns slopes, for it is a mix of the within-group slope and between-group slope.

The best choice is often dependent on the context of the study and its peculiar research questions. Hofman and Gavin (1998) identify four paradigms of interpretation when I have variables that span multiple levels of analysis (*cross-level paradigms*):

i) *Incremental*. Group level variables act as main effects in the prediction of individual-level outcomes. The aim is to investigate the

influence of group-level variables on individual-level outcomes after controlling for various individual-level predictors. Does the group level provide an incremental prediction on individual outcome over and above individual-level predictors?
With this aim, one should use GMC or CWG + group means of individual vars. However, in contextual models, the two are helpful.

ii) *Mediational*. Group level variables influence individual behaviours and attitudes only indirectly through other mediating mechanisms. These influences can consist of shared perceptions, *e.g*. social information shared in networks. In this case, either GMC or CWG plus means are suitable.

iii) *Moderational*. Group level variables moderate the relationship between two individual-level variables. Hofman and Gavin (*ibidem*) purpose the example of how the organisational context can influence the magnitude of the relationship between two individual-level variables. This paradigm estimates slope variance across groups. CWG is appropriate here.

iv) *Separate*. Separate structural models for both the within-group and between-group components of the outcome variable. In this paradigm, the variance in the outcome measure needs to be partitioned into its within and between components. CWG is the right choice. [3]

[3] A few words on three-level hierarchical models: CGM creates a composite score, that is, the sum of 3 terms (Brincks 2012): a deviation of the raw score from the level-2 group mean; a deviation of the level-2 group mean from the level-3 group mean; a deviation of the level-3 group mean from the grand mean. CWC deviates the raw score from the mean across individuals in the level-2 group. To centere around level-3 is not the right choice when there is significant variability in the outcome due to level-2 clustering.

Summary table

Centering	**What I measure**	**Notes**	**Cross-level paradigm**
RAS or GMC on all levels	Effect of individual X on Y. At the school level, the effect of school average on Y, once I account for the effect of individual X.	Not suitable for isolating the effects of each level. Variables from different levels remain correlated among them.	Incremental
CWG on individual X, RAS or GMC on school X.	Effect of individual X on Y. At the school level, the direct effect of school average on Y. In practice, I fit a within school regression on the individual level	The individual-level effect does not correlate with the school average. Suitable to isolate the school effect.	Separate

APPENDIX 5- The Grade Repetition game

Introduction

In his seminal paper on signalling games, Spence (1973) writes that “in most job markets the employer is not sure of the productive capabilities of an individual at the time he hires him. Nor will this information necessarily become available to the employer immediately after hiring” (p.356). Rather than true productivity, employers observe “a plethora of personal data” (p.357). Spence refers to those characteristics that cannot be manipulated by the subject (*e.g.* sex, ethnicity, social origin) as 'indexes', whereas he names 'signals' those on which the subject influences (*e.g.* education). Individuals can invest in a signal by paying a signalling cost. Spence assumes that signalling costs negatively correlate with productive ability, and hence are a useful proxy for the employer. Sometime after the hiring, the employer will learn individual productivity. Upon his experience on hired employees, the employer shapes its beliefs over the relation between signals and indexes, and productivity. Therefore, he offers (and revises) its wage schedule. Through this feedback loop, Spence claims that employers bear self-confirming beliefs. That is, employers' expectations, having an impact on employees’ behaviours, realise after the hiring. Importantly, employees know both their real productivity and the offered wage schedule. By taking this data into account, they make their investments in signalling.

I can profitably model GR by framing teachers' decision in a similar informational structure. (In fact, I might extend the model to interpret scholastic evaluation interactions generally).

The fundamental imbalance of information regards students' latent 'ability', which teachers cannot directly observe. I am agnostic about the exact characterisation of the latent 'ability'. A minimum definition should encompass its multidimensionality (cognitive and non-cognitive traits), and its evolution along the life course due to the parental, social and scholastic environment investments.

Teachers are trained to observe a rich set of latent ability's signals: homework, classwork, behaviour in the classroom, etc. I shall refer to them as the observed (and recorded) academic performance.

Each school set a learning threshold, which works as a reference point to teachers in their evaluation decisions, and, ultimately, in GR decision. The quantity and quality of the required scholastic performance vary across learning contexts. (As in Bowles and Gintis (2001), an agent such a school can prompt diverse sets of traits. The main driver in learning goals differences are high school tracks.

The objective of teachers is twofold. On the one hand, through the implementation of specific teaching strategies, they aim to foster the ability of their students. I shall refer to this objective as the teaching goal. On the other hand, teachers must select students over their ability. For each grade, teachers check if students assimilated the learning goals. I shall refer to this objective as the screening goal. The further students' progress in the educational system and the more crucial the screening goal becomes at the expense of the teaching goal. In the Italian system, the screening goal turns to be particularly relevant from high school.

Emerging literature accords to teachers' expectations toward students a crucial role in the educational outcomes (Steel 1997; Ferguson 2003; Rosenthal and Jacobson 1968; Darmon 2012). Expectations are likely to exert a direct effect on students' performance. Students tend to comply with teachers' expectations if they perceive (either negative or positive) feedback. (Expectations can also affect teachers' evaluation in GR decisions).

Literature also underlines that teachers are the best predictors of a student career. Teachers are likely to keep some beliefs about the effectiveness of their predictions. After the decision to retain or to promote, teachers collect pieces of information about the educational outcomes. They might directly observe students they retained, or they receive some feedback from colleagues; moreover, teachers can observe retained students in their classes. Hence, teachers observe whether scholastic performance has improved or not. Based upon these observations, teachers adjust their belief about the relation between students' latent ability and ii) their observed performance and ii) some other relevant characteristics. If their predictions end up being accurate, they hold self-confirming beliefs.

For their part, students are likely to perceive teachers' feedback and consequently adapt their course of action.

This paper aims to model this interaction between teachers and students drawing upon Spence's seminal paper. I call the model the Grade Repetition Game. In the grade repetition game, actors' beliefs, preferences and courses of action emerge endogenously from the interaction between teachers and students.

I first present a somewhat unrealistic model of the grade repetition game. In the world it describes, there are only two kind of students (strong and weak) and investments on performance have no impact on latent ability. After that, I relax these two assumptions. I model ability as a continuous variable influenced by performance investments. I then introduce, following Spence's argument, indexes, namely, social class and migration background, and I explore the consequences of this development. I show how an academic environment could be performance-enhancing for some while being performance-discouraging for others. To conclude, I explore the case when the grade repetition game is dynamic, and when it includes high school choice.

The Grade Repetition Game

The grade repetition game is a signalling game. In signalling games, a player A learns a private piece of information and might send a signal to a player B. Player B does not observe the information directly, but observes the signal he receives from player A. In the grade repetition game, player A is a student, and player B are the (risk-neutral) teachers. Teachers are called to teach and screen students.

To start, let us assume that the (unrealistic) scenario in which the student learns his type *t* from a finite set:

$$T = \{s; w\} \quad (1)$$

Where *s* and *w* represent a student with strong and weak latent ability, respectively. The probability distribution π over T is the probability of each type *t*. Notably, $\pi(T)$ is common knowledge, that is, both players are aware of the types' distribution in the population of students.

Having learnt his type, the student chooses to invest in a specific quantity/quality of scholastic performance *m*, which works as a signal to teachers. As underlined above, I assume *m* to be scholastic performance in its complex multidimensionality (*e.g.* marks,

test scores, behaviour in the classroom, etc.). Remember that the cost of m correlates negatively with latent ability. For the sake of clarity, I assume real numbers as an example, such that:

$$C_w = m \tag{2}$$

$$C_s = m/2 \tag{3}$$

However, I can think of any quantity such that $C_w > C_s$. The second assumption is that the signal is not productive; that is, it does not have any impact on students' 'ability' (his type). However, as I see later, this assumption can be relaxed.

Having heard/observed the student's signal m, teachers (player B) chose a response r from a finite set R:

$$R = \{gr; pr\} \tag{4}$$

Where gr and pr are grade repetition and promotion, respectively. If teachers opt for grade repetition, the student repeats the year, and he must repeat the grade. Both types of students accord strictly higher utility to the promotion, such that. If teachers choose to promote the student, he advances in its educational career. According to their teaching goal, teachers aim to get students to reach the learning target for any specific grade. Under this respect, they rely on grade repetition as a threat to exert effort from students. According to their screening goal, teachers want to select only those students who, at the end of the year, have reached the learning target. In facts, in this world, teachers want to retain a weak student, while they want to promote a strong student. Teachers know the types of distribution $\pi(T)$. More importantly, they hold some beliefs about the distribution of types given the message. Based upon these pieces of information and beliefs, they develop a strategy, $\theta(r, m)$, that is, the probability distribution of their response given the received signal.[1]

[1] Remember that I am now following the simplistic version of the retention game, where performance investments do not affect latent ability. One could say that their teaching goal is absent because they by no means can affect students' latent ability.

In what follows, I solve the simple version of the grade repetition game by exploring three kinds of equilibria, which correspond to three different sets of beliefs on the side of teachers. I explore the implications of each equilibrium.

Separating equilibrium

In this first equilibrium, I assume teachers believe that there is a quantity of signal m^* such that if $m < m^*$, the student is a weak student with probability 1; whereas if $m \geq m^*$, the student is a strong student with probability 1. Accordingly, if they see $m < m^*$, they retain; if $m \geq m^*$, they promote.

Drawing upon teachers' set of beliefs and signalling costs, students invest in their signal (scholastic performance). Two clarifications are in order at this point. Even without investment, I can assume that latent ability would automatically produce a 'base level' of signal, which is generally higher than zero. The interpretation of this would be that even without any effort or investment, a student reaches a specific scholastic performance. Such an extension is not of any harm to the model. Moreover, in the dynamic extension of the grade repetition game, I maintain that the repeated interaction between teacher and students can push the student to $m = 0$ (see the specific section below).

Nevertheless, in the simple grade repetition game, it is easy to see that the two rational alternatives are either $m = 0$ or $m = m^*$Indeed, under or above m^*, each investment would be a waste of resources. Weak students set $m = 0$ if their signalling costs exceed the 'promotion premium':

$$m^* > (pr - gr) \tag{5}$$

Proof:

$$gr > pr - m^* \tag{5a}$$

$$m^* > pr - gr \tag{5b}$$

While strong students set $m = m^*$ if the promotion premium is higher than their signalling costs:

$$m^* < 2(pr - gr) \tag{6}$$

Proof:

$$gr < pr - m^*/2 \tag{6a}$$

$$m^*/2 < pr - gr \tag{6b}$$

$$m^* < 2(pr - gr) \tag{6c}$$

Teachers' prior beliefs are hence confirmed, provided that:

$$(pr - gr) > m^* > 2(pr - gr) \tag{7}$$

I notice that between these two quantities, there are infinite equilibria, that is, teachers can profitably set m^* between these two quantities and still receiving confirmation of their beliefs from students' behaviour.

However, these equilibria are not the equivalent for the general welfare and aggregate outcomes. In particular, the higher m^*, the more demanding students' investment to avoid grade repetition. In particular, the farther $m^*$is from the lower bound $(pr - gr)$, the greater the number of resources that must be invested by the strong student, with losses in the outcome.

In respect of individual welfare, I can rank signalling equilibria according to the non-signalling equilibrium. The non-signalling equilibrium is the equilibrium that I would have if signalling were not possible. In the grade repetition game, the non-signalling equilibrium corresponds to the 'social promotion regime', in which students automatically advance to the next grade.[2]

[2] When the outcome is a marginal gain from a continuous variable, the receiving player (e.g. the employer) would correspond to each sender (e.g. the employee) the average expected marginal utility (e.g. salary). See Spence (1973).

The payoff for this automatic promotion is the average expectation by teachers:

$$\pi_w(gr) + (1 - \pi_w)(pr) \tag{8}$$

where $\pi_w \in 0 < \pi_w < 1$is the expected proportion of weak students in the population. The smaller the proportion of weak students, the lower is the average payoff. I can interpret this as the perceived value of the promotion from students and teachers. In social promotional regimes, the fact to be promoted is naturally less worthy than in performance-based promotional regimes. Regardless of m^*, weak students strictly prefer the not-signalling equilibrium to any signalling one:

$$\pi_w(gr) + (1 - \pi_w)(pr) > gr \tag{9}$$

(proof: the left branch of the equation is always greater). Strong students instead prefer non-signalling equilibrium when:

$$m^* > 2\pi_w(pr - gr) \tag{10}$$

Or:

$$\pi_w < m^*/[2(pr - gr)] \tag{10a}$$

Proof:

$$\pi_w(gr) + (1 - \pi_w)(pr) > pr - m^*/2 \tag{10b}$$

$$m^*/2 > pr - \pi_w(gr) - (1 - \pi_w)(pr) \tag{10c}$$

$$m^*/2 > pr - \pi_w(gr) - pr + \pi_w(pr) \tag{10d}$$

$$m^*/2 > \pi_w(pr - gr) \tag{10e}$$

These two equations underline the relation between the proportion of weak students and the performance threshold set by teachers. Please notice that, within the range $(pr - gr) > m^* > 2(pr - gr)$ there are big enough $m^*$and small enough π_w such that strong students would also prefer the non-signalling equilibrium.

It is worth to notice that (10e) is equal to the ratio C_w/C_s. Hence, I can rewrite equation (10e) to find for which values of π_w a signalling equilibrium is preferred to strong students:

$$\pi_w > C_w/C_s \tag{10f}$$

The interpretation would be that strong students benefit when they are a 'minority', that is when their number does not exceed the ratio of the marginal signalling costs of the two groups. The more numerous strong students are in the population; the more expensive the accumulation of scholastic performance must be for weak students. When this is not true, the two types would both prefer social promotional regime.

It is useful to spend a few words to stress this last point. For the system of incentive produced by the grade repetition threat to be effective, the actors in the grade repetition game must believe that, on the one hand, strong students are fewer and, on the other, that a weak student should struggle to reach learning goals, that is, to successfully invest in scholastic performance. In the case one of the two believe drops, teachers would need a new system of incentives to exert effort from students and therefore to fulfil their teaching goal as teachers.

"Performance-discouraging" equilibrium

I describe the set of teachers' beliefs in this equilibrium as follows:

$$\begin{cases} P(w|m < m^*) = q \\ P(s|m < m^*) = 1 - q \\ P(s|m \geq m^*) = 1 \end{cases} \tag{11}$$

If teachers observe $m < m^*$, they believe it is a weak student with probability q and strong with probability $1 - q$; whereas when the observe $m \geq m^*$, they think the student is strong with probability 1. As usual, there are only two rational alternatives for students: invest such as $m = m^*$, or not invest at all ($m = 0$).

In this scenario, teachers promote if they see $m \geq m^*$. On the contrary, when they observe $m < m^*$, they adopt a mixed strategy and retain with probability q and promoted

with probability $1 - q$. For students, the payoff for scholastic performance investment is the gain for promotion minus each type's signalling costs. When they do not invest ($m = 0$) students get:

$$q(gr) + (1 - q)pr \tag{12}$$

Note that the expected value of $q(gr) + (1 - q)pr$ in the performance discouraging scenario is greater than the payoff for grade repetition gr in the separating scenario. The interpretation should be that, since the threshold is high, the perceived cost of failure is lower. When $m^* > 2q(pr - gr)$, strong students prefer not to invest, and both groups end up retained.

Proof (for the strong student):

$$q(gr) + (1 - q)pr > pr - m^*/2 \tag{12a}$$

$$m^*/2 > pr - q(gr) - (1 - q)pr \tag{12b}$$

$$m^*/2 > -q(gr) + q(pr) \tag{12c}$$

Performance-enhancing" equilibrium

I describe the set of teachers' beliefs in this equilibrium as follows:

$$\begin{cases} P(w|m < m^*) = 1 \\ P(w|m \geq m^*) = q \\ P(s|m \geq m^*) = 1 - q \end{cases} \tag{13}$$

If teachers observe $m < m^*$, they believe it is a weak student with probability 1; whereas when the observe $m \geq m^*$, they think the student is weak with probability q and strong with probability $1 - q$. This time, teachers opt for a mixed strategy when they observe $m \geq m^*$: they retain with probability q and promote with probability $1 - q$. As usual, there are only two rational alternatives for students: invest such as $m = m^*$, or not

invest at all ($m = 0$). For both students, the resulting payoff from non-investment is *gr*. The payoff when $m = m^*$ the following:

$$q(gr) + (1 - q)pr \tag{14}$$

Note that, inversely to the performance discouraging scenario, the expected value of promotion $q(gr) + (1 - q)pr$ here is smaller than the payoff for promotion pr in the separating scenario. The share of weak students spoils the promotion payoffs. The higher the share, the lower its value. The lower the threshold, the lower the cost of the investment. Such fact implies that when $m^* < (1 - q)(pr - gr)$, both individuals rationally opt for the investment.

Proof (for the weak student):

$$q(gr) + (1 - q)pr - m^* > gr \tag{14a}$$

$$m^* < q(gr) + (1 - q)pr - gr \tag{14b}$$

$$m^* < q(gr - pr) + pr - gr \tag{14c}$$

Relaxing the assumptions

I am interested now in two extensions of the Grade repetition Game. The first one affects the types of students. Let us drop the simplistic strong-weak partition in favour of a continuous index for latent ability, so that:

$$f(a) = N(\bar{a}, \sigma^2) \tag{15}$$

While $f(a)$ is common knowledge, teachers do not observe the individual ability a_i. Additionally, teachers hold beliefs about the distribution of ability given performance, $p(m_i, a_i)$. I assume *m* and *a* to be positively correlated. Teachers can expect a specific amount of performance m^* to be the profitable threshold a^* for any mixed set of beliefs, such that:

$$if\ m_i < m^*: \begin{cases} P(a_i < a^*) = q \\ P(a_i \geq a^*) = 1 - q \end{cases} \quad (16)$$

$$if\ m_i \geq m^*: \begin{cases} P(a_i < a^*) = 1 - z \\ P(a_i \geq a^*) = z \end{cases} \quad (16a)$$

With $q, z \in 0.5{,}1$. (This last condition is to respect the negative correlation between true ability and signalling costs, that is, teachers generally should not expect higher performance by weaker students).

A further step is to allow investments in performance (signalling) to be productive, that is, to influence latent ability. To allow investments to be productive means to reckon the value of the teaching action (the teaching goal I described above). In this respect, the results of the grade repetition game hold as far as signalling costs negatively correlate with performance (*i.e.* performance positively correlate with latent ability). If investments are (believed as) too productive, and the efficacy of investments as (believed as) independent to latent ability, performance carries no signalling power for teachers.[3]

Therefore, for any ability distribution $f(a) = N(\bar{a}, \sigma^2)$, I can observe the three kind of equilibria described above:

When $q = 1$ and $z = 1$, I have a separating equilibrium if:

$$zc_{a_i<a^*}(pr - gr) > m^* > qc_{a_i\geq a^*}(pr - gr) \quad (17)$$

Where $c_{a_i<a^*}$ is the signalling cost for the first individual under a^*, and $c_{a_i\geq a^*}$ is the signalling cost for the first individual above a^*.

When $q < 1$, for any

$$m^* > qc_{a_i\geq a^*}(pr - gr) \quad (17a)$$

[3] Interestingly, from this derives that, in a scholastic system that select over performance, latent ability must be considered as not infinitely malleable, and performance as strictly linked to ability.

students prefer not to invest, and I have a performance discouraging equilibrium.

When $z < 1$, for any

$$m^* < zc_{a_i<a^*}(pr - gr) \tag{17b}$$

students prefer to invest, and I have a performance-enhancing equilibrium.

The impact of indexes

I shall follow Spence (1974) to show how unmodifiable characters can affect opportunity sets, even if preferences and signalling costs are the same. Let assume students can belong to two different social backgrounds:

$$SB = \{h; l\} \tag{18}$$

Where h is high and l is low. (Notably, no substantial changes in the reasoning if I use native vs migrant, or a mix of more than one index).

Although it is not the focus of this paper, the grade repetition game potentially accounts for economic constraints in investments and primary effects of social class (Boudon 1974 *sensu*). At the aggregate level, I would observe fewer retained high social class kids if I allow the two classes varying for costs and ability distribution, which is very likely to be right in the real world.

Nevertheless, I am interested here in showing how variables that relate neither to costs nor to ability can affect the probability to get retained. Hence, I would consider ability as equally distributed within the two groups, such that:

$$f(a_h) = f(a_l) \tag{19}$$

Teachers now observe both performance (as a signal) and social background (as an index). Even when the beliefs are the same for high and low, the social background is distinguishable by teachers, and hence they estimate m^* independently for each group. For example, I have a separating equilibrium if:

$$zc_{a_i<a^*}(pr-gr) < m_h^* < qc_{a_i\geq a^*}(pr-gr) \quad (19a)$$

$$zc_{a_i<a^*}(pr-gr) < m_l^* < qc_{a_i\geq a^*}(pr-gr) \quad (19b)$$

Where m_h^* and m_l^* is the threshold for high and low social background respectively. Notably, within this range, it is possible to imagine:

$$zc_{a_i<a^*}(pr-gr) < m_h^* < m_l^* < qc_{a_i\geq a^*}(pr-gr) \quad (19c)$$

The fact that the two estimations are independent implies that for two kids with the same ability a_i, teachers can pretend different performance within the range of the infinite separating equilibria. Why would teachers differentiate the thresholds? Let us imagine two low performing kids at the end of the year. Teachers now need to screen them; that is, they must decide whether they are ready to pass the grade. They might ask: which is the likelihood that the student will catch-up independently from the repetition? If teachers expect parental support to vary among social classes, they are likely to believe that high SES kids can gain for promotion even when they hold a lower performance, when compared to their low SES peers. High SES kids- teachers might think- can benefit from a broader set of home resources that will allow them to recover and come back on tracks the following year. On the contrary, low SES kids need to repeat the year in order to build a more solid base upon which pile further knowledge and skills. I shall refer to this as the parental support expectation effect. Given two equally performing kids, one can repeat the year because he is expected to receive lower support at home.

Let us make a further step now. What happens when the distribution of teachers' beliefs changes between the two groups? For any ability distribution $f(a_h) = f(a_l)$, I can describe a different set of beliefs.

For example, teachers may hold performance discouraging beliefs only when they evaluate low social background kids:

$$zc_{a_i<a^*}(pr-gr) < m_h^* < qc_{a_i\geq a^*}(pr-gr) < m_l^* \quad (19d)$$

In this picture, high social background pupils benefit from a separating equilibrium, whereas it is rational for low social background kids not to invest and repeat the year. A second potential scenario is when teachers are performance-enhancing only towards high social background kids:

$$m_h^* < zc_{a_i<a^*}(pr-gr) < m_l^* < qc_{a_i\geq a^*}(pr-gr) \qquad (19e)$$

that implies that low social background kids would face a separating equilibrium, whereas high social background pupils would benefit from a performance-encouraging environment and pass the grade at the end of the year.

Eventually, I can also imagine a third scenario, which is the combination of the two above and represent the highest possible level of teachers' bias:

$$m_h^* < zc_{a_i<a^*}(pr-gr) < qc_{a_i\geq a^*}(pr-gr) \; < m_l^* \qquad (19f)$$

Different scenarios can emerge if teachers think that the two groups would not invest in the same way on their performance, given ability distribution and cost of signalling. Unbalanced beliefs can emerge from observed past under investments, or from biased perception. At any rate, once achieved, they might persist for reasons endogenous to the informational structure of the evaluation setting. With Spence, I shall refer to these scenarios as to lower equilibrium traps.

I can build a typology of environment based upon teachers' set of beliefs. The table below compares the environment for the usual two social groups (low and high), assuming the same ability distribution between the groups.

		High SES		
		Enhancing	Separating	Discouraging
Low SES	Enhancing	$m_h^* \leq m_l^* < zc_{a_i<a^*}(pr-gr) < qc_{a_i\geq a^*}(pr-gr)$		
	Separating	$m_h^* < zc_{a_i<a^*}(pr-gr) < m_l^* < qc_{a_i\geq a^*}(pr-gr)$	$zc_{a_i<a^*}(pr-gr) < m_h^* \leq m_l^* < qc_{a_i\geq a^*}(pr-gr)$	
	Discouraging	$m_h^* < zc_{a_i<a^*}(pr-gr) < qc_{a_i\geq a^*}(pr-gr) \; < m_l^*$	$zc_{a_i<a^*}(pr-gr) < m_h^* < qc_{a_i\geq a^*}(pr-gr) < m_l^*$	$zc_{a_i<a^*}(pr-gr) < qc_{a_i\geq a^*}(pr-gr) < m_h^* \leq m_l^*$

In three out of six possible intersections, high social background kids prefer to invest, and they all pass the grade; while it is the case for the low social background only in 1 out of six scenarios. *Vice versa*, half of the times the beliefs of the teachers discourage low SES kids in investing. They will hence prefer grade repetition.

In the cases presented above, the expectations of teachers have a direct effect on performance investments on the side of students. When students face lower expectations as a group, they will comply with the teachers' beliefs. In this way, teachers' beliefs self-confirm through the action of students, that is: students do not invest if they are low SES, and do not invest if they are high SES. When indexes turn into variation in beliefs, the incentive of grade repetition and promotion can fail to assure both the teaching and the screening goal.

Wrapping up

- The scholastic evaluation characterises by imperfect information on the side of teachers that do not observe directly ability
- Students invest in performance so to show their ability. Performance is believed to be positively correlated to ability and negatively correlated to signalling costs
- Teachers hold (self-confirming) beliefs about the distribution of ability given performance
- Some groups might experience a lower equilibrium trap that can persist over time for it continuously confirms teachers' beliefs
- The informational structure of the evaluation can foster discrimination by differentiating thresholds for specific groups of students (involved mechanism: EXPECTATION ON PARENTAL SUPPORT)
- The informational structure of the evaluation can foster discrimination by enhancing or discouraging performance for specific groups (involved mechanism: FAILURE AS MOTIVATIONAL SCHEME)

Expansion 1: repeated interaction (or why performance investment can be different from both zero and m^*)

Critique: The grade repetition game considers rational to invest at the equilibrium level m^*, or not to invest at all. In the real world, I instead see students much more scattered along with the school performance (imaginary) line.

By construction, the model predicts only two courses of actions on the side of students: to invest, or not to invest. I label as non-rational students who invest either above the threshold, or between zero and the threshold.

The first case (that of over investments) raises fewer concerns. When performance investments are productive, they acquire an intrinsic value that outruns that of the mere promotion to the next grade. Especially for high ability kids, or for those who have more permissive economic constraints, to invest in performance can foster their ability.

I can expect the intrinsic value of performance investment to induce all students to end up with $m^* \neq 0$. Nevertheless, this is not enough to explain real-world observations. Those who 'make an effort' are likely to be considered worthy of passing the grade, for they signal a positive attitude toward learning targets. Psycho-sociological literature shows that teachers are more likely to punish a lack of effort\investment, rather than a missed result.

To convincingly address this issue, I maintain that one can profitably break down the grade repetition game into repeated interactions between teachers and students. The first interaction is 'blind' on the two sides. I mean that the student, having learned his type, chooses how much invest in performance without being able to anticipate the required m^* (or having a very approximate idea of it). After his choice, teachers approve or turndown his level of investments, for example by awarding a good or bad grade, or by giving them other kinds of feedbacks. At a certain point, he can realise the effort does not worth the price, and he stops to invest. The result would be a performance $m^* \neq 0$ but below the threshold, and the last response by teachers being grade repetition.

Expansion 2: school choice (or why grade repetition rate is not zero?)

Critique: In educational systems that leave school choices to families, if students have complete information over their ability, their signalling costs and their payoffs are given teachers' strategy, why they should enrol in schools where they are very likely to end up repeating the year?

Let us assume students can choose freely high schools. Different thresholds and sets of beliefs characterise schools. Following the model, students are aware of their latent ability and, given the threshold and the set of beliefs, they can anticipate costs and payoffs of any course of action in each school. The expected aggregate outcome would be a grade repetition rate equal to zero because students would enrol in a school where they are sure to avoid promotion. Given any threshold m^*, I would find all weak students in enhancing the environment, and all strong students in separating equilibria. No one would enrol in a discouraging environment. However, it is not the case in the real word: many students enrol in schools and get retained.

From above, I know one can make sense of the unrealistic predictions of the grade repetition game by allowing constraints to varying among groups. Students can experience constraints in their choice. These might be economic or time constraints. At any rate, likely, some are not able to opt for a school in which their level of ability can assure a promotion.

Moreover, students can experience informational barriers. Students can have imperfect information over teachers' beliefs and strategy. Students can have a very approximate idea of the required m^*. If students have biased knowledge, they might incur into unanticipated scholastic failure. That is, they can enrol in an environment that reveals as discouraging.[4]

Eventually, the model with more than one school should acknowledge the utility of the 'downward' option. In other words, the expected benefits from the choice of the school (no matter what outcome) might be higher than the promotion in another with a lower threshold. In other words, students might prefer to take the risk. I shall note that the

[4] Informational biases over ability and signalling costs are likely to go in the direction of reinforcing inequalities. Abilities: low (high) SES kids are likely to under (over) estimate their ability and over (under) estimates their signalling costs.

grade repetition game goes against the relative risk aversion hypothesis. Given ability, each further step is more dangerous for low SES students. They need to show higher performance to gain their promotion, and, more often than their advantaged peers, they are discouraged by endogenous mechanisms. Rather than endogenous preferences, endogenous mechanisms of selection are here crucial.

Nevertheless, anticipation on the side of student likely has a considerable impact on schools. On the one hand, schools that discourage everyone from investing in performance are expected to disappear over time. To pull students in, schools must either decrease their threshold m^* or update their beliefs about the ability distribution given an observed performance. On the other, schools that promote too effortlessly establish a *de facto* social promotion regime. When signalling costs drop and everyone can reach the learning targets, scholastic performance loses its power as a signal for ability.

On the aggregate level, the expected distribution of ability and performance in the population of students matters jointly and can make the system of schools shift from more demanding to lower demanding thresholds. Consider the observed shifts in promotional standards due to the massification of schools (cf. chapter on Italy). Before, I observe a discouraging environment in which the risk of failure was considerable even for high performing students; the gain of promotion was high, while the cost of failure limited. After, I observe a progressive shift toward more separating or even enhancing environments for many students. The gain for promotion has narrowed down, whereas the cost of failure has turned as considerable. On the one hand, this led to lower educational levels (primary and middle schools) to shift toward a social promotional regime. On the other, at the aggregate level, the thresholds are lower, and the failure rate drops. More importantly, because the high ability students are (believed as) fewer, the cost of failure has considerably risen.

List of Figures

List of Tables

Bibliography

Abbiati, Giovanni, Gianluca Argentin, and Tiziano Gerosa. "Different teachers for different students? Evidence on teacher-student matching and its consequences in the Italian case." Politica economica 33.1 (2017): 12-58.

Adnett, Nick, and Peter Davies. 1999. "Schooling Quasi-Markets: Reconciling Economic and Sociological Analyses." *British Journal of Educational Studies* 47 (3): 221–234.

Agirdag, O., M. Van Houtte, and P. Van Avermaet. 2012. "Why Does the Ethnic and Socio-Economic Composition of Schools Influence Math Achievement? The Role of Sense of Futility and Futility Culture." *European Sociological Review* 28 (3): 366–78.

Akerlof, George A., and Rachel E. Kranton. 2002. "Identity and Schooling: Some Lessons for the Economics of Education." *Journal of Economic Literature* 40 (4): 1167–1201.

Albornoz, Facundo, Samuel Berlinski, and Antonio Cabrales. 2017. "Motivation, Resources, and the Organization of the School System." *Journal of the European Economic Association*, jvx001.

Alet, Élodie, Liliane Bonnal, and Pascal Favard. 2013. "Repetition: Medicine for a Short-Run Remission." *Annals of Economics and Statistics*, no. 111/112: 227.

Alexander, Karl L., Doris R. Entiwisle, and Susan L. Dauber. n.d. "Grade Retention: Lingering Questions." In *On the Success of Failure. A Reassesment of the Effects of Retention in the Primary Grades.* Accessed January 21, 2017.

Allen, C. S., Q. Chen, V. L. Willson, and J. N. Hughes. 2009. "Quality of Research Design Moderates Effects of Grade Retention on Achievement: A Meta-Analytic, Multilevel Analysis." *Educational Evaluation and Policy Analysis* 31 (4): 480–99.

Allot, Nicholas. 2002. "Relevance and Rationality."

Ananat, Elizabeth Oltmans, et al. Children left behind: The effects of statewide job loss on student achievement. No. w17104. National Bureau of Economic Research, 2011.

Anderson, Lorin W. 2018. "A Critique of Grading: Policies, Practices, and Technical Matters." *Education Policy Analysis Archives* 26 (April): 49.

Apple, Michael W. 1988. "Facing the Complexity of Power: For a Parallelist Position in Critical Educational Studies." *Bowles and Gintis Revisited. Correspondance and Contradiction in Educational Theory*, 112–130.

Arrow, Kenneth J. "Models of job discrimination." Racial discrimination in economic life 83 (1972).

Aschaffenburg, Karen, and Ineke Maas. "Cultural and educational careers: The dynamics of social reproduction." American sociological review (1997): 573-587.

Auger, Julie. 2003. “Le Redoublement Des Sujets.” *Journal of French Language Studies* 13 (3): 381–404.

Autin, Frédérique, Anatolia Batruch, and Fabrizio Butera. n.d. “The Function of Selection of Assessment Leads Evaluators to Artificially Create the Social Class Achievement Gap,” 20.

Ayalon, Hanna, and Idit Livneh. 2013. “Educational Standardization and Gender Differences in Mathematics Achievement: A Comparative Study.” *Social Science Research* 42 (2): 432–445.

Azzolini, Davide, and Carlo Barone. 2013. “Do They Progress or Do They Lag behind? Educational Attainment of Immigrants’ Children in Italy: The Role Played by Generational Status, Country of Origin and Social Class.” *Research in Social Stratification and Mobility* 31 (March): 82–96.

Azzolini, Davide, Philipp Schnell, and John R. B. Palmer. 2012. “Educational Achievement Gaps between Immigrant and Native Students in Two ‘New’ Immigration Countries: Italy and Spain in Comparison.” Edited by Alícia Adserà and Marta Tienda. *The ANNALS of the American Academy of Political and Social Science* 643 (1): 46–77.

Baert, Stijn, Bart Cockx, and Matteo Picchio. 2015. “Modeling the Effects of Grade Retention in High School.”

Baker, Dean, and Mark Weisbrot. 1994. “The Logic of Contested Exchange.” *Journal of Economic Issues* 28 (4): 1091–1114.

Baker, Eva L. 2016. “Research to Controversy in 10 Decades.” *Educational Researcher* 45 (2): 122–133.

Bakshi, S., and Zhiwu Chen. n.d. “American Economic Association Is Collaborating with JSTOR to Digitize, Preserve and Extend Access to The American Economic Review.” Accessed March 23, 2017.

Baldino, R., and T. Cabral. 2013. “The Productivity of Students’ Schoolwork: An Exercise on Marxist Rigour.” *Journal for Critical Education Policy Studies* 11 (4): 1–15.

Bali, Valentina A., Dorothea Anagnostopoulos, and Reginald Roberts. 2005. “Toward a Political Explanation of Grade Retention.” *Educational Evaluation and Policy Analysis* 27 (2): 133–155.

Ball, Richard A. 1979. “The Dialectical Method: Its Application to Social Theory.” *Social Forces* 57 (3): 785–798.

Ball, Stephen J. 1993. “Education Markets, Choice and Social Class: The Market as a Class Strategy in the UK and the USA.” *British Journal of Sociology of Education* 14 (1): 3–19.

Ballarino, Gabriele. 2013. “Istruzione, Formazione Professionale, Transizione Scuola-Lavoro. Il Caso Italiano in Prospettiva Comparata.” *Rapporto Di Ricerca IRPET*.

Ballarino, Gabriele, Carlo Barone, and Nazareno Panichella. 2014. “Social Background and Education in Occupational Attainment in 20th Century Italy.”

Bandura, Albert. "Self-efficacy." The Corsini encyclopedia of psychology (2010): 1-3.

Barban, Nicola, and Michael J. White. 2011. “Immigrants’ Children’s Transition to Secondary School in Italy.” *International Migration Review* 45 (3): 702–26.

Barone, C., Ruud Luijkx, and A. Schizzerotto. 2008. “Elogio Dei Grandi Numeri.” *Il Lento Declino Delle Disuguaglianze Scolastiche in Italia*, 21–44.

Barone, Carlo. 2005a. “È Possibile Spiegare Le Disuguaglianze Di Apprendimento Mediante La Teoria Del Capitale Culturale?” *Polis* 19 (2): 173–202.

———. 2005b. “La Teoria Della Scelta Razionale e La Ricerca Empirica. Il Caso Delle Disuguaglianze Educative.” *Rassegna Italiana Di Sociologia* 46 (3): 411–446.

———. "Cultural capital, ambition and the explanation of inequalities in learning outcomes: A comparative analysis." Sociology 40.6 (2006): 1039-1058.

Barone, K Barg, M Ichou "Relative risk aversion models: How plausible are their assumptions?", OSC Papers, 2019

Barone, Carlo, and Lucia Ruggera. 2015. “Social Background and Educational Attainment. Italy in Comparative Perspective.” *Scuola Democratica* 6 (2): 321–342.

Barone, Carlo, and Herman G van de Werfhorst. 2011. “Education, Cognitive Skills and Earnings in Comparative Perspective.” *International Sociology* 26 (4): 483–502.

Bartus, Tamás. n.d. “Marginal Effects and Extending the Blinder-Oaxaca.” Accessed March 17, 2017.

Bassok, Daphna, Jenna E. Finch, RaeHyuck Lee, Sean F. Reardon, and Jane Waldfogel. 2016. "Socioeconomic Gaps in Early Childhood Experiences: 1998 to 2010." *AERA Open* 2 (3): 233285841665392.

Battistin, Erich, and Antonio Schizzerotto. 2018. "Threat of Grade Retention, Remedial Education and Student Achievement: Evidence from Upper Secondary Schools in Italy." *Empirical Economics*, May.

Bautier, Élisabeth, Bernard Charlot, and Jean-Yves Rochex. "Entre apprentissages et métier d'élève: le rapport au savoir." L'école, l'état des savoirs (2000): 179-188.

Becker, Dominik. 2012. "The Impact of Teachers' Expectations on Students' Educational Opportunities in the Life Course." *Pygmalion's Long Shadow–Determinants and Outcomes of Teachers' Evaluations*.

Becker, Gary S. "Human capital theory." Columbia, New York 1964 (1964).

———. "The economics of discrimination." (1957): 1976.

Belot, Michèle, and Vincent Vandenberghe. 2014. "Evaluating the 'Threat' Effects of Grade Repetition: Exploiting the 2001 Reform by the French-Speaking Community of Belgium." *Education Economics* 22 (1): 73–89.

Berchialla, Paola, Franco Cavallo, Sabina Colombini, Patrizia Lemma, Gianfranco De Simone, and Fondazione Giovanni Agnelli. 2011. *Associazione Delle Caratteristiche Individuali, Del Benessere Psico-Fisico e Del Clima Di Classe Con Gli Outcomes Scolastici a 11 e 13 Anni*. Fondazione «Giovanni Agnelli», Working Paper.

Bernardi, Fabrizio. 2014. "Compensatory Advantage as a Mechanism of Educational Inequality: A Regression Discontinuity Based on Month of Birth." *Sociology of Education* 87 (2): 74–88.

Björklund, Anders, and Kjell G. Salvanes. 2011. "Education and Family Background: Mechanisms and Policies." In *Handbook of the Economics of Education*, 3:201–247. Elsevier.

Black, Dan A., Jeffey A. Smith, Berger, Mark C, and Brett J Noel. 2003. "Is the Threat of Reemployment Services More Effective Than the Services Themselves." *The American Economic Review* 93 (4).

Blanchard, Marianne, and Joanie Cayouette-Remblière. Sociologie de l'école. La Découverte, 2016.

Bleyaert, Barbara. 2009. "The Effects of Retention on Drop-Out and Graduation Rates. Research Brief." *Education Partnerships, Inc.*

Blundell, Richard, Lorraine Dearden, and Barbara Sianesi. 2001. “Estimating the Returns to Education: Models, Methods and Results.”

Bodovski, Katerina. 2010. “Parental Practices and Educational Achievement: Social Class, Race, and *Habitus*.” *British Journal of Sociology of Education* 31 (2): 139–56.

Bodovski, Katerina, and George Farkas. 2008. “‘Concerted Cultivation’ and Unequal Achievement in Elementary School.” *Social Science Research* 37 (3): 903–19.

Boero, Gianna, Tiziana Laurenti, and Robin Naylor. 2005. “An Econometric Analysis of Student Withdrawal and Progression in Post-Reform Italian Universities.”

Bol, T., J. Witschge, H. G. Van de Werfhorst, and J. Dronkers. 2014. “Curricular Tracking and Central Examinations: Counterbalancing the Impact of Social Background on Student Achievement in 36 Countries.” *Social Forces* 92 (4): 1545–72.

Bol, Thijs. 2015. “Has Education Become More Positional? Educational Expansion and Labour Market Outcomes, 1985–2007.” *Acta Sociologica* 58 (2): 105–20.

Bol, Thijs, and Herman G. van de Werfhorst. 2013. “Educational Systems and the Trade-Off between Labor Market Allocation and Equality of Educational Opportunity.” *Comparative Education Review* 57 (2): 285–308.

Bonvin, P., G. Bless, and M. Schuepbach. 2008. “Grade Retention: Decision-Making and Effects on Learning as Well as Social and Emotional Development.” *School Effectiveness and School Improvement* 19 (1): 1–19.

Borodankova, Olga, and Ana Sofia Almeida Coutinho. Grade Retention during Compulsory Education in Europe: Regulations and Statistics. Education, Audiovisual and Culture Executive Agency, European Commission. Available from EU Bookshop, 2011.

Borooah, Vani K., and Sriya Iyer. 2005. “The Decomposition of Inter-Group Differences in a Logit Model: Extending the Oaxaca-Blinder Approach with an Application to School Enrolment in India.” *Journal of Economic and Social Measurement* 30 (4): 279–293.

Boudon, Raymond. 1997. “What Theory of Rationality for the Social Sciences: The ‘Cognitivist Model.’” *Társadalom És Gazdaság Közép-És Kelet-Európában/Society and Economy in Central and Eastern Europe*, 75–105.

Boudon, Raymond. L'inégalité des chances: la mobilité sociale dans les sociétés industrielles. Armand Colin, 1973.

Bourdieu, Pierre. 2002. “The Forms of Capital.” In *Readings in Economic Sociology*, edited by Nicole Woolsey Biggart, 280–91. Oxford, UK: Blackwell Publishers Ltd.

———. 2011. “The Forms of Capital.(1986).” *Cultural Theory: An Anthology*, 81–93.

Bowles, Samuel, and Herbert Gintis. Schooling in capitalist America: Educational reform of economic life. Routledge & Kegan Paul, 1976.

Bowles, S., and H. Gintis. 2000. "Walrasian Economics in Retrospect." *The Quarterly Journal of Economics* 115 (4): 1411–39.

Bowles, Samuel. 1972. "Schooling and Inequality from Generation to Generation." *Journal of Political Economy* 80 (3, Part 2): S219–S251.

———. 2001. "Individual Interactions, Group Conflicts, and the Evolution of Preferences." *Social Dynamics* 155: 190.

Bowles, Samuel, and Herbert Gintis. 2000. "Does Schooling Raise Earnings by Making People Smarter." *Meritocracy and Economic Inequality,* 118–136.

———. 2002. "The Inheritance of Inequality." *Journal of Economic Perspectives* 16 (3): 3–30.

Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001a. "Incentive-Enhancing Preferences: Personality, Behavior, and Earnings." *The American Economic Review* 91 (2): 155–158.

———. 2001b. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature* 39 (4): 1137–1176.

Bowman, Lisa J. 2005. "Grade Retention: Is It a Help or Hindrance to Student Academic Success?" *Preventing School Failure: Alternative Education for Children and Youth* 49 (3): 42–46.

Breen, Richard, and John H. Goldthorpe. 1997. "Explaining Educational Differentials towards a Formal Rational Action Theory." *Rationality and Society* 9 (3): 275–305.

———. 1999. "Class Inequality and Meritocracy: A Critique of Saunders and an Alternative Analysis1." *The British Journal of Sociology* 50 (1): 1–27.

Brincks, Ahnalee. n.d. "The Implications of Centering in a Three-Level Multilevel Model," 142.

Bropy, Jere, International Institute for Educational Planning, and International Academy of Education. 2006. *Grade Repetition*. Paris; Brussels: International Institute for Educational Planning; International Academy of Education.

Burawoy, Michael. 2012. "The Roots of Domination: Beyond Bourdieu and Gramsci." *Sociology* 46 (2): 187–206.

Burawoy, Michael, and Erik Olin Wright. n.d. "Sociological Marxism in Jonathan Turner (Ed), Handbook of Sociological Theory (Plenum: Forthcoming)."

Cadigan, Doris, Doris R. Entwisle, Karl L. Alexander, and Aaron M. Pallas. 1988. “First-Grade Retention among Low Achieving Students: A Search for Significant Predictors.” *Merrill-Palmer Quarterly (1982-)*, 71–88.

Cameron, Stephen V., and James J. Heckman. 1998. “Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males.” *Journal of Political Economy* 106 (2): 262–333.

Cappellari, Lorenzo. 2004. “High School Types, Academic Performance and Early Labour Market Outcomes.”

Cappellari, Lorenzo, and Giorgio Brunello. 2007. “Qualità e Rendimento Economico Dell’istruzione Universitaria Italiana1.” *Focus* 34 (3).

Cardano, Mario. n.d. “Il Sociologo e Le Sue Muse.” *Rassegna Italiana Di Sociologia* 32 (2): 181–223.

Carlana, Michela. n.d. “Stereotypes and Self-Stereotypes: Evidence from Teachers’ Gender Bias,” 48.

Carlana, Michela, Eliana La Ferrara, and Paolo Pinotti. n.d. “Goals and Gaps: Educational Careers of Immigrant Children,” 52.

Carneiro, Pedro, and James J. Heckman. 2002. “The Evidence on Credit Constraints in Post-Secondary Schooling.” *The Economic Journal* 112 (482): 705–734.

Carneiro, Pedro Manuel, and James J. Heckman. 2003. “Human Capital Policy.”

Catsambis, Sophia. 1998. “Expanding the Knowledge of Parental Involvement in Secondary Education: Effects on High School Academic Success. Report No. 27.”

Caucutt, Elizabeth, and Lance Lochner. 2012. “Early and Late Human Capital Investments, Borrowing Constraints, and the Family.” w18493. Cambridge, MA: National Bureau of Economic Research.

Cavalli-Sforza, Luigi L., and Marcus W. Feldman. "Cultural versus biological inheritance: phenotypic transmission from parents to children.(A theory of the effect of parental phenotypes on children's phenotypes)." American journal of human genetics 25.6 (1973): 618.

Charles, Maria. 2003. “Deciphering Sex Segregation Vertical and Horizontal Inequalities in Ten National Labor Markets.” *Acta Sociologica* 46 (4): 267–287.

Charles, Maria, and Karen Bradley. 2009. “Indulging Our Gendered Selves? Sex Segregation by Field of Study in 44 Countries1.” *American Journal of Sociology* 114 (4): 924–976.

Checchi, Daniele. 2008. “Il Passaggio Dalla Scuola Media Alla Scuola Superiore.” *INVALSI Report, University of Milan*.

———. 2010. “Uguaglianza Delle Opportunità Nella Scuola Secondaria Italiana.” FGA Working Paper.

Checchi, Daniele, and Valentino Dardanoni. 2003. “Mobilità Sociale.” *Enciclopedia Treccani*.

Checchi, Daniele, and Luca Flabbi. 2006. “Mobilità Intergenerazionale e Decisioni Scolastiche in Italia.” *Scelte Individuali e Vincoli Strutturali. Sistema Scolastico e Disuguaglianza Sociale, Bologna, Il Mulino*.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. “Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates.” *American Economic Review* 104 (9): 2593–2632.

Choi, Soobin, and Yun-Kyung Cha. 2019. “Integration Policy in Education and Immigrant Students’ Patriotic Pride in Host Countries: A Cross-National Analysis of 24 European Countries.” *International Journal of Inclusive Education*, February, 1–15.

Cloninger, C Robert, John Rice, and Theodore Reich. n.d. “Multifactorial Inheritance with Cultural Transmission and Assortative Mating II. A General Model of Combined Polygenic and Cultural Inheritance,” 23.

Cockx, Bart Leo Wim, and Dejemeppe, Muriel. n.d. “Is the Notification of Monitoring a Threat to the Unemployed?: A Regression Discontinuity Approach.” *CESifo Working Paper, No. 2042*. Accessed April 27, 2018.

Cohen, Gerald A. 1979. “The Labor Theory of Value and the Concept of Exploitation.” *Philosophy & Public Affairs*, 338–360.

———. 1982. “Reply to Elster on Marxism, Functionalism and Game Theory.” *Theory and Society* 11 (4): 483–495.

Cohen, Robert. 2016. “You Are Not a Marxist on Your Own.” *Socialism and Democracy*, 1–6.

Coleman, James S. "Social capital, human capital, and investment in youth." Youth unemployment and society 34 (1994).

Collins, Courtney A., and Li Gan. 2013. “Does Sorting Students Improve Scores? An Analysis of Class Composition.” National Bureau of Economic Research.

Collins, Randall. 1971. “Functional and Conflict Theories of Educational Stratification.” *American Sociological Review* 36 (6): 1002.

Cosnefroy, Olivier, and Thierry Rocher. "Le redoublement au cours de la scolarisé obligatoire: nouvelles analyses, mêmes constats." Éducation et formations 70 (2004): 73-82.

Covay, Elizabeth, and William Carbonaro. 2010. “After the Bell: Participation in Extracurricular Activities, Classroom Behavior, and Academic Achievement.” *Sociology of Education* 83 (1): 20–45.

Covington, Martin V. 2000. “Goal Theory, Motivation, and School Achievement: An Integrative Review.” *Annual Review of Psychology* 51 (1): 171–200.

Crahay, Marcel. 2004. “Peut-on conclure à propos des effets du redoublement ?” *Revue française de pédagogie* 148 (1): 11–23.

Crosby, Ned, and John C. Hottinger. 2011. “The Citizens Jury Process.” *The Book of the States*, 321–325.

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach. "Estimating the technology of cognitive and noncognitive skill formation." Econometrica 78.3 (2010): 883-931.

Darling-Hammond, Linda, and Elle Rustique-forrester. 2005. “The Consequences of Student Testing for Teaching and Teacher Quality.” *Yearbook of the National Society for the Study of Education* 104 (2): 289–319.

Darmon, Muriel. "Sélectionner, élire, prédire: le recrutement des classes préparatoires." Sociétés contemporaines 2 (2012): 5-29.

Davies, Richard, Eskil Heinesen, and Anders Holm. 2002. “The Relative Risk Aversion Hypothesis of Educational Choice.” *Journal of Population Economics* 15 (4): 683–713.

Deary, Ian J., Steve Strand, Pauline Smith, and Cres Fernandes. 2007. “Intelligence and Educational Achievement.” *Intelligence* 35 (1): 13–21.

Dehejia, Rajeev H, and Sadek Wahba. n.d. “Propensity Score-Matching Methods for Nonexperimental Causal Studies.” *THE REVIEW OF ECONOMICS AND STATISTICS*, 94.

Deole, Sumit S. 2013. “Economic Analysis of Grade Promotion and Grade Retention Policies.” *Browser Download This Paper*.

Di Pietro, Giorgio, and Peter Urwin. 2006. “Education and Skills Mismatch in the Italian Graduate Labour Market.” *Applied Economics* 38 (1): 79–93.

DiPrete, Thomas A., and Gregory M. Eirich. 2006. “Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments.” *Annual Review of Sociology* 32 (1): 271–97.

Ditton, Hartmut, Jan Krüsken, and Magdalena Schauenberg. "Bildungsungleichheit—der Beitrag von Familie und Schule." Zeitschrift für Erziehungswissenschaft 8.2 (2005): 285-304.

Dompnier, Benoît, Pascal Pansu, and Pascal Bressoux. 2007. “Social Utility, Social Desirability and Scholastic Judgments: Toward a Personological Model of Academic Evaluation.” *European Journal of Psychology of Education* 22 (3): 333–350.

Doucet, Andrea. 2000. “‘There’s a Huge Gulf between Me as a Male Carer and Women’’: Gender, Domestic Responsibility, and the Community as an Institutional Arena.’” *Community, Work & Family* 3 (2): 163–184.

Draelants, Hugues. 2006. “Le Redoublement Est Moins Un Problème Qu’une Solution. Comprendre l’attachement Social Au Redoublement En Belgique Francophone.”

Duncan, Greg J. 2008. “When to Promote, and When to Avoid, a Population Perspective.” *Demography* 45: 763–784.

Dunifon, Rachel, and Greg J. Duncan. 1998. “Long-Run Effects of Motivation on Labor-Market Success.” *Social Psychology Quarterly* 61 (1): 33.

Dunkake, Imke, and Claudia Schuchart. "Stereotypes and teacher characteristics as an explanation for the class-specific disciplinary practices of pre-service teachers." Teaching and Teacher Education 50 (2015): 56-69.

Duru-Bellat, M. "Les inégalités scolaires." Genèse et mythes. Paris: PUF, Education et formation (2002).

Dutta, Shantanu, Mark Bergen, Jan B. Heide, and George John. 1995. “Understanding Dual Distribution: The Case of Reps and House Accounts.” *Journal of Law, Economics, & Organization*, 189–204.

Eboatu. 2017. “Comparative Study of the Impact of Class Repetition and Mass Promotion on Students’ Academic Achievement in Anambra State.” European Scientific Journal 13 (28).

Edling, Christofer. 2000. “Rational Choice Theory and Quantitative Analysis. A Comment on Goldthorpe’s Sociological Alliance.” *European Sociological Review* 16 (1): 1–8.

Egan, Owen, and Peter Archer. "The accuracy of teachers’ ratings of ability: A regression model." American Educational Research Journal 22.1 (1985): 25-34.

Ehmke, Timo, Barbara Drechsel, and Claus H. Carstensen. 2010. “Effects of Grade Retention on Achievement and Self-Concept in Science and Mathematics.” *Studies in Educational Evaluation* 36 (1–2): 27–35.

Eide, Eric R., and Mark H. Showalter. 2001. “The Effect of Grade Retention on Educational and Labor Market Outcomes.” *Economics of Education Review* 20 (6): 563–576.

Einhorn, Hillel J., and Robin M. Hogarth. 1985. “Ambiguity and Uncertainty in Probabilistic Inference.” *Psychological Review* 92 (4): 433.

Elster, Jon. 1982. “The Case for Methodological Individualism.” *Theory and Society* 11 (4): 453–482.

———. 1989. “Social Norms and Economic Theory.” *Journal of Economic Perspectives* 3 (4): 99–117.

———. 1991. “Rationality and Social Norms.” *European Journal of Sociology* 32 (01): 109–129.

Erikson, Robert, and John H. Goldthorpe. 2002. “Intergenerational Inequality: A Sociological Perspective.” *Journal of Economic Perspectives* 16 (3): 31–44.

Erikson, Robert, and Jan O. Jonsson, eds. Can education be equalized?: The Swedish case in comparative perspective. Westview Press, 1996.

Fairlie, Robert W. 2005. “An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models.” *Journal of Economic and Social Measurement* 30 (4): 305–316.

Farkas, George. 2003. “Cognitive Skills and Noncognitive Traits and Behaviors in Stratification Processes.” *Annual Review of Sociology* 29 (1): 541–62.

Farkas, George, and Kurt Beron. 2004. “The Detailed Age Trajectory of Oral Vocabulary Knowledge: Differences by Class and Race.” *Social Science Research* 33 (3): 464–97.

Farkas, George, Robert P. Grobe, Daniel Sheehan, and Yuan Shuan. 1990. “Cultural Resources and School Success: Gender, Ethnicity, and Poverty Groups within an Urban School District.” *American Sociological Review* 55 (1): 127.

Feinstein, Leon. 2003. “Inequality in the Early Cognitive Development of British Children in the 1970 Cohort.” *Economica* 70 (277): 73–97.

Feldman, Marcus W., Sarah P. Otto, and Freddy B. Christiansen. "Genes, culture, and inequality." Meritocracy and economic inequality (2000): 61-85.

Ferguson, Ronald F. 2003. “Teachers’ Perceptions and Expectations and the Black-White Test Score Gap.” *Urban Education* 38 (4): 460–507.

Ferreira, Maria, Bart H H Golsteyn, and Sergio Parra-Cely. 2018. “The Effect of Grade Retention on Secondary School Performance: Evidence from a Natural Experiment,” 47.

Fertig, Michael. 2004. “Shot across the Bow, Stigma or Selection? The Effect of Repeating a Class on Educational Attainment.”

Flabbi, Luca. 2001. “Secondary Schooling Decisions in Italy.” *Rivista Di Politica Economica* 91 (7/8): 85–114.

Flacher, David, and Hugo Harari-Kermadec. 2013. “Tuition Fees, Self-Esteem and Social Heterogeneity.” *Education Economics* 21 (2): 191–210.

Flynn, Timothy M. 1989. “The Effects of Grade Retention on Middle School Students’ Academic Achievement, School Adjustment and School Attendance.” *Middle School Research Selected Studies* 14 (1): 1–6.

Foureaux Koppensteiner, Martin. 2014. “Automatic Grade Promotion and Student Performance: Evidence from Brazil.” *Journal of Development Economics* 107 (March): 277–90.

Frederick, Carl B., and Robert M. Hauser. 2008. “Have We Put an End to Social Promotion? Changes in School Progress among Children Aged 6 to 17 from 1972 to 2005.” *Demography* 45 (3): 719–740.

Freeman, Jennifer, and Brandi Simonsen. 2015. “Examining the Impact of Policy and Practice Interventions on High School Dropout and School Completion Rates A Systematic Review of the Literature.” *Review of Educational Research* 85 (2): 205–248.

Frouillou, Leïla, and Léonard Moulin. n.d. “Article En Révision Dans Population.” Accessed October 23, 2016.

Fullan, Michael, and Jan J. Loubser. 1972. “Education and Adaptive Capacity.” *Sociology of Education* 45 (3): 271.

Gabay-Egozi, L., Y. Shavit, and M. Yaish. 2010. “Curricular Choice: A Test of a Rational Choice Model of Education.” *European Sociological Review* 26 (4): 447–63.

Gabay-Egozi, Limor, Yossi Shavit, and Meir Yaish. 2015. “Gender Differences in Fields of Study: The Role of Significant Others and Rational Choice Motivations.” *European Sociological Review* 31 (3): 284–97.

Gambetta, Diego. Were they pushed or did they jump? Individual decision mechanisms in education. Cambridge University Press, 1987.

Gamoran, Adam. 1992. “Synthesis of Research: Is Ability Grouping Equitable?” *Educational Leadership* 50: 11–11.

Gamoran, Adam, and Robert D. Mare. 1989. “Secondary School Tracking and Educational Inequality: Compensation, Reinforcement, or Neutrality?” *American Journal of Sociology* 94 (5): 1146–83.

Garfinkel, Harold. 1956. “Conditions of Successful Degradation Ceremonies.” *American Journal of Sociology*, 420–424.

Gershenson, Seth, Stephen B Holt, and Nicholas W Papageorge. n.d. “Who Believes in Me? The Effect of Student-Teacher Demographic Match on Teacher Expectations,” 45.

Gershenson, Seth, and Laura Langbein. 2015. “The Effect of Primary School Size on Academic Achievement.” *Educational Evaluation and Policy Analysis* 37 (1_suppl): 135S-155S.

Geven, Sara, Anatolia Batruch, and Herman van de Werfhorst. "Inequality in Teacher Judgements, Expectations and Track Recommendations: A Review Study." (2018).

Gibb, Sheree J., David M. Fergusson, and L. John Horwood. 2008. “Gender Differences in Educational Achievement to Age 25.” *Australian Journal of Education* 52 (1): 63–80.

Gintis, Herbert. 1995. “The Political Economy of School Choice.” *Teachers College Record* 96: 492–492.

Glaeser, Edward L., and Denise DiPasquale. 1998. “Incentives and Social Capital: Are Homeowners Better Citizens?”.

Glock, Sabine, and Julia Karbach. "Preservice teachers’ implicit attitudes toward racial minority students: Evidence from three implicit measures." Studies in Educational Evaluation 45 (2015): 55-61.

Goldhaber, Dan, Lesley Lavery, and Roddy Theobald. 2015. “Uneven Playing Field? Assessing the Teacher Quality Gap between Advantaged and Disadvantaged Students.” *Educational Researcher* 44 (5): 293–307.

Goldschmidt, Pete, and Jia Wang. 1999. “When Can Schools Affect Dropout Behavior? A Longitudinal Multilevel Analysis.” *American Educational Research Journal* 36 (4): 715–738.

Goldthorpe, John H. 1996. “The Quantitative Analysis of Large-Scale Data-Sets and Rational Action Theory: For a Sociological Alliance.” *European Sociological Review* 12 (2): 109–126.

———. 2001. “Causation, Statistics, and Sociology.” *European Sociological Review* 17 (1): 1–20.

———. 2014. “The Role of Education in Intergenerational Social Mobility: Problems from Empirical Research in Sociology and Some Theoretical Pointers from Economics.” *Rationality and Society* 26 (3): 265–289.

Good, Thomas L. "Two decades of research on teacher expectations: Findings and future directions." Journal of teacher education 38.4 (1987): 32-47.

Goos, Mieke, Brigitte Maria Schreier, Heidi Maria Eduard Knipprath, Bieke De Fraine, Jan Van Damme, and Ulrich Trautwein. 2013. “How Can Cross-Country Differences in the Practice of Grade Retention Be Explained? A Closer Look at National Educational Policy Factors.” *Comparative Education Review* 57 (1): 54–84.

Gottfredson, Denise C., Carolyn M. Fink, and Nanette Graham. 1994. “Grade Retention and Problem Behavior.” *American Educational Research Journal* 31 (4): 761–784.

Graaf, Nan Dirk De, Paul M. De Graaf, and Gerbert Kraaykamp. 2000. “Parental Cultural Capital and Educational Attainment in the Netherlands: A Refinement of the Cultural Capital Perspective.” *Sociology of Education* 73 (2): 92.

Gregory, Anne, and Francis Huang. 2013. “It Takes a Village: The Effects of 10th Grade College-Going Expectations of Students, Parents, and Teachers Four Years Later.” *American Journal of Community Psychology* 52 (1–2): 41–55.

Griffith, Catherine Aurentz, John Wills Lloyd, Kathleen Lynne Lane, and Melody Tankersley. 2010. “Grade Retention of Students During Grades K–8 Predicts Reading Achievement and Progress During Secondary Schooling.” *Reading & Writing Quarterly* 26 (1): 51–66.

Grissom, James B., and Lorrie A. Shepard. 1989. “Structural Equation Modeling of Retention and Overage Effects on Dropping Out of School.”

Guryan, Jonathan, Erik Hurst, and Melissa Kearney. n.d. “Parental Education and Parental Time with Children,” 52.

Hadjar, Andreas, Sabine Krolak-Schwerdt, Karin Priem, and Sabine Glock. 2014. “Gender and Educational Achievement.” *Educational Research* 56 (2): 117–125.

Hallinan, Maureen T. n.d. “Ability Grouping and Student Learning,” 47.

Hannum, Emily, and Claudia Buchmann. 2003. “The Consequences of Global Educational Expansion.” In *American Academy of Arts and Sciences*. Citeseer.

Hansen, Karsten T., James J. Heckman, and Kathleen J. Mullen. 2004. “The Effect of Schooling and Ability on Achievement Test Scores.” *Journal of Econometrics* 121 (1–2): 39–98.

Hanushek, Eric A. 1992. “The Trade-off between Child Quantity and Quality.” *Journal of Political Economy* 100 (1): 84–117.

Hanushek, Eric A, and Ludger Woessmann. n.d. “Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries,” 27.

Hao, Lingxin, and Wei-Jun Jean Yeung. 2015. “Parental Spending on School-Age Children: Structural Stratification and Parental Expectation.” *Demography* 52 (3): 835–60.

Harden, K. Paige, Benjamin W. Domingue, Daniel W. Belsky, Jason D. Boardman, Robert Crosnoe, Margherita Malanchini, Michel Nivard, Elliot M. Tucker-Drob, and Kathleen Mullan Harris. 2019. “Genetic Associations with Mathematics Tracking and Persistence in Secondary School.” Preprint. Genetics.

Hartlaub, Vanessa, and Thorsten Schneider. 2012. “Educational Choice and Risk Aversion: How Important Is Structural vs. Individual Risk Aversion?” SOEPpapers on multidisciplinary panel data research.

Haug, Wolfgang Fritz. 1981. “Structural Hegemony.” *ARGUMENT* 23 (SEP): 628–648.

———. 2006. “Marx’s Learning Process. Against Correcting Marx with Hegel.” *Rethinking Marxism* 18 (4): 565–584.

———. 2016. “Nella Fossa Dei Leoni. Si Può Leggere Il Capitale Di Marx a Partire Dalle Tesi Su Feuerbach?” *Materialismo Storico*, no. 1–2: 75–91.

Healy, Kieran. 2017. “Fuck Nuance.” *Sociological Theory* 35 (2): 118–127.

Heckman, James J. 1995. “Lessons from the Bell Curve.” *Journal of Political Economy* 103 (5): 1091–1120.

Heckman, James J., and Tim Kautz. 2012. “Hard Evidence on Soft Skills.” *Labour Economics* 19 (4): 451–64.

Helliwell, John F., and Robert D. Putnam. Education and social capital. No. w7121. National Bureau of Economic Research, 1999.

Herbert J. Walberg, Arthur J. Reynolds, Margaret C. Wang. n.d. *Can Unlike Students Learn Together?: Grade Retention, Tracking, and Grouping*.

Herrnstein, Richard J., and Charles A. Murray. The Bell Curve: Reshaping of American Life by Differences in Intelligence. Free Press, 1994.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. “Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference.” *Political Analysis* 15 (3): 199–236.

Hofmann, David A., and Mark B. Gavin. 1998. “Centering Decisions in Hierarchical Linear Models: Implications for Research in Organizations.” *Journal of Management* 24 (5): 623–41.

Holm, Anders, Mads Meier J\a eger, and others. 2005. “Relative Risk Aversion and Social Reproduction in Intergenerational Educational Attainment: Application of a Dynamic Discrete Choice Model.” In *Summer Meeting of ISA-RC28 (Social Stratification and Mobility), Los Angeles, August*, 18–21.

Hombres, Béatrice d’. 2010. “Inequality in Tertiary Education Systems: Which Metric Should We Use for Measuring and Benchmarking?” *World Bank with Funding from the Bank Netherlands Partnership Program (BNPP)*.

Hong, Guanglei, and Bing Yu. 2007. “Early-Grade Retention and Children’s Reading and Math Learning in Elementary Years.” *Educational Evaluation and Policy Analysis* 29 (4): 239–61.

———. 2008. “Effects of Kindergarten Retention on Children’s Social-Emotional Development: An Application of Propensity Score Method to Multivariate, Multilevel Data.” *Developmental Psychology* 44 (2): 407–21.

House, Ernest R. 1998. “The Predictable Failure of Chicago’s Student Retention Program.” In *Rethinking Retention to Help All Students Succeed’conference, November, Chicago, Illinois*. Citeseer.

Hughes, Jan N., Qian Cao, Stephen G. West, Paula Allee Smith, and Carissa Cerda. 2017. “Effect of Retention in Elementary Grades on Dropping out of School Early.” *Journal of School Psychology* 65 (December): 11–27.

Hughes, Jan N., Katie A. Gleason, and Duan Zhang. 2005. “Relationship Influences on Teachers’ Perceptions of Academic Competence in Academically at-Risk Minority and Majority First Grade Students.” *Journal of School Psychology* 43 (4): 303–20.

Ichou, Mathieu, and Louis-André Vallet. 2013. “Academic Achievement, Tracking Decisions, and Their Relative Contribution to Educational Inequalities.” In *Determined to Succeed?*, edited by Michelle Jackson, 116–48. Stanford University Press.

Ikeda, Miyako, and Emma García. 2013. “Grade Repetition: A Comparative Study of Academic and Non-Academic Consequences.” *OECD Journal. Economic Studies* 2013 (1): 269.

“International Encyclopedia of Teaching and Teacher Education - 2nd Edition.” n.d. Accessed April 29, 2018.

Ishitani, Terry T. 2006. “Studying Attrition and Degree Completion Behavior among First-Generation College Students in the United States.” *Journal of Higher Education*, 861–885.

Jackson, C. Kirabo. 2012. “Non-Cognitive Ability, Test Scores, and Teacher Quality: Evidence from 9th Grade Teachers in North Carolina.” w18624. Cambridge, MA: National Bureau of Economic Research.

Jackson, Michelle, ed. Determined to succeed?: performance versus choice in educational attainment. Stanford University Press, 2013.

Jackson, Gregg B. 1975. “The Research Evidence on the Effects of Grade Retention.” *Review of Educational Research* 45 (4): 613.

Jacob, Brian A., and Lars Lefgren. 2002. “Remedial Education and Student Achievement: A Regression-Discontinuity Analysis.” NBER Working Paper.

———. 2009. “The Effect of Grade Retention on High School Completion.” *American Economic Journal: Applied Economics* 1 (3): 33–58.

Jean-Jacques, Paul. 2004. “Les Apports de La Recherche Sur l’impact Du Redoublement Comme Moyen de Traiter Les Difficultés Scolaires Au Cours de La Scolarité Obligatoire.” *Haut Conseil de l’évaluation de l’école*.

Jensen, Arthur. 1969. “How Much Can We Boost IQ and Scholastic Achievement.” *Harvard Educational Review* 39 (1): 1–123.

Jensen, Arthur R. "Environment, heredity, and intelligence." (1969).

Jimerson, Shane R. 2001. “Meta-Analysis of Grade Retention Research: Implications for Practice in the 21st Century.” *School Psychology Review* 30 (3): 420.

Jimerson, Shane R., Gabrielle E. Anderson, and Angela D. Whipple. 2002. “Winning the Battle and Losing the War: Examining the Relation between Grade Retention and Dropping out of High School.” *Psychology in the Schools* 39 (4): 441–57.

Jimerson, Shane R., Phillip Ferguson, Angela D. Whipple, Gabrielle E. Anderson, and Michael J. Dalton. 2002. “Exploring the Association between Grade Retention and Dropout: A Longitudinal Study Examining Socio-Emotional, Behavioral, and Achievement Characteristics of Retained Students.” *The California School Psychologist* 7 (1): 51–62.

Jimerson, Shane R., Sarah M. W. Pletcher, Kelly Graydon, Britton L. Schnurr, Amanda B. Nickerson, and Deborah K. Kundert. 2006. “Beyond Grade Retention and Social Promotion: Promoting the Social and Academic Competence of Students.” *Psychology in the Schools* 43 (1): 85–97.

Jonsson, Jan O. 1999. “Explaining Sex Differences in Educational Choice An Empirical Assessment of a Rational Choice Model.” *European Sociological Review* 15 (4): 391–404.

Jussim, Lee, Stacy L. Robustelli, and Thomas R. Cain. "Teacher expectations and self-fulfilling prophecies." Handbook of motivation at school 349380 (2009).

Kaineman and Tversky. "Prospect theory: An analysis of decision under risk." Econometrica 47.2 (1979): 363-391.

Kaushal, Neeraj, Katherine Magnuson, and Jane Waldfogel. "How is family income related to investments in children’s learning?" (2011): 187-206.

Kloosterman, R., N. Notten, J. Tolsma, and G. Kraaykamp. 2011. “The Effects of Parental Reading Socialization and Early School Involvement on Children’s Academic Performance: A Panel Study of Primary School Pupils in the Netherlands.” *European Sociological Review* 27 (3): 291–30.

Kloosterman, Rianne, and Paul M. de Graaf. 2010. “Non-promotion or Enrolment in a Lower Track? The Influence of Social Background on Choices in Secondary Education for Three Cohorts of Dutch Pupils.” *Oxford Review of Education* 36 (3): 363–84.

Koppensteiner, Martin Foureaux. 2014. “Automatic Grade Promotion and Student Performance: Evidence from Brazil.” *Journal of Development Economics* 107: 277–290.

Kornrich, Sabino. 2016. “Inequalities in Parental Spending on Young Children: 1972 to 2010.” *AERA Open* 2 (2): 233285841664418.

Kreft, Ita G.G., Jan de Leeuw, and Leona S. Aiken. 1995. “The Effect of Different Forms of Centering in Hierarchical Linear Models.” *Multivariate Behavioral Research* 30 (1): 1–21.

Labaree, David F. n.d. “Setting the Standard: Alternative Policies for Student Promotion.” *Harvard Educational Review*, 23.

Lamont, Michele, and Annette Lareau. 1988. “Cultural Capital: Allusions, Gaps and Glissandos in Recent Theoretical Developments.” *Sociological Theory* 6 (2): 153.

Lareau, Annette. 1987. “Social Class Differences in Family-School Relationships: The Importance of Cultural Capital.” *Sociology of Education* 60 (2): 73.

Lareau, Annette, and Elliot B Weininger. n.d. “Cultural Capital in Educational Research: A Critical Assessment,” 40.

Lauen, Douglas Lee, and S. Michael Gaddis. 2016. “Accountability Pressure, Academic Standards, and Educational Triage.” *Educational Evaluation and Policy Analysis* 38 (1): 127–147.

Lavecchia, Adam M., Heidi Liu, and Philip Oreopoulos. 2016. “Behavioral Economics of Education: Progress and Possibilities.” In *Handbook of the Economics of Education*, 5:1–74. Elsevier.

Lecouteux, Guilhem, and Léonard Moulin. 2015. “To Gain or Not to Lose? Tuition Fees for Loss Averse Students.” *Economics Bulletin* 35 (2): 1005–1019.

Lerner, Abba P. 1972. “The Economics and Politics of Consumer Sovereignty.” *The American Economic Review* 62 (1/2): 258–266.

Levin, M. n.d. “Accelerated Schools for Disadvantaged Students,” 4.

Lewis, Oscar. La vida: A Puerto Rican family in the culture of poverty-San Juan and New York. Vol. 13. New York: Random House, 1966.

Lindahl, Erica. "Are teacher assessments biased?–evidence from Sweden." Education economics 24.2 (2016): 224-238.

Lindenberg, Siegwart. 1990. “Homo Socio-Oeconomicus: The Emergence of a General Model of Man in the Social Sciences.” *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift Für Die Gesamte Staatswissenschaft* 146 (4): 727–748.

Link, Bruce G., and Jo C. Phelan. "Conceptualizing stigma." Annual review of Sociology 27.1 (2001): 363-385.

Lipnevich, Anastasiya A., and Richard D. Roberts. 2012. “Noncognitive Skills in Education: Emerging Research and Applications in a Variety of International Contexts.” *Learning and Individual Differences* 22 (2): 173–177.

Liston, Daniel P., and Kevin Murray. 2015. “Schooling in Capitalism: Navigating the Bleak Pathways of Structural Fate.” *Educational Theory* 65 (3): 245–264.

Lizardo, Omar. 2004. “The Cognitive Origins of Bourdieu’s Habitus.” *Journal for the Theory of Social Behaviour* 34 (4): 375–401.

Lorence, Jon. 2006. “Retention and Academic Achievement Research Revisited from a United States Perspective.” *International Education Journal* 7 (5): 731–777.

Lorence, Jon, A. Gary Dworkin, Laurence A. Toenjes, Antwanette N. Hill, Andrew Rotherham, and Lorrie A. Shepard. 2002. “Grade Retention and Social Promotion in Texas, 1994-99: Academic Achievement among Elementary School Students.” *Brookings Papers on Education Policy*, no. 5: 13–67.

Lorenzoni, Franco. 2017. “I Voti e Le Bocciature Fanno Male Agli Studenti e Alla Scuola.” Internazionale. February 4, 2017.

Lovaglio, Pietro Giorgio, and Gloria Ronzoni. 2007. “Dropout Nelle Scuole Superiori: L’impatto Della Tipologia Istituzionale.” *Statistica Applicata* 19: 77–96.

Lucas, Samuel R. 2001. “Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects.” *American Journal of Sociology* 106 (6): 1642–1690.

Lunn, Anna, and Sabino Kornrich. 2018. “Family Investments in Education during Periods of Economic Uncertainty: Evidence from the Great Recession.” *Sociological Perspectives* 61 (1): 145–63.

Lynch, Matthew. 2013. “Alternatives to Social Promotion and Retention.” *Interchange* 44 (3–4): 291–309.

Mahjoub, Mohamed-Badrane. 2017. “The Treatment Effect of Grade Repetitions.” *Education Economics* 25 (4): 418–32.

Manacorda, Marco. 2008. *The Cost of Grade Retention*. London: Centre for Economic Performance, London School of Economics and Political Science.

Manzo, Gianluca. 2013. “Is Rational Choice Theory Still a Rational Choice of Theory? A Response to Opp.” *Social Science Information* 52 (3): 361–382.

———. 2014. *Analytical Sociology: Actions and Networks*. John Wiley & Sons.

Marcoux, Géry, and Marcel Crahay. 2008. “Mais Pourquoi Continuent-Ils à Faire Redoubler? Essai de Compréhension Du Jugement Des Enseignants Concernant Le Redoublement.” *Revue Suisse Des Sciences de l’éducation* 30 (3): 501–518.

Mariano, Lou, and Paco Martorell. 2015. “The Causal Effects of Grade Retention on Behavioral Outcomes.” *Society for Research on Educational Effectiveness*.

Marsh, Herbert W., et al. "The big-fish–little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research." Educational psychology review 20.3 (2008): 319-350.

Martin, Andrew J. 2011. “Holding Back and Holding behind: Grade Retention and Students’ Non-Academic and Academic Outcomes.” *British Educational Research Journal* 37 (5): 739–63.

Martini, Alberto, and Marco Sisti. Valutare il successo delle politiche pubbliche. Bologna: Il Mulino, 2009.

Martini, Angela. 2008. “L’accountability Nella Scuola.” *Fondazione Giovanni Agnelli*, no. 8.

———. 2014. “La Pubblicità Dei Risultati Della Valutazione Esterna Dei Livelli Di Apprendimento Raggiunti Dalle Scuole: Pro e Contro, Alla Luce Delle Esperienze Internazionali.” FGA Working Paper.

Matteucci, Maria Cristina, Carlo Tomasetto, Patrizia Selleri, and Felice Carugati. 2008. “Teacher Judgments and Pupils’ Causal Explanations: Social Valorization of Effort-Based Explanations in School Context.” *European Journal of Psychology of Education* 23 (4): 421–432.

Mayer, Susan E., Ariel Kalil, Philip Oreopoulos, and Sebastian Gallegos. 2015. “Using Behavioral Insights to Increase Parental Engagement: The Parents and Children Together (PACT) Intervention.” National Bureau of Economic Research.

McCoy, Ann R., and Arthur J. Reynolds. 1999. “Grade Retention and School Performance: An Extended Investigation.” *Journal of School Psychology* 37 (3): 273–298.

McCoy, Shannon K., and Brenda Major. "Priming meritocracy and the psychological justification of inequality." Journal of experimental social psychology 43.3 (2007): 341-351.

McEneaney, Lizabeth H., and John W. Meyer. "The content of the curriculum." Handbook of the Sociology of Education. Springer, Boston, MA, 2000. 189-211.

McGovern, Dr Mark, Dr Slawa Rokicki, and UCD Geary. n.d. “The Great Recession, Household Income, and Children’s Test Scores,” 22.

McGrew, Ken. 2011. “A Review of Class-Based Theories of Student Resistance in Education Mapping the Origins and Influence of Learning to Labor by Paul Willis.” *Review of Educational Research* 81 (2): 234–266.

McGuigan, Martin, Sandra McNally, and Gill Wyness. 2012. “Student Awareness of Costs and Benefits of Educational Decisions: Effects of an Information Campaign. CEE DP 139.” *Centre for the Economics of Education*.

Meisels, Samuel J., and Fong-Ruey Liaw. 1993. “Failure in Grade: Do Retained Students Catch Up?” *The Journal of Educational Research* 87 (2): 69–77.

Meissel, Kane, et al. "Subjectivity of teacher judgments: Exploring student characteristics that influence teacher judgments of student ability." Teaching and Teacher Education 65 (2017): 48-60.

Merton, Robert K. 1987. “Three Fragments from a Sociologist’s Notebooks: Establishing the Phenomenon, Specified Ignorance, and Strategic Research Materials.” *Annual Review of Sociology* 13 (1): 1–29.

Merton, Robert K. "Theself-fulfillingprophecy." TheAntioch Review8 (2) (1948): 193-210.

Micklewright, John, Mark Pearson, and Stephen Smith. "Unemployment and early school leaving." The Economic Journal 100.400 (1990): 163-169.

Mingat, Alain. 1998. “The Strategy Used by High-Performing Asian Economies in Education: Some Lessons for Developing Countries.” *World Development* 26 (4): 695–715.

Ministero della Pubblica Istruzione. 2000. *La Dispersione Scolastica: Una Lente Sulla Scuola*. Ministero della Pubblica Istruzione.

Moulin, Léonard. 2014. “Frais d’inscription Dans l’enseignement Supérieur: Enjeux, Limites et Perspectives.” ENS Cachan.

N’tchougan-Sonou, Christina H. 2001. “Automatic Promotion or Large-Scale Repetition—Which Path to Quality?” *International Journal of Educational Development* 21 (2): 149–62.

N’tchougan-Sonou, Christina H. 2001. “Automatic Promotion or Large-Scale Repetition—Which Path to Quality?” *International Journal of Educational Development* 21 (2): 149–162.

O’Mara, Alison J., Herbert W. Marsh, Rhonda G. Craven, and Raymond L. Debus. 2006. “Do Self-Concept Interventions Make a Difference? A Synergistic Blend of Construct Validation and Meta-Analysis.” *Educational Psychologist* 41 (3): 181–206.

Palladino, Benedetta Emanuela, Annalaura Nocentini, Enrica Ciucci, and Ersilia Menesini. 2011. “I Diversi Volti Del Fallimento Scolastico: Ruolo Dei Predittori Individuali e Relazionali.” *Eta Evolutiva* 33 (98): 95.

Pansu, Pascal, Nicole Dubois, and Benoît Dompnier. 2008. “Internality-Norm Theory in Educational Contexts.” *European Journal of Psychology of Education* 23 (4): 385–397.

Parsons, Jacquelynne E., Terry Adler, and Judith L. Meece. 1984. “Sex Differences in Achievement: A Test of Alternate Theories.” *Journal of Personality and Social Psychology* 46 (1): 26.

Passeron, Jean Claude, and Pierre Bourdieu. "La reproduction." Eléments pour (1970).

Pato, Ana Henriques. 2014. “Dualism Wave-Particle and the Value of the Scientific Theory in Bohr. The Necessary Consideration of Materialist Dialectics.” *Journal of Humanities of Valparaiso*, no. 3: 23–31.

Paul, J J. 1997. “Le Redoublement à l’école: Une Maladie Universelle?” *International Review of Education* 43 (5–6): 611–627.

Pedulla, Joseph J., Peter W. Airasian, and George F. Madaus. "Do teacher ratings and standardized test results of students yield the same information?" American Educational Research Journal 17.3 (1980): 303-307.

Peterson, Sarah E., James S. DeGracie, and Carol R. Ayabe. 1987. “A Longitudinal Study of the Effects of Retention/Promotion on Academic Achievement.” *American Educational Research Journal* 24 (1): 107–118.

Pietro, Giorgio Di, and Andrea Cutillo. 2006. “University Quality and Labour Market Outcomes in Italy.” *Labour* 20 (1): 37–62.

Pintrich, Paul R. n.d. “Multiple Goals, Multiple Pathways: The Role of Goal Orientation in Learning and Achievement,” 12.

Piraino, Patrizio. 2007. “Comparable Estimates of Intergenerational Income Mobility in Italy.” *The BE Journal of Economic Analysis & Policy* 7 (2).

Pisati, Maurizio, and Antonio Schizzerotto. 1999. “Pochi Promossi, Nessun Bocciato. La Mobilità Di Carriera in Italia in Prospettiva Comparata e Longitudinale.” *Stato e Mercato* 19 (2): 249–280.

Pit-ten Cate, Ineke M., Sabine Krolak-Schwerdt, and Sabine Glock. 2016. “Accuracy of Teachers’ Tracking Decisions: Short- and Long-Term Effects of Accountability.” *European Journal of Psychology of Education* 31 (2): 225–43.

Plank, Stephen B., Stefanie DeLuca, and Angela Estacion. 2008. “High School Dropout and the Role of Career and Technical Education: A Survival Analysis of Surviving High School.” *Sociology of Education* 81 (4): 345–370.

Powdthavee, Nattavudh, and Anna Vignoles. 2009. “The Socioeconomic Gap in University Dropouts.” *The BE Journal of Economic Analysis & Policy* 9 (1).

Pozzoli, Dario. 2009. “The Transition to Work for Italian University Graduates.” *Labour* 23 (1): 131–169.

Raisa, Akifyeva, and Alieva Alisa. "The influence of student ethnicity on teacher expectations and teacher perceptions of warmth and competence." Psychology in Russia: State of the art 11.1 (2018).

Rauber, Michael. n.d. “Noncognitive Skills and Success in Life: The Importance of Motivation and Self-Regulation,” 32.

Reardon, Sean F. 2018. “The Widening Academic Achievement Gap Between the Rich and the Poor.” In *Inequality in the 21st Century*, by David Grusky, edited by David B. Grusky and Jasmine Hill, 1st ed., 177–89. Routledge.

Reay, Diane, and Stephen J. Ball. 1997. “Spoilt for Choice’: The Working Classes and Educational Markets.” *Oxford Review of Education* 23 (1): 89–101.

Ress, Anna, and Davide Azzolini. 2014. “Primary and Secondary Effects of Social Background on Educational Attainment in Italy. Evidence from an Administrative Dataset.” *Italian Journal of Sociology of Education* 6 (1).

Rikowski, Glenn. 1997. “Scorched Earth: Prelude to Rebuilding Marxist Educational Theory.” *British Journal of Sociology of Education* 18 (4): 551–574.

Roderick, Melissa. 1994. “Grade Retention and School Dropout: Investigating the Association.” *American Educational Research Journal* 31 (4): 729–759.

Rodrıguez, Germán. 2012. “Cumulative Incidence.”

Romito, Marco. 2014. “Provenienze sociali e orientamento scolastico. L’«accompagnamento alla scelta» della scuola superiore.” *Etnografia e ricerca qualitativa*, no. 3: 481–504.

Rosen, Jeffrey A., and Jeffrey A. Rosen. 2010. *Noncognitive Skills in the Classroom: New Perspectives on Educational Research*. RTI Press Book Series. Research Triangle Park, NC: RTI Press.

Rosenbaum, James E. 2011. “The Complexities of College for All: Beyond Fairy-Tale Dreams.” *Sociology of Education* 84 (2): 113–17.

Rosenbaum, James E., Stefanie DeLuca, Shazia R. Miller, and Kevin Roy. 1999. “Pathways into Work: Short- and Long-Term Effects of Personal and Institutional Ties.” *Sociology of Education* 72 (3): 179–96.

Rosenthal, Robert, and Lenore F. Jacobson. "Teacher expectations for the disadvantaged." Scientific American 218.4 (1968): 19-23.

Ross, Catherine E., and Beckett A. Broh. 2000. “The Roles of Self-Esteem and the Sense of Personal Control in the Academic Achievement Process.” *Sociology of Education* 73 (4): 270.

Rözer, Jesper, and Herman van de Werfhorst. n.d. “Inequalities in Educational Opportunities by Socioeconomic and Migration Background: A Comparative Assessment Across European Societies,” 72.

Ryan, James E. 2004. “The Perverse Incentives of the No Child Left behind Act.” *NYUL Rev.* 79: 932.

Sanders, William L., and June C. Rivers. 1996. “Cumulative and Residual Effects of Teachers on Future Student Academic Achievement.”

Santagati, Mariagrazia. 2015. “Researching Integration in Multiethnic Italian Schools. A Sociological Review on Educational Inequalities.” *Italian Journal of Sociology of Education* 7 (10/2015): 294–334.

Sayer, Derek. 1975. “Method and Dogma in Historical Materialism.” *The Sociological Review* 23 (4): 779–810.

Schizzerotto, Antonio, and Carlo Barone. Sociologia dell'istruzione. Il mulino, 2006.

Schmid, Carol L. 2001. “Educational Achievement, Language-Minority Students, and the New Second Generation.” *Sociology of Education* 74: 71.

Schneider, Hastings, and LaBriola. n.d. “Income Inequality and Class Divides in Parental Investments.” *American Sociological Review*, 33.

Schunck, Reinhard. 2016. “Cluster Size and Aggregated Level 2 Variables in Multilevel Models. A Cautionary Note.” *Methods, Data, Analyses, Vol 10, No 1 (2016)*, July.

Schwandt, Thomas A. 1996. “Farewell to Criteriology.” *Qualitative Inquiry* 2 (1): 58–72.

Schwerdt, Guido, and Martin R West. n.d. “The Effects of Early Grade Retention on Student Outcomes over Time: Regression Discontinuity Evidence from Florida,” 53.

Searle, John. 2013. “Theory of Mind and Darwin’s Legacy.” *Proceedings of the National Academy of Sciences* 110 (Supplement 2): 10343–10348.

Seth Gershenson. n.d. “Linking Teacher Quality, Student Attendance, and Student Achievement.” Accessed September 28, 2018.

Shepard, Lorrie A., and Mary Lee Smith. 1989. *Flunking Grades: Research and Policies on Retention*. Education Policy Perspectives. The Falmer Press, Taylor & Francis, Inc.,.

Sherman, Jennifer, and Elizabeth Harris. 2012. “Social Class and Parenting: Classic Debates and New Understandings: Social Class and Parenting.” *Sociology Compass* 6 (1): 60–71.

Shirasu, Maitê Rimekká, and Ronaldo de Albuquerque. 2015. “Determinantes Da Evasão e Repetência Escolar No Ensino Médio Do Ceará.” *Revista Econômica Do Nordeste* 46 (4): 117–136.

Silberglitt, Benjamin, James J. Appleton, Matthew K. Burns, and Shane R. Jimerson. 2006. “Examining the Effects of Grade Retention on Student Reading Performance: A Longitudinal Study.” *Journal of School Psychology* 44 (4): 255–70.

Spence, Michael. 1973. “Job Market Signaling.” *The Quarterly Journal of Economics* 87 (3): 355.

Sprietsma, Maresa. "Discrimination in grading: Experimental evidence from primary school teachers." Empirical economics 45.1 (2013): 523-538.

Stearns, Elizabeth, Stephanie Moller, Judith Blau, and Stephanie Potochnick. 2007. “Staying Back and Dropping out: The Relationship between Grade Retention and School Dropout.” *Sociology of Education* 80 (3): 210–240.

Steele, Claude M. 1997. “How Stereotypes Shape Intellectual Identity and Performance.” *American Psychologist,* 17.

Stevens, Ann Huff, and Jessamyn Schaller. "Short-run effects of parental job loss on children's academic achievement." Economics of Education Review 30.2 (2011): 289-299.

Stocke, V. 2007. “Explaining Educational Decision and Effects of Families’ Social Class Position: An Empirical Test of the Breen Goldthorpe Model of Educational Attainment.” *European Sociological Review* 23 (4): 505–19.

Streeck, Wolfgang, and Ettore Recchi. 1994. “Vincoli Benefici: Sui Limiti Economici Dell’attore Razionale.” *Stato e Mercato,* 185–213.

Sullivan, Alice. 2002. “Bourdieu and Education: How Useful Is Bourdieu’s Theory for Researchers?” *Netherlands Journal of Social Sciences* 38 (2): 144–166.

Swartz, David L. 2003. “From Correspondence to Contradiction and Change: Schooling in Capitalist America Revisited.” In *Sociological Forum,* 18:167–186. Springer.

Swidler, Ann. 1986. “Culture in Action: Symbols and Strategies.” *American Sociological Review* 51 (2): 273.

Gershenson, Seth, and Nicholas Papageorge. "The power of teacher expectations: How racial bias hinders student attainment." Education Next 18.1 (2018): 64-71.

Thompson, Charles L., and Elizabeth K. Cunningham. 2000. “Retention and Social Promotion: Research and Implications for Policy. ERIC Digest Number 161”.

Timmermans, Anneke C., Hester de Boer, and Margaretha PC van der Werf. "An investigation of the relationship between teachers’ expectations and teachers’ perceptions of student attributes." Social Psychology of Education 19.2 (2016): 217-240.

Tingle, Lynne R., Jason Schoeneberger, and Bob Algozzine. 2012. “Does Grade Retention Make a Difference?” *The Clearing House: A Journal of Educational Strategies, Issues and Ideas* 85 (5): 179–85.

Todd, Petra E., and Weilong Zhang. 2017. “A Dynamic Model of Personality, Schooling, and Occupational Choice.”

Tomchin, Ellen M., and James C. Impara. 1992. “Unraveling Teachers’ Beliefs about Grade Retention.” *American Educational Research Journal* 29 (1): 199–223.

Tonello, Marco. 2009. “A New Approach to the Economics of Education: The Case of the English Schooling System.”

Triventi, Moris. 2019. “Are Children of Immigrants Graded Less Generously by Their Teachers than Natives, and Why? Evidence from Student Population Data in Italy.” *International Migration Review*, October, 019791831987810.

Tropf, Felix C., S. Hong Lee, Renske M. Verweij, Gert Stulp, Peter J. van der Most, Ronald de Vlaming, Andrew Bakshi, et al. 2017. “Hidden Heritability Due to Heterogeneity across Seven Populations.” *Nature Human Behaviour* 1 (10): 757–65.

Van de Werfhorst, Herman G., and Jonathan J.B. Mijs. 2010. “Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective.” *Annual Review of Sociology* 36 (1): 407–28.

Van Ewijk, Hans. "Maatschappelijk werk in een sociaal gevoelige tijd." (2010).

Weiner, Bernard. "The classroom as a courtroom." Social Psychology of Education 6.1 (2003): 3-15.

Wenz, Sebastian E., Melanie Olczyk, and Georg Lorenz. "Measuring Teachers' Stereotypes in the NEPS." (2016).

Werfhorst, Herman G. van de. 2018. “Early Tracking and Socioeconomic Inequality in Academic Achievement: Studying Reforms in Nine Countries.” *Research in Social Stratification and Mobility* 58 (December): 22–32.

Werfhorst, Herman G. van de, and Saskia Hofstede. 2007. “Cultural Capital or Relative Risk Aversion? Two Mechanisms for Educational Inequality Compared.” *The British Journal of Sociology* 58 (3): 391–415.

West, Martin R. 2012. “Is Retaining Students in the Early Grades Self-Defeating.” *CCF Brief* 49.

“When Students Repeat Grades or Are Transferred out of School: What Does It Mean for Education Systems?” n.d. Accessed January 21, 2017.

Wigfield, Allan, and Jacquelynne S. Eccles. 2000. “Expectancy–Value Theory of Achievement Motivation.” *Contemporary Educational Psychology* 25 (1): 68–81.

William Carbonaro. 2005. “Tracking, Students’ Effort, and Academic Achievement.” *Sociology of Education* 78 (1): 27–49.

Willson, Victor L., and Jan N. Hughes. 2006. “Retention of Hispanic/Latino Students in First Grade: Child, Parent, Teacher, School, and Peer Predictors.” *Journal of School Psychology* 44 (1): 31–49.

———. 2009. “Who Is Retained in First Grade? A Psychosocial Perspective.” *The Elementary School Journal* 109 (3): 251–66.

Wóbmann, Ludger. 2000. “Institutions of the Education System and Student Performance: The International Evidence.” In *Sept. 2000 Conference of ISNIE, Available from Http://Www. ISNIE2000. Org*.

Woolcock, Michael. 1998. “Social Capital and Economic Development: Toward a Theoretical Synthesis and Policy Framework.” *Theory and Society* 27 (2): 151–208.

Wright, Margaret O’Dougherty, and Ann S. Masten. 2005. “Resilience Processes in Development.” In *Handbook of Resilience in Children*, edited by Sam Goldstein and Robert B. Brooks, 17–37. Boston, MA: Springer US.

Wu, Wei, Stephen G. West, and Jan N. Hughes. 2008. “Short-Term Effects of Grade Retention on the Growth Rate of Woodcock–Johnson III Broad Math and Reading Scores.” *Journal of School Psychology* 46 (1): 85–105.