UNIVERSITÀ DELLA CALABRIA

UNIVERSITA' DELLA CALABRIA

DIPARTIMENTO DI ECONOMIA, STATISTICA E FINANZA

Scuola di Dottorato in

SCIENZE ECONOMCHE E AZIENDALI

Indirizzo

ECONOMIA APPLICATA

CICLO

XXVII

TITOLO TESI

The Impact of Parents' Background and Students' Age on Educational Outcomes in Italian Primary School: Evidence from INVALSI Data

Settore Scientifico Disciplinare SECS-P02 Politica Economica

Direttore: Ch.mo Prof. Patrizia Ordine

Firma____________________

Supervisore: Ch.mo Prof. Patrizia Ordine

Firma____________________

Dottorando: Dott.ssa Daniela Sposato

Firma____________________

A.A. 2013/2014

Abstract

La tesi di dottorato si propone l'obiettivo di indagare alcune determinanti dei risultati scolastici nella scuola primaria italiana. Quali variabili di outcomes sono utilizzati i punteggi in Italiano e Matematica conseguiti dagli alunni nell'ambito dei test somministrati dall'INVALSI (Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di Formazione).

Il primo capitolo, intitolato ***Parents' Background Effects on Students' Achievements: Evidence from Italian Primary School***, analizza l'impatto del background familiare sulla risultati scolastici in seconda e quinta elementare, controllando per caratteristiche dello studente, della scuola e territoriali. Al fine di gestire la presenza di dati mancanti nei sets di dati è stimato un modello di regressione lineare con imputazioni multiple. Le stime mostrano un forte impatto del background dei genitori sui test scores, impatto che tende a persistere durante la scuola primaria. Ulteriori specificazioni senza imputazioni e includendo effetti fissi di scuola, di classe e provinciali stimano effetti dello status socio-economico dei genitori sulla performance scolastica coerenti con le stime di base. I risultati consentono di capire quanta parte delle diseguaglianze nei livelli di apprendimento è generata dallo sfondo familiare, stimolando una riflessione sul ruolo delle scuole e delle autorità governative nel ridurre le disparità sociali che sorgono sin dai primi anni di istruzione e che si ripercuoteranno sul conseguimento del titolo di studio e, in tempi successivi, sul mercato del lavoro.

Il secondo capitolo, dal titolo ***"Gift of Time" and "Family Gift": The Effect of Early School Entry on Pupils Performance***, presenta un'analisi dell'impatto dell'ingresso anticipato a scuola sui risultati scolastici in seconda e quinta elementare. La procedura empirica è disegnata per districare l'effetto dell'età di ingresso a scuola (*Gift of Time*) da possibili fattori non osservabili (*Family Gift*) che influenzano sia la decisione di iscrizione anticipata sia i risultati accademici. Nello specifico, il problema della selezione sulle "non osservabili" è affrontato implementando un Regression Discontinuity Design comparando, in base alla soglia età, la performance degli studenti "anticipatari" con quelli aventi simile età ma "regolari". I risultati suggeriscono che coloro che entrano in anticipo a scuola raggiungono migliori performance rispetto agli studenti regolari, ossia beneficiano del *Family Gift*. Tuttavia, dopo aver controllato per la decisione dei genitori di iscrivere i figli a scuola in anticipo, ossia dopo aver neutralizzato il *Family Gift*, le stime mostrano che gli studenti anticipatari raggiungono minori punteggi ai test INVALSI. L'impatto non svanisce durante la scuola elementare.

Il terzo e ultimo capitolo, ***Students "in advance" and Peer Age Effect***, stima l'effetto della composizione della classe in termini di età sui rendimenti degli studenti frequentanti la scuola elementare in Italia. La strategia identificativa utilizza i cambiamenti verificatisi negli ultimi anni nella normativa di iscrizione al primo anno scolastico evidenziando che, in assenza di sistematica assegnazione degli alunni alle classi per età anticipata di ingresso a scuola, è possibile stimare un effetto non distorto del condividere la classe con studenti "in anticipo". I risultati indicano che la proporzione di studenti anticipatari più giovani nella classe ha un impatto positivo sui test scores individuali. L'effetto differisce per gruppo di età, con incidenza maggiore per gli altri studenti anticipatari più piccoli e decrescente all'aumentare dell'età.

La tesi è redatta in lingua inglese.

Index

CHAPTER 3

Students "in advance" and Peer Age Effect

Introduction

The thesis aims to analyze some determinants of students' outcomes in Italian primary school by using data from the National Institute for the Educational Evaluation of Instruction and Training (INVALSI).

First, I focus on the impact of parents' background on students' achievements. Then, I address to understand how the "individual" and "peer" age operates on school performance.

The choice of giving attention to primary school stems from the following reasons.

Education is a key to social and economic success and is the basis of intergenerational mobility in educational qualifications and income. Hence, primary school is the stage from which social inequalities are projected to the next academic stages. It follows the importance to act in primary schools to mitigate social inequalities among students and to promote equality of opportunity.

Socioeconomic status of students' family is relevant because the individual has not chosen his or her family background and thus cannot be held accountable for any impact of parents' background on his or her status during adulthood. Therefore, more important is family background for educational achievements, less is equality of opportunity.

In a child development perspective, researchers focus on how human capital accumulation is affected by early childhood resources. Specifically, studies with this aim try to understand what types of parental resources or inputs are important for children's development, why and when they are important, arguing that parental resources such as educational qualification and occupational status, or book at home, are relevant in early childhood and therefore in primary school. Home and family context are crucial for skill acquisition at this stage, while in the teenage years (secondary school), a spontaneous tendency to stratify the school system occurs, as students self-select in the choice of school type.

Beside family background, another important factor that determines school success is the children's "readiness" for school. This opens several questions concerning the optimal age to start school. Indeed, if a child is not be mature enough to school learning, to enroll him/her before he/she is ready may turn out less productivity. This implies lower performance, lower probabilities to gain a higher educational qualification and as a consequence lower earnings. If, instead, a child albeit younger is ready for education, he/she should start formal learning as soon as possible to begin accumulating skills sooner and enter the labor market earlier. Moreover, better skills implies higher education level and this latter mean higher wages.

It is easy to understand that verifying the effect of ***individual age at school entry*** is a crucial issue for parents, educators and governments. Similarly, identifying the effect of ***classroom peer age composition*** is an important issue occurring from the first year of schooling.

Classroom are formed by children with different age. A child's ability to accumulate human capital is affected by his/her characteristics – including age – and by characteristics of his/her peers – including age. The way through which the classmates' age can operates is differentiated. First, a positive spillover concerning the presence of youngest pupils can be due to teaching methods, which redirect more attention towards pupils. Second, peer age effect can be due to a learning spillover between classmates as youngest children could be more able than other pupils in the same classroom. Third, the presence of a certain proportion of youngest pupils in the classroom could create a more disciplined school environment generating a positive effect on the entire classroom. Finally, youngest pupils could experience more learning difficulties with a negative effect on their peers' educational outcomes.

Starting from these motivations, in the thesis I would like to analyze Italian primary education raising the following main research questions:

- Has parents' socioeconomic and cultural background a strong impact on students' achievements? Do this effect persist during primary school?
- Is the age at school entry an important factor for pupils' performance? In particular, do younger entrants achieve better or lower test scores compared to older entrants?
- Does the age of a child's peer affect the child's cognitive achievement? And specifically, does the proportion of youngest students in the classroom affect individual performance?

Employing data from INVALSI, which include the universe of pupils (about 500.000 per grade) attending primary school in Italy, this research take place in the scientific panorama in important ways. First, previous studies exploit the relationship between socioeconomic status and students' achievements using data from international surveys – PISA, PIRLS or TIMSS – which cover between 4.000 and 8.000 pupils per grade. I employ instead data which are more representative of all Italian pupils allowing a higher precision in estimates. Second, with reference to the first research question, this is the first attempt in the context of Italy to understand if the impact of parents' background on educational outcomes persists during primary school by exploiting the same cohort of students over two repeated cross-section. Third, in analyzing the effect of individual age at school entry, I present an original empirical procedure designed to disentangle the treatment effect (entering at school "in advance") from possible unobserved confounding factors affecting both enrollment decisions and educational outcomes. Finally, in the interest to explore how social interactions at class level affect academic achievements, I focus on peer age effects, adding a novel contribution in a topic – peer age – for which litte literature exists.

The thesis consists of three chapters. Chapter 1 analyzes the impact of parents' background on students' achievements. Charter 2 examines the effect of early age at school entry on pupils performance while Chapter 3 focuses on peer age effects on educational outcomes.

The analysis of parents' background effects on students' performance is conducted by using the cohort of pupils attending the 2nd grade in 2008/2009, who are substantially the same students attending the 5th grade in 2011/2012. As a selection problem related to voluntary participation in INVALSI assessment in 2008/2009 could bias parents' socio-economic effects on scores at the 2nd grade, I provide robustness checks by replicating results on data of pupils attending the 2nd grade in 2011/2012 when tests are administered in all Italian schools. I estimate a multiple linear regression in which student achievements are measured by Rasch test scores in Reading and Mathematics while parents' background is measured by country of birth, educational attainment and occupational status proxies. I control for other student traits and for school and territorial characteristics which are likely to affect educational performance.

To deal with missing values issue, a multiple imputation model is estimated. Findings show that parents' background is strongly positively related to students' educational achievements. Additional specifications without imputations and exploiting within-between school and provincial variation lead to very similar results. This additional evidence, which is consistent in sign and magnitude to basic estimates, lends strong support to the causal impact of parents' background on individual achievements both in Reading and in Mathematics.

Chapter 2 provides a comprehensive analysis of the effect of early school entry on educational outcomes by employing an empirical procedure designed to disentangle the effect of regular entry (*Gift of Time*) from possible unobserved confounding factors (*Family Gift)* affecting both enrollment decision and schooling outcome. I face the selection issue on unobservables by using a Regression Discontinuity Design. In this way exogenous age thresholds are used to compare children with similar age but different educational choices.

The analysis is carried out on pupils attending both the 2nd and the 5th grade of primary school in the scholastic year 2011/2012. Educational outcomes measures are represented by Normalized test scores in Reading and Mathematics. Results suggest that pupils who enroll in advance are peculiar as they perform better than regular ones with almost identical age. After neutralizing this "schooling ability" effect, I find that pupils in advance perform worse than regular ones. Hence, a severe distortion in the evaluation of the true effect of early entry arises when neglecting unobserved characteristics driving the early school decision. Since pupils in advance tend to be selected according to their schooling ability, the real impact of early entry on schooling performance is underestimated. After getting rid of selection bias, results show that pupils in advance present educational penalties. This effect goes on for the entire path of primary education.

The last Chapter of thesis focuses in estimating peer age effects on schooling outcomes of Italian pupils by exploiting changes in Italian enrollment rules over the last few year. I employ INVALSI data on students attending both the 2nd and the 5th grade of primary school in 2012/2013. As robustness check I use data on 8th grade too, i.e. on the last year of middle school.

I tackle the main barrier overcoming in estimating peer effects on student achievement: the *selection problem*. Students could be sorted by parents or by schools into peer group with other students with similar characteristics. Thus, the selection problem reflects all unobserved characteristics that may confound peer effect estimate. I use an empirical procedure allowing to understand if there is selection in classroom formation, arguing that in the absence of pupils sorting by early age at school entry, it is possible to estimate the "true" peer age effect. Results suggest that the proportion of youngest students "in advance" in the classroom has a positive impact on child's achievements measured by Normalized and Rasch test scores. Additional empirical evidence shows that the effect on individual scores of sharing the classroom with youngest pupils "in advance" differs by students' age group.

CHAPTER 1

Parents' Background Effects on Students' Achievements: Evidence from Italian Primary School

Abstract: This chapter examines the impact of parents' background on students' performance by employing INVALSI data on pupils attending both the 2nd and the 5th grade in Italian primary schools. Findings from a multiple linear regression model show that parents' background is strongly positively related to students' educational achievements. More interesting, this impact does not fade away but seems to persist during primary school. To deal with missing values issue, a multiple imputation model is estimated. Additional specifications without imputations and exploiting within-between school and provincial variation lead to very similar results. This additional evidence, which is consistent in sign and magnitude to basic estimates, lends strong support to the causal impact of parents' background on individual achievements both in Reading and in Mathematics. Results allow to understand how much of the inequality in educational achievements is due to socio-economic status of students' family. This has policy implications both for schools and for governments. These authorities, in fact, should fight the huge social disparities in terms of education opportunities and improve social mobility.

Keywords*:* educational achievements, parents' background, primary schools, equality of educational opportunity.

JEL Codes*:* I20, I24.

1.1 Background and Motivations

People’s educational achievement is positively correlated with their parents’ education or with other indicators of their parents’ socioeconomic status (SES). This pattern has fascinated many scholars, with early seminal contributions in sociology by James Coleman (1966) in the so-called *Coleman Report*, and in economics by Gary Becker (1964). Since this two seminal contributions, the nature of the relationship between socioeconomic status and student achievement has been debated for decades, with the most influential arguments appearing in *Equality of Educational Opportunity* (Coleman et al., 1968) and *Inequality* (Jencks et al., 1973) in the United States of America, and a number of commissioned inquiries in Australia (Commission of Inquiry into Poverty, 1976; Karmel, 1973).

The topic has also arisen frequently in policy debates: all are agree that, in a democratic society, socioeconomic inequalities in educational outcomes should be minimal, and most democratic societies have adopted policies aimed at reducing the impact of family background on educational achievements.

Today the topic is central in social science, and there is no doubt that research in this area has intensified during recent decades and even during the past few years. In fact, it is well known in the specialized literature that student background has a big predictive power of school results: higher is the SES of student’s family, higher is his academic achievement. So, the social, economic, and cultural background of pupils and their families exerts a very important influence on knowledge and skills that pupils acquire in their school career, even in primary school.

Students from low-SES homes are disadvantaged in schools because they lack an academic home environment, which influences their academic success at school. Then, as schooling affects an individual’s life chances, children from disadvantaged families often become adults with lower income and lower job status. These adults’ economic and educational disadvantages in turn affect their children’s schooling, and so on throughout later generations (Chiu and Xihua, 2008).

So, the analysis of differences in students skills is crucial because current differences will set the stage for future differences both in terms of competence of the adult class and in terms of income and thus of inequality.

For these reasons, in today’s society, socioeconomic inequalities in education are an important issue for both researchers and policy-makers. From a political perspective, socioeconomic inequalities in educational outcomes should be minimal to ensure to all individuals equality of opportunity in income and wealth.

From a researchers prespective, these are interested in this topic because has been demonstrated in countless studies that family socioeconomic status is the most powerful predictor of students’ performance, and seems to hold no matter what measure of status is used (parents’ occupation, parents’ education, family income, or some combination of these) (Boocock, 1972).

Researchers are therefore agree that education is a key to social and economic success and serves as a major mechanism for intergenerational transmission of socioeconomic status (see, for example, Haveman and Wolfe, 1995; Kane, 2004; Mare, 1992; Neckerman and Torche, 2007).

The relationship between social background and educational outcomes is well-established. Typically, the relationship is strong and positive, wherein higher socioeconomic status is associated with better educational outcomes, where educational outcomes can be represented by intermediate results such as test scores, or by educational attainment, or by income. In other words, students from privileged social backgrounds have on average higher test scores (Organisation for Economic Co-operation and Development [OECD], 2007; Perie et al., 2005), are more likely to complete secondary education (Renzulli and Park, 2000; Polidano et al., 2013), are more likely to complete college (Buchmann and Di Preite, 2006) and are more likely to attend university (Blossfeld and Shavit, 1993; Connor and Dewson, 2001; Lee, 1999; Terenzini et al., 2001) compared to their less-privileged peers.

Several empirical studies, moreover, have identified family background as one of the strongest predictors of students' educational attainment. For example, Binder (1998) explores the importance of parental schooling for children's educational attainment in Mexico; Broaded and Liu (1996) study the relationship between family background and educational attainment in Cina; Lauer (2003) analyzes the impact of family background, cohort and gender on educational attainment in France and Germany; McIntos and Munk (2007) examine the role of scholastic ability and family background variables in the determination of educational attainment in Denmark.

The effect of family background on educational attainment has been also studied by Datcher (1982), Teachman (1987), Pong and Post (1991), Gorman and Pollitt (1993), Wojtkiewicz and Donato (1995), Peraita and Sánchez (1998), Ermish and Francesconi (2001), Brunello and Checchi (2005), De Haan and Plug (2006), Holmlund et al (2008), Nam and Huang (2009), Sen and Clemente (2010), Huang (2013).

Abundant is also evidence of earnings returns to quantitative measures of education, like years of schooling (see, for instance, Corcoran, 1976; Card, 1999; Ashenfelter et al., 1999). More specifically, Solon (1999) studies the relationship in earnings between parents and their children; Currie and Thomas (2001) examine interactions between socio-economic status, children's test scores, and future wages and employment, Liu et al. (2001) analyze the effects of family background on worker's wage in Taiwan; Arias et al. (2004) investigate the role of race, family background and education in earnings inequality; Ordine and Rose (2012) analyze the impact of university quality, family background and mismatch on wages of young Italian graduates.

Therefore, a large economic literature on intergenerational transmission has emphasized the importance of family background for children's educational outcomes measured by educational attainment or income. However, intermediate outcomes such as grades and test scores are often studied because they are useful to plan educational interventions.

In this study, I focus on intermediate results measured by the test scores achieved by students of primary school for three main reasons.

The first is related to ***equality of opportunity***. Equality of opportunity for all citizens is a major concern in all open societies (Roemer, 1998). An important foundation for the future civil, social, and economic opportunities of citizens is laid in the education system. Here, family background is relevant because the individual has not chosen his or her family background and thus cannot be held accountable for any impact of family background on his or her status during adulthood. Therefore, more important is family background for educational achievements, less is equality of opportunity.

As the importance of educational performance for future income and productivity of individuals and societies has been documented by a large literature (Bishop, 1992; Card, 1999), I support the thesis on the basis of which it is necessary to act in primary schools to mitigate inequalities and promote equality of opportunity. Moreover, because in secondary school there is a spontaneous tendency to stratify the school system as students self-select in the choice of school (Checchi, 2003), the impact of the family background on students' performance can be study better in primary school. This kind of reasoning raises several research questions, the basic question is: *is family background* – in the broad sense that incorporates factors not chosen by individuals – *a determinant of educational achievements in primary school?* If so, *is this impact attenuated during school career*?

The second reasons for this research is the ***child development*** perspective. Here, the focus is on how human-capital accumulation is affected by early childhood resources. Studies with this focus address the questions: *what types of parental resources or inputs are important for children's development, why are they important and when are they important?* In this chapter I argue that for children development the main parental resources are parents' educational attainment, parents' occupational status and parents' ethnicity. These factors are primarily important in early childhood and therefore in primary school because at this stage the home and family context are crucial for skill acquisition, while in the teenage years (secondary school), as already said, there is a spontaneous tendency to stratify the school system due to the students' choice of the type of secondary school. In adulthood, instead, family is still essential but the work context becomes more relevant together with community and social relationships that start playing a significant role (Braga and Checchi, 2008).

Finally, the third motivation for this research is related to the choice to study the Italian case. Recent international surveys (PISA, PIRLS, TIMSS) have tested the ***level of student preparation from various countries, with a significant delay of the Italians***. The surveys also agree in describing a reality differentiated within Italy, in which southern students show a lower preparation in all subjects under investigation (ability to understand a text, Mathematics, science, problem solving) (Montanaro, 2008).

Several studies on geographical differences in terms of academic achievement have demonstrated that Italian school system has a great territorial complexity, which translates into an enlargement of differences and education inequalities among regions and macro-areas. In accordance with this, we would presume that primary education is the stage from which social inequalities are projected to the next academic stages and then that primary school has the responsibility to reduce social inequalities among students to get benefit from education.

In light of this premise, this research aims to ***analyze the impact of parents' background on students performance*** and ***to investigate the effect of territorial inequalities on educational achievements in Italian primary school.*** Specifically, I address to verify:

1) *if parents' socioeconomic and cultural background has a strong impact on students' achievements;*
2) *if this impact goes on between the 2th and the 5th grade of primary school;*
3) *if the preparation of pupils reveals wide territorial differences within the country.*

I use data from INVALSI (National Insititute for the Educational Evaluation of Instruction and Training) and estimate a multiple linear regression in which student achievements are measured by Rasch test scores in Reading and Mathematics while parents' background is measured by country of birth, educational attainment and occupational status proxies. I control for other student traits and for school and territorial characteristics which are likely to affect educational performance.

The chapter is organized as follows. Section 2 provides an overview of the Italian educational system. Section 3 reviews the literature on the impact of family background on student performance, with a special focus on previous researcheas employed in the context of Italy. Section 4 describes data and variables. Section 5 presents the estimation model and discusses methodological issues while section 6 provides empirical results. Robustness checks are provided in Section 7. The chapter ends with concluding remarks in Section 8.

1.2 Overview of the Education System in Italy

The Italian Constitution requires the state to offer a public school system to all citizens and permits the co-existence of state and non-state schools[1]. Overall responsibility for education lies with the Ministry of Education, Universities and Research (MIUR), which operates centrally and is responsible for organizing the various education levels, public schools, and curricula. At the local level, regional school offices delegate responsibilities to provincial and municipal authorities. In 1999, the School Autonomy Reform introduced a degree of decentralization that delegated a number of important administrative and management functions to schools. Schools have autonomy with regard to didactics, organization, research, experimentation, and development[2].

The most recent school reform, starting from 2008, has reorganized the education system, which currently is structured into preprimary education followed by two education cycles[3]:

- ***Preprimary education***: this stage enrolls for children ages 3–6 and is not compulsory. It can be divided in Daycare Centres (for children 0-3 years old) and Nursery School (for children 3-6 years old).

[1] Costituzione della Repubblica Italiana (1947), Art.34 [Constitution of the Italian Republic]. Legge del 3 Febbraio 2006, n. 27, "*Conversione in legge, con modificazioni, del decreto-legge 5 dicembre 2005, n.250, recante misure urgenti in materia di università, beni culturali ed in favore di soggetti affetti da gravi patologie, nonché in tema di rinegoziazione di mutui*;" Art. 1-bis. Norme in materia di scuole non statali [Rules on non-state schools].

[2] Decreto del Presidente della Repubblica del 8 Marzo 1999, n. 275, *"Regolamento recante norme in materia di autonomia delle istituzioni scolastiche"* [Regulation on the school autonomy].

[3] Legge del 30 Ottobre 2008, n. 169, "*Disposizioni urgenti in materia di istruzione e università"* [Urgent measures about instruction and university]; Decreto del Presidente della Repubblica del 20 Marzo 2009, n. 89, "*Revisione dell'assetto ordinamentale, organizzativo e didattico della scuola dell'infanzia e del primo ciclo di istruzione"* [Reform of the organization of the pre-primary school and of the first cycle of education]; Decreto del Presidente della Repubblica del 15 Marzo 2010, n. 87 e 88, "*Regolamenti di riordino dei licei, degli istituti tecnici e degli istituti professionali"* [Regulation about the reform of lyceums, technical institutes and vocational training institutes].

- ***First cycle of education:*** this cycle lasts eight years, it is compulsory and is divided into two levels: primary education (lasting five years) for students ages 6–11; and lower secondary education (lasting three years) for students ages 11–14.
- ***Second cycle of education:*** this cycle lasts five years, the first two of which are compulsory, and includes two possible types: upper secondary and initial vocational training. Upper secondary education is governed by State (is under the jurisdiction of the MIUR), lasts five years, and is for students ages 14–19. These schools include lyceums, technical institutes, and vocational training institutes. There are six different types of lyceum, eleven different types of technical institutes, and six types of vocational training institutes. Initial vocational training, governed by regional authorities in regional and private vocational training centers, lasts three years and is for young people (ages 14–16) who have completed the first cycle of education.

Students move from primary to lower secondary school on the basis of a positive evaluation at the end of their final year of primary education (Grade 5); there is no state examination at this level.

At the end of lower secondary school (Grade 8), all students take a state examination. If they obtain an overall grade higher than six out of ten, they obtain the certificate needed to enter upper secondary education.

Passing the upper secondary school state examination is required for access to higher education in universities and higher education courses (for example, AFAM – Higher Education in Art and Music).

At both primary and secondary education levels, periodic assessments of student learning are carried out by teachers, and students receive numerical grades on a ten-point scale based on these assessments[4]. At the primary level, student grades are accompanied by written analytical assessments. At the end of each school year, the teachers of each class meet in a class council and assign final grades to each student. A grade of six out of ten (equivalent to "satisfactory") is the minimum passing grade.

Students also receive periodic and annual evaluations of their conduct, which also is expressed as a numerical grade on a ten-point scale. If the class council gives a conduct grade lower than six, the student can not advance to the next grade; if such a grade is given in Grade 8, the student cannot take the final lower secondary school examination.

Students with learning difficulties identified before the fourth grade have personalized study plans and engage in remedial activities in class during normal lesson times. For students younger than eighth grade who have learning difficulties, MIUR establishes personalized study plans and afternoon remedial courses that the student would take with their own class teacher.

The language of instruction is Italian, which is the official language of Italy. The state recognizes and safeguards twelve linguistic minorities found in certain regions of the country[5], but only four of these are

[4] Decreto-Legge del 1 Settembre 2008, n. 204, *"Disposizioni urgenti in materia di istruzione e università;" Art. 2 e 3: Valutazione del comportamento degli studenti; Valutazione del rendimento scolastico degli studenti* [Student assessment].

[5] Albanian, Catalan, German, Greek, Slovenian, Croatian, French, Provencal, Occitan, Friulan, Ladin, and Sardinian. To ensure the learning of these minority languages, Italian schools have the autonomy to determine their own approaches to teaching the language and cultural traditions of local communities, though this also depends on requests from students' parents.

legally recognized[6]: French (in the region of Valle d'Aosta), Slovenian (in Friuli-Venezia Giulia); German (in the province of Bolzano), and Ladin (in Trentino-Alto Adige and the autonomous province of Trento). The Statute of Trentino-Alto Adige, for example, requires that instruction in schools in the province of Bolzano be conducted in German for students who have German as their mother tongue, by teachers whose mother tongue also is German. In statefunded schools in the region of Valle d'Aosta, the Statute of the Valle d'Aosta provides that the number of hours per week dedicated to French instruction be equal to those dedicated to teaching Italian, and that certain other subjects also may be taught in French.

Teachers of all school levels currently receive their initial training at universities. Primary school teachers are generalists, even if they acquire responsibility for a certain disciplinary field. Secondary school teachers are specialists and, starting from the lower secondary level (Grade 6), it is mandatory to have a degree related to the subject taught. In order to teach Mathematics or science, teachers must have either a Mathematics or science degree (e.g., biology, life science, or geology), respectively. New education programs introduced with Decree 249/2010 include the acquisition of English language skills to intermediate proficiency (the B2 examination level) and of digital technology skills.

Primary and lower secondary schools are generally equipped with ***instructional materials and tools*** to support teaching the various school subjects; nonetheless, the textbook is still the main instructional tool used. Beginning in the 2011–12 school year, schools will be required to adopt textbooks usable exclusively in a downloadable or mixed media format.

Primary ***school textbooks*** are free of charge for families, and MIUR establishes a limit for expenditures for lower secondary school student textbooks. Schools are encouraged to set up laboratories and other specially equipped spaces such as libraries, gymnasiums, and science and music labs; and schools are responsible for purchasing instructional materials, instruments, and equipment, according to each school's budget.

In recent years, significant investments have been made in the Italian school system to promote and develop ***technology use*** in education. There have been various information and communication technology (ICT) interventions at the national level that aim to promote and disseminate "best practices" for using technology. These interventions have included students, but mostly have involved teachers (e.g., ICT teacher training organized nationally). Many Italian schools have a computer room, often with an Internet connection and sometimes a portable or installed Interactive Multimedia Whiteboard (IMW)[7].

---

[6] Legge del 15 Dicembre 1999, n. 482, *"Norme in materia di tutela delle minoranze linguistiche storiche"* [Norms with regard to the safeguard of historical linguistic minorities].

[7] The IMW is a white panel the size of a normal blackboard that is connected to a video projector and a computer. Any writing or drawing on the whiteboard can be saved on a computer, printed, placed on the school website, or sent by e-mail to colleagues or students unable to attend class. The IMW is particularly useful for students with specific learning disorders and for the integration of non-EU students, hospitalized students, and students with restricted freedom of movement.

1.3 Literature review

Since the Coleman Report (Coleman et al., 1966), numerous studies have shown that student socio-economic background is strongly associated with academic achievements. Existing literature is very wide so the overview presented in this Section is not exhaustive. Anyway, a brief review of the cross-country analysis and country case studies is provided. A special focus is then given to the previous studies on the impact of family background on Italian students' achievements, as Italy is the focus of the present research.

Regarding ***cross-country analysis***, interesting is the recent contribute by Perry and McConney (2013) who study school socioeconomic status and student outcomes in Reading and Mathematics through a comparison of Australia and Canada; they find that the relationship is substantially stronger in Australia than in Canada.

Another comparation between two countries is provided by Wang (2004). He compares family background factors and Mathematics success of Hong Kong and US students, finding that Hong Kong students outperformed their US counterparts in Mathematics scores and that Hong Kong has advantages in half and US about one-fifth of the family background factors.

Then, Chiu and Xihua (2008) analyze family and motivation effects on students' Mathematics achievement 41 countries. Marks (2006) provides evidence from 30 countries on the between- and within-school differences in student performance due to socio-economic background. Marks et al. (2006) explain socioeconomic inequalities in student achievement in 30 countries, pointing out the role of home and school factors.

Baker et al. (2002) show a cross-national analysis on the relationship between socio-economic status, school quality and national economic development; 36 developed and developing nations are involved in the study.

Fuchs and Wößmann (2008) examine in 31 countries tested in Reading literacy to prove that international differences in student performance are due to the family background too. In the same year, Nonoyama-Tarumi (2008) provides a cross-national estimates of the effects of family background on student achievements, 40 countries are included in the analysis.

Wößmann contributes in the literature with three studies. In the first one, he focuses on 39 countries to provide international evidence on the relationship between schooling resources, educational institutions and student performance (Wößmann, 2003). The second one, provides evidence on the impact of family Background and Student Achievement in United States and in 17 Western European Countries (Wößmann, 2004), while in the third he studies the impact of family background and schooling policies on student performance in 5 East Asia economies (Wößmann, 2005).

With reference to the individual ***country case studies***, many studies on the impact of family background on student achievement are conducted in the context of *United States of America*. For instance, Perl (1973) studies the relationship between family background, secondary school expenditure, and student ability; Datcher-Loury (1989) focuses on family background and school achievement among low income blacks; Tate (1997) provides evidence on some determinants in Mathematics achievement: race-ethnicity, socioeconomic status, gender, and language proficiency. Then, Rumberger and Palardy (2005) demonstrate the impact of student socio-economics background on academic achievement in high school; Davis-Kean (2005) demonstrates the influence of parents' education and family income on US's student achievement, demonstrating the existence of an indirect role of parental expectations and the home environment.

Significant contributions are also related to the Australian context. See, for example, Rothman (2003) or, more recently, the contributions by McConney and Perry which provide evidence that school socioeconomic status (SES) is consistently associated with substantial increases in science and Mathematics performance in Australia (McConney and Perry, 2010a, 2010b; Perry and McConney 2010a, 2010b).

Other empirical evidences are then provided for Belgium (Opdenakker and Van Damme, 2001), Scotland (Willms, 1986), Sri Lanka (Aturupane et al., 2013), Philippines (Bernardo, 2009), Sweden (Björklund et al., 2003), Louisiana (Caldas and Bankston, 1997), Hong Kong (Ming Chiu and Sui Chu Ho, 2006), Canada (Chow, 2004), Uganda (Currie, 1977), Germany (Fertig, 2003), Chile (Gubbins et al., 2006), Czech Republic (Matěakejů and Straková, 2005), Great Britain (Currie and Duncan, 2001).

Finally, some studies are related to the context of Italy. Next Section provides a detailed review.

1.3.1 A new contribute among previous studies in the context of Italy

Since the 2000s, several studies have analyzed the impact of socioeconomic and family background on students' achievement in Italy. Some of these studies have point out territorial divide across macro-areas in terms of students' performance.

The first contribute dates 2004, when Checchi analyzes the distribution of competences in Italy in 2000, using data of the OECD Programme for International Student Assessment (PISA). With particular emphasis on the skills of the 15-year-olds students, he shows that they differ significantly by type of secondary school attended and by geographical area of residence: skills are higher in lyceums and in the North of Italy. The author also estimates the determinants that underlie the process of skills training. At the individual level these appear largely attributable to the family, described not only through the parents' education, but also through books in the home, the presence of cultural activities, participation in family discussions and support received from relatives for homework. Checchi (2004) therefore shows that the distribution of skills acquired reflects for the most part behind the family environment, which is also responsible for the choice of different secondary schools for their children.

Checchi and Peragine (2005) provide a methodology to measure opportunity inequality, comparing two Italian macro-regions, South and Centre-North, and using two applications. The first one studies the effect of the family background on individual earnings, and is directly connected to the human capital approach; the second one analyzes the effect of family background on the distribution of cognitive abilities among students, and therefore refers to inequalities existing before the access to the labor market. The analysis is based on data from PISA 2000, a survey conducted to assess Reading ability of 15-year-olds students.

According to the results, parents' education play a great role in the level of individual achievements, and this effect is stronger in the South than in the North. So, the main findings are that the less developed regions in the South are characterized by greater disparities at the global level and by greater incidence of inequality of opportunity.

Bratti et al. (2006), using data from PISA 2003, estimate the impact that a set of variables produces on the level of Mathematical skills of 15-year-olds students in Italy. The variables are grouped into three blocks: individual variables, related to the student, his behavior and the family of origin; variables at the school level; territorial variables.

From the analysis of the first group of variables, it is clear that students with high skills levels are the children of parents who hold occupations prestigious, have books and computers in the home, as well as being provided with other durable goods.

With reference to the variables at school level, it is observed that the teaching style that informs the relationship between teachers and students of authoritarian character seems to characterize better student performance, and that an indirect measure of the resources available at the school level, given by the number of Internet-connected computers in the school, is positively correlated with students' performance.

Finally, at the level of territorial variables, the only significant effect of resources on students' skills seem to be associated to equipment and buildings, while not reflected spending on staff. The authors also found significant effects relating to the situation of the labor market: the probability of employment is highly correlated with students' performance. Furthermore, there are negative effects related to housing and the presence of foreign nationals. Contrary to what was expected, the large urban centers seem to exert a negative impact.

Bratti et al. (2007) investigate the existence and the size of territorial differences in 15-year-olds Italian students' Mathematical competencies. Their analysis use a data set that merges the 2003 wave of the OECD PISA with territorial data collected from several statistical sources and with administrative school data collected by the Italian Ministry of Education. Authors consider three different groups of educational inputs: individual characteristics (mainly family background), school types and available resources, and territorial features related to labor market, cultural resources and aspirations.

In addition to the standard gradient represented by parental education and occupation, they find that student sorting across school types also plays a significant role. Among the local factors measured at province level, they find a significant impact of buildings maintenance and employment probabilities. Finally, they find that most of the North-South divide (75%) is accounted for by differences in endowments, while the remaining share of variance is related to different school processes across regions.

Montanaro (2008) collects results from the main national (INVALSI 2005-2006) and international surveys (PIRLS 2001, PIRLS and TIMSS 2003) to illustrate the differentiated territorial nature of the Italian education system. After confirming the North-South divide in terms of educational effectiveness, the author highlights the factors that may explain these disparities: a significant portion of the differences between North and South is attributable to students from disadvantaged families.

Braga and Checchi (2010) evaluate the effectiveness of regional training systems through the study of the evolution of Italian regional disparities skills in transition from primary to secondary education. They use PIRLS 2001 and PISA 2006 data, focusing on the common area of the two detection: Reading skills.

The main finding are three. First, is confirmed the duality of the Italian school system with the southern regions that record a worse performance. Second, regional disparities do not seem to recede during the school career. Finally, in regions with a positive value added[8] there is a mitigation of impacts of the characteristics of context and background on skills training.

Checchi and Radaelli (2010) analyze the factors that influence the choice of secondary school during the three editions of the PISA survey (2000, 2003, 2006) and then study the relationship between family environment and acquisition of skills in 15-year-olds Italian students. The authors also provide evidence of the skills gap for macro-geographic area and type of secondary school, demonstrating major skills in the lyceums of Nord-East area.

With reference to the choice of the type of secondary school, it appears the increase in the number of students enrolled in lyceums between 2000 and 2006. To capture the impact of the family environment in which the choice of the type of school has matured, the authors consider the level of academic qualification and occupational prestige associated with the parent "more educated", and an index that captures the presence of educational resources at home. In all three years considered, the family background is statistically significant at 1%. It is also proved the hypothesis of the relationship between family environment and students' performance in the years 2000, 2003 and 2006.

Since in PISA data there is the problem of the absence of information about income, Checchi and Radaelli (2010) impute to the families of PISA 2006 survey, a family income from another data set – EUSILC 2004. Then, they repeat the regression analysis for 2006 by introducing the variables "disposable household income" and "liquidity constraints", demonstrating that the latter influence the choice of secondary school and student performance.

Benadusi et al. (2010) address the issue of school equity in Italy, using PISA 2006 data and presenting two types of analysis. In the first, they provide an overall view of the performance of 15-year-olds students of upper secondary school, showing that the highest scores were obtained by students who attend lyceums in the North of Italy. The second analysis focuses on the determinants of performance in science: it appears that the influence of family background – as measured by ESCS – on the results is greater in the Centre, lower in the North, with the South and the Islands in the middle position.

Berchialla et al. (2011) employ the data on 11 and 13 years old Italian students sampled in the 2009/10 edition of the international survey Health Behavior in School-aged Children (HBSC) to investigate the determinants of the probability to be held back in school. They find that both at the first and at the third grade of Italian lower secondary school, the typical profile of those who show a delay in the course of study is that of a boy, born to parents with low educational attainment and income and, in most of the cases, with

[8] Value Added is measured by the difference in competences between the secondary school and primary school.

an immigrant origin[9]. Delays are also associated with risk behaviors such as drinking alcohol and smoking. They find mixed evidence across grades as regards the impact of physical wellbeing and that of possible conflicts with parents, while the class climate (proxied by the quality of relationship with peer students) shows no effect on the probability of being held back. Matching at the class level the HBSC data with the INVALSI test scores in Italian language and Math, Berchialla et al. (2011) show that delays in the course of study and low achievement share the same socio-cultural determinants.

Ferrer-Estaban (2011) analyzes the effect of territorial inequalities on educational opportunities in the Italian lower secondary school. Through hierarchical regression models, which allow both observing the heterogeneity between provinces, and accounting for structural and economic differences between macro-areas, Ferrer-Estaban (2011) sets out some aggregated factors influencing academic performance: beyond the traditional North-South differences, cross-province factors such as social heterogeneity between classes within schools, social segregation of schools, and the rate of teachers in precarious employment, are observed to adversely affect the Reading scores of students.

De Simone (2013) provides new evidence on the learning divides in the final grade of Italian lower-secondary school. He investigates the determinants of cognitive achievement in Math and science as measured by test scores at grade 8 of the 2007 edition of the Trends in International Mathematics and Science Study (TIMSS). In order to circumvent cumulative effects of education, De Simone (2013) employs a pseudo-panel approach to link achievements of the cohort of 8th-graders in 2007 with those of the same cohort of students in the 2003 edition of TIMSS, when they were enrolled in the grade 4. This allowed to distinguish the responsibilities of primary and lower-secondary education in generating the learning gaps observed just before the selection into upper-secondary school tracks.

Results reveal that the gender gap in Math observed at the grade 8 should actually be ascribed to primary education, while responsibilities on the gap in science are shared by the two school levels. On the other hand, in both subject, the largest part of the learning divide due to family background originates at the lower secondary school. We also find that, although foreign-origin students are more prone than their native peers to be held back, they show a spectacular recovery at the lower-secondary school, once the entry level of competence is taken into account.

From 2004, therefore, several contributions exist in the literature panorama, until the recent paper by De Simone, dated 2013. From the analysis of literature, we can see that researchers used different instruments and data collection to demonstrate that family background affects students' performance in Italy.

In Table 1 I report a detailed overview of previous studies in the context of Italy, with the aim to point out the most used data and school levels involved. In Table 2 I outline the most family background and students' performance measures used.

[9] For *school delay* means the phenomenon that concerns students enrolled in classes lower than their chronological age. Students who lag behind in school are not only those who were not admitted to the next class or to the next grade (the so-called 'rejected'). Within this category includes those who, for various reasons, were enrolled in classes in which the average age of students is less than their other: generally it comes to first-generation immigrants and young people with cognitive delay or special psico and physical disorders.

Following empirical findings of previous studies, in this study I estimate the relationship between parents' background and students achievements in Italian primary schools, adding a new contribute to literature in two important ways. First, I employ a much broader national sample: the majority of previous researches uses data from international surveys – PISA, PIRLS or TIMSS – which include between 4.000 and 8.000 students every year per each grade. I use, instead, INVALSI data as I consider them ideal data because the sample size is very large for each cohort group, allowing a higher precision in estimates.

Second, this is the first attempt to exploit information on educational achievements collected for the same cohort of students in primary school over two repeated cross-sections. Basic idea is to consider the cohort of pupils attending the 2nd grade in 2008/2009, who are substantially tha same pupils attending the 5th grade in 2011/2012. But, in comparing the two cohorts of students, a sample size issue emerges. In fact, while in 2008/2009 INVALSI tests were not mandatory for all schools, starting from the next scholastic year the assessment became compulsory for all educational institutions so for 2011/2012 I can employ the entire universe of pupils attending the 5th grade. This means that schools voluntarily participated in assessment of the 2nd grade in 2008/2009 are selected and, as a consequence, the effect of parents' background on scores that I find could be affected by this kind of selection bias. As robustness checks I replicate findings by using pupils attending the 2nd grade in 2011/2012. Results lead to very similar effects of parents' socio-economic status on students' performance, suggesting not only that the approach of considering two cohorts of students over two repeated cross-sections can be a suitable substitute for system-level analysis when longitudinal data are unavailable, but in particular providing evidence that the impact of family background on scores is strong independently from sample size and schools involved in INVALSI assessment.

1.4 Data and related issues

1.4.1 Data source and sample

I collected data from the National Institute for the Educational Evaluation of Instruction and Training (INVALSI – Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione), an Italian Institute having the task to carry out regular and systematic assessments on the knowledge and skills of students with the aim to improve the quality of the education system[10].

Since 2008/2009 school year, the INVALSI, through the National Service for the Evaluation of Education and Training (SNV – Servizio Nazionale di Valutazione), has carried out a systematic survey on students' performance every academic year. For 2008/2009 scholastic year, the national assessment involved pupils

[10] Decreto-Legge del 19 Novembre 2004, n. 286, *"Istituzione del Servizio nazionale di valutazione del sistema educativo di istruzione e di formazione, nonché riordino dell'omonimo istituto, a norma degli articoli 1 e 3 della Legge 28 Marzo 2003, n. 53"* [Establishment of the national evaluation of the education and training, as well as namesake school reorganization, in accordance with articles 1 and 3 of Law 28 March 2003, n. 53]

attending the 2nd and the 5th year of primary school, in which was added the 1st and the 3rd year of lower secondary school in 2009/2010 school year, and the 2nd year of upper secondary school in 2010/2011[11].

The evaluation survey focuses on Reading and Mathematics competences, and tests are elaborated according to curricular objectives of each grade, as well as taking into account several frameworks from international evaluation surveys (PIRLS and PISA for Reading, and TIMSS, PISA and NCTM for Maths). Reading competences in Italian are structured into three main sections: 1) Reading comprehension of a narrative text, 2) Reading comprehension of expository text, and 3) grammatical knowledge and skills. The contents of the Mathematics test are divided instead into four areas: 1) Numbers, 2) Space and figures, 3) Data and forecasts, 4) Relations and functions. The last area is not subject of evaluation at the 2nd grade of primary school, where the test is limited to the first three.

The present research focuses its attention both on Reading and Mathematics competences.

The choice to use INVALSI data rather PISA or TIMSS data come from the sample size: INVALSI datasets contain more observations than other two international surveys. In INVALSI assessment, in fact, are involved all schools and all classes – with exception of the academic year 2008/2009 for which participation in survey was voluntary – while in PISA and TIMSS assessments are involved random samples (between 4.000 and 8.000 students every year per each grade). So, INVALSI data are considered ideal data because the sample sizes are very large for each cohort group, allowing a higher precision in estimates.

The INVALSI databases are single cross-sections of observations from different grades and school years. This does not allow to estimate panel data models, but only cross-section estimations per each grade.

As one purpose of this research is to verify if the impact of family background on educational performance perists or is attenuated during the school career, cross-section regressions are estimated and compared. Specifically, in the basic idea to ”follow” the same students during primary schools, the study sample consists of the cohort of students attending the 2nd grade in 2008/2009, who are substantially the same students attending the 5th grade in 2011/2012. But, while in 2008/2009 participation in survey was voluntary and each school decided whether to join or not, for the school year 2011/2012 the National survey involved all schools and all classes. Consequently, 2008/2009 and 2011/2012 assessments do not involve the same number of students. Tha sample of the 2nd grade (2008/2009) covers in fact about 150.000 pupils while data of 5th grade (year 2011/2012) cover around 490.000 pupils. Hence, a selection problem arise for the 2nd grade and it’s evident that this issue could bias the effect in which I’m interested in.

To tackle this problem I provide, as robustness, further empirical evidence by employing data on pupils attending the 2nd grade in 2011/2012 (about 480.000 observations). In this way I can better “compare” two educational grades having more homogeneous size samples. However, if I find the same magnitude of the impact of parents’ background on educational outcomes, I can assert the effect actually exists and is strong independently from sample size and schools involved in assessment.

Going back to the numerosity of samples, I would like to underline, moreover, that in primary school, given the age of pupils involved, tests are taken in two separate days to avoid the fatigue effect, so, both at

[11] For the third class of the lower secondary school is taken into account the evaluation of learning faced by students on the occasion of the State examination at the end of the first cycle.

the 2nd and 5th grade, the number of observations in Reading achievements dataset differs from that in Mathematics achievements dataset. This is certainly due to the absence of students in two days in which tests are performed.

INVALSI data contains useful information on pupils participating in assessment, for example students' gender, students' country of birth, regularity in the studies, parents' country of birth, educational attainment and occupational status. Datasets also include school and territorial variables. At geographical level, for each student observed, macro-area and region information are provided. At school level, I have data on school weekly hours and *sample school*. The presence of a wide number of variables allow to study the effect of family background on pupil' achievements controlling for students, school and geographic characteristics which are likely to influence educational outcomes.

For each year and grade, *sample schools* have been identified. In these schools, tests administration is conducted in the presence of an external observer in order to ensure the correctness of administration and their execution. In this way, INVALSI try to control opportunistic behaviors that allow pupils to provide correct answers not by virtue of their skills, but because copied from other students or from books and other sources (*student cheating*) or, even, more or less explicitly suggested by teachers (*teacher cheating*).

The proper conduct during the test, i.e. compliance with the protocol of administration, is so essential to ensure that answers given by students can be considered as truly reliable and, therefore, indicative of their effective skills.

1.4.2 Educational Outcomes Variables and Covariates

As dependent variables I use ***student achievements both in Reading and Mathematics***, measured by ***Rasch test scores***. Rasch test scores are useful measures because they consider both the skills of the student and the difficulty of the item. Indeed, the *Rasch model*, is a psychometric model for analyzing categorical data, such as answers to questions on an assessment or questionnaire responses, as a function of the trade-off between a) the respondent's abilities, attitudes or personality traits and b) the item difficulty (Rasch, 1960).

The possibility of ordering both the difficulty of items and the ability of students is very important from an interpretation point of view because it allows to understand which and how many are students who show learning levels above or below the difficulty of a given question or a set of questions, and consequently to understand what these students know or are able to do.

With regard to explanatory variables, 4 groups of covariates are identified. The first one refers to ***students characteristics***, and includes the following dimensions: *gender, country of birth, regularity in the studies*. The second group of variables is represented by ***school characteristics***: *sample school and school weekly hours* are the dimensions. The third group refers to ***parents' background***, and includes *country of birth, educational qualification and occupational status of the father and the mother*. Concerning explanatory factors related to the fourth group, ***territorial characteristics***, the empirical model consider the *macro-geographical area* dimension.

For each dimensions, I generated one or more binary variables and included them in the regression model.

Therefore, all explanatory variables are dichotomous, while only the dependent variable – Rasch test score – is a continuous measure. For details of explanatory dummy variables, see Table 3.

The descriptive statistics of educational outcome variables are presented in Table 4. Frequencies of covariates by grade and subject of assessment are instead reported in Table 5.

On average, test score achieved in 2008/2009 by students attending the 2nd grade is 0.033 in Reading, and 0.346 in Mathematics. Test score in Reading ranges from -5.390 to 3.622 while test score in Mathematics ranges from -4.654 to 4.038.

Rasch test score in Reading in 2011/2012 at 5th grade is 0.101 points on average with a range from -5.824 to 4.112. The average test score in Mathematics, instead, is -0.333 points with a range from -4.722 to 4.875.

Students' gender is substantially balanced. In all years involved in the analysis, males are approximately just 1% more than females. Almost all students are Italian (between 82% and 90% in each year of the sample, per each grade). A very small percentage (between approximately 1% and 2%) are born in an European Union country, or in another European country or elsewhere. Moreover, students are mostly in regularity with their studies: the percentage of regular students is in fact between 83% and 95%.

Regarding to school characteristics, school weekly hours is up to 30 hours for more than half of schools. The school is a *sample school* approximately in 70% of case in 2008/2009, while in 2011/2012 the percentage reaches 94%.

The majority of pupils assessed have parents born in Italy. Father's educational qualification is lower secondary certificate for 21-25% of students, while upper secondary school diploma is the qualification of 27-30% of students' fathers. The situation of mothers is opposite: a higher percentage (between 25-29%) of pupils have a mother having an upper secondary school diploma, followed by lower secondary school diploma (the percentage is between 24-25%). Few parents have a university degree or a postgraduate qualification, a vocational secondary school diploma and, very small percentages, have a primary school certificate or another qualification higher than diploma such as Fine Arts Academy and Conservatory.

Looking at parents' occupational status, in each year considered, 22-23% of fathers are laborers or members of cooperatives, 15-16% are teachers, employers or militaries in career, 14-16% are self-employed workers and 7-10% are professional employee or freelancers. Less than 10% of the pupils' fathers have another occupational status. About 30% of mothers are homemakers, between 17-19% are laborers or members of cooperatives. All others occupational status detect few percentages for mothers.

Finally, concerning the macro-geographical area, in 2008/2009, about 47% of students assessed live in the South and Islands, followed by 36% who live in the North and by 17% in the Centre.

In 2011/2012, instead, 43% of students live in North, followed by 39% who live in South and Islands and by 18% in the Centre.

An important issue emerging from Table 4 and 5 is the presence of *Missing Values*.

In fact, all explanatory variables contain missing values, with the only exception of variables related to dimensions *Sample school* and *Macro-geographical area*. The percentage of missing data ranges from

4.35% to 33.09%; the only negligible percentages of missing values (between 0.05% and 0.22%) are those relating to *Gender* and *Regularity in the studies* in dataset of the year 2011/2012 (see Table 5).

With regard to dependent variables – Rasch test score in Reading and Mathematics – INVALSI datasets do not include missing values in 2008/2009 while in 2011/2012 the percentage is 5.06% in Reading dataset and 6.10% in Mathematics dataset (see Table 4).

1.5 Estimation model and Conceptual Framework

To analyze the relationship between students' achievements and their parents' background, I estimate a multiple linear regression model, controlling for students, schools and territorial characteristics which are likely to affect students' performance. The regression function has the following form:

$$Y_{ics} = \beta_0 + \beta_1 ST_{ics} + \beta_2 SC_{ics} + \beta_3 B_{ics} + \beta_4 T_{ics} + \varepsilon_{ics} \qquad (1)$$

where Y is the Rasch test score either in Reading or in Mathematics of student i in the class c in the school s; ST, SC, T are groups of student, school and territorial characteristics variables respectively; B is a group of parents' background proxies, and ε is an error term.

β_0 is the intercept and $\beta_1, \beta_2, \beta_3, \beta_4$ are the parameter vectors to be estimated in the regression.

I run four separate regressions, as I consider both Reading and Mathematics Rasch test scores and two educational grades – 2nd grade in the scholastic year 2008/2009 and 5th grade in 2011/2012.

Figure 1 provide a conceptual framework of the regression model. I expect a relationship between each group of variables included in the model and students' achievements, better identified as follows.

With reference to *Student Characteristics*, I first expect that differences in *gender* reflect differences in Reading and Mathematics scores. According to some previous studies, I wait for a higher score in Reading for females while for a higher score in Mathematics for males. Then, I attend that non-native speaker students experience an educational disadvantage, in the following terms: *i)* Italian students perform better than non-native speaker students, *ii)* students born in the European Union or in another European country non EU perform less well than other students born in another continent. I expect to find interesting findings about *regularity in the studies,* which allow to understand if pupils "in advance" get a higher o lower score than those who are regular.

As proxies of *School Characteristics*, I control for *school size*, expecting a negative relationship with student performance, and for *school weekly hours*, but in this case the direction in which the effect goes could be twofold. On the one hand, I would expect that a higher weekly hours corresponds to a greater performance of students but, on the other hand, too many hours could reduce attention and pupils learning level. To taking into account the *cheating* phenomenon, I control for the variable *Sample school,* indicating the schools in which the test is taken in the presence of an external observer. In this way I control for the opportunistic behaviors of students and teachers which could invalidate scores and then the effective individual skills.

Finally, I am interested in checking the sign and the significance of territorial dummy variables. Albeit is well know the territorial divide between the Centre-North and the South of Italy in terms for instance of GDP, employment rate, and so on, I would like to verify if it also reflects students skills.

The group of variables of main interest remains *Parents' Background.* I would like to prove that students with a higher socio-economic background perform better than those having a low socio-economic status. As proxies for socio-economics status, I use country of birth, educational qualification and occupational status of parents which could strongly affect their children educational achievements. First, parents of students from some ethnic backgrounds may have more difficulty to assist their children's schooling than other parents. I assume that: *i)* if parents are Italian, students perform better; *ii)* if parents are born in UE or in another European country, their children get higher scores than those having parents born in another continent. Second, more educated parents get certainly more educated children. In other words, educational level achieved by students' parents is strongly positively related to students' educational performance: students with parents having at least a lower secondary school certificate get better scores than those with parents having a primary school certificate. And finally, parents with a higher *Occupational status* may provide their offspring a better learning environment both in terms of educational resources such as books and computer at home or other home possessions, and in terms of the possibility to attend extra-school cultural activities.

1.5.1 Handling missing data

As seen in Tables 4 and 5, variables of INVALSI datasets include a substantial amount of missing data. Missing values are an important issue that has to be faced carefully in order to reduce estimation bias. There are many different methods to handling missing data. Some of the most popular methods involve ad-hoc deletion or replacement of missing data. These methods typically edit missing data to produce a complete data set and are attractive because they are easy to implement. However, these methods have serious drawbacks (Little and Schenker, 1995; Graham and Hofer, 2000; Graham et al., 1997; Schafer and Graham, 2002). For example, handling missing data by eliminating cases with missing data (*listwise deletion* or *complete case analysis*) will bias results if the remaining cases are not representative of the entire sample. This method is the default in most statistical software.

In the case of the present research, dropping all students with a missing value on at least one variable would delete the information available on the other explanatory variables for these students, and it would introduce bias.

Another very simple approach is to replace missing values with the sample mean. While popular, ***mean imputation*** produces distributions that have far too many cases at the mean: because the same value is being substituted for each missing case, this method artificially reduces the variance of the variable in question. More importantly, mean imputation can often produce estimates that are more biased than those from complete case analysis (Little and Rubin, 2002).

Another possibility is to perform a ***conditional mean imputation***, that is, rather than imputing the sample mean, use the mean from cases that are similar to the case with the missing values in important ways. Replacing missing values with predicted values from a regression analysis of the complete data is a form of conditional mean imputation. What these methods of imputation have in common is that the imputed values are completely determined by a model applied to the observed data, in other words, they contain no error. This tends to reduce variance, and can distort relationships among variables.

An alternative approach is to incorporate some error into the imputed values. The values imputed in imputation are draws from a distribution, in other words, they inherently contain some variation. A limitation of ***single imputation*** is that it treats imputed values as though they were observed, which is not the case, imputations are only estimates. As a result, standard analyses of a single imputation will tend to overstate our confidence in the parameter estimates, that is, the standard errors are too small.

Multiple imputation addresses this problem by introducing an additional form of error based on variation in the parameter estimates across the imputations, so called, between imputation error.

In multiple imputation, missing values for any variable are predicted using existing values from other variables. The predicted values, called "imputes", are substituted for the missing values, resulting in a full data set called an "imputed data set." This process is performed multiple times, producing multiple imputed data sets (hence the term *multiple imputation*). Standard statistical analysis is carried out on each imputed data set, producing multiple analysis results. These analysis results are then combined to produce one overall analysis.

In order to handle missing data in the present research, multiple imputation method is used because it represents a good balance between quality of results and easy of use.

Moreover, literature consider this approach the best. Graham et al. (2003), for instance, refer to traditional methods (listwise deletion, mean imputation, conditional mean imputation and single imputation) as "unacceptable methods".

The performance of multiple imputation in a variety of missing data situations has been well-studied and it has been shown to perform favorably (Graham et al., 1997; Graham and Schafer, 1999; Schafer and Graham, 2002). Multiple imputation has been shown to produce unbiased parameter estimates which reflect the uncertainty associated with estimating missing data (for detail, see Rubin, 1987, 1996).Taking into account for missing-data uncertainty, the method does not underestimate the variance of estimates like single imputation methods. Further, multiple imputation has been shown to be robust to departures from normality assumptions and provides adequate results in the presence of low sample size or high rates of missing data.

Regarding to the number of imputed datasets to be generated, the recommendation is for 3 to 5 but the choice depends on the specific case (see Graham, Olchowski and Gilreath, 2007, for a detailed discussion). To estimate the model of the present research, I generated 5 imputations.

1.6 Empirical Results

<u>*Basic Results*</u>

Multiple imputation estimation results are reported in Table 6. I find a strong impact of parents' background on students' achievements both in Reading and in Mathematics.

First, I provide evidence that students with parents born in Italy or in a country of the European Union perform better than students with parents born in another continent. Then, more important, all regression coefficients of *parents' educational qualification* variables are positive and statistically significant at level of 1%. If parents' qualification is higher than the primary school certificate, students gain an advantage in test scores both in Reading and in Mathematics. Moreover, coefficients are higher in higher qualifications. This confirm that the educational level achieved by parents is strongly positively related to their children educational achievements. Results are statistically significant both for father and for mother educational qualification, with coefficients slightly higher for father proxies.

The effect of the *parents' occupational status* is also strong. If parents work, their children gain an educational advantage than in the case of unemployed parents. Regression parameters are higher for higher occupational status (e.g. manager, university lecturer, officer, professional employee or freelancer, teacher, employee, military in career). Just one coefficient, instead, goes in the "wrong" direction: students having the mother employed as entrepreneur or landowner perform lower in Mathematics at the 5th grade than students having an unemployed mother.

The joint significant tests for parents' background dimensions – country of birth, educational qualification and occupational status – provide evidence that there is a statistically significant relationship between parents' background and students' educational achievements measured by Rasch test scores in Reading and Mathematics. In fact, the probabilities of the *F statistic* for all parents' background variables are <0.01.

The impact of socio-economic status seems to persist during primary school. Coefficients of all parents' background variables are statistically significant at the level of 1% both at the 2nd and at the 5th grade, and the size of parameters is generally slightly increasing over the two school level considered.

<u>*Other results*</u>

Results show that students' family characteristics have a strong effect on educational performance. But also students' personal, school and territorial characteristics are related to achievements.

With regard to *gender gap* boys lag behind girls in Reading test scores both at the 2nd and 5th grade, while in Mathematics boys perform better at both grades. Holding the effects of all other explanatory variables constant, on average, males perform 0.052 points lower than females in Reading test score at the 2nd grade, and 0.178 points lower at the 5th grade. In Mathematics test, instead, males perform on average 0.064 points higher than females at the 2nd grade and 0.095 points higher at the 5th grade. All coefficients are statistically significant at level of 1%. We can note the both in Reading and in Mathematics, the gender gap increase over the years. At the level of significance of 1% *students born in Italy* or in the European Union get a higher score than students born in another continent, both at the 2nd and 5th grade. Gap rises over the grades. Regression coefficients associated to students born in the EU are not statistically significant in the Reading and Mathematics test at 2nd grade.

I obtain interesting results with regard to *Regularity in the studies*. Not only students in delay seem to get worse at school, but also students “in advance”. This means that youngest students in the classroom perform lower than regular students. Coefficients are statistically significant at the level of 1%. For example, on average, students in advance attending the 2nd grade in 2008/2009, get 0.139 points lower than regular students in Reading and 0.065 points less in Mathematics. The difference reaches 0.156 points in Reading test for students in advance attending the 5th grade in 2011/2012 while for Mathematics I find no statistically significant coefficient at the 5th grade.

As proxies for school characteristics, three explanatory variables are included in the regression model. First, I control for school size, finding a positive parameter for both grades and test subjects. The second control is the variable *Sample school*, which shows a negative statistically significant coefficient at the level of 1% in all regressions. This means that in schools in which the tests are carried out in the presence of an external observer, students get lower scores both in Reading and in Mathematics, and both at the 2nd and 5th grade.

If *school weekly hours* is from 31 to 39 hours or equal to 40 hours, students perform worse than those attending school up to 30 hours. This is true for both grades but in Reading test scores only. Regarding to achievements in Mathematics, instead, students attending an extended or full time schooling get higher scores than students attending regular school weekly hours. Only exception is represented by students attending the 5th grade in 2011/2012, for which coefficient of dummy variable “from 31 to 39 hours” is negative.

Results seems do not confirm the traditional and severe *territorial divide* in favor of North and Centre of Italy in the educational system. Students living in the North and attending the 5th grade seems to perform better than those living in the South and Islands both in Reading and in Mathematics. At the 2nd grade, this advantage for northern Italy results in Reading test scores only. In Mathematics, instead, students living in the North seems to perform lower than southern students: holding the effects of all other explanatory variables constant, on average, pupils living in the North get 0.377 points less than southern students in Mathematics.

Students living in the Centre of Italy perform better than southern students only in Reading and at the 5th grade. Regression results for Mathematics, 5th grade, and for both subject, 2nd grade, show instead a disvantage for students living in the Centre respect to southern pupils. In the Centre of Italy, students get 0.016 points less in Reading than in the South and Islands. The magnitude is higher in Mathematics: pupils from the Centre get 0.351 points less at the 2nd grade and 0.319 points less at the 5th grade than those living in the southern Italy. Almost a reversal, then, of the territorial divide when considering the students’ Mathematics performance as a measure of divide.

1.7 Robustness Checks

1.7.1 A first robustness: listwise deletion estimates and differences among Italian schools

To support the reliability of estimates on the causal impact of parents' background on students' achievements, I present some robustness checks to basic analysis.

First, I replicate results without handling missing data, i.e. employing the listwise deletion. Second, I include School Fixed Effects to control for unobserved students characteristics that differ between school and influence student outcomes. In this way, the basic model (1) becomes as follows:

$$Y_{ics} = \beta_0 + \beta_1 ST_{ics} + \beta_2 SC_{ics} + \beta_3 B_{ics} + \beta_4 T_{ics} + FE_s + \varepsilon_{ics} \qquad i,c,s=1,\ldots,n \qquad (2)$$

Eq. (2) differ from the basic regression model for the inclusion of a set of School Fixed Effects (FE_s).

Considering pupils attending the 2nd grade in 2008/2009 I don't have information on classroom code and provices so in estimates presented in this Section I include School Fixed Effect only. A futher robustness will be provided afterwards by using data on students attending the 2nd grade in 2011/2012.

If parents' background variables are correlated with unobserved determinants of achievement that are constant within schools and differ between schools – such as school environment, school reputation, teacher ability – then the inclusion of fixed effects will improve bias in estimates of β_3. So, controlling for FE_s I can check if previous estimates are biased or if, indeed, they are reliable.

I replicate results with School Fixed Effects for both listwise deletion and multiple imputation models.

Table 7 shows findings on the impact of parents' background on Reading test scores while Table 8 provides results for Mathematics test scores. Covariates for student, school and territorial characteristics are included in the regressions and go in the expected direction so are not reported for simplicity.

Both Tables present results from four specifications per grade. Column (1) and (2) report listwise deletion estimates, without and with School FE respectively; column (3) and (4) provide multiple imputation estimates, without and with School FE respectively.

Additional specifications (1) (2) and (4), presented as robustness checks, confirm that basic estimations reported in column (3) are reliable. Alternative regressions, in fact, lead to very similar results. The sign, the significance and also the magnitude of parents' background variables are consistent to those of previous estimates so I avoid to present comments for each specification. I would like to assert, instead, that this additional evidence lends strong support to the causal impact of parents' background on individual achievements both in Reading and in Mathematics. This impact seems to persist during primary school.

1.7.2 Is the impact of parents' background on scores affected by selection at the 2nd Grade?

Basic idea of analysis presented in this chapter is to evaluate whether parents' background affects educational outcomes and whether this impact persists during primary education. If so, the role of both governments and schools in reducing social disparities at first stage of education becomes relevant, as educational achievement differences translate into income gap over life.

I use data on pupils attending the 2nd grade in 2008/2009, who represent the same cohort of pupils attending the 5th grade few years after, in 2011/2012. But, as already discussed in previous Sections, in 2008/2009 the participation in INVALSI tests was voluntary so sample cover a smaller number of schools than in 2011/2012, when assessment became mandatory for all schools and classrooms. Hence, a selection problem arise for the 2nd grade and parents' background effect on students' scores that I find could suffer from selection bias.

To tackle this issue, in this Section I present additional empirical evidence to check and to confirm that family socio-economic background actually affects individual performance. I employ data on pupils attending the 2nd grade in 2011/2012. First, I check the sign and the significance of parents' background proxies. Then, I compare coefficients with those of regressions runned for the 5th grade of school year 2011/2012. In this way I can better "compare" two educational grades having more homogeneous size samples and argue whether parents' background effects on scores reduce or persist during primary school.

There is more. Considering data of 2011/2012, I have information on both classrooms codes and provinces so I can include in regressions, in addition to School Fixed Effect, also Classroom and Provincial Fixed Effects. Through these additional specifications I can also check if both differences within schools – i.e. between classrooms – and between provinces matter in determining the impact of students' family socio-economic status on their educational performance.

I report results in Tables from 9 to 12. Specifically, Tables 9 and 10 present estimates on Rasch test scores in Reading at the 2nd and 5th grade respectively, whereas Tables 11 and 12 show findings for Rasch test score in Mathematics. Each Tables present estimations without and with multiple imputations. For each I present results with no fixed effects (column 1), with school FE (column 2), with classroom FE (column 3) and finally with provincial FE (column 4).

I find the same magnitude of the impact of parents' background on educational outcomes both not considering differences between schools, classrooms and provinces, and taking into account School, Class and Province FE. Hence, I can assert that: *i)* the effect actually exists; *ii)* it is strong independently from sample size and schools involved in assessment; *iii)* it is relevant independently from differences between and within schools and across provinces. Parents' background effect does not fade away but persists during primary education.

1.7.3 Estimating family socio-economic background on pupils' scores in a subsample of schools controlled for "cheating"

In this Section I report some robustness checks to take into account the so-called "cheating" phenomenon. In all previous estimates I control for Sample Schools representing educational institutions in which INVALSI tests took place in the presence of an external observer. If on the one hand pupils could cheat by consulting books or others educational source and by suggesting answers among themselves, on the other hand teachers could suggest the correct answers to pupils and/or allow them to collaborate during the test. The presence of cheating affects the quality of results risking improperly to “bloat” test scores of those who behave incorrectly.

The precence of an external observer could guarantee the correctness of the test and hence a reliable representation of students learning levels. But few schools only are selected to control for cheating. Hence, a way to check if results are biased by incorrect behaviors during the test could be replicate estimates in *Sample schools* only. These schools do not reveal incorrect behaviors (see INVALSI Report, 2012). For schools no sample, instead, INVALSI found anomalies related to cheating phenomenon, especially in some regions of the South of Italy (Molise, Campania, Calabria e Sicilia) and also in the Centre (Lazio) (see INVALSI Report, 2012). However, cheating is reduced in the last few years, also thanks to an information campaign carried out in these regions, in partnership with Ministry of Education, Universities and Research (MIUR).

To check the reliability of estimates, I replicate results for the subsample of schools controlled for cheating. In Table 13 I report findings for basic estimation model, i.e. for basic idea to “follow” the same cohort of students across primary schools. Specifically, I present results of parents’ background effects on Rasch test score in Reading for pupils attending the 2nd grade in 2008/2009 and the 5th grade in 2011/2012, showing that the impact of parents’ background on test scores is also strong in *Sample schools* (see Table 13). Findings for Mathematics and for sample of pupils attending the 2nd Grade in 2011/2012 lead to very similar coefficients and statistical significance. Hence, I can assert that previous results are not invalidated by cheating phenomenon and I can confirm that parents’ socio-economic status is an important determinant of educational achievements of Italian pupils, with a higher impact of father proxies.

Finally, to check findings concerning the territorial divide in scores, I report coefficients of macro-geographical area variables estimated in the subsample of schools controlled for cheating, in which incorrect behaviors seem do not occur. Results for students attending the 2nd grade in 2008/2009 and the 5th grade in 2011/2012 are reported in Table 14 and are consistent with those commented in Section 6 and presented in Table 6. In Reading, students living in the North of Italy achieve higher test scores both at the 2nd and 5th grade whereas in Mathematics pupils living in the South and Islands perform better than northern students at the 2nd grade but not at the end of primary schools. For students living in the Centre, test scores are higher in Reading but lower in Mathematics with respect to southern pupils (see Table 14).

When I replicate regressions focusing on pupils attending both the 2nd and the 5th grade in 2011/2012, I obtain an advantage for northen students both in Reading and in Mathematics at the 5th grade. At the 2nd school level, instead, pupils living in the North of Italy perform better than southern in Reading but not in Mathematics, consistently with previous findings. Also for pupils living in the Centre, Rasch test score in Mathematics are lower with respect to those living in the South and Islands (see Table 15).

To conclude, I can argue that at first years of primary school (2nd grade), southern pupils perform better in Mathematics than those living in the North of Italy. They achieve higher Mathematics scores at the 5th grade too with respect to pupils living in the Centre. The advantage in Reading scores is, instead, for northen students in both grades.

1.8 Concluding remarks and policy implications

In recent years, great attention has been devoted (and many resources have been invested) in the measurement of educational outcomes based on standardized tests with the aim to evaluate the effectiveness and efficiency of the different educational systems. In most of classifications drawn up based on results of international surveys on students' skills (PISA, TIMSS, PIRLS) Italy is always in rather low positions.

We might be tempted to attribute the poor performance of students to the presence of a lacking cultural environment: if these students living in households with poor education, they do not receive enough support and family pressures to achieve good results at school. Moreover, parents with a high occupational status have more resources to provide a better environment for their children to do well in school.

This chapter provide an analysis of Italian students' achievements in Reading and Mathematics tests. Results are in line with the existing evidence on the impact of family background on students' performance.

The main findings are as follows. Regarding to gender gap in test scores, Italian boys lag behind girls at all grades in Reading while they perform better in Mathematics. Students regular in their studies get better than those in delay. An increment of school weekly hours results in a lower individual performance.

With reference to territorial characteristics, North-South territorial divide in educational performance in favor of the North of Italy seems do not fully take place. While in Reading pupils living in the North perform better in both grades of primary school than southern students, in Mathematics pupils living in the South and Islands get higher scores with respect to northern pupils at the 2nd grade, and higher scores than those living in the Centre at the 5th grade.

The educational level of parents is the most fundamental factor in explaining the child's success at school. In fact, the educational attainment achieved by students' parents is strongly positively related to students' educational performance. The effect of the parents' occupational status is also strong.

The study, moreover, proves that the impact of parents' background on students' achievements does not reduce during the primary school, and in particular between the 2nd and the 5th grade, but persist.

Results are robust as confirmed by using both reduced and full sample of pupils attending primary education in Italy. The impact of parents' background on individual test scores is relavent independently from sample sizes and from differences between and within schools and across provinces. Moreover, findings seems not to be biased by cheating phenomena.

This analysis focuses on learning divides across social-groups in primary school showing that intergenerational educational persistence and social immobility originates in the early stages of the schooling process. Students with an advantaged family background perform better in Reading and Mathematics at grade 2 and 5. This certainly translates into social inequalities along upper secondary school and then in labor market.

There is no doubt that, in recent years, the analysis of intergenerational mobility and the role of family background in forming human capital accumulation in the next generations has become a very active research field in economics. Given the central role of students' educational performance for the future economic prospects of societies, the evidence presented through this research may reveal interesting aspects for educational and social policies in Italy. There are, indeed, clear policy implications if socio-economic and educational family resources prove to be largely responsible for socioeconomic inequalities in education.

A causal relationship between better educated parents and children indicates schooling externalities, and may have distributional consequences as well. If inherited abilities drive the academic success of children in school, then inequality in opportunity would merely be a reflection of the existing gene pool, leaving scant room for pro-education policies. If, on the other hand, parents' education is primarily responsible for the child's success in school, then improving the educational achievement would not only increase education and reduce the inequality in educational opportunity for future generations, but also affect their level and distribution of income. Thus, the causal intergenerational effect of schooling is informative about spill-over effects and indicates a broad range of returns to educational investments, and the implications for public policy are therefore enormous (Holmlund et al., 2008).

So, schools should fight the huge social disparities in terms of education opportunities and improve social mobility. But also governments have to take action. For example, they could increase funding to students with a low socio-economic status providing financial support in the form of scholarships, allowances for textbooks and other educational materials, and high tax deductions for educational expenses.

To conclude, designing better strategies to promote life – long learning is an important policy issue since it may enhance social inclusion and, at the same time, it can reduce marginalisation of segments of the population and increase socio-economic cohesion (Braga and Checchi, 2008).

References

Ashenfelter O., Harmon C. and Oosterbeek H. (1999), A Review of Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias, *Labour Economics,* 6(4): 453-470.

Aturupane H., Glewwe P. and Wisniewski S. (2013), The impact of school quality, socioeconomic factors, and child health on students' academic performance: evidence from Sri Lankan primary schools, *Education Economics*, 21(1): 2-37.

Baker D., Goesling B. and Letendre G. (2002), Socio-economic Status, School Quality and National Economic Development: A Crossnational Analysis of the 'Heyneman-Loxley Effect' on Mathematics and Science Achievement, *Comparative Education Review*, 46(3): 291-312.

Barro R.J. (2001), Human capital and growth, *American Economic Review*, 91(2): 12-17.

Becker G. (1964), *Human Capital. A Theoretical and Empirical Analysis, with Special Reference to Education*, NBER: Chicago.

Benadusi L., Fornari R. and Giancola O. (2010), La questione dell'equità scolastica in Italia. Analisi delle performance in scienze dei quindicenni della scuola secondaria superiore, *FGA Working Paper* No. 26.

Berchialla P., Cavallo F., Colombini S., De Simone G. and Lemma P. (2011), Associazione delle caratteristiche individuali, del benessere psico-fisico e del clima di classe con gli outcomes scolastici a 11 e 13 anni, *FGA Working Paper* No. 41.

Bernardo A.B.I. (2009), Filipino students' reported parental socialization of academic achievement by socioeconomic group, *Psychological Reports*, 105: 427-436.

Bishop J.H. (1992), The Impact of Academic Competencies on Wages, Unemployment, and Job Performance, *Carnegie-Rochester Conference Series on Public Policy*, 37: 127-194.

Björklund A., Lindahl M. and Sund K. (2003), Family Background and School Performance During a Turbulent Era of School Reforms, *Swedish Economic Policy Review*, 10(2): 111-136.

Björklund A. and Salvanes K.G. (2010), Education and family background: Mechanisms and policies, *IZA Discussion Paper* No. 5002.

Blossfeld H.P. and Shavit Y. (1993), Persisting barriers: Changes in educational opportunities in thirteen countries, in Shavit Y., Blossfeld H.P., *Persistent inequality*, Boulder, CO: Westview.

Boocock S.S. (1972), *An introduction to the sociology of learning*, Boston: Houghton-Mifflin.

Braga M. and Checchi D. (2010), Sistemi scolastici regionali e capacità di sviluppo delle competenze. I divari dalle indagini Pirsl e Pisa, *La Rivista delle Politiche Sociali*, 3: 1-25.

Bratti M., Checchi D. and Filippin A. (2006), Il Mezzogiorno nella conoscenza. Competenze degli studenti italiani quindicenni, *Rivista delle Politiche Sociali*, 4: 93-112.

Bratti M., Checchi D. and Filippin A. (2007), Territorial Differences in Italian Students' Mathematical Competencies: Evidence from PISA 2003, *Giornale degli Economisti e Annali di Economia*, 66(3): 299-335.

Caldas S.J. and Bankston III C. (1997), Effect of school population socioeconomic status on individual academic achievement, *Journal of Educational Research*, 90(5): 269-277.

Card D. (1999), The Causal Effect of Education on Earnings, in: Ashenfelter O., Card D., *Handbook of Labor Economics*, 3A: 1801-1863, Amsterdam: Elsevier.

Checchi D. (2004), Da dove vengono le competenze scolastiche?, *Stato e Mercato*, 72: 413-453.

Checchi D. and Peragine V. (2005), Regional Disparities and Inequality of Opportunity: The Case of Italy, *Journal of Economic Inequality*, 8: 429-450.

Checchi D. and Radaelli S. (2010), Scelte scolastiche e ambiente familiare, in Checchi D. (2010), *Immobilità diffusa. Perché la mobilità intergenerazionale è così bassa in Italia*, Il Mulino.

Chiu M.M. and Xihua Z. (2008), Family and motivation effects on Mathematics achievement: Analyses of students in 41 countries, *Learning and Instruction*, 18(4): 321-336.

Chow H.P.H. (2004), The Effects of Ethnic Capital and Family Background on School Performance: A Case Study of Chinese-Canadian Adolescents in Calgary, *The Alberta Journal of Educational Research*, 50(3): 321-326.

Coleman J., Campbell E., Hobson C., McPartland J., Mood A., Weinfeld F. and York R. (1966), *Equality of educational opportunity*, Washington, DC: US Government Printing Office.

Collins C., Kenway J. and McLeod J. (2000), Factors Influencing the Educational Performance of Males and Females in School and their Initial Destinations after leaving School, *Report* Commissioned by the Analysis and Equity Branch of the Department of Education, Training and Youth Affairs, Commonwealth of Australia.

Commission of Inquiry into Poverty (1976), *Poverty and education in Australia. Fifth main report*, (R.T. Fitzgerald, Commissioner), Canberra: AGPS.

Connor H. and Dewson S. (2001), *Social class and higher education: Issues affecting decisions on participation by lower social class groups*, London, UK: HMSO.

Currie J. (1977), Family Background, Academic Achievement and Occupational Status in Uganda, *Comparative Education Review*, 21(1): 14-28.

Currie J. and Duncan T. (2001), Early Test Scores, School Quality and SES: Longrun Effects on Wage and Employment Outcomes, *Research in Labor Economics*, 20: 103-132.

Datcher-Loury L. (1989), Family Background and School Achievement among Low Income Blacks, *The Journal of Human Resources*, 24(3): 528-544.

Davis-Kean P.E. (2005), The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment, *Journal of Family Psychology*, 19(2): 294-304.

De Simone G. (2013), Render unto Primary the Things which are Primary's. Inherited and Fresh Learning Divides in Italian Lower Secondary Education, *Economics of Education Review*, 35: 12-23.

Ferrer-Esteban G. (2011), Beyond the traditional territorial divide in the Italian Education System. Effects of system management factors on performance in lower secondary school, *FGA Working Paper* No. 43.

Fertig M. (2003), Who's to blame? The determinants of German students' achievement in the PISA 2000 study, *IZA Discussion Paper* No. 739, Institute for the Study of Labor, Bonn.

Fertig M. and Schmidt C.M. (2002), The role of background factors for Reading literacy: straight national scores in the PISA 2000 study, *IZA Discussion Paper* No. 545, Institute for the Study of Labor, Bonn.

Fuchs T. and Wößmann L. (2008), What accounts for international differences in student performance? A re-examination using PISA data, *The Economics of Education and Training Studies in Empirical Economics*: 209-240.

Gubbins F.V., Dois C.A. and Alfaro S.M. (2006), Factores que influyen en el buen rendimiento escolar de niños y niñas que viven en condiciones familiares de pobreza.

Haveman R. and Wolfe B. (1995), The Determinants of Children Attainments: A Review of Methods and Findings, *Journal of Economic Literature*, 33(4): 1829-1878.

Hedges L.V., Laine R.D. and Greenwald R. (1994), Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes, *Educational Researcher,* 23(3): 5-14.

Hong H.K. (2012), Trends in Mathematics and science performance in 18 countries: Multiple regression analysis of the cohort effects of TIMSS 1995-2007, *Education Policy Analysis Archives*, 20(33): 1-23.

Jencks C., Smith M., Acland H., Bane M.J., Cohen D., Gentis H., Heynes B. and Michelson S. (1972), *Inequality: A Reassessment of the Effect of Family and Schooling in America*, New York: Basic Books.

Kane T.J. (2004), College-coing and inequality, in Neckerman K.M., *Social inequality*, New York: Russell Sage.

Karmel P. (1973), *Schools in Australia. Report of the Interim Committee*, Canberra: AGPS.

Lee J.B. (1999), How do students and families pay for college?, in King J.E., *Financing a college education: How it works, how it's changing*, Phoenix, AZ: American Council on Education.

Mare R.D. (1992), Trends in the process of social stratification, *Contemporary Sociology*, 21(5): 654-658.

Marjoribanks K. (1979), *Families and their learning environments: An empirical analysis*, London: Routledge and Kegan Paul.

Marks G.N. (2006), Are between- and within-school differences in student performance largely due to socio-economic background? Evidence from 30 countries, *Educational Research*, 48(1): 21-40.

Marks G.N., Cresswell J. and Ainley J. (2006), Explaining socioeconomic inequalities in student achievement: The role of home and school factors, *Educational Research and Evaluation*, 12(2): 105-128.

Matěakejů P. and Straková J. (2005), The role of the family and the school in the reproduction of educational inequalities in the post-Communist Czech Republic, *British Journal of Sociology of Education*, 26(1): 17-40.

McConney A. and Perry L.B. (2010a), Science and Mathematics achievement in Australia: the role of school socioeconomic composition in educational equity and effectiveness, *International Journal of Science and Mathematics Education*, 8(3): 429-452.

McConney A. and Perry L.B. (2010b), Socioeconomic status, self-efficiency, and Mathematics achievement in Australia: a secondary analysis, *Educational Research for Policy and Practice*, 9(2): 77-91.

Ming Chiu M. and Sui Chu Ho E. (2006), Family Effects on Student Achievement in Hong Kong, *Asia Pacific Journal of Education*, 26(1): 21-35.

Montanaro P. (2008), I divari territoriali nella preparazione degli studenti italiani: evidenze dalle indagini nazionali e internazionali, *Occasional Paper* No. 14, Questioni di Economia e Finanza, Banca d'Italia.

Neckerman K.M. and Torche F. (2007), Inequality: Causes and consequences, *Annual Review of Sociology*, 33: 335-357.

Noel S. and De Broucker P. (2001), Intergenerational inequities: A comparative analysis of the influence of parents' educational background on length of schooling and literacy skills, in Hutmacher W., Cochrane D., Bottani N. (Eds.), *In pursuit of equity in education: Using international indicators to compare equity policies* (pp. 277–298), Dordrecht: Kluwer Academic.

Nonoyama-Tarumi Y. (2008), Cross-National Estimates of the Effects of Family Background on Student Achievement: A Sensitivity Analysis, *International Review of Education*, 54(1): 57-82.

Opdenakker M.C. and Van Damme J. (2001), Relationship between school composition and characteristics of school process and their effect on Mathematics achievement, *British Educational Research Journal*, 27(4): 407-432.

Ordine P. and Rose G. (2012), The effect of family background, university quality and educational mismatch on wage: an analysis using a young cohort of Italian graduates, *Education Economics*, iFirst Art: 1-25.

Organisation for Economic Co-operation and Development (2004), *Learning for tomorrow's world: First results from PISA* 2003, Paris, France: Author.

Organisation for Economic Co-operation and Development (2005), *School factors related to quality and equity: Results from PISA 2000*, Paris, France: Author.

Organisation for Economic Co-operation and Development, (2007), *PISA 2006: Science competencies for tomorrow's world*, vol. 1. Paris, France: Author.

Perie M., Moran R. and Lutkus A.D. (2005), *NAEP 2004 trends in academic progress: Three decades of student performance in Reading and Mathematics*, Washington, DC: US Department of Education.

Perl L.J. (1973), Family background, secondary school expenditure, and student ability, *Journal of Human Resource*, 8: 156-180.

Perry L.B. and McConney A. (2010a), Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003, *Teachers College Record*, 112(4): 1137-1162.

Perry L.B. and McConney A. (2010b), School Socio-Economic Composition and Student Outcomes in Australia: Implications for Educational Policy, *Australian Journal of Education*, 54(1): 72-85.

Perry L.B. and McConney A. (2013), School socioeconomic status and student outcomes in Reading and Mathematics: A comparison of Australia and Canada, *Australian Journal of Education*, 57(2): 124-140.

Rasch G. (1960), *Probabilistic models for some intelligence and attainment tests*. Oxford, England: Nielsen & Lydiche.

Renzulli J.S. and Park S. (2000), Gifted dropouts: The who and the why, *Gifted Child Quarterly*, 44: 261-271.

Roemer J.E. (1998), *Equality of Opportunity*, Cambridge, MA: Harvard University Press.

Rothman S. (2003), The changing influence of socioeconomic status on student achievement: recent evidence from Australia, *Paper presented at the American Educational Research Association*, Chicago.

Rubin D.B. (1987), *Multiple Imputation for Nonresponse in Surveys*, New York: Wiley.

Rubin D.B. (1996), Multiple imputation after 18+ years, *Journal of the American Statistical Association*, 91:473-489.

Rumberger R.W. and Palardy G.J. (2005), Does segregation still matter? The impact of student composition on academic achievement in high school, *Teachers College Record*, 107(9): 1999-2045.

Schafer J.L. (1997), *Analysis of Incomplete Multivariate Data*, Boca Raton, FL: Chapman & Hall/CRC.

Schütz G., Ursprung H.W. and Woessmann L. (2005), Education policy and equality of opportunity, *IZA Discussion Papers* No. 1906.

Sirin S.R. (2005), Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3): 417-453.

Solon G. (1999), Intergenerational Mobility in the Labor Market, in: Ashenfelter O., Card D., *Handbook of Labor Economics*, 3A: 1801-1863, Amsterdam: Elsevier.

Steinmayr R., Dinger F.C. and Spinath B. (2010), Parents' education and children's achievement: The role of personality, *European Journal of Personality*, 24: 535-550.

Sui-Chu E.H. and Willms J.D. (1996), Effects of parental involvement on eighth-grade achievement. *Sociology of Education*, 69(2): 126-141.

Tate W.F. (1997), Race-ethnicity, SES, gender, and language proficiency trends in Mathematics achievement: An update, *Journal for Research in Mathematics Education*, 28(6): 652-679.

Terenzini P.T., Cabrera A.F. and Bernal E.M. (2001), *Swimming against the tide: The poor in higher education*, New York, NY: College Board.

Turmo A. (2004), Scientific Literacy and Socio-economic Background among 15-year-olds- A Nordic Perspective, *Scandinavian Journal of Educational Research*, 48(3): 287-305.

Wang D.B. (2004), Family background factors and Mathematics success: A comparison of Chinese and US students, *International Journal of Educational Research*, 41: 40-54.

White K.R. (1982), The Relation Between Socioeconomic Status and Academic Achievement, *Psychological Bulletin*, 91(3): 461-481.

Willms J.D. (1986), Social class segregation and its relationship to pupils' examination results in Scotland. *American Sociological Review*, 51(2): 224-241.

Willms J.D. (2006), Learning Divides: Ten Policy Questions About the Performance and Equity of Schools and Schooling Systems, *UNESCO Institute for Statistics Working Paper* No. 5.

Wößmann L. (2003), Schooling resources, educational institutions and student performance: the international evidence, *Oxford Bulletin of Economics and Statistics*, 65(2): 117-170.

Wößmann L. (2004), How Equal Are Educational Opportunities? Family Background and Student Achievement in Europe and the United States, *IZA Discussion paper series* No. 1284.

Wößmann L. (2005), Educational Production in East Asia: The Impact of Family Background and Schooling policies on Student Performance, *German Economic Review*, 6(3): 331-353.

Appendix of Tables and Figures

Table 1 – Overview of Previous Studies in Italy

Author	Year	Type of contribute		Language		Research Profile Studied			Data				School level		
		Paper	**Scientific publication**	**Italian**	**English**	**Territorial differences in students' competences**	**Impact of family background on students' performance**	**Impact of family backgr. on the choice of secondary school**	**PISA**	**PIRLS**	**TIMSS**	**Other**	**Primary school (Grade 4)**	**Lower Secondary school (Grade 8)**	**Upper Secondary school (15-year-olds students)**
Checchi	2004		■	■		■	■	■	■ (2000)						■
Checchi and Peragine	2005		■		■	■	■		■ (2000)						■
Bratti et al.	2006		■	■		■	■		■ (2003)						■
Bratti et al.	2007		■		■	■	■		■ (2003)						■
Montanaro	2008	■		■		■	■		■ (2003)	■ (2001)	■ (2003)	■ Invalsi 2005/06	■	■ (Grade 6 too)	■ (1^ and 3^ Secondary School)
Braga and Checchi	2010		■	■		■	■		■ (2006)	■ (2001)			■		■
Checchi and Radaelli	2010		■ (part of a book)	■		■	■	■	■ (2000 2003 2006)			■ Eusilc 2004			■
Benadusi et al.	2010	■		■		■	■		■ (2006)						■
Berchialla et al.	2011	■		■		■	■					■ Hbsc e Invalsi 2009/10		■ (Grade 6 too)	
Ferrer-Estaban	2011	■			■	■	■					■ Invalsi 2009/10		■ (Grade 6 only)	
De Simone	2013		■		■		■				■ (2003 2007)		■	■	

Table 2 – Family Background and Students' Performance Measures

Author	Year	Family Background Measures												Students' performance Measures		
		Parents' education	**Parents' occupation**	**Number of parents at home**	**Computers at home**	**Books at home**	**Family wealth[1] / Home possessions**	**Presence of cultural activities**	**Participation in family discussion / Dialogue with parents**	**Family support in homework**	**Disposable household income**	**Liquidity constraints[2]**	**(ESCS)[3]**	**Test scores in Reading**	**Test scores in Maths**	**Test scores in science**
Checchi	2004	■	■			■	■	■	■	■				■	■	■
Checchi and Peragine	2005	■												■		
Bratti et al.	2006	■	■		■	■	■								■	
Bratti et al.	2007	■	■		■	■	■								■	
Montanaro	2008												■	■	■	■
Braga and Checchi	2010	■	■		■[4]	■[4]							■	■		
Checchi and Radaelli	2010	■	■		■[4]	■[4]					■	■			■	
Benadusi et al.	2010												■			■
Berchialla et al.	2011	■		■					■					■	■	
Ferrer-Estaban	2011												■	■		
De Simone	2013	■				■									■	■

1 It is a variable reconstructed from the information provided in relation to the presence in the family of a room dedicated to the children, a dishwasher, a training software, a connection to internet, as well as the number of mobile phones, televisions, computers, cars and bathrooms at home.

2 Liquidity constraints refers to the inability to deal with unexpected expenses or a week of holidays in a year.

3 ESCS (Index of Economic, Social and Cultural status) is a synthetic index calculated by the OECD on the basis of occupational prestige of the parents, the education of parents, cultural resources whose family has and material resources owned.

4 The authors use the terms *educational resources*, *material resources*, *cultural resources*, but they are substantially related to books and computers at home.

Table 3 – Description of Variables used in the Analysis

Dependent Variables:	***RASCH TEST SCORES IN READING AND MATHS***	
Independent Variables:		
Group	*Dimensions*	*Dummy Variables*
Student characteristics	Gender	Male Female
	Country of birth	Italy European Union European Country no EU Other
	Regularity in the studies	Regular In advance In delay
School characteristics	Sample school	Sample school School no sample
	School weekly hours	Up to 30 hours From 31 to 39 hours 40 hours
Parents' background	Father's/Mother's country of birth	Italy European Union European Country no EU Other
	Father's/Mother's educational qualification	Primary school certificate Lower secondary school certificate Vocational secondary school diploma (3 years of study) Upper secondary school diploma Another qualification higher than diploma (Fine Arts Academy, Conservatory, etc.) University degree or Postgraduate qualification
	Father's/Mother's employment status	Unemployed Homemaker Manager, university lecturer, officer Entrepreneur, landowner Professional employee or freelancer (doctor, lawyer, psychologist, researcher, etc.) Self-employed worker (trader, farmer, craftsman, mechanic, etc.) Teacher, employee, military in career Laborer, services personnel, member of cooperatives Retired worker
Territorial characteristics	Macro-geographical area	North Centre South and Islands

Table 4 – Descriptive Statistics of Dependent Variables

Variable	School year	Grade	Obs	Mean	Std. Dev.	Min	Max
Rasch test score in Reading	2008/2009	2nd	153,052	0.033	1.204	-5.390	3.622
Rasch test score in Maths	2008/2009	2nd	153,092	0.346	1.216	-4.654	4.038
Rasch test score in Reading	2011/2012	5th	489,581*	0.101	1.040	-5.824	4.112
Rasch test score in Maths	2011/2012	5th	489,279*	-0.333	1.274	-4.722	4.875

* The number of observations includes observations for which there are *missing values*.
The percentage of missing values in the Reading dataset is 5.06% while in the Maths dataset is 6.10%.

Table 5 – Frequencies of Explanatory Variables

	2008/2009 – Grade 2				2011/2012 – Grade 5			
	Reading		Mathematics		Reading		Mathematics	
	Frequency	%	Frequency	%	Frequency	%	Frequency	%
Gender								
Male	66,458	43.42%	66,469	43.42%	246,024	50.25%	246,326	50.34%
Female	65,093	42.53%	65,105	42.53%	242,546	49.54%	242,661	49.60%
Missing Values	21,501	14.05%	21,518	14.06%	1,011	0.21%	292	0.05%
Country of birth								
Italy	126,024	82.34%	126,037	82.33%	439,101	89.69%	438,705	89.66%
European Union	1,951	1.27%	1,954	1.28%	10,220	2.09%	10,224	2.09%
European Country no EU	1,122	0.73%	1,123	0.73%	6,512	1.33%	6,526	1.33%
Other	1,414	0.92%	1,416	0.92%	11,089	2.26%	11,183	2.29%
Missing Values	22,541	14.73%	22,562	14.74%	22,659	4.63%	22,641	4.63%
Regularity in the studies								
Regular	127,505	83.31%	127,521	83.30%	463,574	94.69%	463,828	94.80%
In advance	1,941	1.27%	1,943	1.27%	8,453	1.73%	8,486	1.73%
In delay	2,048	1.34%	2,052	1.34%	16,487	3.37%	16,653	3.40%
Missing Values	21,558	14.09%	21,576	14.09%	1,067	0.22%	312	0.06%
Sample school								
Sample school	109,753	71.71%	109,759	71.69%	458,704	93.69%	458,410	93.69%
School no sample	43,299	28.29%	43,333	28.31%	30,877	6.31%	30,869	6.31%
Missing Values	-	-	-	-	-	-	-	-
School weekly hours								
Up to 30 hours	85,626	55.95%	85,603	55.92%	317,897	64.93%	317,261	64.84%
From 31 to 39 hours	15,900	10.39%	15,910	10.39%	26,054	5.32%	25,912	5.30%
40 hours	26,005	16.99%	26,034	17.01%	124,339	25.40%	124,815	25.51%
Missing Values	25,521	16.67%	25,545	16.69%	21,291	4.35%	21,291	4.35%
Father's country of birth								
Italy	116,427	76.07%	116,439	76.06%	393,659	80.41%	393,283	80.38%
European Union	2,757	1.80%	2,761	1.80%	14,313	2.92%	14,305	2.92%
European Country no EU	3,147	2.06%	3,150	2.06%	12,773	2.61%	12,782	2.61%
Other	4,913	3.21%	4,916	3.21%	22,958	4.69%	23,136	4.73%
Missing Values	25,808	16.86%	25,826	16.87%	45,878	9.37%	45,773	9.36%
Mother's country of birth								
Italy	114,175	74.60%	114,174	74.58%	389,646	79.59%	389,212	79.55%
European Union	4,160	2.72%	4,163	2.72%	18,365	3.75%	18,390	3.76%
European Country no EU	3,913	2.56%	3,924	2.56%	14,943	3.05%	14,935	3.05%
Other	5,321	3.48%	5,329	3.48%	26,638	5.44%	26,867	5.49%
Missing Values	25,483	16.65%	25,502	16.66%	39,989	8.17%	39,875	8.15%
Father's educational qualification								
Primary school certificate	5,122	3.35%	5,120	3.34%	15,695	3.21%	15,617	3.19%
Lower secondary school certificate	42,004	27.44%	42,031	27.45%	144,466	29.51%	144,334	29.50%
Vocational secondary school diploma	8,905	5.82%	8,907	5.82%	36,819	7.52%	36,830	7.53%
Upper secondary school diploma	33,175	21.68%	33,185	21.68%	121,332	24.78%	121,010	24.73%
Another qualification higher than diploma	1,363	0.89%	1,366	0.89%	6,224	1.27%	6,218	1.27%
University degree or Postgraduate qualification	11,837	7.73%	11,839	7.73%	45,285	9.25%	45,159	9.23%
Missing Values	50,646	33.09%	50,644	33.08%	119,760	24.46%	120,111	24.55%

(Table 5 - continued on the next page)

	2008/2009 – Grade 2				2011/2012 – Grade 5			
	Reading		Mathematics		Reading		Mathematics	
	Frequency	%	Frequency	%	Frequency	%	Frequency	%
Mother's educational qualification								
Primary school certificate	4,748	3.10%	4,746	3.10%	14,370	2.94%	14,303	2.92%
Lower secondary school certificate	36,289	23.71%	36,306	23.72%	123,604	25.25%	123,459	25.23%
Vocational secondary school diploma	8,556	5.59%	8,555	5.59%	36,944	7.55%	36,970	7.56%
Upper secondary school diploma	39,091	25.54%	39,110	25.55%	140,944	28.79%	140,738	28.76%
Another qualification higher than diploma	2,194	1.43%	2,193	1.43%	10,001	2.04%	9,938	2.03%
University degree or Postgraduate qualification	13,163	8.60%	13,165	8.60%	51,049	10.43%	50,861	10.40%
Missing Values	49,011	32.02%	49,017	32.02%	112,669	23.01%	113,010	23.10%
Father's employment status								
Unemployed	4,336	2.83%	4,339	2.83%	21,440	4.38%	21,428	4.38%
Homemaker	143	0.09%	143	0.09%	4,446	0.91%	4,429	0.91%
Manager, university lecturer, officer	4,104	2.68%	4,110	2.68%	12,840	2.62%	12,763	2.61%
Entrepreneur, landowner	6,143	4.01%	6,144	4.01%	20,099	4.11%	20,017	4.09%
Professional employee or freelancer	10,809	7.06%	10,807	7.06%	48,965	10.00%	48,790	9.97%
Self-employed worker	21,314	13.93%	21,327	13.93%	80,553	16.45%	80,365	16.43%
Teacher, employee, military in career	24,371	15.92%	24,376	15.92%	72,881	14.89%	72,889	14.90%
Laborer, services personnel, member of cooperatives	34,109	22.29%	34,130	22.30%	112,483	22.98%	112,476	22.99%
Retired worker	-	-	-	-	3,685	0.75%	3,680	0.75%
Missing Values	47,723	31.17%	47,716	31.17%	112,189	22.92%	112,442	22.98%
Mother's employment status								
Unemployed	3,734	2.44%	3,740	2.44%	17,245	3.52%	17,265	3.53%
Homemaker	46,093	30.12%	46,096	30.11%	53,437	31.34%	153,282	31.33%
Manager, university lecturer, officer	1,380	0.90%	1,381	0.90%	4,885	1.00%	4,884	1.00%
Entrepreneur, landowner	1,799	1.18%	1,802	1.18%	8,365	1.71%	8,294	1.70%
Professional employee or freelancer	6,992	4.57%	6,990	4.57%	29,255	5.98%	29,156	5.96%
Self-employed worker	7,603	4.97%	7,605	4.97%	30,287	6.19%	30,230	6.18%
Teacher, employee, military in career	26,546	17.34%	26,560	17.35%	94,310	19.26%	94,110	19.23%
Laborer, services personnel, member of cooperatives	13,326	8.71%	13,340	8.72%	52,241	10.67%	52,248	10.68%
Retired worker	-	-	-	-	762	0.16%	759	0.16%
Missing Values	45,579	29.78%	45,578	29.77%	98,794	20.18%	99,051	20.24%
Macro-geographical area								
North	54,451	35.58%	54,475	35.58%	212,479	43.40%	212,097	43.35%
Centre	26,187	17.11%	26,171	17.09%	88,246	18.02%	88,386	18.06%
South and Islands	72,414	47.31%	72,446	47.33%	188,856	38.58%	188,796	38.59%
Missing Values	-	-	-	-	-	-	-	-

Table 6 – Multiple Imputation Estimation Results

	RASCH TEST SCORES			
	Reading		**Mathematics**	
Explanatory Variables	**2008/2009 Grade 2**	**2011/2012 Grade 5**	**2008/2009 Grade 2**	**2011/2012 Grade 5**
STUDENT CHARACTERISTICS				
Gender				
Male	-0.052***	-0.178***	0.064***	0.095***
(Omitted Variable: Female)				
Country of birth				
Italy	0.211***	0.248***	0.108**	0.177***
European Union	-0.041	0.099***	-0.002	0.243***
European Country no European Union	-0.100*	-0.003	-0.055	0.082***
(Omitted Variable: Other)				
F test for Country of Birth	*51.06****	*270.40****	*10.47****	*63.12****
Regularity in the studies				
In advance	-0.139***	-0.156***	-0.065**	0.065***
In delay	-0.050*	-0.390***	0.012	-0.207***
(Omitted Variable: Regular)				
F test for Regularity in the studies	*13.78****	*964.74****	*3.20***	*186.37****
F TEST for STUDENT CHARACTERISTICS	***40.93****	***1315.18****	***22.00****	***228.76****
SCHOOL CHARACTERISTICS				
School size	0.002***	0.001***	0.002***	0.006***
Sample School				
Sample school	-0.139***	-0.150***	-0.448***	-0.217***
(Omitted Variable: School no sample)				
School weekly Hours				
From 31 to 39 hours	-0.021**	-0.029***	0.034***	-0.049***
40 hours	-0.073***	-0.063***	0.016*	0.007*
(Omitted Variable: Up to 30 hours)				
F test for School weekly Hours	*39.76****	*154.99****	*5.14***	*23.04****
F TEST for SCHOOL CHARACTERISTICS	***92.00****	***287.48****	***781.59****	***4302.99****
PARENTS' BACKGROUND				
Father's country of birth				
Italy	0.263***	0.250***	0.212***	0.125***
European Union	0.183***	0.176***	0.121***	-0.018
European Country no European Union	0.010	0.084***	0.104***	-0.038*
(Omitted Variable: Other)				
F test for Father's country of birth	*63.15****	*206.57****	*30.88****	*90.13****
Mother's country of birth				
Italy	0.175***	0.155***	0.107***	0.097***
European Union	0.182***	0.097***	0.112***	-0.008
European Country no European Union	0.066*	0.010	0.032	-0.050**
(Omitted Variable: Other)				
F test for Mother's country of birth	*22.63****	*122.86****	*8.11****	*59.60****
F test for Parents' country of birth	*109.00****	*480.72****	*45.85****	*176.50****
Father's educational qualification				
Lower secondary school certificate	0.195***	0.203***	0.127***	0.120***
Vocational secondary school diploma	0.294***	0.285***	0.196***	0.136***
Upper secondary school diploma	0.372***	0.382***	0.248***	0.266***
Another qualification higher than diploma	0.360***	0.349***	0.260***	0.209***
University degree or Postgraduate qualification	0.453***	0.486***	0.300***	0.423***
(Omitted Variable: Primary school certificate)				
F test for Father's education	*90.47****	*515.66****	*45.70****	*285.29****
Mother's educational qualification				
Lower secondary school certificate	0.196***	0.225***	0.081***	0.037***
Vocational secondary school diploma	0.300***	0.334***	0.121***	0.092***
Upper secondary school diploma	0.435***	0.468***	0.206***	0.224***
Another qualification higher than diploma	0.456***	0.475***	0.196***	0.235***
University degree or Postgraduate qualification	0.580***	0.613***	0.307***	0.383***
(Omitted Variable: Primary school certificate)				
F test for Mother's education	*181.75****	*1024.18****	*58.84****	*362.01****
F test for Parents' education	*234.22****	*1209.22****	*87.45****	*526.84****

(Table 6 - continued on the next page)

Explanatory Variables	RASCH TEST SCORES			
	Reading		Mathematics	
	2008/2009 Grade 2	2011/2012 Grade 5	2008/2009 Grade 2	2011/2012 Grade 5
Father's employment status				
Homemaker	-0.056	0.119***	0.024	-0.173***
Manager, university lecturer, officer	0.260***	0.220***	0.152***	0.195***
Entrepreneur, landowner	0.121***	0.157***	0.096***	0.082***
Professional employee or freelancer	0.173***	0.187***	0.067**	0.147***
Teacher, employee, military in career	0.176***	0.202***	0.083***	0.144***
Self-employed worker	0.139***	0.161***	0.077***	0.099***
Laborer, services personnel, member of cooperatives	0.096***	0.115***	0.033	0.061***
Retired worker	-	0.126***	-	0.057**
(Omitted Variable: Unemployed)				
F test for Father's occupation	*15.94****	*86.38****	*7.50****	*52.85****
Mother's employment status				
Homemaker	0.023	0.038***	0.080***	0.059***
Manager, university lecturer, officer	0.136***	0.067***	0.104***	0.039*
Entrepreneur, landowner	0.103***	0.040**	0.086**	-0.088***
Professional employee or freelancer	0.104***	0.087***	0.073***	0.103***
Teacher, employee, military in career	0.130***	0.136***	0.113***	0.137***
Self-employed worker	0.098***	0.076***	0.091***	0.103***
Laborer, services personnel, member of cooperatives	0.024	0.039***	0.035*	0.0004
Retired worker	-	-0.035	-	0.007
(Omitted Variable: Unemployed)				
F test for Mother's occupation	*18.72****	*62.42****	*7.27****	*63.26****
F test for Parents' occupation	*19.74****	*88.17****	*8.68****	*69.93****
F TEST for PARENTS' BACKGROUND	***252.32*****	***1157.35*****	***86.62*****	***532.82*****
TERRITORIAL CHARACTERISTICS				
North	0.043***	0.039***	-0.377***	0.062***
Centre	-0.016*	0.035***	-0.351***	-0.319***
(Omitted Variable: South and Islands)				
F TEST for TERRITORIAL CHARACTERISTICS	***26.14*****	***60.99*****	***1290.79*****	***2837.82*****
Constant	-1.382***	-1.335***	-0.326***	-1.729***
Imputations	5	5	5	5
Number of Obs	153.052	489.581	153.092	489.279

Note: * p<0.1; ** p<0.05; *** p<0.01

Table 7 – Estimation Results on Rasch Test Scores in Reading

	2008/2009 – Grade 2				2011/2012 – Grade 5			
	NO Imputations		YES Imputations		NO Imputations		YES Imputations	
Explanatory Variables	(1) NO FE	(2) School FE	(3) NO FE	(4) School FE	(1) NO FE	(2) School FE	(3) NO FE	(4) School FE
PARENTS' BACKGROUND								
Father's country of birth								
Italy	0.229***	0.207***	0.263***	0.235***	0.222***	0.224***	0.250***	0.245***
European Union	0.187***	0.158***	0.183***	0.155***	0.175***	0.169***	0.176***	0.173***
European Country no European Union	0.020	-0.005	0.010	-0.016	0.086***	0.084***	0.084***	0.082***
(Omitted Variable: Other)								
F test for Father's country of birth	*30.50****	*29.95****	*63.15****	*63.02****	*113.90****	*116.60****	*206.57****	*187.38****
Mother's country of birth								
Italy	0.139***	0.165***	0.175***	0.187***	0.147***	0.143***	0.155***	0.150***
European Union	0.148***	0.181***	0.182***	0.199***	0.095***	0.089***	0.097***	0.095***
European Country no European Union	0.068*	0.103***	0.066*	0.093***	0.017	0.010	0.010	0.010
(Omitted Variable: Other)								
F test for Mother's country of birth	*9.26****	*13.46****	*22.63****	*28.79****	*73.18****	*72.12****	*122.86****	*112.79****
F test for Parents' country of birth	*44.60****	*41.79****	*109.00****	*97.12****	*250.17****	*216.44****	*480.72****	*342.60****
Father's educational qualification								
Lower secondary school certificate	0.183***	0.174***	0.195***	0.184***	0.197***	0.185***	0.203***	0.189***
Vocational secondary school diploma	0.282***	0.280***	0.294***	0.280***	0.278***	0.266***	0.285***	0.269***
Upper secondary school diploma	0.355***	0.369***	0.372***	0.363***	0.370***	0.365***	0.382***	0.368***
Another qualification higher than diploma	0.344***	0.335***	0.360***	0.341***	0.344***	0.341***	0.349***	0.341***
University degree or Postgraduate qualification	0.443***	0.473***	0.453***	0.446***	0.476***	0.482***	0.486***	0.475***
(Omitted Variable: Primary school certificate)								
F test for Father's education	*90.15****	*115.88***	*90.47****	*100.40****	*479.81****	*484.96****	*515.66****	*455.06****
Mother's educational qualification								
Lower secondary school certificate	0.193***	0.192***	0.196***	0.183***	0.228***	0.204***	0.225***	0.198***
Vocational secondary school diploma	0.307***	0.304***	0.300***	0.285***	0.334***	0.306***	0.334***	0.302***
Upper secondary school diploma	0.434***	0.441***	0.435***	0.416***	0.464***	0.442***	0.468***	0.437***
Another qualification higher than diploma	0.459***	0.470***	0.456***	0.442***	0.475***	0.446***	0.475***	0.439***
University degree or Postgraduate qualification	0.594***	0.608***	0.580***	0.555***	0.610***	0.589***	0.613***	0.578***
(Omitted Variable: Primary school certificate)								
F test for Mother's education	*179.30****	*209.71****	*181.75****	*171.75****	*852.00****	*823.57****	*1024.18****	*896.81****
F test for Parents' education	*222.01****	*235.69****	*234.22****	*209.72****	*1063.40****	*954.70****	*1209.22****	*890.65****
Father's employment status								
Homemaker	-0.030	0.020	-0.056	-0.030	0.084***	0.065***	0.119***	0.100***
Manager, university lecturer, officer	0.264***	0.232***	0.260***	0.220***	0.210***	0.194***	0.220***	0.200***
Entrepreneur, landowner	0.134***	0.110***	0.121***	0.091***	0.152***	0.140***	0.157***	0.143***
Professional employee or freelancer	0.184***	0.155***	0.173***	0.139***	0.187***	0.169***	0.187***	0.166***
Teacher, employee, military in career	0.186***	0.155***	0.176***	0.141***	0.200***	0.183***	0.202***	0.184***
Self-employed worker	0.144***	0.110***	0.139***	0.103***	0.160***	0.143***	0.161***	0.142***
Laborer, services pers., member of cooperatives	0.095***	0.064***	0.096***	0.063***	0.111***	0.094***	0.115***	0.097***
Retired worker	-	-	-	-	0.119***	0.110***	0.126***	0.116***
(Omitted Variable: Unemployed)								
F test for Father's occupation	*18.56****	*17.01****	*15.94****	*13.18****	*71.74****	*61.74****	*86.38****	*68.98****
Mother's employment status								
Homemaker	0.012	0.012	0.023	0.024	0.027***	0.034***	0.038***	0.042***
Manager, university lecturer, officer	0.107***	0.150***	0.136***	0.152***	0.055***	0.076***	0.067***	0.088***
Entrepreneur, landowner	0.081**	0.065*	0.103***	0.084**	0.026*	0.033**	0.040**	0.051***
Professional employee or freelancer	0.099***	0.115***	0.104***	0.110***	0.069***	0.077***	0.087***	0.091***
Teacher, employee, military in career	0.114***	0.121***	0.130***	0.130***	0.119***	0.123***	0.136***	0.136***
Self-employed worker	0.087***	0.086***	0.098***	0.090***	0.057***	0.061***	0.076***	0.075***
Laborer, services pers., member of cooperatives	0.013	0.020	0.024	0.028	0.025**	0.025**	0.039***	0.040***
Retired worker	-	-	-	-	-0.024	-0.041	-0.035	-0.035
(Omitted Variable: Unemployed)								
F test for Mother's occupation	*15.59****	*19.77****	*18.72****	*20.45****	*52.08****	*53.78****	*62.42****	*65.58****
F test for Parents' occupation	*19.42****	*20.08****	*19.74****	*18.46****	*72.61****	*66.85****	*88.17****	*77.59****
F TEST for PARENTS' BACKGROUND	***217.57****	***177.76****	***252.32****	***181.89****	***921.05****	***650.47****	***1157.35****	***618.84****
Number of Obs	95.969	95.969	153.052	153.052	319.288	319.288	489.581	489.581
Number of groups	-	3.758	-	5.042	-	6.057	-	7.555

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Multiple imputation estimations include 5 imputations. 3) Regressions presented in columns (2) and (4) include standard errors adjusted for clusters in School Code – i.e. for "Number of groups" as reported in the Table. 4) All specifications include student, school and territorial characteristics as defined in Table 3.

Table 8 – Estimation Results on Rasch Test Scores in Mathematics

Explanatory Variables	2008/2009 – Grade 2				2011/2012 – Grade 5			
	NO Imputations		YES Imputations		NO Imputations		YES Imputations	
	(1) NO FE	(2) School FE	(3) NO FE	(4) School FE	(1) NO FE	(2) School FE	(3) NO FE	(4) School FE
PARENTS' BACKGROUND								
Father's country of birth								
Italy	0.201***	0.162***	0.212***	0.167***	0.111***	0.092***	0.125***	0.100***
European Union	0.123***	0.094**	0.121***	0.092***	0.012	0.064***	-0.018	0.045***
European Country no European Union	0.078*	0.048	0.104***	0.055*	-0.025	0.028	-0.038*	0.010
(Omitted Variable: Other)								
F test for Father's country of birth	*20.23****	*18.63****	*30.88****	*30.32****	*42.36****	*24.11****	*90.13****	*48.96****
Mother's country of birth								
Italy	0.047	0.063***	0.107***	0.095***	0.101***	0.085***	0.097***	0.079***
European Union	0.057	0.086***	0.112***	0.104***	0.029	0.086***	-0.008	0.062***
European Country no European Union	0.010	0.075**	0.032	0.073**	-0.024	0.030*	-0.050**	0.014
(Omitted Variable: Other)								
F test for Mother's country of birth	*1.51*	*2.82***	*8.11****	*8.11****	*36.50****	*21.04****	*59.60****	*24.05****
F test for Parents' country of birth	*21.91****	*19.35****	*45.85****	*41.84****	*99.01****	*45.50****	*176.50****	*73.83****
Father's educational qualification								
Lower secondary school certificate	0.127***	0.136***	0.127***	0.118***	0.119***	0.125***	0.120***	0.106***
Vocational secondary school diploma	0.189***	0.193***	0.196***	0.177***	0.136***	0.182***	0.136***	0.151***
Upper secondary school diploma	0.246***	0.267***	0.248***	0.232***	0.269***	0.249***	0.266***	0.220***
Another qualification higher than diploma	0.248***	0.229***	0.260***	0.218***	0.231***	0.216***	0.209***	0.181***
University degree or Postgraduate qualification	0.301***	0.331***	0.300***	0.284***	0.431***	0.361***	0.423***	0.323***
(Omitted Variable: Primary school certificate)								
F test for Father's education	*41.84****	*66.00****	*45.70****	*52.29****	*276.00****	*265.97****	*285.29****	*224.31****
Mother's educational qualification								
Lower secondary school certificate	0.078***	0.090***	0.081***	0.080***	0.041***	0.105***	0.037***	0.077***
Vocational secondary school diploma	0.121***	0.139***	0.121***	0.122***	0.095***	0.147***	0.092***	0.119***
Upper secondary school diploma	0.202***	0.225***	0.206***	0.201***	0.226***	0.259***	0.224***	0.221***
Another qualification higher than diploma	0.202***	0.226***	0.196***	0.195***	0.248***	0.264***	0.235***	0.219***
University degree or Postgraduate qualification	0.306***	0.335***	0.307***	0.293***	0.388***	0.378***	0.383***	0.333***
(Omitted Variable: Primary school certificate)								
F test for Mother's education	*50.90****	*82.16****	*58.84****	*68.11****	*330.25****	*323.13****	*362.01****	*343.48****
F test for Parents' education	*76.35****	*108.58****	*87.45****	*91.21****	*503.15****	*350.42****	*526.84****	*341.24****
Father's employment status								
Homemaker	-0.012	0.021	0.024	0.072	-0.240***	0.142***	-0.173***	0.057***
Manager, university lecturer, officer	0.146***	0.160***	0.152***	0.150***	0.184***	0.173***	0.195***	0.163***
Entrepreneur, landowner	0.093***	0.079***	0.096***	0.079***	0.073***	0.112***	0.082***	0.096***
Professional employee or freelancer	0.067**	0.073***	0.067**	0.068***	0.142***	0.117***	0.147***	0.113***
Teacher, employee, military in career	0.078***	0.072***	0.083***	0.076***	0.141***	0.135***	0.144***	0.127***
Self-employed worker	0.071***	0.061***	0.077***	0.066***	0.098***	0.107***	0.099***	0.097***
Laborer, services pers., member of cooperatives	0.030	0.022	0.033	0.027	0.052***	0.061***	0.061***	0.058***
Retired worker	-	-	-	-	0.044*	0.065***	0.057**	0.070***
(Omitted Variable: Unemployed)								
F test for Father's occupation	*6.71****	*9.45****	*7.50****	*9.83****	*72.12****	*44.61****	*52.85****	*37.60****
Mother's employment status								
Homemaker	0.076***	0.056***	0.080***	0.055***	0.066***	0.024***	0.059***	0.029***
Manager, university lecturer, officer	0.093**	0.155***	0.104***	0.124***	0.045**	0.061***	0.039*	0.059***
Entrepreneur, landowner	0.064*	0.049	0.086**	0.057**	-0.091***	0.031**	-0.088***	0.016
Professional employee or freelancer	0.077***	0.099***	0.073***	0.077***	0.104***	0.042***	0.103***	0.056***
Teacher, employee, military in career	0.108***	0.124***	0.113***	0.110***	0.145***	0.099***	0.137***	0.103***
Self-employed worker	0.085***	0.099***	0.091***	0.085***	0.105***	0.066***	0.103***	0.070***
Laborer, services pers., member of cooperatives	0.036	0.042**	0.035*	0.035**	-0.003	-0.001	0.0004	0.005
Retired worker	-	-	-	-	0.056	0.012	0.007	0.007
(Omitted Variable: Unemployed)								
F test for Mother's occupation	*5.88****	*13.17****	*7.27****	*11.57****	*67.13****	*43.35****	*63.26****	*42.64****
F test for Parents' occupation	*7.41****	*12.98****	*8.68****	*12.56****	*86.49****	*49.63****	*69.93****	*46.00****
F TEST for PARENTS' BACKGROUND	***71.33****	***83.52****	***86.62****	***84.97****	***490.17****	***186.27****	***532.82****	***190.33****
Number of Obs	96.002	96.002	153.092	153.092	315.256	315.256	489.279	489.279
Number of groups	-	3.759	-	5.042	-	6.041	-	7.570

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Multiple imputation estimations include 5 imputations. 3) Regressions presented in columns (2) and (4) include standard errors adjusted for clusters in School Code – i.e. for "Number of groups" as reported in the Table. 4) All specifications include student, school and territorial characteristics as defined in Table 3.

Table 9 – Estimation results on Rasch Test Scores in Reading at the 2nd Grade (2011/2012 INVALSI Data)

Explanatory Variables	Estimations without Multiple Imputations				Estimations with Multiple Imputations			
	(1) NO FE	(2) School FE	(3) Class FE	(4) Prov FE	(1) NO FE	(2) School FE	(3) Class FE	(4) Prov FE
PARENTS' BACKGROUND								
Father's country of birth								
Italy	0.247***	0.246***	0.240***	0.248***	0.262***	0.258***	0.250***	0.263***
European Union	0.147***	0.160***	0.171***	0.146***	0.167***	0.178***	0.181***	0.169***
European Country no European Union	0.069***	0.075***	0.086***	0.068***	0.065***	0.072***	0.074***	0.064***
(Omitted Variable: Other)								
F test for Father's country of birth	*153.70****	*147.74****	*145.79****	*158.04****	*251.98****	*221.43****	*234.63****	*257.96****
Mother's country of birth								
Italy	0.214***	0.193***	0.184***	0.213***	0.215***	0.196***	0.189***	0.214***
European Union	0.127***	0.107***	0.096***	0.126***	0.113***	0.099***	0.092***	0.114***
European Country no European Union	0.057***	0.052***	0.042**	0.059***	0.047***	0.044***	0.039***	0.049***
(Omitted Variable: Other)								
F test for Mother's country of birth	*134.38****	*116.34****	*120.43****	*133.62****	*176.23****	*149.00****	*158.52****	*175.33****
F test for Parents' country of birth	*431.11****	*307.98****	*350.11****	*430.27****	*665.63****	*427.57****	*529.90****	*663.58****
Father's educational qualification								
Lower secondary school certificate	0.172***	0.167***	0.156***	0.170***	0.176***	0.169***	0.155***	0.175***
Vocational secondary school diploma	0.241***	0.231***	0.221***	0.240***	0.248***	0.236***	0.219***	0.248***
Upper secondary school diploma	0.333***	0.332***	0.323***	0.335***	0.338***	0.329***	0.311***	0.339***
Another qualification higher than diploma	0.308***	0.305***	0.295***	0.310***	0.310***	0.304***	0.288***	0.312***
University degree or Postgraduate qualification	0.429***	0.433***	0.424***	0.434***	0.434***	0.425***	0.404***	0.436***
(Omitted Variable: Primary school certificate)								
F test for Father's education	*332.93****	*361.96****	*385.90****	*346.64****	*308.35****	*314.21****	*300.84****	*325.67****
Mother's educational qualification								
Lower secondary school certificate	0.163***	0.161***	0.168***	0.155***	0.163***	0.153***	0.151***	0.156***
Vocational secondary school diploma	0.236***	0.238***	0.242***	0.229***	0.243***	0.233***	0.227***	0.237***
Upper secondary school diploma	0.372***	0.377***	0.379***	0.369***	0.382***	0.372***	0.360***	0.377***
Another qualification higher than diploma	0.391***	0.390***	0.392***	0.386***	0.396***	0.380***	0.369***	0.391***
University degree or Postgraduate qualification	0.529***	0.538***	0.535***	0.527***	0.538***	0.525***	0.506***	0.534***
(Omitted Variable: Primary school certificate)								
F test for Mother's education	*592.37****	*631.09****	*663.24****	*606.24****	*690.82****	*679.48****	*667.58****	*701.25****
F test for Parents' education	*741.58****	*683.34****	*782.53****	*758.74****	*787.33****	*672.60****	*699.52****	*798.21****
Father's employment status								
Homemaker	0.013	0.054**	0.059***	0.044**	0.040**	0.061***	0.056***	0.065***
Manager, university lecturer, officer	0.154***	0.172***	0.167***	0.154***	0.164***	0.168***	0.156***	0.163***
Entrepreneur, landowner	0.105***	0.110***	0.107***	0.101***	0.109***	0.103***	0.096***	0.105***
Professional employee or freelancer	0.129***	0.133***	0.126***	0.121***	0.129***	0.125***	0.114***	0.122***
Self-employed worker	0.123***	0.122***	0.113***	0.114***	0.126***	0.118***	0.106***	0.119***
Teacher, employee, military in career	0.166***	0.176***	0.169***	0.159***	0.163***	0.165***	0.154***	0.157***
Laborer, services personnel, member of cooperatives	0.075***	0.078***	0.072***	0.066***	0.076***	0.073***	0.065***	0.070***
Retired worker	0.090***	0.083***	0.083***	0.085***	0.082***	0.073***	0.072***	0.077***
(Omitted Variable: Unemployed)								
F test for Father's occupation	*46.57****	*52.93****	*54.19****	*43.89****	*48.74****	*48.77****	*49.84****	*47.01****
Mother's employment status								
Homemaker	0.006	0.008	0.005	0.005	0.001	0.001	0.003	0.001
Manager, university lecturer, officer	0.075***	0.101***	0.098***	0.087***	0.078***	0.096***	0.098***	0.087***
Entrepreneur, landowner	0.011	0.043***	0.043***	0.026	0.006	0.033**	0.040***	0.018
Professional employee or freelancer	0.079***	0.085***	0.079***	0.082***	0.079***	0.079***	0.077***	0.081***
Self-employed worker	0.050***	0.058***	0.052***	0.054***	0.049***	0.051***	0.049***	0.051***
Teacher, employee, military in career	0.117***	0.120***	0.114***	0.117***	0.110***	0.111***	0.109***	0.111***
Laborer, services personnel, member of cooperatives	0.009	0.015	0.012	0.009	0.008	0.009	0.013	0.007
Retired worker	0.099	0.096	0.069	0.095	0.032	0.044	0.023	0.034
(Omitted Variable: Unemployed)								
F test for Mother's occupation	*55.80****	*57.13****	*58.59****	*56.12****	*60.59****	*61.92****	*61.54****	*62.32****
F test for Parents' occupation	*61.42****	*65.37****	*65.83****	*59.66****	*65.08****	*63.51****	*64.55****	*69.05****
F TEST for PARENTS' BACKGROUND	***755.72*****	***523.82*****	***648.33*****	***762.27*****	***920.88***	***576.19*****	***687.73*****	***935.60*****
Number of Obs	326.231	326.231	326.231	326.231	480.541	480.541	480.541	480.541
Number of groups	-	6.027	22.380	103	-	7.169	27.919	103

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Multiple imputation estimations include 5 imputations. 3) Regressions presented in columns (2), (3) and (4) include standard errors adjusted for clusters in School Code, Classroom Code and Provincial Code respectively – i.e. for "Number of groups" reported in the Table. 4) All specifications include student, school and territorial characteristics as defined in Table 3.

Table 10 – Estimation results on Rasch Test Scores in Reading at the 5th Grade (2011/2012 INVALSI Data)

Explanatory Variables	Estimations without Multiple Imputations				Estimations with Multiple Imputations			
	(1) NO FE	(2) School FE	(3) Class FE	(4) Prov FE	(1) NO FE	(2) School FE	(3) Class FE	(4) Prov FE
PARENTS' BACKGROUND								
Father's country of birth								
Italy	0.222***	0.224***	0.218***	0.224***	0.250***	0.245***	0.237***	0.251***
European Union	0.175***	0.169***	0.166***	0.173***	0.176***	0.173***	0.172***	0.177***
European Country no European Union	0.086***	0.083***	0.078***	0.085***	0.084***	0.082***	0.074***	0.084***
(Omitted Variable: Other)								
F test for Father's country of birth	*113.98****	*116.64****	*120.23****	*116.62****	*206.46****	*187.07****	*187.10****	*215.65****
Mother's country of birth								
Italy	0.147***	0.143***	0.145***	0.149***	0.155***	0.150***	0.150***	0.157***
European Union	0.095***	0.089***	0.102***	0.095***	0.097***	0.094***	0.103***	0.098***
European Country no European Union	0.017	0.010	0.024	0.008	0.010	0.010	0.022	0.006
(Omitted Variable: Other)								
F test for Mother's country of birth	*73.00****	*71.83****	*69.55****	*80.47****	*122.48****	*111.99****	*110.07****	*131.11****
F test for Parents' country of birth	*249.99****	*216.10****	*231.53****	*259.54****	*479.90****	*340.96****	*376.76****	*485.07****
Father's educational qualification								
Lower secondary school certificate	0.196***	0.185***	0.172***	0.191***	0.203***	0.190***	0.177***	0.199***
Vocational secondary school diploma	0.277***	0.265***	0.252***	0.274***	0.285***	0.270***	0.254***	0.282***
Upper secondary school diploma	0.370***	0.364***	0.351***	0.368***	0.382***	0.368***	0.350***	0.379***
Another qualification higher than diploma	0.343***	0.340***	0.329***	0.343***	0.349***	0.341***	0.325***	0.348***
University degree or Postgraduate qualification	0.475***	0.481***	0.469***	0.475***	0.486***	0.475***	0.454***	0.484***
(Omitted Variable: Primary school certificate)								
F test for Father's education	*476.19****	*481.91****	*527.88****	*486.55****	*516.42****	*458.00****	*492.64****	*527.06****
Mother's educational qualification								
Lower secondary school certificate	0.228***	0.203***	0.191***	0.221***	0.225***	0.199***	0.188***	0.218***
Vocational secondary school diploma	0.334***	0.305***	0.287***	0.327***	0.334***	0.302***	0.283***	0.327***
Upper secondary school diploma	0.464***	0.441***	0.421***	0.459***	0.468***	0.437***	0.415***	0.462***
Another qualification higher than diploma	0.474***	0.445***	0.425***	0.467***	0.475***	0.440***	0.417***	0.468***
University degree or Postgraduate qualification	0.610***	0.587***	0.563***	0.605***	0.613***	0.579***	0.548***	0.606***
(Omitted Variable: Primary school certificate)								
F test for Mother's education	*848.58****	*820.26****	*819.22****	*854.58****	*910.58****	*900.19****	*924.02****	*917.56****
F test for Parents' education	*954.53****	*949.05****	*902.28****	*960.04****	*924.92****	*895.44****	*905.56****	*931.02****
Father's employment status								
Homemaker	0.084***	0.064***	0.079***	0.089***	0.119***	0.101***	0.107***	0.129***
Manager, university lecturer, officer	0.209***	0.193***	0.188***	0.199***	0.220***	0.201***	0.194***	0.213***
Entrepreneur, landowner	0.151***	0.140***	0.134***	0.139***	0.157***	0.143***	0.139***	0.149***
Professional employee or freelancer	0.186***	0.168***	0.160***	0.173***	0.187***	0.166***	0.161***	0.178***
Self-employed worker	0.160***	0.142***	0.136***	0.147***	0.161***	0.142***	0.136***	0.152***
Teacher, employee, military in career	0.199***	0.183***	0.176***	0.189***	0.202***	0.184***	0.177***	0.195***
Laborer, services personnel, member of cooperatives	0.111***	0.094***	0.090***	0.098***	0.115***	0.097***	0.094***	0.106***
Retired worker	0.118***	0.110***	0.118***	0.105***	0.126***	0.116***	0.118***	0.116***
(Omitted Variable: Unemployed)								
F test for Father's occupation	*71.21****	*61.19****	*64.13****	*66.41****	*86.41****	*69.43****	*72.57****	*75.68****
Mother's employment status								
Homemaker	0.027***	0.034***	0.034***	0.031***	0.038***	0.042***	0.041***	0.040***
Manager, university lecturer, officer	0.055***	0.076***	0.081***	0.064***	0.066***	0.088***	0.090***	0.074***
Entrepreneur, landowner	0.026*	0.033**	0.036**	0.033**	0.040**	0.050***	0.054***	0.047**
Professional employee or freelancer	0.069***	0.077***	0.078***	0.070***	0.088***	0.091***	0.089***	0.088***
Self-employed worker	0.057***	0.060***	0.059***	0.057***	0.075***	0.074***	0.073***	0.075***
Teacher, employee, military in career	0.118***	0.122***	0.122***	0.121***	0.136***	0.136***	0.132***	0.138***
Laborer, services personnel, member of cooperatives	0.025**	0.025***	0.025***	0.022**	0.039***	0.039***	0.039***	0.038***
Retired worker	-0.024	-0.041	-0.008	-0.020	-0.035	-0.033	-0.017	-0.034
(Omitted Variable: Unemployed)								
F test for Mother's occupation	*51.61****	*53.37****	*53.96****	*52.85****	*62.54****	*65.86****	*66.75****	*63.78****
F test for Parents' occupation	*71.99****	*66.31****	*68.54****	*70.18****	*88.26****	*78.04****	*81.60****	*88.05****
F TEST for PARENTS' BACKGROUND	***901.84****	***646.42****	***731.77***	***899.51****	***957.33****	***621.64****	***801.82****	***925.98****
Number of Obs	319.288	319.288	319.288	319.288	489.581	489.581	489.581	489.581
Number of groups	-	6.057	22.623	103	-	7.555	29.184	103

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Multiple imputation estimations include 5 imputations. 3) Regressions presented in columns (2), (3) and (4) include standard errors adjusted for clusters in School Code, Classroom Code and Provincial Code respectively – i.e. for "Number of groups" reported in the Table. 4) All specifications include student, school and territorial characteristics as defined in Table 3.

Table 11 – Estimation results on Rasch Test Scores in Mathematics at the 2nd Grade (2011/2012 INVALSI Data)

Explanatory Variables	Estimations without Multiple Imputations				Estimations with Multiple Imputations			
	(1) NO FE	(2) School FE	(3) Class FE	(4) Prov FE	(1) NO FE	(2) School FE	(3) Class FE	(4) Prov FE
PARENTS' BACKGROUND								
Father's country of birth								
Italy	0.170***	0.160***	0.150***	0.166***	0.180***	0.176***	0.165***	0.176***
European Union	0.114***	0.133***	0.136***	0.110***	0.120***	0.143***	0.144***	0.124***
European Country no European Union	0.044**	0.045**	0.041**	0.041*	0.045**	0.056***	0.047***	0.042**
(Omitted Variable: Other)								
F test for Father's country of birth	*66.90****	*60.79****	*63.84****	*65.10****	*105.24****	*95.59****	*101.90****	*103.21****
Mother's country of birth								
Italy	0.146***	0.118***	0.115***	0.140***	0.146***	0.125***	0.124***	0.140***
European Union	0.103***	0.078***	0.068***	0.102***	0.097***	0.084***	0.077***	0.097***
European Country no European Union	0.035*	0.036**	0.039**	0.037*	0.013	0.015	0.024*	0.013
(Omitted Variable: Other)								
F test for Mother's country of birth	*56.55****	*40.88****	*44.49****	*50.46****	*79.97****	*67.07****	*71.76****	*75.06****
F test for Parents' country of birth	*198.81****	*123.33****	*151.77****	*183.57****	*300.24****	*198.76****	*259.50****	*271.66****
Father's educational qualification								
Lower secondary school certificate	0.151***	0.149***	0.149***	0.153***	0.147***	0.141***	0.136***	0.148***
Vocational secondary school diploma	0.209***	0.206***	0.206***	0.213***	0.206***	0.198***	0.190***	0.208***
Upper secondary school diploma	0.294***	0.294***	0.295***	0.297***	0.293***	0.284***	0.272***	0.293***
Another qualification higher than diploma	0.264***	0.256***	0.257***	0.266***	0.275***	0.263***	0.253***	0.274***
University degree or Postgraduate qualification	0.387***	0.397***	0.395***	0.393***	0.389***	0.384***	0.364***	0.392***
(Omitted Variable: Primary school certificate)								
F test for Father's education	*247.82****	*295.25****	*356.98****	*257.29****	*270.33****	*291.18****	*293.19****	*281.65****
Mother's educational qualification								
Lower secondary school certificate	0.131***	0.153***	0.166***	0.138***	0.126***	0.134***	0.137***	0.128***
Vocational secondary school diploma	0.184***	0.217***	0.226***	0.196***	0.191***	0.203***	0.200***	0.195***
Upper secondary school diploma	0.313***	0.347***	0.358***	0.325***	0.314***	0.323***	0.318***	0.317***
Another qualification higher than diploma	0.317***	0.352***	0.360***	0.329***	0.305***	0.318***	0.310***	0.309***
University degree or Postgraduate qualification	0.441***	0.481***	0.486***	0.457***	0.442***	0.452***	0.438***	0.448***
(Omitted Variable: Primary school certificate)								
F test for Mother's education	*400.13****	*513.63****	*601.74****	*427.49****	*477.77****	*578.73****	*601.06****	*490.05****
F test for Parents' education	*519.23****	*556.44****	*703.59****	*543.64****	*583.19****	*577.08****	*641.77****	*602.98****
Father's employment status								
Homemaker	-0.001	0.078***	0.072***	0.034*	0.022	0.067***	0.060***	0.048**
Manager, university lecturer, officer	0.128***	0.166***	0.150***	0.145***	0.129***	0.146***	0.136***	0.140***
Entrepreneur, landowner	0.108***	0.135***	0.130***	0.117***	0.111***	0.117***	0.113***	0.117***
Professional employee or freelancer	0.113***	0.140***	0.127***	0.120***	0.105***	0.115***	0.107***	0.106***
Self-employed worker	0.113***	0.134***	0.119***	0.118***	0.111***	0.116***	0.108***	0.112***
Teacher, employee, military in career	0.142***	0.175***	0.164***	0.151***	0.134***	0.149***	0.142***	0.137***
Laborer, services personnel, member of cooperatives	0.065***	0.086***	0.073***	0.068***	0.062***	0.070***	0.063***	0.062***
Retired worker	0.064**	0.092***	0.097***	0.085***	0.045	0.067**	0.071***	0.056*
(Omitted Variable: Unemployed)								
F test for Father's occupation	*33.71****	*50.20****	*54.43****	*36.28****	*34.01****	*41.96****	*45.54****	*40.66***
Mother's employment status								
Homemaker	0.015	0.005	0.002	0.005	0.014*	0.007	0.008	0.008
Manager, university lecturer, officer	0.075***	0.106***	0.110***	0.084***	0.074***	0.096***	0.102***	0.081***
Entrepreneur, landowner	-0.009	0.029*	0.027*	0.006	-0.013	0.020	0.030**	0.001
Professional employee or freelancer	0.058***	0.069***	0.067***	0.064***	0.059***	0.064***	0.065***	0.063***
Self-employed worker	0.051***	0.061***	0.054***	0.055***	0.054***	0.056***	0.053***	0.057***
Teacher, employee, military in career	0.110***	0.115***	0.114***	0.112***	0.108***	0.111***	0.110***	0.110***
Laborer, services personnel, member of cooperatives	0.004	0.010	0.008	0.006	0.007	0.010	0.014*	0.008
Retired worker	0.004	-0.000	-0.053	0.013	-0.029	-0.014	-0.034	-0.023
(Omitted Variable: Unemployed)								
F test for Mother's occupation	*44.47****	*56.32****	*67.77****	*49.02****	*43.18****	*53.81****	*62.03****	*48.34****
F test for Parents' occupation	*47.34****	*60.80****	*69.48****	*50.65****	*46.58****	*54.65****	*62.49****	*51.63****
F TEST for PARENTS' BACKGROUND	***493.66*****	***382.17*****	***545.07*****	***514.21*****	***594.48*****	***414.94*****	***551.44*****	***568.08*****
Number of Obs	323.918	323.918	323.918	323.918	482.018	482.018	482.018	482.018
Number of groups	-	5.998	22.308	103	-	7.178	28.061	103

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Multiple imputation estimations include 5 imputations. 3) Regressions presented in columns (2), (3) and (4) include standard errors adjusted for clusters in School Code, Classroom Code and Provincial Code respectively – i.e. for "Number of groups" reported in the Table. 4) All specifications include student, school and territorial characteristics as defined in Table 3.

Table 12 – Estimation results on Rasch Test Scores in Mathematics at the 5th Grade (2011/2012 INVALSI Data)

Explanatory Variables	Estimations without Multiple Imputations				Estimations with Multiple Imputations			
	(1) NO FE	(2) School FE	(3) Class FE	(4) Prov FE	(1) NO FE	(2) School FE	(3) Class FE	(4) Prov FE
PARENTS' BACKGROUND								
Father's country of birth								
Italy	0.111***	0.092***	0.088***	0.096***	0.126***	0.099***	0.099***	0.112***
European Union	0.012	0.064***	0.068***	0.027	-0.019	0.045***	0.054***	-0.006
European Country no European Union	-0.025	0.027	0.032*	0.024	-0.036*	0.011	0.012	0.004
(Omitted Variable: Other)								
F test for Father's country of birth	*42.35****	*24.12****	*23.27****	*23.96****	*91.17****	*47.64****	*54.81****	*51.63****
Mother's country of birth								
Italy	0.100***	0.085***	0.078***	0.086***	0.099***	0.078***	0.069***	0.091***
European Union	0.030	0.086***	0.082***	0.040**	-0.007	0.060***	0.056***	0.011
European Country no European Union	-0.024	0.030*	0.029*	0.023	-0.046**	0.012	0.013	0.007
(Omitted Variable: Other)								
F test for Mother's country of birth	*35.91****	*20.99****	*20.09****	*19.10****	*60.13****	*23.24****	*22.27****	*17.28****
F test for Parents' country of birth	*98.12****	*45.46****	*52.60****	*55.93****	*178.28****	*71.41****	*89.20****	*101.75****
Father's educational qualification								
Lower secondary school certificate	0.117***	0.125***	0.112***	0.118***	0.120***	0.107***	0.099***	0.117***
Vocational secondary school diploma	0.134***	0.181***	0.167***	0.177***	0.137***	0.152***	0.140***	0.162***
Upper secondary school diploma	0.264***	0.249***	0.233***	0.273***	0.266***	0.221***	0.205***	0.264***
Another qualification higher than diploma	0.225***	0.216***	0.197***	0.251***	0.209***	0.183***	0.163***	0.221***
University degree or Postgraduate qualification	0.422***	0.360***	0.343***	0.427***	0.422***	0.324***	0.302***	0.414***
(Omitted Variable: Primary school certificate)								
F test for Father's education	*264.46****	*264.61****	*315.96****	*306.91****	*283.64****	*226.23****	*252.09****	*243.06****
Mother's educational qualification								
Lower secondary school certificate	0.040***	0.105***	0.099***	0.083***	0.038***	0.077***	0.073***	0.065***
Vocational secondary school diploma	0.093***	0.146***	0.141***	0.125***	0.093***	0.119***	0.114***	0.113***
Upper secondary school diploma	0.222***	0.259***	0.250***	0.257***	0.224***	0.222***	0.209***	0.242***
Another qualification higher than diploma	0.245***	0.263***	0.265***	0.273***	0.235***	0.219***	0.216***	0.249***
University degree or Postgraduate qualification	0.382***	0.377***	0.366***	0.419***	0.382***	0.335***	0.314***	0.399***
(Omitted Variable: Primary school certificate)								
F test for Mother's education	*322.01****	*323.59****	*432.77****	*370.41****	*359.84****	*347.25****	*402.10****	*380.06****
F test for Parents' education	*485.25****	*351.71****	*525.90****	*556.44****	*523.99****	*344.86****	*473.87****	*534.99****
Father's employment status								
Homemaker	-0.243***	0.141***	0.128***	0.145***	-0.173***	0.059***	0.057***	0.081***
Manager, university lecturer, officer	0.178***	0.173***	0.165***	0.201***	0.196***	0.164***	0.153***	0.210***
Entrepreneur, landowner	0.069***	0.112***	0.103***	0.141***	0.083***	0.096***	0.089***	0.133***
Professional employee or freelancer	0.137***	0.116***	0.108***	0.150***	0.146***	0.114***	0.105***	0.155***
Self-employed worker	0.096***	0.106***	0.099***	0.117***	0.099***	0.097***	0.091***	0.117***
Teacher, employee, military in career	0.138***	0.135***	0.130***	0.148***	0.143***	0.128***	0.121***	0.153***
Laborer, services personnel, member of cooperatives	0.051***	0.061***	0.056***	0.057***	0.061***	0.058***	0.055***	0.065***
Retired worker	0.043*	0.065***	0.066***	0.058***	0.056**	0.072***	0.067***	0.074**
(Omitted Variable: Unemployed)								
F test for Father's occupation	*70.09****	*44.38****	*51.94****	*48.63****	*52.65****	*38.16****	*44.95****	*41.35****
Mother's employment status								
Homemaker	0.066***	0.024***	0.023***	0.022**	0.059***	0.028***	0.031***	0.032***
Manager, university lecturer, officer	0.042*	0.061***	0.062***	0.084***	0.040*	0.058***	0.064***	0.083***
Entrepreneur, landowner	-0.091***	0.031**	0.028**	0.039**	-0.086***	0.014	0.020*	0.017
Professional employee or freelancer	0.102***	0.042***	0.045***	0.074***	0.102***	0.056***	0.059***	0.087***
Self-employed worker	0.105***	0.066***	0.065***	0.093***	0.103***	0.069***	0.069***	0.099***
Teacher, employee, military in career	0.142***	0.099***	0.095***	0.123***	0.136***	0.103***	0.099***	0.130***
Laborer, services personnel, member of cooperatives	-0.002	-0.001	-0.001	-0.001	0.000	0.004	0.008	0.007
Retired worker	0.060	0.012	0.038	0.039	0.003	0.011	0.037	-0.014
(Omitted Variable: Unemployed)								
F test for Mother's occupation	*65.16****	*43.15****	*47.86****	*49.31****	*62.43****	*42.89****	*46.94****	*48.71****
F test for Parents' occupation	*84.11****	*49.59****	*58.23****	*58.26****	*69.20****	*46.45****	*54.33****	*57.68****
F TEST for PARENTS' BACKGROUND	***465.09****	***188.83***	***344.97****	***491.10****	***530.41****	***190.73****	***342.49****	***560.85****
Number of Obs	315.256	315.256	315.256	315.256	489.279	489.279	489.279	489.279
Number of groups	-	6.041	22.489	103	-	7.570	29.340	103

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Multiple imputation estimations include 5 imputations. 3) Regressions presented in columns (2), (3) and (4) include standard errors adjusted for clusters in School Code, Classroom Code and Provincial Code respectively – i.e. for "Number of groups" reported in the Table. 4) All specifications include student, school and territorial characteristics as defined in Table 3.

Table 13 – Estimation Results on Rasch Test Scores in Reading: SAMPLE SCHOOLS

	2008/2009 – Grade 2				2011/2012 – Grade 5			
	NO Imputations		YES Imputations		NO Imputations		YES Imputations	
Explanatory Variables	(1) NO FE	(2) School FE	(3) NO FE	(4) School FE	(1) NO FE	(2) School FE	(3) NO FE	(4) School FE
PARENTS' BACKGROUND								
Father's country of birth								
Italy	0.327***	0.312***	0.323***	0.310***	0.240***	0.228***	0.264***	0.251***
European Union	0.275***	0.268***	0.253***	0.242***	0.211***	0.192***	0.230***	0.192***
European Country no European Union	0.114	0.085	0.011	-0.001	0.118*	0.135*	0.071	0.090
(Omitted Variable: Other)								
F test for Father's country of birth	*16.64****	*14.71****	*30.74****	*27.39****	*8.84****	*7.05****	*19.52****	*15.83****
Mother's country of birth								
Italy	0.126**	0.156***	0.149***	0.184***	-0.011	0.010	0.052	0.081**
European Union	0.212***	0.240***	0.201***	0.236***	-0.120**	-0.096	0.007	0.004
European Country no European Union	0.043	0.073	0.056	0.091	-0.119*	-0.114*	-0.066	-0.048
(Omitted Variable: Other)								
F test for Mother's country of birth	*4.43****	*5.62****	*6.03****	*9.00****	*4.13****	*4.23****	*3.51***	*5.87****
F test for Parents' country of birth	*22.68****	*23.86****	*43.49****	*43.18****	*10.89****	*41.43****	*27.59****	*22.27****
Father's educational qualification								
Lower secondary school certificate	0.178***	0.177***	0.195***	0.191***	0.201***	0.198***	0.202***	0.186***
Vocational secondary school diploma	0.261***	0.268***	0.281***	0.279***	0.284***	0.277***	0.285***	0.263***
Upper secondary school diploma	0.386***	0.400***	0.399***	0.399***	0.394***	0.390***	0.388***	0.371***
Another qualification higher than diploma	0.366***	0.363***	0.382***	0.376***	0.407***	0.385***	0.388***	0.361***
University degree or Postgraduate qualification	0.520***	0.540***	0.513***	0.513***	0.502***	0.487***	0.494***	0.456***
(Omitted Variable: Primary school certificate)								
F test for Father's education	*35.67****	*39.27****	*36.31****	*38.71****	*43.15****	*54.85****	*54.78****	*49.17****
Mother's educational qualification								
Lower secondary school certificate	0.201***	0.198***	0.236***	0.224***	0.212***	0.163***	0.207***	0.165***
Vocational secondary school diploma	0.281***	0.275***	0.331***	0.313***	0.352***	0.299***	0.333***	0.283***
Upper secondary school diploma	0.471***	0.468***	0.504***	0.485***	0.439***	0.381***	0.433***	0.380***
Another qualification higher than diploma	0.471***	0.461***	0.515***	0.491***	0.433***	0.388***	0.420***	0.378***
University degree or Postgraduate qualification	0.635***	0.642***	0.665***	0.647***	0.612***	0.560***	0.596***	0.544***
(Omitted Variable: Primary school certificate)								
F test for Mother's education	*59.95****	*70.30****	*63.98****	*63.35****	*63.64****	*71.68****	*78.15****	*65.50****
F test for Parents' education	*77.92****	*77.4****	*82.92****	*75.05****	*87.28****	*8.14****	*110.86****	*87.30****
Father's employment status								
Homemaker	0.243	0.280	0.136	0.159	0.041	0.103	0.052	0.119
Manager, university lecturer, officer	0.386***	0.340***	0.322***	0.292***	0.256***	0.242***	0.241***	0.210***
Entrepreneur, landowner	0.237***	0.190***	0.181***	0.145***	(omitted)	(omitted)	0.151	0.136
Professional employee or freelancer	0.283***	0.246***	0.222***	0.199***	0.217***	0.213***	0.197***	0.180***
Teacher, employee, military in career	0.281***	0.263***	0.227***	0.215***	0.211***	0.206***	0.192***	0.169***
Self-employed worker	0.232***	0.192***	0.185***	0.162***	0.203***	0.178***	0.178***	0.146***
Laborer, services pers., member of cooperatives	0.162***	0.137***	0.124***	0.110***	0.146***	0.128***	0.131***	0.104***
Retired worker	-	-	-	-	0.174***	0.188***	0.152**	0.149***
(Omitted Variable: Unemployed)								
F test for Father's occupation	*9.25****	*7.17****	*7.57****	*6.12****	*8.38****	*7.00*	*5.93****	*4.13****
Mother's employment status								
Homemaker	-0.010	-0.021	0.002	0.001	0.026	0.029	0.039	0.048*
Manager, university lecturer, officer	0.153**	0.172**	0.161**	0.165**	0.089	0.131**	0.078	0.125**
Entrepreneur, landowner	0.063	0.088	0.082	0.098	(omitted)	(omitted)	0.102	0.114
Professional employee or freelancer	0.139***	0.147***	0.129**	0.136**	0.122***	0.107***	0.130***	0.114***
Teacher, employee, military in career	0.134***	0.119***	0.139***	0.130***	0.147***	0.141***	0.157***	0.152**
Self-employed worker	0.092*	0.085*	0.103*	0.098*	0.078**	0.068*	0.090***	0.087***
Laborer, services pers., member of cooperatives	0.038	0.023	0.042	0.038	0.022	0.023	0.041	0.044
Retired worker	-	-	-	-	-0.004	0.068	0.077	0.143
(Omitted Variable: Unemployed)								
F test for Mother's occupation	*8.24****	*7.73****	*8.66****	*7.38****	*7.91****	*7.28****	*6.40****	*5.91****
F test for Parents' occupation	*9.88****	*8.64****	*9.16****	*7.65****	*9.37****	*7.99****	*7.09****	*5.54****
F TEST for PARENTS' BACKGROUND	***84.66*****	***72.22*****	***97.49*****	***76.33*****	***83.80*****	***57.91*****	***88.14*****	***59.27*****
Number of Obs	26.560	26.560	43.299	43.299	20.964	20.964	30.877	30.877
Number of groups	-	801	-	1.057	-	1.192	-	1.365

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Multiple imputation estimations include 5 imputations. 3) Regressions presented in columns (2) and (4) include standard errors adjusted for clusters in School Code – i.e. for "Number of groups" as reported in the Table. 4) All specifications include student, school and territorial characteristics as defined in Table 3.

Table 14 – Comparison of Territorial Divide in Scores between Full Sample and Schools Sampled to Control for "Cheating": Basic Estimation Model

	READING							
	2008/2009 – Grade 2				2011/2012 – Grade 5			
	NO Imputations		YES Imputations		NO Imputations		YES Imputations	
	Full Sample	Sample Schools	Full Sample	Sample Schools	Full Sample	Sample Schools	Full Sample	Sample Schools
TERRITORIAL CHARACTERISTICS								
North	0.058***	0.186***	0.043***	0.172***	0.040***	0.100***	0.039***	0.091***
Centre	-0.002	0.126***	-0.016*	0.116***	0.033***	0.085***	0.035***	0.067***
(Omitted Variable: South and Islands)								
Number of Obs	95.969	26.560	153.052	43.299	319.288	20.964	489.581	30.877
	MATHEMATICS							
	2008/2009 – Grade 2				2011/2012 – Grade 5			
	NO Imputations		YES Imputations		NO Imputations		YES Imputations	
	Full Sample	Sample Schools	Full Sample	Sample Schools	Full Sample	Sample Schools	Full Sample	Sample Schools
TERRITORIAL CHARACTERISTICS								
North	-0.334***	-0.075***	-0.377***	-0.088***	0.050***	0.122***	0.062***	0.110***
Centre	-0.331***	-0.093***	-0.351***	-0.101***	-0.363***	-0.208***	-0.319***	-0.214***
(Omitted Variable: South and Islands)								
Number of Obs	96.002	26.580	153.092	43.333	315.256	20.802	489.279	30.869

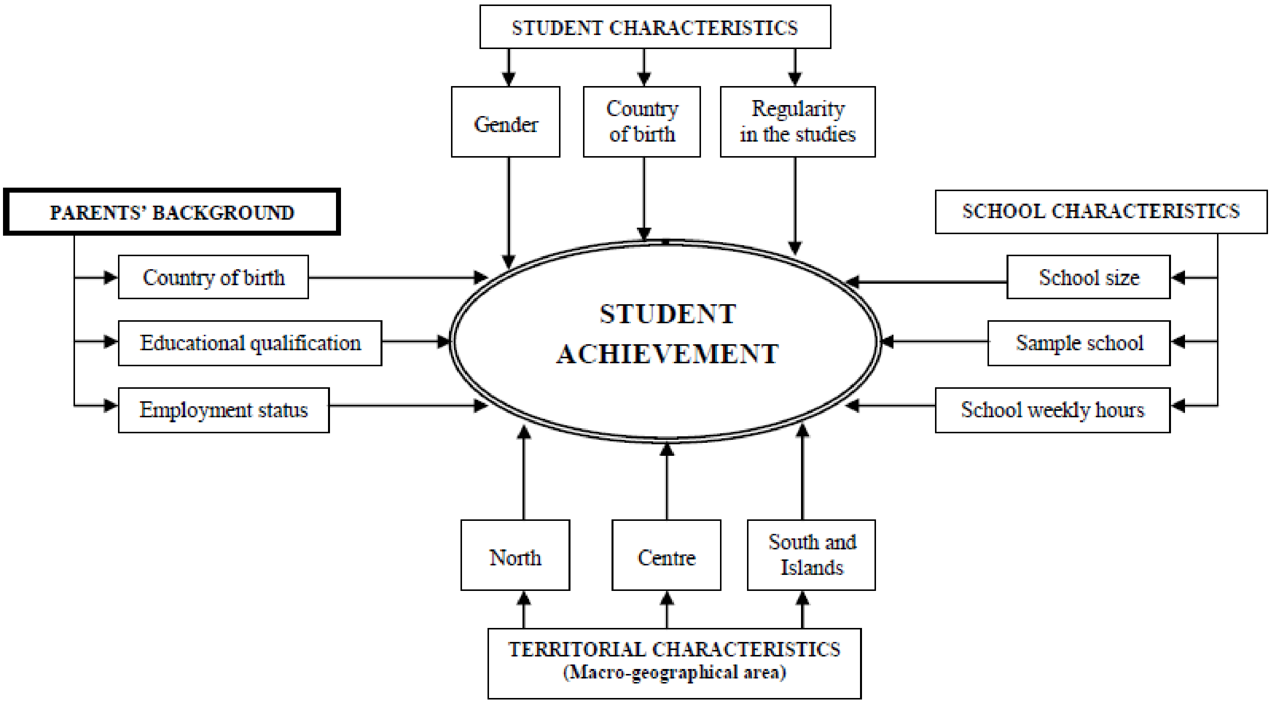
Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Multiple imputation estimations include 5 imputations. 3) All specifications include student, school and parents' background characteristics as defined in Table 3.

Table 15 – Comparison of Territorial Divide in Scores between Full Sample and Schools Sampled to Control for "Cheating": 2011/2012 Data

	READING							
	Grade 2				Grade 5			
	NO Imputations		YES Imputations		NO Imputations		YES Imputations	
	Full Sample	Sample Schools	Full Sample	Sample Schools	Full Sample	Sample Schools	Full Sample	Sample Schools
TERRITORIAL CHARACTERISTICS								
North	-0.060***	0.099***	-0.077***	0.109***	0.041***	0.101***	0.038***	0.088***
Centre	-0.039***	0.116***	-0.041***	0.130***	0.033***	0.085***	0.035***	0.070***
(Omitted Variable: South and Islands)								
Number of Obs	326.231	21.634	480.541	31.526	319.288	20.964	489.581	30.877
	MATHEMATICS							
	Grade 2				Grade 5			
	NO Imputations		YES Imputations		NO Imputations		YES Imputations	
	Full Sample	Sample Schools	Full Sample	Sample Schools	Full Sample	Sample Schools	Full Sample	Sample Schools
TERRITORIAL CHARACTERISTICS								
North	-0.280***	-0.043**	-0.310***	-0.041***	0.055***	0.128***	0.070***	0.109***
Centre	-0.176***	-0.010	-0.183***	-0.001	-0.362***	-0.206***	-0.319***	-0.214***
(Omitted Variable: South and Islands)								
Number of Obs	323.918	21.558	482.018	31.758	315.256	20.802	489.279	30.869

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Multiple imputation estimations include 5 imputations. 3) All specifications include student, school and parents' background characteristics as defined in Table 3.

Figure 1 – Conceptual framework



CHAPTER 2

"Gift of Time" and "Family Gift":

The Effect of Early School Entry on Pupils Performance

Abstract: This chapter provides a comprehensive analysis of the effect of early school entry on educational outcomes using Normalized test score data on Italian pupils. The empirical procedure is designed to disentangle the effect of regular entry (*Gift of Time*) from possible unobserved confounding factors (*Family Gift)* affecting both enrollment decision and schooling outcome. I tackle the issue of selection on unobservables by using a Regression Discontinuity Design so that exogenous age thresholds are used to compare children with similar age but different educational choices. Estimates suggest that pupils who enroll in advance are peculiar in the sense that they perform better than regular ones with almost identical age. After neutralizing this "schooling ability" effect, I find that anticipating pupils present severe penalties in test scores. Findings have policy implications for parents, which struggle with the question of whether they should send their children to school as soon as they are eligible, and for governments, which can change cutoff birth date for first enrollment into school.

Keywords*:* Age at school entry, primary school, Normalized test scores.

JEL codes: I20, H52.

2.1 Introduction

In the past, child development researchers have often argued that children's "readiness" for school is an important factor that determines school success. However, there is a considerable debate in the research community regarding how school readiness can be measured. In the absence of any consensus, researchers have traditionally used chronological age as the standard to evaluate it. There are two dominant viewpoints surrounding the entrance age debate. On the one hand, basic human capital theory suggests that children should start formal learning as soon as possible: the earlier children enroll at school, the sooner they begin accumulating skills (Bedard and Duhey, 2012). Moreover, some authors suggest that young children are more receptive for learning than older ones and believe that school provides the nurturing environment that helps to promote children's learning and development (Datar, 2006). On the other hand, child developmentalists have stressed that age and human capital could be complement so that young children might not be mature enough to learn complex material in the school environment (Mayer and Knutson, 1999). In other words, children need the *Gift of Time* and general out-of-school experience to be able to better perform in school. As a consequence, enrolling a pupil before he/she is ready for the rigor of formal education may turn out to be less productive than waiting until he/she is more mature. The school entrance age also has an impact on lifetime earnings. Individuals who start school in advance enter the labor market earlier, and can collect the returns of their human capital investments over a longer time horizon (Fredriksson and Öckert, 2006). Conversely, children entering the labor market one year later could be more likely to have the necessary skills and maturity to succeed in school and therefore to learn more in each grade. In this perspective, postponing school entrance implies better skills, which may provide higher wages (Elder and Lubotsky, 2009).

The identification of the effect of age at school entry is not an easy task. The main reason is that parental decisions to delay or expedite their child's school entry are almost certainly related to both households' and pupils' characteristics which can simultaneously affect schooling outcomes through several channels. It follows that the evaluation of the causal effect (if any) of entry age on schooling outcomes, i.e., the presence of a *Gift of Time* is particularly problematic because of the presence of potential *Family Gift* shaping pupils' cognitive and schooling ability.

Having these caveats in mind, in this work I address the following research questions:

1) Do younger entrants achieve lower test scores compared to older entrants, i.e., does exist a *Gift of Time*?

2) Is the evaluation of the achievement gap biased by unobservable characteristics, i.e., does exist a *Family Gift*?

3) Do these differences in achievement scores persist during primary school?

The identification of these points is achieved throughout a strategy designed to disentangle the treatment effect of early entry from possible unobserved confounding factors affecting both enrollment and schooling outcomes.

I deal with selection on unobservables bias by means of a Regression Discontinuity Design so that exogenous age-thresholds are used to compare pupils with almost identical age but different educational choices. The empirical analysis is carried out on data containing the universe of students who attend primary school in Italy. Measures concerning pupils' performance are based on Normalized National Tests Score in Mathematics and Reading implemented by the National Institute for the Educational Evaluation of Instruction and Training (INVALSI). Results suggest that students who enroll in advance perform worse than regular ones. I point out that a severe distortion in the evaluation of the true effect of early entry arises when neglecting unobserved characteristics driving the early school decision. Since pupils in advance tend to be selected according to their schooling ability, the real impact of early entry on schooling performance is underestimated. After I get rid of selection bias, I find that anticipating pupils perform substantially worse than regular ones. This effect proves to be particularly scarring since it lasts for the entire path of primary education.

The chapter proceeds as follow. The next Section provides a review of the main recent studies. Section 3 describes how the research adds to the existing literature and gives an intuition of my identification strategy. Section 4 highlights data source and variables used in the analysis. Section 5 presents the derivation of empirical strategy while Section 6 discusses the main results as well as several robustness and falsification exercises. Some concluding remarks are addressed in Section 7.

2.2 Literature Review

Since the '80s several studies have exploited the variation in school entry age to identify its effect on educational performance showing that there is a disadvantage to early entry (Maddux, 1983; Uphoff and Gilmore, 1985)[1].

Typically, the outcome variable examined in the literature is children's achievement test scores in the primary grades. Most of studies focuses on within grade comparisons of performance of older and younger school entrants who differ in birth dates within the year (for a review see Stipek, 2003). The evidence from this literature suggests that youngest students have lower test scores compared to oldest students in the same grade (Sweetland and De Simone, 1987; Jones and Mandeville, 1990; Sharp, 1995; Strøm, 2004; Datar, 2006; Elder and Lubotsky, 2009; McEwan and Shapiro, 2008; Crawford et al., 2010; Ponzo and Scoppa, 2014) and are more likely to repeat a grade (Elder and Lubotsky, 2009; McEwan and Shapiro, 2008). Only few studies provide evidence that youngest students achieve higher test scores than oldest ones (Leuven et al., 2010; Robertson, 2011). Economists have also shown interest in the effects of age at school start on educational attainments and wages. This literature provides mixed results. Some studies find that older entrants attain slightly less education (Angrist and Krueger, 1991, 1992; Fertig and Kluve, 2005; Dobkin and

[1] Although the interest in this subject has grown since the '80s, a first contribution dates from the early '30s, when the SUMMIT New Jersey school system was interested in determining which students to admit into first grade. To help answer to this question, Bigelow (1934) studied the achievement of 127 fourth graders in the school system finding that children who were older when they began first grade were less likely to repeat one of the first three grades and also tended to score higher on the achievement test.

Ferreira, 2010) and lower labor market outcomes (Angrist and Krueger, 1991). In contrast, Fredriksson and Öckert (2006) and Kawaguchi (2011) find evidence in support of higher educational attainment and wages for students who enter school at an older age. The positive association between these variables is also provided by Bedarh and Duhey (2012).

Instead of discussing in details this impressive wide literature, a schematic summary is provided in Tables 1 and 2. In these tables, I pose focus on the effect of age on the evaluated outcome variables. Moreover, I emphasize methods of analysis and results. I remark that most of these studies use quarter of birth or legal entry age as instruments to deal with the endogeneity issue. This approach has been recently criticized by Barua and Lang (2009) who show that the quarter of birth and the legal entry age instrument give biased estimates of the policy-relevant local average treatment effect (LATE) because of the failure of the monotonicity assumption. These authors propose an instrument that satisfies the monotonicity assumption and gives a consistent estimates of the policy-relevant LATE showing that the effect of school entry age on educational attainment appears to be very close to zero. As things stand, evidence of the entry age effect on schooling outcomes appear to be far from being well defined.

2.3 Insights and Identification Procedure

The research follows the recent attempts of Crawford et al. (2010) and Dobkin and Ferreira (2010) to provide estimates not derived by instrumental variables techniques. I present an original empirical procedure designed to disentangle the treatment effect (having entered the primary school in advance) from possible unobserved confounding factors affecting enrollment decisions as well as schooling outcomes.

The Italian normative setting regulating access to primary education allows to address this point. Italian primary schools usually start in September. In a given year (say year *t*) all pupils who are 6 years old and those who will be 6 years old by December 31st *must* start school in September. Then, the law also *permits* enrollment to pupils who will be aged 6 by April 30th in *t+1*. Crucially, this is only an opportunity (it is not mandatory) and it is apparent that self-selection into schools may be related to potentially unobserved characteristics. At this stage, it is important to remark that also part of the pupils who are aged 6 in *t* are affected by a selection problem. This is true for those who became 6 years old between January 1st and April 30th in year *t* since these are pupils whose parents decided not to send them at school in advance in *t-1*. Therefore, among scholars in the same class, only those aged 6 between May 1st and December 31st in year *t* do not suffer from self-selection into education.

This institutional setup implies that the effect on schooling outcomes of anticipating the entry age of one year cannot be estimated by comparing scores of pupils who turned 6 in the first and in the fourth quarter of year *t*, i.e. using quarter of birth and instrumental variables technique, since the former suffer from selection problems. In my case, the non-selected group contains pupils with only seven months of age differences and this source of variation could be not sufficient to assess the effect of entering school one year earlier, especially in the case in which the age-related penalties are particularly important between pupils with almost one year of age difference. In other words, in my case an IV estimation of the entry age effect obtained using quarter of birth as an instrument for actual age on a seven-month period may be very uninformative about the effect of entering school a full year earlier.

To overcome this problem, my strategy proceeds as follows.

Consider pupils who started school in *t* and will be aged 6 in April of year *t+1*. In order to compare their schooling outcome with that of their classmates who became 6 years old in May of *t* an Average Treatment on the Treated (ATT) estimation procedure may be implemented. The ATT estimator gives a parameter β^{*}_{ATT} which provides an estimate of the effect on schooling outcomes of entering school 1 year and 1 month earlier. However, while the group of older pupils does not suffer from any enrollment selection, pupils in advance are selected according to their parents' choice. Therefore, the Conditional Independence Assumption (CIA) required to obtain unbiased ATT estimates is likely to fail and, consequently, the estimated parameter is potentially biased since:

$$\beta^{*}_{ATT} = \underbrace{\beta^{*}}_{\textit{unbiased effect of early entry on scores}} + \underbrace{\beta^{\#}}_{\textit{effect of unobserved confounders on scores}} \qquad (1)$$

Notwithstanding, from eq. (1) it appears that the unbiased parameter of early entry on pupil's performance (β^{*}) can still be evaluated. This requires that the effect of unobserved components on test scores ($\beta^{\#}$) is firstly estimated and then expunged from β^{*}_{ATT}. I estimate $\beta^{\#}$ by relying on a Regression Discontinuity Design (RDD) which evaluates difference in scores of pupils aged 6 in December of year *t* with respect to those aged 6 in January of year *t+1*. As far as one month age difference on schooling performance is negligible, differences in test scores between these two groups of pupils should only reflect unobserved heterogeneity related to selection issues. Following this strategy I can firstly estimate the mean effect of unobserved confounders on test score. Then, I can evaluate the effect of entering primary school one year earlier on test scores using β^{*}_{ATT} and $\beta^{\#}$.

2.4 Data and Descriptive Statistics

Data used in this work have been collected by the INVALSI, which yearly assesses students' knowledge in Reading (Italian Language) and Mathematics through the National Service for the Evaluation of Education and Training (SNV)[2]. Tests are administered in the primary school (Grade 2 and 5), in the lower secondary school (Grade 6 and 8), and in the upper secondary school (Grade 10).

[2] Reading test is divided into three main sections: 1) Reading comprehension of a narrative text, 2) Reading comprehension of expository text, and 3) Grammatical knowledge and skills. Mathematics test is divided into four areas: 1) Numbers, 2) Space and figures, 3) Data and forecasts, 4) Relations and functions. At the 2nd grade of primary school, the maths test is limited to the first three areas.

For the purpose of the present study, I employ data of primary education – both the 2nd and the 5th grade – of the school year 2011/2012[3]. I use information on about 500,000 pupils.

Data set of the INVALSI contains Normalized[4] tests scores in Reading and Mathematics. On top of that, useful data on personal, family and schooling background of students, gender, date and country of birth of pupils, country of birth, occupational status and educational level of their parents and territorial characteristics are provided. This rich data set gives me the opportunity of controlling for many relevant observable characteristics which have not been considered in recent studies on the field. Among others, Crawford et al. (2010) realize the limits imposed by information contained in their available data set. Moreover, since INVALSI identifies each year a number of schools where the test is done in the presence of an external observer, it is possible to control for the phenomenon of cheating. All variables used in the analysis are described in Table 3.

Rather than focusing on descriptive statistics of all variables, I prefer to report mean and standard deviation of the dependent variables used in the analysis – Normalized test scores in Reading and Mathematics – by date of birth and by parental background. In Table 4 (Column I), at the 2nd grade there is an advantage for pupils enrolled in advance compared with regular students only in Mathematics test scores. This gap does not fade away during school. Concerning language test scores, advanced pupils appear to perform as good as regular ones. Interestingly, if I split the group of regular pupils in order to untangle those who could enroll in advance but have decided to enroll regularly (i.e. those aged 6 in the first four months of year *t*), I detect some additional insights. In particular, in Column II of Table 4 regulars appear to have lower Mathematics test scores than both the oldest and the youngest pupils. For Reading, regulars perform worse than older pupils only. This preliminary evidence, which requires further investigations, shows that when dealing with Mathematics tests both the age and the selection effect could be present, while for Reading skills age proves to be more important than selection.

Turning to family background, in Table 5 and Table 6 we can observe that pupils with more educated parents have a higher score than those with low educated parents: mean test score increases with educational qualifications of both father and mother. Mean score gap reaches approximately 10 points considering students with parents who have a high level of education with respect to those who have parents with a low level of education. Some difference also arises across pupils from families that are heterogeneous in terms of income. Students in low-income families perform worse than students in medium- and high-income ones. This is possibly due to a better cultural environment for children in non-disadvantaged environments providing them ample opportunities to develop their cognitive and language skills.

[3] Norms regulating early school entry discussed in Section 3 have been introduced in 2006 by the so called Fioroni reform. As a consequence, children in grades higher than the 5th cannot be used for my analysis since they started primary school before 2006. For pupils who attend the 2nd grade in 2011/2012, the law concerning their first school enrollment is the Ministerial Circular n. 4/2010: School enrollment for school year 2010/2011. For pupils who attend the 5th grade in 2011/2012, the law concerning their first school enrollment is the Ministerial Circular n. 74/2006: School enrollment for school year 2007/2008.

[4] Normalized scores are computed by INVALSI starting from raw scores gain by students taking the test and calculating them on a range from 0 to 100.

2.5 The Empirical Framework

To examine the effect on schooling performance of one year difference in age at school entry, I start by estimating the Average Treatment effect on the Treated (ATT). In the presence of potential selection on *observables*, the ATT can be consistently estimated by running OLS on a sample of pupils who started school in year *t* and become 6 years old in either May of year *t* or April of year *t+1* according to the following framework which includes variables that may potentially affect the outcome as well as treatment participation:

$$Y_{ics} = \alpha + \beta Age_{ics} + \gamma StudC_{ics} + \delta SchoolC_{ics} + \qquad (2)$$
$$+ \eta ParentsC_{ics} + \theta GeographicC_{ics} + \varepsilon_{ics}$$

In eq. (2) Y is performance measured by Normalized test score in either Reading or Mathematics of student i in the class c in the school s; *StudC*, *SchoolC*, *ParentsC*, *GeographicC* are vectors of student, school, parent's socioeconomic background and geographic characteristics respectively, as defined in Table 3, which for the sake of simplicity from now on will be indicated all together as z; ε is the error term. *Age* is a variable taking the value 1 for pupils aged 6 in April of year *t+1* and 0 otherwise. The estimated parameter β associated to this variable gives us the treatment effect i.e.:

$$\beta^*_{ATT} = E[Y_{ics}|z_{ics}, Age_{ics} = 1] - E[Yics|z_{ics}, Age_{ics} = 0] \qquad (3)$$

However, in the presence of selection on *unobservables* the ATT estimators is given by:

$$E[Y_{ics}|z_{ics}, Age_{ics} = 1] - E[Yics|z_{ics}, Age_{ics} = 0] = \beta^* + (Selection\ Effect) \qquad (4)$$

where β^* indicates the unbiased estimator of β in eq. (2). To obtain unbiased estimates of the treatment on schooling outcomes, the selection effect should be differentiated out from β^*_{ATT}, i.e.:

$$\beta^* = \beta^*_{ATT} - (Selection\ Effect) \quad (5)$$

Albeit the approach contained in eq. (5) grounds on a clear-cut identification procedure, its application requires a non trivial evaluation of the selection effect[5]. In my case I can consistently estimate the selection effect using a RDD approach. The idea is as follows. I start by evaluating scores along pupils' month of birth and I investigate whether any discontinuity arises at the threshold imposed by the Italian normative setting for mandatory school entrance. The cutoff point is posed between children who will be 6 years old in December of year *t* and those who will be aged 6 in January of year *t+1*. The main assumption is that whether a child is born in December or in January is completely random. Then, the only difference arising between these two groups is that those born in January have been selected by their family to be enrolled at school. In this way, it is possible to assess the effect of selection on schooling outcomes and then to evaluate the pure effect of early schooling through eq. (5), i.e., I can identify β^*.

[5] For a complete discussion, see Cameron and Trivedi (2005), p. 845.

2.6 Findings

2.6.1 Main Results

In Table 7 I start by presenting preliminary OLS estimates of eq. (2) using all pupils in each grade. Coefficients associated to controlling variables are significant, going in the expected direction and I avoid to present long comments on quite standard results. In this case regressor of main interest is *"Students in Advance"* which in this case is a dummy variable taking the value 1 for pupils who were 6 years old between January 1st and April 30th in year *t+1* and 0 for all others who were 6 during year *t*. The former get on average, a score of 2.043 points less than regular students in Reading and 1.862 points less in Mathematics at the 2nd grade. At the 5th grade, the gap reduces to 0.842 points less in Reading and -0.901 point less in Mathematics. As amply discussed, I cannot give any causal interpretation on these results. To make one step further, I start by estimating the ATT using only a sample of pupils who were 6 years old in May of year *t* and those who were aged 6 in April of year *t+1*. In this case, the interest is on the effect of one year early school entry on test scores. I use a Propensity Score Matching Method which allows to control for all observed variables that are likely to affect the treatment. The ATT is evaluated using the nearest-neigbor matching estimators[6]. I repeat this procedure for Reading and Mathematics test scores and for both the 2nd and the 5th grade and the results are reported in Table 8. Interestingly, all coefficients confirm previous findings, that is, the presence of a significant penalty for pupils who entered primary school one year earlier. Results are robust and obtained using about 25,000 observations for each specification.

As I have discussed, the interpretation of the β^{*}_{ATT} parameter must be very cautious since, albeit the use of propensity score implies that selection on observables is not present in my data, potential unobserved factors may drive pupils into the treatment. These unobserved components can affect test scores so that I do not have an identification of the *Gift of Time* effect. Indeed, *Family Gift* may affect cognitive schooling ability as well as early enrollment decisions leading to a biased estimation of the early school entry effect. To tackle the issue of selection on unobservables I adopt the RDD strategy.

Figure 1, 2, 3 and 4 contain a graphical illustration of the RDD estimates for Reading and Mathematics test scores at the 2nd and at the 5th grade respectively.

Estimates of the selection effect (β^{*}_{RDD}) are provided in Table 9. As it appears in all figures, a significant effect arises around the threshold highlighting a *positive* selection effect. The results are confirmed for all grades and for both Mathematics and Reading test scores. In addition, albeit I report results arising from a single bandwidth, I remark that the dimension of my sample size is such that different bandwidths are likely to generate almost identical outcomes. This finding – which is robust at 1% significance level as reported in Table 9 – highlights that in Italy those children who anticipate schooling enrollment are actually different with respect to the average of their regular peers and, in particular, they appear to be selected on the basis of characteristics which positively affect schooling outcomes. In other words anticipating pupils benefit from a positive *Family Gift*.

[6] Estimates of the propensity score are available from the author. I remark that different matching procedures (Kernel - different types - and Stratification) yield almost identical point esimates for the ATT.

I can now turn attention to the presence of the *Gift of Time* effect. Table 10 contains differences between β^*_{ATT} and β^*_{RDD} coefficients providing unbiased estimates of schooling performance of pupils anticipating of one year school entry.

I detect severe penalty for anticipating pupils which would be underestimated if the selection bias were not considered. Penalties are present in both Reading and Mathematics and persist during the entire primary education path.

2.6.2 Can the Gift of Time Redeem the Family Gift?

In this Section I present an empirical exercises to further inspect the presence of a *Gift of Time*. Consider an RDD design where pupils aged 6 in January of year *t+1* are compared with those born in November-December of year *t*. In this case, I am constructing a comparison group which includes pupils older than those used in my previous RDD since in this case only those born in December were considered. This approach – based on the assumption that the month of birth is random – yields the possibility of inspecting whether pupils that are on average two month older than the "selected ones" are able to close the score gap. This procedure can be repeated by keeping fixed the treated group (pupils in advance born in January *t+1*) and comparing them with pupils aged 6 in the period October-December of year *t*; in the period September-December, and so on. In this way I can check if a *Gift of Time* actually exists since, in this case, I should observe that the RDD parameters are decreasing in the average age of the control group.

Estimates are reported in Table 11 and a graphical illustration is also provided in Figures 5-8. The results go in the expected direction. When pupils get older, they perform better in test scores compared with pupils in advance since the gap between selected and unselected is decreasing in age. Interestingly, if I consider Reading test scores for pupils at the second grade (Figure 5) age proves to be particularly important since the selection effect disappears after 6 months and a "pivotal-point" arises: the selection effect becomes negative, highlighting that at the 2nd grade reading skills are particularly sensitive to the *Gift of Time*. The same decreasing path arises for Mathematics at the second grade (Figure 6). However, albeit decreasing, in this case the selection effect remains positive and statistically significant highlighting that when mathematical reasoning and logical-skills are required differences between selected pupils and the average population cannot be completely redeemed by the *Gift of Time*. The same is true for both Reading and Mathematics scores at the 5th grade (Figure 7 and 8 respectively). Overall, the results show that a *Gift of Time* actually exists since selection becomes less important along age. In addition, the *Family Gift* appear to be important and long-lasting during primary school.

2.6.3 Robustness Check and Falsification Exercises

In this Section I present a robustness and a falsification exercise to check the reliability of RDD estimation of the selection effect. The idea is to present an alternative identification strategy which is based on entry-age provincial variation arising for pupils at the 5th grade. In this case, I can provide an estimation of the selection effect which relies on a very different construction of treated and control groups as well as an alternative estimation procedure. Finally, a falsification test is also discussed.

An alternative identification procedure can be constructed since the Italian legislation allows for autonomous setting of school entry-age in the provinces of Trento and Bolzano who are recognized by the Italian Constitutional Law as two Special and Autonomous Provinces. The provincial legislation in Trento and Bolzano can depart from the National Law by setting different threshold for mandatory schooling. In 2006, the province of Trento has fixed mandatory schooling for those children who become 6 years old by August 31st of the year in which school starts, allowing for *optional* enrollment of pupils who reach the age of 6 in the period September 1st - December 31st (L.P. 5/2006). It is then possible to use this provincial variation to build up an alternative identification strategy of the selection effect. In particular I can compare selected pupils resident in the province of Trento with all other pupils in Italy who must start school if aged 6 between September and December 2007[7]. More precisely, I am interested in a sort of difference-in-differences estimator whose intuition is graphically provided in Figure 9. In this figure I draw a negative relation between test scores and month of birth and a negative gap between pupils from Trento and those from the rest of Italy. From this graph it is easy to gather that in order to estimate the selection effect on test scores some steps are required. In particular, I need to evaluate: *i)* the difference in test scores between pupils born in the period January 1st - August 31st resident in Trento and those born in the same period who are resident in the rest of Italy; *ii)* the difference in test scores between pupils born in the period September 1st - December 31st resident in Trento and those born in the same period who are resident in the rest of Italy; *iii)* the difference between these two differences.

Whether a selection effect is present I should detect an improvement in test score performance for pupils resident in Trento aged 6 from September 2007 with respect to the performance of the Trento pupils aged 6 before September 2007 when compared with their peers from the rest of Italy. In this way I have an alternative estimate of the *Family Effect*.

In addition, this identification strategy can be supported by a placebo test implemented by using the Autonomous Province of Bolzano. Indeed, in 2006 the provincial law fixing entry-age did not depart from the national legislation (L.P. 40/2006) hence there is scope for a falsification exercise.

Table 12 contains the results. The OLS estimators applied to the 5th grade pupils resident in the Province of Trento for both Reading and Mathematics confirm the presence of a positive selection effect. In this case, pupils in advance perform better in terms of test scores with respect to the performance of all children in the

[7] Pupils attending the 5th grade in 2011/2012 are enrolled at first grade in 2007/2008. For this scholastic year, in Trento the law in force regulating the access to primary school is the L.P. 5/2006.

same province. Interestingly, if I replicate the empirical exercise using pupils born between September 1st and December 31st in the province of Bolzano, I do not detect any significant parameter.

2.6.4 A further Robustness: The Role of Cheating

In this last Section I address concerns that may arise because of possible unfair behavior adopted from both students and teachers when tests have been submitted. Indeed, while the INVALSI tests pupils performance, a target of the analysis is teachers' evaluation. Consequently, it could be reasonable to think they could help their students when doing tests and allow them to suggest each other correct answers, in order to achieve a better evaluation for the entire classroom. This cheating behavior could induce severe distortions in results casting some doubts on the reliability of my study.

To tackle the issue, I replicate previous analyses using only a particular subsample of schools from data known as *sample schools*. These are schools who have been selected by the SNV to undertake the entire testing procedure in the presence of an external supervisor. Estimates of the β^*_{ATT} and β^*_{RDD} are reported in Table 13.

It is important to remark that albeit I have drastically reduced the sample size, my previous finding are widely confirmed, showing a significant average penalty for anticipating school of one year of -4.0 and -8.0 points for Reading and Mathematics at the 2nd grade and of -3.0 points for both Reading and Mathematics at the 5th grade.

2.7 Concluding Remarks

In this chapter I examine the effect of age at school entry on Italian Normalized test scores exploiting the peculiar Italian normative setting. Unlike other studies, I deal with selection on unobservable by estimating the potential selection bias comparing pupils who *should* start school in year *t* and pupils who have the *opportunity* to start school in that year. Through this strategy I am able to estimate unbiased effect of starting primary school one year earlier on test scores. I provide results which are consistent with most of the existing literature, i.e., the youngest children in a classroom have scores lower than their older classmates. The unbiased effect of early schooling on test scores is negative both in Reading and in Mathematics and, more interestingly, it tends to persist during primary school. This evidence is not based on instrumental variables techniques whose robustness has been heavily questioned in the literature. I point out that a severe distortion in the evaluation of early entry arises when neglecting the effect of unobserved characteristics driving school entry decisions. In particular, in the presence of a positive *Family Gift* leading best pupils to enter school in advance, the penalty imposed by early school entry is substantially underestimated.

The question concerning at which age a child should start school is a controversial topic in education policy. Governments could change cutoff birth date for first enrollment into school, weighting penalties of being younger at school entry against the costs for parents in terms of child care and delayed entrance in the labor market. My work contributes to this debate posing a word of warning on the magnitude of the skill gap and on its persistence during the entire primary education track. Further researches should be devoted to understand if this gap is actually bridged in the long run.

References

Angrist J.D. and Krueger A.B. (1991), Does Compulsory School Attendance Affect Schooling and Earnings?, *The Quarterly Journal of Economics*, 106(4): 979-1014.

Angrist J.D. and Krueger A.B. (1992), The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples, *Journal of the American Statistical Association*, 87(418): 328-336.

Barua R. and Lang K. (2009), School Entry, Educational Attainment and Quarter of Birth: A Cautionary Tale of Late, *NBER Working Paper* No. 15236.

Bauer P.C. and Riphahn R.T. (2009), Age at school entry and intergenerational educational mobility, *Economics Letters*, 103: 87-90.

Bedard K. and Dhuey E. (2012), School Entry Policies and Skill Accumulation Across Directly and Indirectly Affected Individuals, *Journal of Human Resources*, 47(3): 643-683.

Bigelow E.B. (1934), School Progress of under-Age Children, *The Elementary School Journal*, 35(3): 186-192.

Cameron A.C. and Trivedi P.K. (2005), *Microeoconometrics: Methods and Applications*, New York: Cambridge University Press.

Crawford C., Dearden L. and Meghir C. (2010), When You are Born Matters: the Impact of Date of Birth on Educational Outcomes in England, *DoQSS Working Paper* No. 10-09.

Datar A. (2006), Does Delaying Kindergarten Entrance Give Children a Head Start?, *Economics of Education Review*, 25: 43-62.

Dobkin C. and Ferreira F. (2010), Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes?, *Economics of Education Review*, 29: 40-54.

Elder T.E. and Lubotsky D.H. (2009), Kindergarten Entrance Age and Children’s Achievement: Impacts of State Policies, Family Background, and Peers, *Journal of Human Resources*, 44(3): 641-683.

Fertig M. and Kluve J. (2005), The Effect of Age at School Entry on Educational Attainment in Germany, *RWI Discussion Papers* No. 27.

Fredriksson P. and Öckert B. (2006), Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labor Market Performance, *IFAU Working Paper* No. 12.

Jones M.M. and Mandeville G.K. (1990), The Effect of Age at School Entry on Reading Achievement Scores Among South Carolina Students, *Remedial and Special Education*, 11(2): 56-62.

Kawaguchi D. (2011), Actual Age at School Entry, Educational Outcomes, and Earnings, *Journal of the Japanese and International Economies*, 25: 64-80.

Leuven E., Lindahl M., Oosterbeekc H. and Webbink D. (2010), Expanding Schooling Opportunities for 4-Year-Olds, *Economics of Education Review*, 29: 319-328.

Maddux C.D. (1983), Early School Entry for the Gifted: New Evidence and Concerns, *Roeper Review*, 5(4): 15-17.

Martin A.J. (2009), Age appropriateness and motivation, engagement, and performance in high school: Effects of age within cohort, grade retention, and delayed school entry, *Journal of Educational Psychology*, 101(1): 101-114.

Mayer S.E. and Knutson D. (1999), Does the Timing of School Affect How Much Children Learn, in Mayer S.E. and Peterson P. (eds.), *Earning and Learning: How school Matters*, Washington D.C.: Brookings Institution Press, 79-102.

McEwan P.J. and Shapiro J.S. (2008), The Benefits of Delayed Primary School Enrollment Discontinuity Estimates Using Exact Birth Dates, *Journal of Human Resources*, 43(1): 1-29.

Ponzo M. and Scoppa V. (2014), The long-lasting effects of school entry age: Evidence from Italian students, *Journal of Policy Modeling*, 36: 578-599.

Robertson E. (2011), The Effects of Quarter of Birth on Academic Outcomes at the Elementary School Level, *Economics of Education Review*, 30: 300-311.

Sharp C. (1995),What's Age Got to Do with It? A Study of Patterns of School Entry and the Impact of Season of Birth on School Attainment, *Educational Research*, 37(3): 251-265.

Spitzer S., Cupp R. and Parke R.D. (1995), School Entrance Age, Social Acceptance, and Self-Perceptions in Kindergarten and 1st Grade, *Early Childhood Research Quarterly*, 10: 433-450.

Stipek D.J. (2003), School Entry Age, in Tremblay R.E., Barr R.G. and Peters R. (Eds.), *Encyclopedia on Early Childhood Development*, Montreal, Quebec: Centre of Excellence for Early Childhood Development, 1-5.

Strøm B. (2004), Student Achievement and Birthday Effects, *Mimeo*, Norwegian University of Science and Technology.

Suggate S.P. (2009), School entry age and reading achievement in the 2006 Programme for International Student Assessment (PISA), *International Journal of Educational Research*, 48(3): 151-161.

Sweetland J.D. and De Simone P.A. (1987), Age of Entry, Sex, and Academic Achievement in Elementary School Children, *Psychology in the Schools*, 24: 406-412.

Uphoff J.K. and Gilmore J. (1985), Pupil Age at School Entrance. How Many are Ready for Success?, *Educational Leadership*, 43: 86-90.

Appendix of Tables and Figures

Table 1 – Literature Review by Author, Outcome Variables, Method and Results

Author	Age variable	Outcome variable	Method	Results
Angrist and Krueger (1991)	Season of birth	Schooling and earnings	OLS 2SLS: Quarter of birth as an instrument for education	Children born in the first quarter of the year have a slightly lower average level of education than children born later in the year; students who are compelled to attend school longer by compulsory schooling laws (the youngest) earn higher wages as a result of their extra schooling.
Angrist and Krueger (1992)	Age at school entry	Educational attainment	2SLS: Quarter of birth as an instrument for entry age	Older entrants tend to attain slightly less education.
Spitzer et al. (1995)	Age at school entry	Social acceptance, self-perceptions and competences	Correlational analyses Chi square analyses Analyses of variance (ANOVA)	Few differences related to school entrance age: teachers' ratings and peer nominations generally describe initial social problems for the youngest children which are overcome by first grade. No deficits in language and math skills for younger children.
Stipek and Byler (2001)	Age at kindergarten entry	Children's academic achievement, social skills, academic engagement, relationship with teachers, and self ratings of academic skill	Analyses of variance (ANOVA) Analyses of covariance (ANCOVA)	Children who enter kindergarten relatively young initially perform less well than their older peers, but this disadvantage disappear within a few years of elementary school. No evidence for age-of-entry effects on teachers' ratings of children's social skills, engagement in academic tasks, or their relationship with teachers.
Strøm (2004)	Enrollment age	Achievement tests in reading	OLS	Students born late in the calendar year achieve significantly lower test scores in reading compared to their oldest classmates. The disadvantage from being the youngest is highest for children with relatively large home and parental resources.
Fertig and Kluve (2005)	a) Age at school entry b) Being deferred, i.e. enrolling at age 7 versus enrolling at age 6	Schooling degree and probability of repeating a grade	a) linear probability b) matching models c) 2SLS: Age at school entry according to the regulation as an instrument for actual age at school entry	a) e b) an older age at school entry is associated with a higher probability to repeat a class, a lower probability to receive a high schooling degree in West Germany, and a) a higher probability to attain a low schooling degree or less in the Eastern part of the country; b) No difference for East Germany; c) no effect of age at school entry on educational performance.
Sandgren and Strøm (2005)	Age composition within classroom	Tests in reading and math	OLS	Being in a class with older peers increases achievement in mathematics, but not in reading. Peer age effect is higher for the late born children than for the early born. Moreover, this peer age effect found in mathematics seems to be most prevalent among the students with low educated parents.
Datar (2006)	Age at school entry	Math and reading test scores	2SLS, two instruments for entrance age: (i) Number of days between a child's 5th birthday and the school's cutoff date, and (ii) State's kindergarten entrance cutoff date	1-year delay in kindergarten entrance is associated with a significant increase in math and reading test scores at kindergarten entry. This initial advantage increases by half a point in math and by 1 point in reading during the first 2 years in school.
Fredriksson and Öckert (2006)	Age at school entry	Education and labour market outcomes	2SLS: Expected age at school entry as an instrument for actual school starting age	Children who start school at an older age do better in school and go on to have more education than their younger peers. The long-run earnings effects are positive but small. However, since starting school later entails the opportunity cost of entering the labour market later, the net earnings effect over the entire life-cycle is negative.
Elder and Lubotsky (2008)	Age at school entry	Test scores; probability of repeating kindergarten, 1st or 2nd grade	OLS 2SLS: Predicted entrance age as an instrument for actual entrance age	Being a year older at the beginning of kindergarten reduces the probability of repeating kindergarten, first, or second grade in primary school. They also find differences in reading and math test scores, but as children progress through school, achievement gaps between older and younger children tend to fade away. The entrance age effect is larger and more persistent among children from higher socioeconomic status families. Having older classmates tends to raise reading and math achievement but also increases the probabilities of repeating a grade.
McEwan and Shapiro (2008)	Delayed school enrollment	Test scores, probability of repeating first grade	OLS 2SLS: Birth dates as instruments for first grade enrollment age	One-year delay decreases the probability of repeating first grade, and increases fourth and eighth grade test scores.

(Table 1 - continued on the next page)

Author	Age variable	Outcome variable	Method	Results
Barua and Lang (2009)	Age at school entry	Educational attainment	2SLS: Effect of requiring a child to enter school in the year she turns six when she would otherwise have entered a year earlier as an instrument for age at school entry	The effect of school entry age on educational attainment is very close to zero.
Bauer and Riphahn (2009)	Age at school entry	Probability that children attend high level secondary schooling given their parents' educational background	Multinomial logit models	Early school entry reduces educational mobility.
Martin (2009)	Age within cohort, grade retention, and delayed school entry	Motivation, engagement, and performance	Structural equation models	Older-for-cohort students and delayed-entry students experience some academic disadvantage in motivation, engagement, and performance. The effects of grade retention are consistently negative.
Suggate (2009)	Age at school entry	Reading achievements	Hierarchical linear regression models	No significant association.
Crawford et al. (2010)	Month of birth	Achievement test scores	Regression discontinuity approach	Younger children perform, on average, significantly worse in national achievement tests than the older peers.
Dobkin and Ferreira (2010)	Age at school entry	Educational attainment and labour market outcomes	Regression discontinuity approach	School entry laws increase educational attainment of students who enter school early, but also lower their academic performance while in school. No evidence that the age at which children enter school effects job market outcomes.
Leuven et al. (2010)	Age at school entry	Language and math scores	OLS	One additional month of time in school increases language and math scores of disadvantaged pupils while for non-disadvantaged pupils there are no effects.
Kawaguchi (2011)	Age at school entry	Test scores, educational attainment, and labour market outcomes	OLS	Older children in a school cohort obtain higher test scores and more education years than their younger peers. This difference in academic outcomes seem to turn into higher annual earnings among males.
Robertson (2011)	Age at school entry	Reading and math test scores, grade retention	OLS 2SLS: Quarter of birth as an instrument for age at school entry	Older students appear to do worse on standardized tests scores than their younger peers for the 3rd and 8th grades, whereas there are no age effects in the 5th grade (OLS). The oldest students achieve higher results on math and reading tests, as well as have lower grade retention (2SLS).
Bedard and Duhey (2012)	Age at school entry	Adult wages	OLS	One-month increase in the minimum school entry age increases wages.
Ponzo and Scoppa (2014)	Age at school entry	Test scores in reading, math and science	2SLS: the "expected age" as an instrument for the student's actual age. The expected age is the age a student shoud have on the basis of his/her month of birth and of the established cut-off date.	Younger children score substantially lower than their older peers at the fourth, the eighth and the tenth grade. The advantage of older students does not dissipate as they grow older.

Table 2 – Literature Review by Author, Outcome Variables, Method and Results in Term of Best Performers

Author	Outcome variable				Method			Best performers	
	Test scores	Educational attainment	Labor market outcomes	Other	OLS	2SLS	Other	Youngest	Oldest
Angrist and Krueger (1991)		■	■		■	■		■	
Angrist and Krueger (1992)		■				■		■	
Spitzer et al. (1995)				■ (Social acceptance, self-perceptions and competences)			■ (Correlational analyses; Chi square analyses; Analyses of variance)		
Stipek and Byler (2001)				■ (Children achievement, social skills, academic engagement, relationship with teachers, and self ratings of academic skill)			■ (Analyses of variance; Analyses of covariance)		■ (Gap disappear within a few years of primary school)
Strøm (2004)	■				■				■
Sandgren and Strøm (2005)	■				■			■ (Peer age effect)	
Fertig and Kluve (2005)		■ (Schooling degree)		■ (Probability of repeating a grade)		■	■ (Linear probability; Matching models)	■	
Datar (2006)	■					■			■
Fredriksson and Öckert (2006)	■	■	■			■			■
Elder and Lubotsky (2008)	■			■ (Probability of repeating a grade)	■	■			■ (Gap fade away through school)
McEwan and Shapiro (2008)	■			■ (Probability of repeating a grade)	■	■			■
Barua and Lang (2009)		■				■			
Bauer and Riphahn (2009)		■ (Probability of attending high level secondary school)					■ (Multinomial logit models)		
Martin (2009)				■ (Motivation and performance)			■ (Structural eq. models)	■	
Suggate (2009)	■						■ (Hierarchical linear regression models)		
Crawford et al. (2010)	■						■ (RDD)		■
Dobkin and Ferreira (2010)		■	■				■ (RDD)	■ (Educational attainment)	
Leuven et al. (2010)	■				■			■ (For disadvantaged pupils)	
Kawaguchi (2011)	■	■	■		■				■
Robertson (2011)	■			■ (Grade retention)	■	■		■ (OLS)	■ (2SLS)
Bedard and Duhey (2012)			■		■				■
Ponzo and Scoppa (2014)	■					■			■

Table 3 – Description of Variables

DEPENDENT VARIABLES:		*Description*	
Normalized test scores in Reading and Mathematics		Continuous variable (scale from 0 to 100)	
REGRESSORS: *Group*	*Dimensions*	*Description*	*Dummy variables*
Student characteristics	Date of birth (Year)	Dummy variable	$Year_{t-n}$ (Delayed students) $Year_{t}$ (Regular students) $Year_{t+1}$ (Students "In advance")
	Date of birth (Four months)	Dummy variable	1st Four months $(January-April)_{t}$ (Students not enrolled "In advance") 2nd Four months $(May-August)_{t}$ (Regular students) 3rd Four months $(September-December)_{t}$ (Regular students) 1st Four months $(January-April)_{t+1}$ (Students "In advance")
	Gender	Dummy variable	Male Female
	Country of birth	Dummy variable	Italy Foreign Country
	Pre-school attendance	Dummy variable	Daycare (yes/no) Kindergarten (yes/no)
School characteristics	School size	Discrete variable	-
	Class size	Discrete variable	-
	Index of Sample school	Dummy variable	Sample school School no sample
	School weekly hours	Dummy variable	Up to 30 hours From 31 to 39 hours 40 hours
Parents' background	Father's/Mother's country of birth	Dummy variable	Italy Foreign Country
	Father's/Mother's educational qualification	Dummy variable	'Low' if educational qualifications are: primary school certificate, lower secondary school certificate, vocational secondary school diploma (3 years of study) 'Medium' if educational qualifications are: upper secondary school diploma, another qualification higher than diploma (Fine Arts Academy, Conservatory, etc.) 'High' if educational qualifications are: university degree or postgraduate qualification
	Father's/Mother's employment status	Dummy variable	Unemployed Homemaker 'Low' if employment statuses are: Laborer, services personnel, member of cooperatives 'Medium' if employment statuses are: Self-employed worker (trader, farmer, craftsman, mechanic, etc.); Teacher, employee, military in career; Retired worker 'High' if employment statuses are: Entrepreneur, landowner; Manager, university lecturer, officer; Professional employee or freelancer (doctor, lawyer, psychologist, researcher, etc.)
Territorial characteristics	Macro-geographical area	Dummy variable	North Centre South and Islands
	Regions	Dummy variable	Abruzzo, Basilicata, Calabria, Campania, Emilia-Romagna, Friuli-Venezia Giulia, Lazio, Liguria, Lombardia, Marche, Molise, Piemonte, Puglia, Sardegna, Sicilia, Toscana, Umbria, Valle d'Aosta, Veneto, Autonomous Province of Bolzano, Autonomous Province of Trento
Interactions	Interaction "Trento*$September_{t}$-$December_{t}$"	Dummy variable	Autonomous Province of Trento*Students born between September and December of the $year_{t}$
	Interaction "Bolzano*$September_{t}$-$December_{t}$"	Dummy variable	Autonomous Province of Bolzano* Students born between September and December of the $year_{t}$

Table 4 – Mean and Std. Dev. of test scores by Date of birth Variables

	Year$_t$ (REGULAR STUDENTS)	Year$_{t+1}$ (STUDENTS IN ADVANCE)	1st Four months$_t$ (STUDENTS NOT ENROLLED IN ADVANCE)	2nd Four months$_t$ (REGULAR STUDENTS)	3rd Four months$_t$ (REGULAR STUDENTS)	1st Four months$_{t+1}$ (STUDENTS IN ADVANCE)
Reading – 2nd Grade	72.65 (17.36)	72.74 (18.10)	74.25 (16.63)	73.18 (17.26)	71.12 (17.78)	72.74 (18.10)
Reading – 5th Grade	79.49 (13.24)	79.79 (13.28)	80.10 (13.05)	79.84 (13.12)	78.73 (13.45)	79.79 (13.27)
Maths – 2th Grade	64.64 (20.87)	66.63 (22.14)	66.23 (20.10)	65.32 (20.81)	62.95 (21.29)	66.63 (22.14)
Maths – 5th Grade	58.32 (21.10)	60.12 (21.80)	58.93 (20.80)	58.88 (21.09)	57.33 (21.28)	60.12 (21.80)

Table 5 – Mean and Std. Dev. of test scores by Father's background

	Educational Qualification			Employment status				
	Low	Medium	High	Unemployed	Homemaker	Low	Medium	High
Reading – 2nd Grade	70.01 (17.93)	75.02 (16.16)	77.94 (15.26)	68.17 (19.85)	68.02 (17.90)	69.84 (17.86)	74.28 (16.60)	75.81 (16.00)
Reading – 5th Grade	77.15 (14.01)	81.64 (11.83)	84.17 (10.77)	74.20 (15.95)	76.65 (14.09)	77.11 (14.00)	80.65 (12.53)	82.36 (11.65)
Maths – 2th Grade	62.17 (21.37)	67.02 (20.00)	69.88 (19.19)	62.07 (23.37)	58.37 (20.21)	61.86 (21.23)	66.30 (20.34)	67.63 (19.89)
Maths – 5th Grade	55.44 (21.41)	61.01 (20.40)	64.64 (19.64)	53.65 (22.86)	54.11 (20.77)	55.14 (21.36)	59.94 (20.78)	61.87 (20.23)

Table 6 – Mean and Std. Dev. of test scores by Mother's background

	Educational Qualification			Employment status				
	Low	Medium	High	Unemployed	Homemaker	Low	Medium	High
Reading – 2nd Grade	69.04 (18.30)	74.58 (16.13)	78.00 (15.15)	70.25 (18.37)	70.78 (18.41)	69.51 (17.26)	75.45 (15.84)	75.92 (15.82)
Reading – 5th Grade	76.44 (14.32)	81.46 (11.78)	84.17 (10.68)	76.68 (14.95)	77.57 (14.45)	77.08 (13.50)	82.02 (11.52)	82.32 (11.36)
Maths – 2th Grade	61.38 (21.69)	66.51 (19.99)	69.74 (19.14)	62.69 (21.73)	63.91 (22.03)	60.47 (20.45)	67.07 (19.67)	67.16 (19.58)
Maths – 5th Grade	54.63 (21.55)	60.66 (20.36)	64.62 (19.59)	54.92 (21.82)	56.92 (22.09)	53.94 (20.51)	61.40 (20.09)	61.25 (19.96)

Table 7 – OLS estimates of the Effect of Year of Birth on Normalized Reading and Math Test Scores

	Reading		Mathematics	
	2nd Grade	5th Grade	2nd Grade	5th Grade
STUDENT CHARACTERISTICS				
Date of birth (Year)				
$Year_{t+1}$ *(Students "In Advance")*	-2.043***	-0.842***	-1.862***	-0.901***
$Year_{t-n}$ *(Delayed Students)*	-3.887***	-5.587***	-2.832***	-4.879***
(Omitted Variable: $Year_t$ – Regular students)				
Gender				
Male	-1.204***	-2.323***	1.017***	2.285***
(Omitted Variable: Female)				
Country of birth				
Italy	2.674***	2.036***	1.912***	1.690***
(Omitted Variable: Foreign country)				
Pre-school attendance				
Daycare	-0.319***	-0.452***	-0.294***	-0.680***
(Omitted Variable: No)				
Kindergarten	2.821***	2.667***	2.842***	2.795***
(Omitted Variable: No)				
SCHOOL CHARACTERISTICS				
School size	-0.002***	0.005***	-0.0002	0.008***
Class size	-0.021***	0.036***	-0.107***	-0.024***
Index of sample school				
Sample school	-4.404***	-2.070***	-6.728***	-5.588***
(Omitted Variable: School no sample)				
School weekly hours				
From 31 to 39 hours	-1.131***	0.084	-0.982***	-0.150
40 hours	-1.108***	-0.757***	-1.079***	-0.691***
(Omitted Variable: Up to 30 hours)				
FAMILY BACKGROUND				
Father's country of birth				
Italy	3.317***	2.089***	2.661***	2.202***
(Omitted Variable: Foreign country)				
Mother's country of birth				
Italy	2.558***	1.566***	2.104***	1.783***
(Omitted Variable: Foreign country)				
Father's educational qualification				
Medium	2.590***	2.294***	2.722***	3.030***
High	3.876***	3.289***	4.332***	4.802***
(Omitted Variable: Low)				
Mother's educational qualification				
Medium	3.327***	3.061***	3.322***	3.887***
High	5.407***	4.452***	5.340***	6.198***
(Omitted Variable: Low)				
Father's employment status				
Unemployed	-1.646***	-2.326***	-1.495***	-1.977***
Homemaker	-0.283	-1.087**	-0.794	0.310
Medium employment status	1.006***	0.779***	0.976***	1.171***
High employment status	0.821***	0.853***	0.772***	1.080***
(Omitted Variable: Low employment status)				
Mother's employment status				
Unemployed	-0.297*	-0.529***	0.030	-0.393*
Homemaker	-0.333***	-0.427***	0.190	0.498***
Medium employment status	1.193***	0.833***	1.694***	2.220***
High employment status	0.691***	0.285***	0.938***	0.856***
(Omitted Variable: Low employment status)				
TERRITORIAL CHARACTERISTICS				
Macro- geographical area				
North	-2.867***	-0.308***	-8.000***	-4.761***
Centre	-2.023***	-0.210***	-5.627***	-3.949***
(Omitted Variable: South and Islands)				
Number of Obs	282.468	276.307	282.742	275.851

Notes: 1) * p<0.1, ** p<0.05, *** p<0.01. 2) Coefficients are estimated with robust standard errors.

Table 8 – Treatment Effect of Early Schooling on Pupil's Performance. ATT nearest neighbor estimates

	Reading		Mathematics	
	2nd Grade	5th Grade	2nd Grade	5th Grade
Treatment Effect (β_{ATT})	-4.344***	-2.313***	-2.962***	-1.751***

Notes: 1) * p<0.1, ** p<0.05, *** p<0.01. 2) ATT Nearest Neighbor uses the nearest-neighbor matching method. 3) Coefficients are estimated with bootstrap standard error. 4) Propensity scores include covariates as in Table 7.

Table 9 – RDD Estimates of Early Schooling on Pupil's Performance

	Reading		Mathematics	
	2nd Grade	5th Grade	2nd Grade	5th Grade
Treatment Effect (β_{RDD})	2.286***	1.693***	4.470***	3.157***

Notes: 1) * p<0.1, ** p<0.05, *** p<0.01. 2) Kernel used: triangle. 3) Cutoff date: January_{t+1}.

Table 10 – Consistent and unbiased Estimates of Early Schooling on Pupil's Performance

	Reading		Mathematics	
	2nd Grade	5th Grade	2nd Grade	5th Grade
Treatment Effect (β_{ATT})	-4.344***	-2.313***	-2.962***	-1.751***
Treatment Effect (β_{RDD})	2.286***	1.693***	4.470***	3.157***
Unbiased Effect (β)	**-6.630*** **	**-4.006*** **	**-7.432*** **	**-4.908*** **

Note: Consistent and unbiased effect is calculated from eq. (5).

Table 11 – Treatment Effect (RDD) of Pupils Grouped by Months of Birth

	Reading		Mathematics	
Cutoff date	2nd Grade	5th Grade	2nd Grade	5th Grade
Cutoff between Dec_t and Jan_{t+1}	2.286***	1.693***	4.470***	3.157***
Cutoff between Nov-Dec_t and Jan_{t+1}	2.477***	1.278***	4.371***	2.899***
Cutoff between Oct-Dec_t and Jan_{t+1}	1.375***	0.968***	3.129***	2.395***
Cutoff between Sept-Dec_t and Jan_{t+1}	1.126***	1.293***	4.155***	2.810***
Cutoff between Aug-Dec_t and Jan_{t+1}	0.876***	0.737***	2.533***	2.037***
Cutoff between Jul-Dec_t and Jan_{t+1}	0.619***	0.597***	2.228***	1.833***
Cutoff between Jun-Dec_t and Jan_{t+1}	0.368***	0.441***	1.927***	1.623***
Cutoff between May-Dec_t and Jan_{t+1}	0.114	0.272***	1.635***	1.391***
Cutoff between Apr-Dec_t and Jan_{t+1}	-	0.183*	-	1.288***
Cutoff between Mar-Dec_t and Jan_{t+1}	-	0.116	1.250***	1.211***
Cutoff between Feb-Dec_t and Jan_{t+1}	-0.343**	0.223	1.164***	1.179***
Cutoff between Jan-Dec_t and Jan_{t+1}	-0.451***	0.069	1.111***	1.190***

Note: * p<0.1, ** p<0.05, *** p<0.01.

Table 12 – OLS estimates for Autonomous Province of Trento and Bolzano

	Reading		Mathematics	
	5th Grade Trento (Treatment)	**5th Grade Bolzano (Placebo)**	**5th Grade Trento (Treatment)**	**5th Grade Bolzano (Placebo)**
Autonomous Province of either Trento or Bolzano	-2.050***	-3.421***	-8.101***	-5.905***
Students born between September-December of the year$_t$	-1.487***	-1.493***	-2.147***	-2.156***
Interaction: Students born between September-December of year$_t$ times Province of either Trento or Bolzano	-1.693***	-0.211	-2.893**	-0.941
Number of Obs	276.304	276.304	275.850	275.850

Notes: 1) * p<0.1, ** p<0.05, *** p<0.01. 2) Coefficients are estimated with robust standard errors.
3) Estimates include covariates as in Table 7, geographic controls include regions rather than macro-areas.

Table 13 – Consistent and Unbiased Estimates of One Year Early School Entry on Pupil's Performance: SAMPLE SCHOOLS

	Reading		Mathematics	
	2nd Grade	**5th Grade**	**2nd Grade**	**5th Grade**
β_{ATT}	**-2.071****	**-2.129****	**-6.080****	**-0.724****
Number of Obs	1,589	1,589	1,576	1,576
β_{RDD}	**2.076****	**1.104****	**2.522*****	**2.333****
Number of Obs	17,865	17,865	17,978	17,978
Unbiased Effect (β)	**-4.147****	**-3.233****	**-8.602****	**-3.057****

Note: Unbiased effect is calculated according to eq. (5).

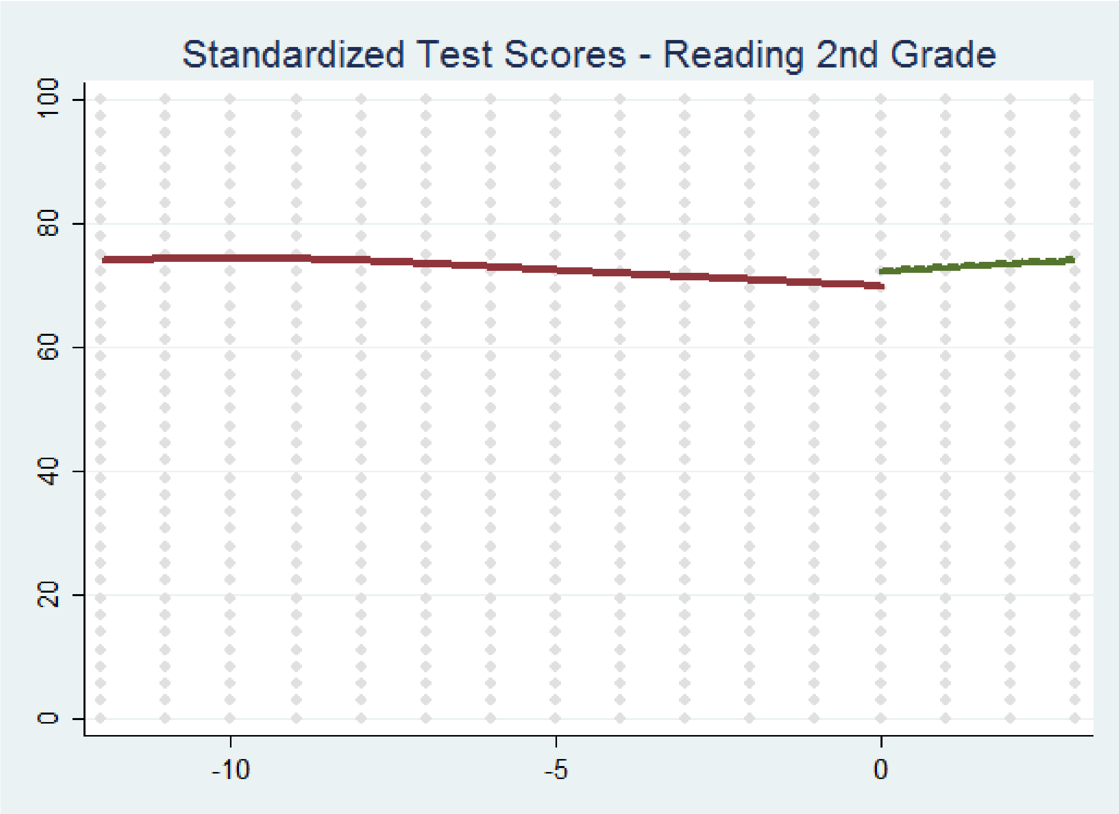


Figure 1 – RDD estimates. The horizontal axis reports months of birth, it ranges from January of the year *t* to April of the year *t+1;* the cutoff is set at January *t+1*. Difference across the margin is statistically significant at 1% level. The vertical axis reports test scores in Reading for all pupils at the 2nd Grade in Italian Schools (282.468 obs.).

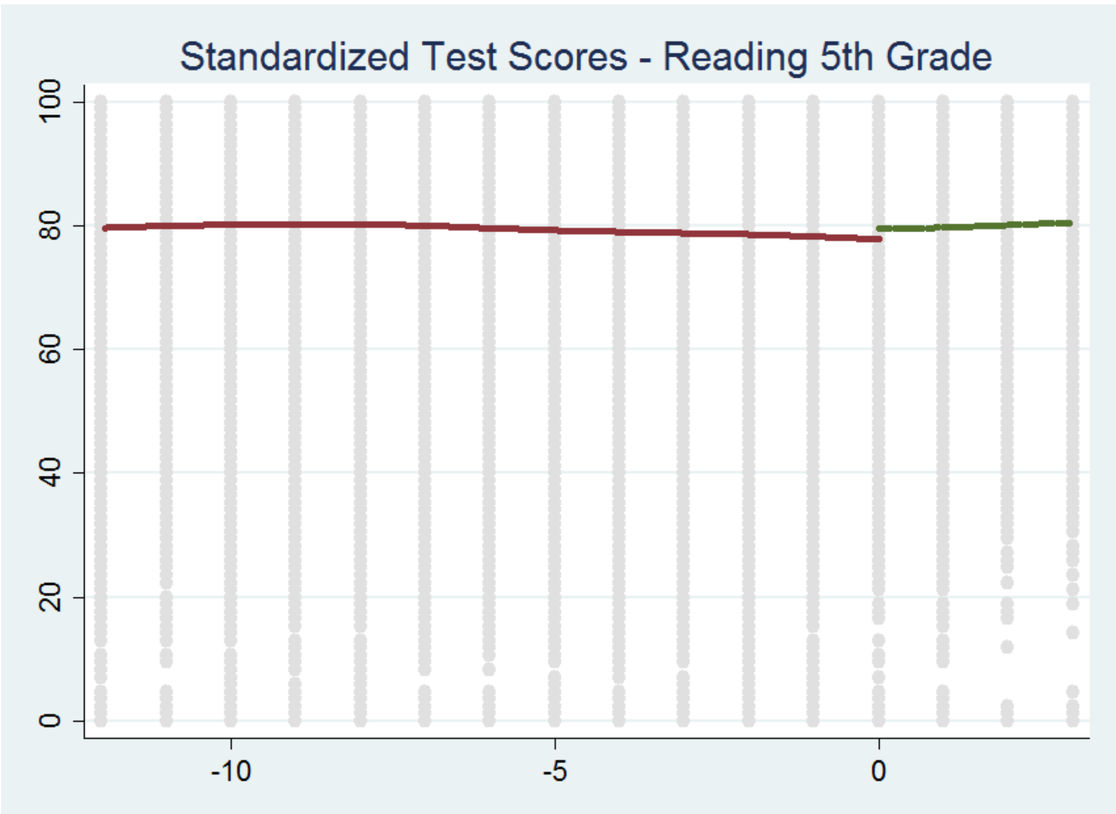


Figure 2 – RDD estimates. The horizontal axis reports months of birth, it ranges from January of the year *t* to April of the year *t+1;* the cutoff is set at January *t+1*. Difference across the margin is statistically significant at 1% level. The vertical axis reports test scores in Reading for all pupils at the 5th Grade in Italian Schools (276.307 obs.).

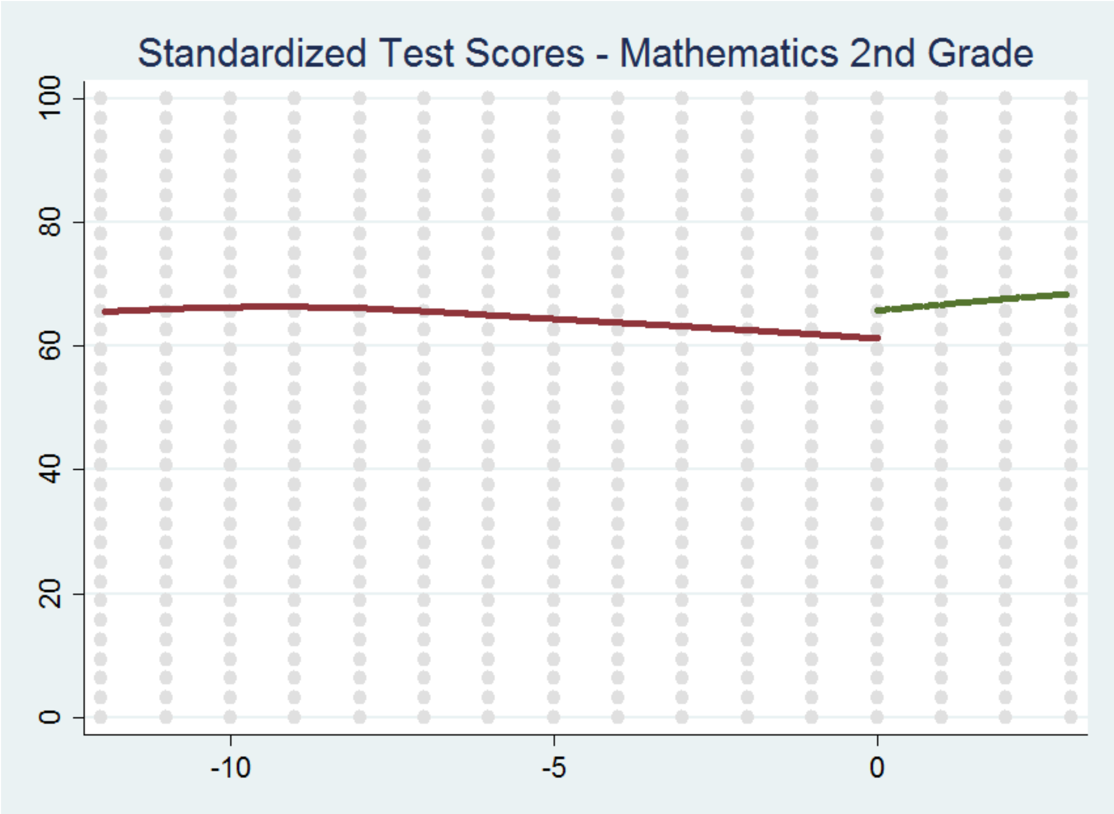


Figure 3 – RDD estimates. The horizontal axis reports months of birth, it ranges from January of the year *t* to April of the year *t+1;* the cutoff is set at January *t+1*. Difference across the margin is statistically significant at 1% level. The vertical axis reports test scores in Mathematics for all pupils at the 2nd Grade in Italian Schools (282.742 obs.).

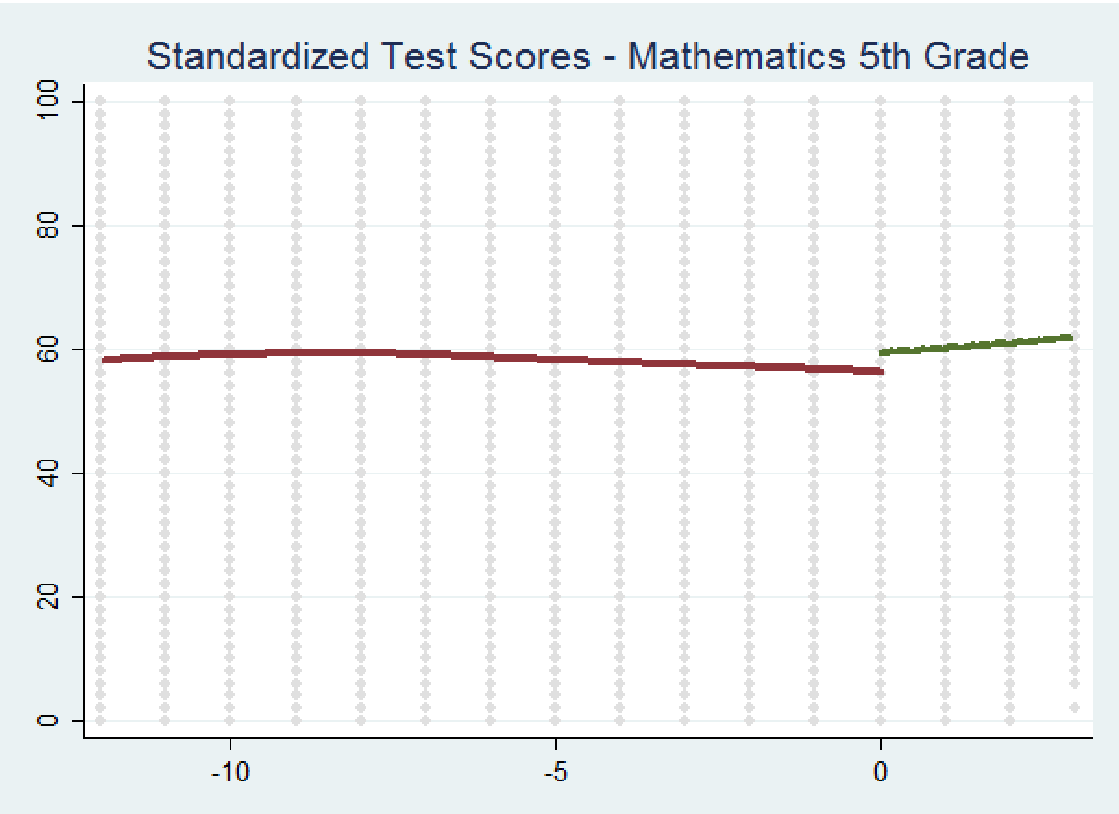


Figure 4 – RDD estimates. The horizontal axis reports months of birth, it ranges from January of the year *t* to April of the year *t+1;* the cutoff is set at January *t+1*. Difference across the margin is statistically significant at 1% level. The vertical axis reports test scores in Mathematics for all pupils at the 5th Grade in Italian Schools (275.851 obs.).

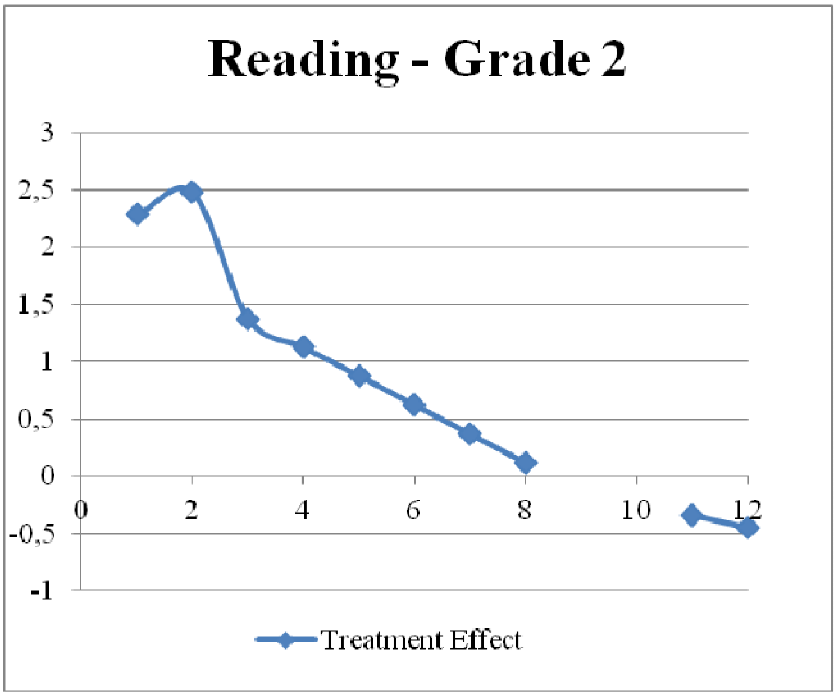


Figure 5 – Trend in the treatment effect of students grouped by months of birth.
The horizontal axis reports date of birth of students by cutoff order as in Table 11.
The vertical axis reports treatment effect of students grouped by months of birth
in Reading at the 2nd Grade in Italian Schools (282.468 obs.).

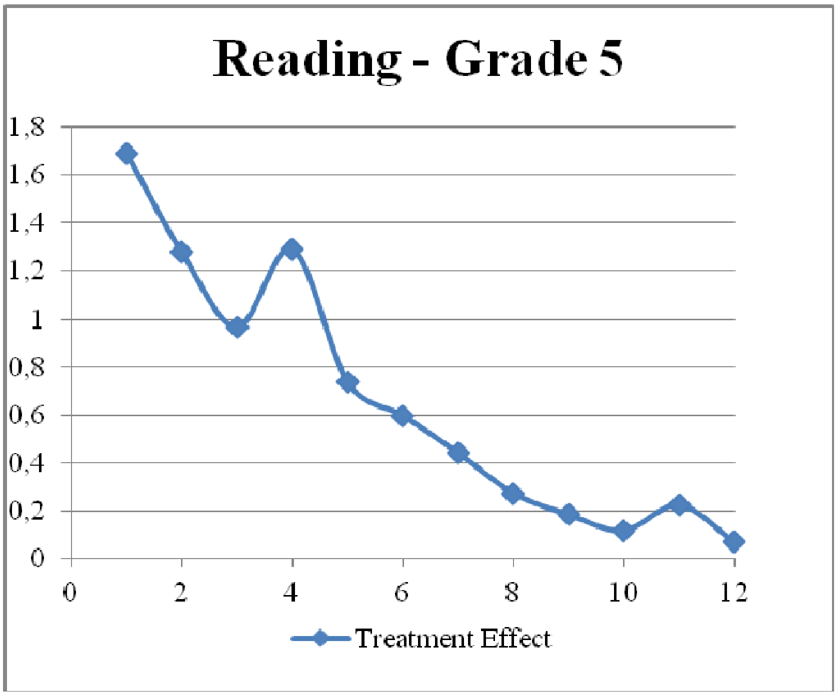


Figure 6 – Trend in the treatment effect of students grouped by months of birth.
The horizontal axis reports date of birth of students by cutoff order as in Table 11.
The vertical axis reports treatment effect of students grouped by months of birth
in Reading at the 5th Grade in Italian Schools (276.307 obs.).

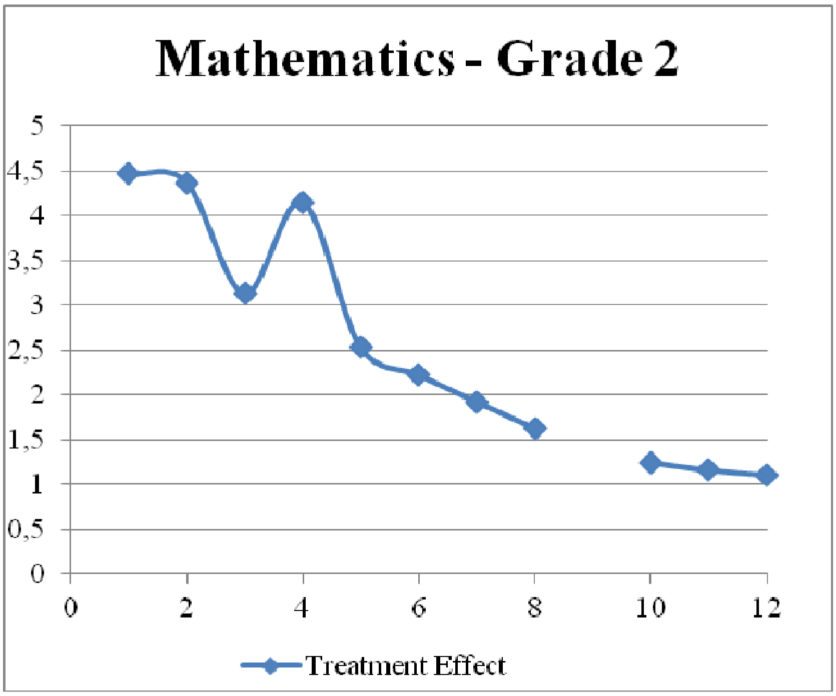


Figure 7 – Trend in the treatment effect of students grouped by months of birth.
The horizontal axis reports date of birth of students by cutoff order as in Table 11.
The vertical axis reports treatment effect of students grouped by months of birth
in Mathematics at the 2nd Grade in Italian Schools (282.742 obs.).

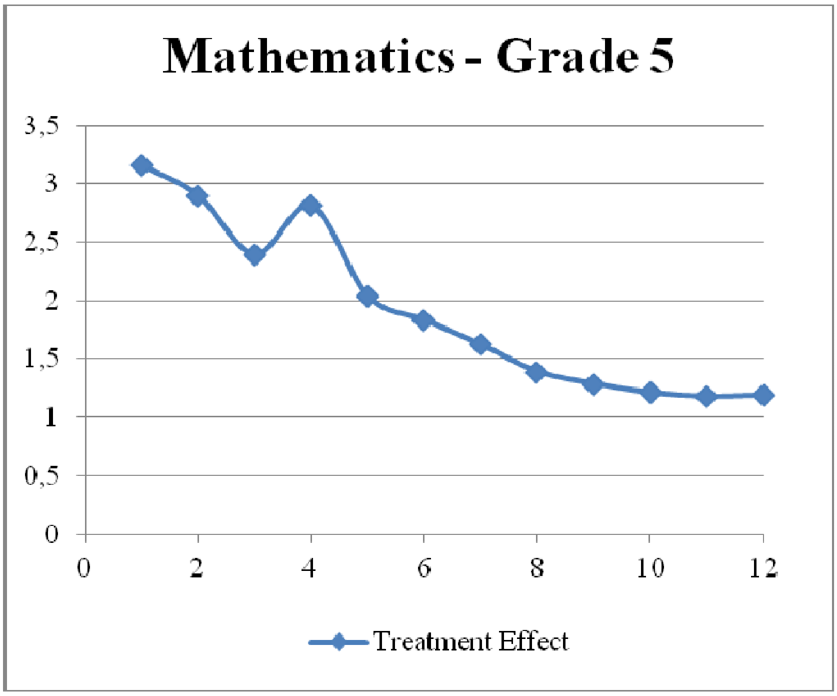


Figure 8 – Trend in the treatment effect of students grouped by months of birth.
The horizontal axis reports date of birth of students by cutoff order as in Table 11.
The vertical axis reports treatment effect of students grouped by months of birth
in Mathematics at the 5th Grade in Italian Schools (275.851 obs.).

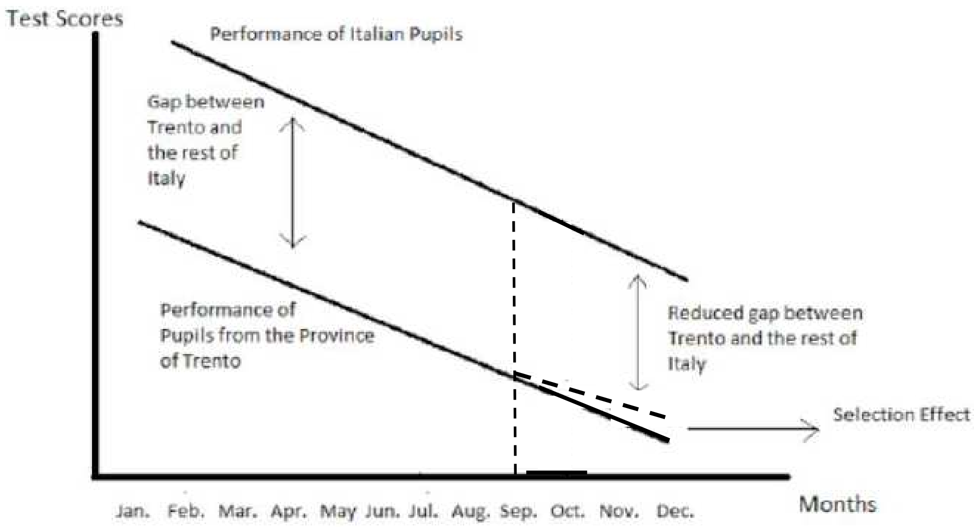


Figure 9 – Robustness Checks: Identification of the Selection Effect

CHAPTER 3

Students “in advance” and Peer Age Effect

Abstract: This chapter estimates peer age effect on schooling outcomes of Italian pupils by exploiting changes in enrollment rules over the last few years. The empirical procedure allows to understand if there is selection in classroom formation, arguing that in the absence of pupils sorting by early age at school entry, it is possible to estimate the “true” peer age effect. Results suggest that the proportion of youngest students “in advance” in the classroom has a positive impact on child’s achievements measured by Normalized and Rasch test scores. Additional empirical evidence shows that the effect on individual scores of sharing the classroom with youngest pupils “in advance” differs by students’ age group.

Keywords*:* Peer age effect, early enrollment, primary school, test scores, selection in classrooms formation.

JEL codes*:* A20, I20, I24.

3.1 Introduction

Human capital production inevitably takes place in classrooms where pupils interact all together, generating what pedagogues call *peer effects*, sociologists *contextual effects* and economists *social externalities* (Vandenberghe, 2002).

The importance of peer effects occurs from the first years of schooling. As argued by Ladd (1990), children must foster positive peer groups early at school in order to become well-adjusted adolescents and adults.

Futhemore, academic achievement and the often corresponding level of educational attainment tend to predict the average earnings an individual may secure over a lifetime. For this reason, isolating peer effects on academic achievement can make a significant contribution to the public debate over education reform.

Despite a growing literature on students' gender, ethnicity, and socioeconomic background peer effects, little literature exists on *peer age effects*. Hence, this research aims to give a novel contribution to social interactions at class level by focusing on the impact of classmates' age on individual achievements, in the awareness that classrooms are formed by children with different age and that a child's ability to accumulate human capital is affected by his/her characteristics – including age – and by characteristics of his/her peers – including age. In addition, it may be influenced by any kind of correlated group-level unobservables.

Specifically, I am interested in estimating the impact of classroom peer age composition on students outcomes in the context of Italy, where rules on first enrollment at primary school permit to have classes composed by pupils' age from 65 to 80 months at school entry.

As youngest children in the classroom may be more or less "ready" to learning, understanding how classroom age composition acts on a student's outcomes becomes an important issue.

The way through which the classmates' age operates can be differentiated. First, the presence of youngest pupils, not mature enough for school, may exert a positive spillover on individual performance because teachers redirect more attention towards pupils. Second, the presence of youngest students could create a more disciplined school environment generating a positive spillover on the entire classroom. Then, youngest children could be more able with respect to other students in the class. In this case, the effect of peer age is due to a learning spillover between classmates. However it should be also considered that youngest pupils could experience more learning difficulties with a negative effect on their peers' educational outcomes.

Understanding the way social interactions affect academic achievement is important for parents, educators, and policymakers, but, in practice, estimating them is a difficult task. Empirical research which seeks to identify peer effects runs into two problems: omitted variable bias due to selection into a group and common teacher effects that influence all members of a group (i.e., correlated effects), and the reflection problem described by Manski (1993).

The main barrier that must be overcome when estimating peer effects on student achievement is, however, the *selection problem*, as this issue reflects all unobserved characteristics that may confound peer effect estimate.

Students could be sorted into peer group with other students with similar characteristics. First, parents send their children into schools based on their job locations or residential preferences – they live or work near the school – or on the basis of peer in the community as well as the quality of the school peers. Second, classes within schools could be formed more o less randomly with respect to family background or other students characteristics. For example, in primary schools, pupils from the same neighborhood or kindergarten could be put in the same classroom. In the middle and upper secondary school, moreover, students could be systematically assigned to classes by abilities – that is by similar score achieved in the previous grade – in order to minimize teaching difficulty.

Another type of selection could take place in the formation of classes. Parents could influence the particular class to which their child is assigned within his/her school considering teachers' quality. If for example, parents believe that a certain teacher is best, they could get their children assigned to his/her class, creating a class in which parents care about to an unusual degree.

If classrooms are formed randomly, the assignment of students to classes does not suffer from selection and, in the absence of teachers sorting too, the peer effect can be easily estimated.

Starting from these considerations, I introduce an identification strategy generating estimates of peer age effects on educational outcomes that are credibly free of selection.

Exploiting changes in Italian enrollment rules over the last few years, I am able to extract the causal impact of peer age group and to understand if the effect of classmates' age is due to selection in classrooms formation.

Thus, the question raised in this research is twofold:

1) Does the age of a child's peer affect the child's cognitive achievement? And specifically, does the proportion of youngest students in the classroom affect individual performance?

2) Is this peer age effect due to selection in classroom formation?

To answer these questions I use data from INVALSI (National Institute for the Educational Evaluation of Instruction and Training), focusing on assessment of pupils attending the 2nd and the 5th grade of primary school.

The chapter is organized as follows. The next Section provides an overview of the previous studies on peer effects in education. Section 3 explains my insights and identification strategy. In section 4 data source and variables used in the analysis are described. Section 5 discusses the empirical framework and Section 6 provides results of peer age effect on scores. A first robustness check is presented in Section 7. Section 8 analyzes the impact of youngest pupils in the classroom on educational achievement by students' age group. Section 9 provides further robustness checks while Section 10 concludes.

3.2 Literature Review

The estimation of peer effects in the classroom and at school has received intense attention in recent years. However, very little literature exists regarding *peer age effect* on educational outcomes.

Researchers have used various approaches to solve peer effects estimation issues.

One common strategy to deal with *the problem of correlated effects* - i.e. the concern that measures of peer achievement may be biased for omitted unobservable characteristics that affect individual achievement - is to implement a fixed effect model. Most studies introduce school fixed effects to address omitted variable bias due to self-selection into a school and so, to take into account non-random assignment of students across schools (e.g., Ammermueller and Pischke, 2009; Boucher et al., 2010; Duflo et al., 2008; McEwan, 2003; Ponzo and Scoppa, 2014). Other researches exploit the availability of large panel administrative datasets to introduce student fixed (e.g., Carman and Zhang, 2012; Hanushek et al., 2003; Lavy et al., 2009), grade-within school and cohort-by-grade effects (e.g., Angrist and Lang, 2004; Hanushek et al., 2003). Finally, some studies use teacher fixed effects to address common teacher influences (e.g., Burke and Sass, 2013; Carman and Zhang, 2012).

The *reflection problem*, arising when a researcher observing the distribution of behaviour in a population tries to infer whether the average behaviour in some group influences the behaviour of individuals that comprise the group[1], is handled using two main strategies. Most papers use instruments to obtain consistent estimates of the endogenous peer effect (e.g., Angrist and Lang, 2004; 2010; Duflo et al., 2008; Foster, 2006; Hoxby, 2000; Kang, 2007; Ponzo and Scoppa, 2014). A second strategy is to use lagged peer achievement as a proxy for current achievement (e.g., Lavy et al., 2009; Lefgren, 2004). Specifications based on lagged peer achievement eliminate the problem of simultaneous equations. This approach requires panel data to be implemented.

In short, various strategies have been proposed to address the issues present in the estimation of peer effects. Most of them rely on strong assumptions that are difficult to motivate and to hold in practice. Moreover, some of them requires panel data.

The majority of studies focusing on peer effects examine the effect of peer ability on students' outcomes. Several researches provide evidence on race, gender and socioeonomic background peer effects too.

Few contributions on peer age effects currently exist. For example, Elder and Lubotsky (2009) focus on the relationship between kindergarten entrance age and school achievements, arguing that school achievement primarily reflects skill accumulation prior to kindergarten, rather than a heightened ability to learn in school among older children. Their results suggest that having older classmates tends to raise Reading and Math achievement but also increases the probabilities of repeating a grade.

Both Leuven and Rønning (2011) and Sandgren and Strøm (2005) find that students in Norway benefit from sharing the classroom with older peers. Leuven and Rønning (2011) conclude that students in multi-grade classrooms perform better than students in single-grade classrooms and attribute this to students benefiting from sharing the classroom with older peers. Sandgren and Strøm (2005) examine whether students with older peers achieve higher score levels in Math and Reading at the 4th grade. They find a positive effect on achievement for male students but not for females.

[1] Manski uses the term "reflection" as he considers the problem similar to that of interpreting the almost simultaneous movements of a person and his reflection in a mirror. Does the mirror image *cause* the person's movements or *reflect* them? Or do the person and image *move together* in response to a common external stimulus? (For details, see Manski, 1993).

Boucher et al. (2010), instead, find that peer's average age have a positive effect on individual test scores by implementing an IV approach, but a negative effect by using the Conditional Maximum Likelihood (CML) method.

More recently, Ponzo and Scoppa (2014) find no effect in Italian school context of a pupil's relative age with respect to the classmates' age. They employ an instrumental variable approach instrumenting average age of classmates with average of their expected age, that is the age a student should have on the basis of the cut-off date according to the enrollment rule.

Table 1 provides a schematic review of the main recent studies on peer effects in schools, highlighting country object of analysis, research focus, data, methods and the main results. In Table 2, instead, I report the peer focus (gender, age, race, ability, socioeconomic background) and the specifications used for peer measures.

3.3 Insights and Identification Strategy

My research follows the approach of Zimmerman (2003) who proposes a way to obtain unbiased peer effects. Specifically, the author examines the relationship between a student's first year college grade point average and certain observable academic characteristics of his/her roommate. Thus, if the housing assignment process is such that these *peers are randomly assigned* or are randomly assigned conditional on information that is known to the researcher, the relationship between the college outcomes of students and the characteristics of their roommates represents a *causal impact of peers that is not invalidated by the hard bias that may occur when students' peers are endogenously determined*.

Since Zimmerman (2003), some studies run OLS to estimate roommate peer effects on freshman student outcomes, exploiting random assignment of students to housing and obtaining unbiased ability rommates effects (see, for example, Winston and Zimmerman, 2003; Stinebrickner and Stinebrickner, 2006; Brunello et al., 2010).

Differently from these authors, which focus on peer ability as a measure of peer quality, I investigate *peer age effect* on individual achievements measured by Normalized and Rasch test scores both in Reading and Mathematics. More important, I don't know beforehand that the assignment is random but I know that classrooms formation should be random after observing "balanced criteria". So, exploiting changes in enrollment rules and using some insights, I single out selection effects in classrooms formation by age at school entry.

The Italian enrollment rules set the age at first school entry. Beside to the national law, the Ministry of Education yearly issues a circular that sets the limit birth date for entry into first grade of primary school.

As I use data on pupils attending both the 2nd and the 5th Grade in the year 2012/2013, I exploit different rules for the corresponding 1st Grade enrollment that allow to implement a novel identification strategy.

My insight is the following (see Figure 1: Conceptual Framework of Identification Strategy). Students attending the 2nd Grade are enrolled in the year 2011/2012. For this academic year, Italian regulations *impose* enrollment at school to pupils who are 6 years old by December 31st of the year when the school starts (from now on, year *t*), and also *permit* enrollment to pupils who will be aged 6 by April 30th of the year $t+1$[2].

Students attending the 5th Grade are enrolled in the year 2008/2009. For this academic year, Italian regulations *impose* enrollment at school to pupils who are 6 years old by August 31st of the year *t*, and also *permit* enrollment to pupils who will be aged 6 by April 30th of year $t+1$[3].

We can note that students "in advance" are those born between January and April of the year $t+1$ for the 2nd Grade, while are those born both between September and December of the year *t* and between January and April of the year $t+1$ for the 5th Grade.

In this framework, I can firstly investigate if the proportion of students "in advance" in the classroom both at the 2nd and at the 5th Grade affects individual performance. This is to say I am interested in the coefficient of proportion of students born between January and April of the year $t+1$, which are the youngest students "in advance" in the class.

Secondly, as a selection problem may arise in estimating classroom peer age effects on scores, I look at the proportion of students attending the 5th Grade, born between September and December of the year *t*, who are pupils in advance for the 5th Grade only, according the enrollment rule. In the absence of any selection in classroom formation by early age, the correspondent coefficient will be not statistically significant.

Consequently, the peer effects of proportion of students born between January and April of the year $t+1$ give us the true peer age effect of youngest students "in advance", as they do not suffer from selection bias.

3.4 Data and Descriptive Statistics

I use data from the INVALSI (National Institute for the Educational Evaluation and Training), which yearly carries out a survey on students' Reading and Mathematics competences through the National Service for the Evaluation of Education and Training (SNV). Assessment is currently realized at the 2nd and 5th Grades – primary school, 6th and 8th Grades – middle school, and Grade 10 – high school, and involve the universe of pupils attending respective grades.

For the next years, a survey on students' performance at last grade of high school (Grade 13) is also planned. A first version of the test has already been tested in a small sample of classes in the scholastic year 2012/2013. It is thought to involve in testing on Grade 13 a large number of schools and classes, but did not arrive at an administration on a universal basis during the school year 2013-14.

In the present study, I focus on primary education and analyze data of both the 2nd and the 5th grade of the school year 2012/2013. I consider students of all Italian Provinces, except the Autonomous Provinces of Trento and Bolzano, which have special enrollment rules for first school year, not included in my identification strategy.

[2] See Circular of the Ministry of Education n.101/2010.
[3] See Circular of the Ministry of Education n.110/2007.

My sample is representative and quite large, as it consists of about 500,000 pupils for the second school level and 480,000 pupils for the fifth school level. Both grades cover about 7,000 schools and 29,000 classrooms.

Data sets of the INVALSI contain a considerable number of variables that allow to control for student-level, school-level, family-level and geographic-level characteristics.

For the purpose of this study, I generated some age measures: as I know the month and the year of birth of the student, I first computed the age at school entry (in months); then I generated peer age variables at classroom level: proportion of students born in each four months, from January-April of the year *t* to January-April of the year *t+1*. I also consider the proportion of students who are born in the year *t-n*, identifying them as delayed students.

As proxies of schooling outcomes, INVALSI provides both Normalized tests scores and Rasch test scores in Reading and Mathematics. I use both test measures. The first one represents the scores achieved by students taking the test. They are called "normalized" as computed in range 0 to 100 starting from "raw" scores. Rasch test scores are instead computed taking into account both the students' skills and the item difficulty, according the Rasch model (Rasch, 1960). In this way, Rasch scores give the opportunity to understand which and how many students show learning levels above or below the difficult of a certain item.

With reference to debate about the true reliability of results achieved by students, it is important to underline that INVALSI is working to strengthen, on the one hand, the action of training and information already started in some areas to spread the culture of evaluation and, on the other hand, to enhance the methods and control measures.

Beside to the traditional presence of external supervisor in some sampled schools, it is added in 2012/2013 the presence of second-level controllers, sent in some schools randomly selected, independently from being already sample schools. The aim is certainly to ensure a greater control of the correctness of tests execution and to avoid both *student cheating* – students copy from each other or from books – and *teacher cheating* – correct answer is suggested by the teacher.

In estimates I control for the "Sample Schools" to guarantee more reliable findings. Then, as robustness checks, I replicate results using as outcome variables scores revised for "cheating" by implementing a "correct factor" computed by the INVALSI[4].

All variables used in the analysis are described in Table 3.

In Table 4 I report the descriptive statistics of outcome variables: Normalized and Rasch test scores in Reading and Mathematics.

Normalized test scores in Reading and Mathematics range from 0 to 100. We can note the mean in Reading is around 64 scores at the 2nd Grade while reaches around 77 scores at the 5th Grade, with an increase of 13 points. In Mathematics, instead, average test score is nearly constant between the two school levels: it decreases by almost one and a half a point going from 60.7 to 59.2. However, mean Normalized test

[4] For details on computing procedure see *Rapporto SNV PN 2013* at *www.invalsi.it*.

score is higher in Reading than in Mathematics both at the 2nd and 5th Grade. In contrast, average Rasch test score is higher in Mathematics than in Reading.

3.5 The Empirical Framework

To exploit peer age effect on scholastic achievements, I use data of Italian students attending both the 2nd and the 5th Grade in the academic year 2012/2013. I start by using a pooled sample of pupils attending both grades and regress potential observable factors that may have an influence on educational outcomes by considering the following equation:

$$Y_{ics} = \alpha + \beta ClassAgeComposition_{ics} + \gamma StudC_{ics} + \delta SchoolC_{ics} + \qquad (1)$$
$$+ \eta ParentsC_{ics} + \theta GeographicC_{ics} + \phi Grade5_{ics} + \varepsilon_{ics}$$

In eq. (1) *Y* denotes individual performance measured by Normalized test score and Rasch test score in either Reading or Mathematics of student *i* in the classroom *c* in the school *s*; *StudC, SchoolC, ParentsC, GeographicC* are vectors of student, school, parents' socioeconomic background and geographic characteristics respectively, as defined in Table 3, which affect the outcome variables; *Grade5* is a dummy variable indicating that student attends the 5th school level, and ε_{ics} the individual error term.

ClassAgeComposition is a matrix of variables related to the age at school entry of pupils in the classroom, better identified as follows: *i)* Proportion of students born in the years *t-n*; *ii)* Proportion of students born between January-April of the year *t*; *iii)* Proportion of students born between May-August of the year *t*; *iv)* Proportion of students born between September-December of the year *t*; *v)* Proportion of students born between January-April of the year *t+1*.

The special focus is on the *Proportion of students born between January and April of the year t+1*. This variable represents the proportion of students "in advance" in the classroom for both grades and, more interesting, *the proportion of youngest students "in advance" in the classroom*.

But we know that, according to the Italian enrollment rule, students attending the 5th grade in 2012/2013 are also *in advance* whether they are born between September and December of the year *t*. These, however, are pupils in advance but not youngest students in the classroom.

In this setting, I am also interested in estimating the eq. (1) separately for grade 2 and 5 by considering the proportion of students in advance in the respective grades.

Specifically, for the 2nd grade I refer to *students in advance* to those born between January and April of the year *t* while for the 5th grade I join the proportion of students born between September-December of the year *t* with that of pupils born between January-April of the year *t+1*.

After estimating both pooled and separate cross-sections of pupils attending the 2nd and the 5th grade of Italian primary school, I answer to the second research question concerning the presence of selection in classroom formation by employing the insights and identification strategy described in Section 3.3. I use a pooled sample of students attending both grades and regress the following equation:

$$Y_{ics} = \alpha + \beta ClassAgeComposition_{ics} + \gamma StudC_{ics} + \delta SchoolC_{ics} + \eta ParentsC_{ics} + \theta GeographicC_{ics} + \phi Grade5_{ics} + \lambda(Grade5*ClassAgeComposition_{ics}) + \psi_{cs} + \varepsilon_{ics} \quad (2)$$

where I add to the parameters of the equation (1) the interaction vector $Grade5*ClassAgeComposition_{ics}$ and ψ_{cs} which is the class level error term. The latter reflects correlated effects, which arise when the peer group is subject to a common influence not modeled directly. Correlated effects give a biased parameter of *ClassAgeComposition* if there are *unobservable* determinants of achievement that vary across classrooms within a school and that are correlated with peer group composition[5].

In particular, if there is sorting in classrooms formation on the basis of *unobserved variables*, the estimated parameter β associated to peer age effect is confounded by correlated effects and will be biased.

I could estimate unbiased peer age effect as follows:

$$\beta^* = \underbrace{\beta^{\#}}_{\text{biased peer age effect on scores}} - \underbrace{(Selection\ Effect)}_{\text{effect of unobserved confounders on scores}} \quad (3)$$

where β^* indicates the unbiased estimator of β in eq. (2). It could be calculated by removing the *Selection Effect* from biased peer age effect on scores ($\beta^{\#}$).

This issue is more easily solved in the presence of *random assignment* of students and teachers to classrooms. Random assignment, in fact, breacks the link between peer characteristics and extraneous effects on the class.

Hence, in the absence of selection on *unobservables* I can estimate an unbiased effect of peer age by running an OLS regression of eq. (2), as $\beta^*= \beta^{\#}$ in (3). In other words, in the absence of the selection effect, the bias from correlated effects is removed and β can be estimated consistently.

Thus, through the identification strategy better described in Section 3.3, I am able to estimate the "true" effect on individual test score of sharing the classroom with youngest students "in advance". I can also prove there is no systematic assignment of youngest students "in advance" to classrooms. The empirical procedure is as follows. I start by evaluating the effect of sharing classrooms with youngest pupils "in advance", both in the Grade 2 and 5. To do this, I check the sign and the significance of the parameter of the regressor *Proportion of students born between January-April of the year t+1*. Then, I look at the *Proportion of students attending the Grade 5 and born between September and December of the year t*. I expect this coefficient is not statistically significant. If so, it is possible to argue there is no systematic assignment of students to classes by early age at school entry. Consequently, I can confirm that the estimated effect of sharing the classroom with youngest students "in advance" on schooling outcomes is the "true" peer age effect.

[5] For example, in some schools, students could be sorted into classrooms by specific characteritics (e.g., by ability). Moreover, a classroom with a certain number of pupils in advance could be assigned a not very able (or a very able) teacher but the ability characteristics of the teacher are not observable. These traits are unobserved to researchers, but influence achievement. This raises the possibility that researchers will confound the influences of unobserved student, teacher and peer characteristics in high-ability classrooms.

3.6 Estimating Peer Age Effect on Scores

3.6.1 Preliminary Findings

With the purpose of estimating the impact of classroom peer age on educational outcomes, I start by considering a pooled sample of students attending both the 2nd and the 5th grade in Italian primary school.

I control for all observed variables that are likely to affect the individual achievement. The estimated coefficients present the sign and the expected significance so I don't report all estimated parameters for simplicity.

First of all I comment on individual age coefficients. Results from Table 5 show that student's age at school entry has a positive effect on both Reading and Mathematics scores: an older pupil performs better than the younger one. All coefficients are in fact positive and statistically significant at the level of 1%. The impact is higher in Reading than in Mathematics. These findings are consistent with those I find in Chapter 2: *students in advance* – the youngest in the classroom – present severe penalties both in Reading and in Mathematics test scores.

Pooled estimates of classroom peer age effects prove that not only the individual age affects educational achievements, but also peer age. Specifically, in the interest of checking the sign and the significance of the Proportion students born between January and April of the year *t+1*, I can assert that *sharing classroom with youngest students "in advance" may arise a positive spillover on individual pupil performance*.

The impact seems to be lower in Mathematics than in Reading. For example, an increase of the proportion of youngest students "in advance" in the classroom determines, at 1% significance level, a higher individual Normalized test score, on average, of 1.139 points in Reading and 0.845 points in Mathematics, and a higher Rasch scores of 0.105 points in Reading and 0.083 points in Mathematics (see Table 5).

Respect to the 2nd grade, in which students in advance are the youngest pupils in the classroom, at the 5th grade students in advance are not only those born in the year *t+1* but also those born between September and December of the year *t*. Hence, we can evaluate the effect of the proportion of students in advance separately for grade 2 and 5 looking at Table 6. Findings show that the proportion of students in advance in the classroom positively affects individual performance at the 2nd grade. The presence of one more youngest student in the classroom generates an increase of on average 3.199 points more in Reading Normalized test scores and 4.706 points more in Mathematics Normalized test scores. Similarly, a pupil achieves on average 0.233 points more in Reading Rasch scores and 0.342 points more in Mathematics Rasch scores. All coefficients are statistically significant at the level of 1% (see Table 6).

I find, instead, that the proportion of students in advance at the 5th grade has a negative impact on individual scores both in Reading and in Mathematics (see Table 6). I think this could be due to the fact that students in advance at the 5th grade include pupils with a higher peer average age, and/or to the fact that the impact of youngest students in the classroom tends to reduce during the school career – primary education in this case.

3.6.2 Inspecting the decrease/disappearence in Classroom Peer Age Effect at the 5th Grade

To investigate if the impact of sharing the classroom with youngest pupils on individual scores tends to decrease or even disappear at the end of primary school, I provide additional estimates in Table 7.

The approach consists of considering, for the 5th grade, not the proportion of students “in advance” – those born between September of the year *t* and April of the year *t+1*, but focusing on the proportion of youngest students “in advance” in the classroom – those born in the year *t+1*.

In a first specification, I estimate the proportion of youngest students “in advance” in both the 2nd and the 5th Grade, i.e. the proportion of pupils born between January and April of the year *t+1* with the aim to check and compare the magnitude of the effect in respective grades.

Results provide evidence that sharing the classroom with youngest students “in advance” at the 2nd Grade generates a positive spillover on individual performance. This effect reduces and becomes negative at the 5th grade or even disappears. Specifically, at the 2nd grade the presence of one more youngest student “in advance” in the classroom results in an increase of on average 3.199 points more in Reading Normalized test scores and 4.706 points more in Mathematics Normalized test scores. The effect becomes not statistically significant at the 5th grade in Reading and decreases to -1.965 in Mathematics. Also in Rasch test scores I find a disappearance of the impact of the proportion of students born between January and April *t+1* in Reading, and a reduction in Mathematics.

In the Specification 2, I repeat the procedure by decomposing the proportion of pupils born between January and April of the year *t+1* into two groups: *i)* students born between January and February *t+1*, *ii)* students born between March and April *t+1*. In this way I can confirm the reduction/disappearance at the 5th Grade of the effect in which I am interested and, important too, I can check if is the proportion of youngest students “in advance” in the classroom that actually affects individual scores.

Looking at Table 7, we can note a decreasing effect of the proportion of students born between January and February *t+1* at the 5th grade with respect to the 2nd grade, and a reduction or disappearance of the impact of the proportion students born between March and April *t+1* on test scores.

More interesting, estimates provide evidence that the proportion of youngest pupils “in advance” in the classroom really influences individual performance. Coefficients associated with the proportion of students born between March-April *t+1* are higher than coefficients of the proportion of pupils born in January-February *t+1*. For example, considering the 2nd grade in which the effect is wider, an increase of the proportion of students born between January and February *t+1* determines, at 1% significance level, a higher individual Normalized test score, on average, of 1.771 points in Reading and 1.763 points in Mathematics. The proportion of pupils born in the last months of limit birth date for optional enrollment – March and April *t+1* – generates a higher spillover on individual scores. The impact is, on average, of 6.373 point more in Reading Normalized test scores and 11.281 points more in Mathematics. A broader effect for the proportion of pupils born in the last two months results in Rasch test scores too.

3.6.3 *Identification of Selection Effect and Unbiased Peer Age Effect*

In this Section I try to identify the Selection Effect, i.e. to verify if there is sorting in classroom formation by exploiting changes in Italian enrollment rules over the last few year, as described in Section 3.3. Estimates are provided in Table 8. We can first note that results on the individual age coefficients are consistent with both those I found in Chapter 2 and those I found in this Chapter in pooled estimates and estimates by grade. Pupils enrolled one year later have better score both in Reading and in Mathematics.

With reference to the *Classroom Peer Age Composition*, we can observe that the proportion of delaying students negatively affects individual scores. The presence of one more delayed student in the classroom generates a decrease of on average 7.896 points less in Reading Normalized test scores and 10.776 points less in Mathematics Normalized test scores. Similarly, a pupil achieves on average 0.334 points less in Reading Rasch scores and 0.571 points less in Mathematics Rasch scores. All coefficients are statistically significant at the level of 1% (see Table 8). The negative impact also results of the previous estimates (Table 5 and 6).

More interesting, looking at Table 8, coefficients associated to the proportion of youngest students "in advance" in the classroom, born between January and April of the year *t+1*, seem to provide evidence of a benefit for those in the same classroom. An increase of the proportion of youngest students "in advance" in the classroom determines a higher individual Normalized test score, on average, of 7.584 points in Reading and 8.999 points in Mathematics. These positive effects, which are robust at 1% significance level, are also found in regressions on Rasch test scores.

Coefficients associated to pupils in advance attending the 5th Grade, who are not the youngest pupils in the classroom – i.e. students born between September and December of the year *t*, have the expected significance. They are not statistically significant for both Reading and Mathematics tests, and for both Normalized and Rasch scores. This means that systematic assignment of students and teachers to classrooms does not seem to take place and to be relevant in determining performance of Italian primary school pupils.

Results are obtained employing the whole universe of pupils attending the Grades 2 and 5 of primary schools in Italy[6]. Data cover all schools and all classrooms. Thus, findings provide evidence that in Italy sharing the classroom with youngest children "in advance" may arise a positive spillover on individual test scores.

3.7 Does the proportion of youngest students "in advance" in the classroom really affect scores? A Robustness Check

To check the reliability of my findings, I make an empirical excercise by simulating changes in the threshold of limit birth date for first enrollment at school. Specifically, I hypothesize a change in the cutoff for the 2nd grade only, keeping fixed the threshold for the 5th grade (August of the year *t*). In this way, I can

[6] Only students attending primary schools in Autonomous Provinces of Trento and Bolzano are excluded from the sample.

confirm previous results, that is: *i)* the proportion of youngest students "in advance" in the classroom really affects individual score; *ii)* there is no selection in classroom formation.

First, I move the cutoff of the 2nd grade from December of the year *t* to February of the year *t+1*. In this simulation, pupils in advance at the 2nd grade are those born either in March or in April of the year *t+1*. Students in advance for the 5th grade are instead those born between September of the year *t* and April of the year *t+1*. Hence, pupils in advance for both grades are born in March or April *t+1*. If there is not sorting in classroom formation, the proportion of students attending the 5th grade born between September *t* and February *t+1* give a not statistically significant parameter (see Figure 2: Simulation 1).

Then, I consider that the cutoff for the 2nd grade is on October of the year *t*. Pupils in advance at the 2nd grade are born between November of the year *t* and April of the year *t+1* while pupils anticipating school entry for the 5th grade remain those born between September *t* and April *t+1*. Thus, pupils in advance for both grades are born between November *t* and April *t+1*. Information about Selection Effect can be extracted by coefficient of regressor *Student born between September and October of the year t*, which I expect is not statistically significant (see Figure 3: Simulation 2).

I also expect the parameter of students in advance in Simulation 1 is higher than that of basic estimates and the coefficient in Simulation 2 is lower. In other words, if pupils "in advance" are those actually youngest (born in March-April of the year *t+1*), the effect of sharing classroom with these students is higher. Peer age effect reduces joining more months of birth for students classified as *in advance*. This because "pupils in advance" include students with a higher average age.

Through these simulations, I would like to provide some robustness checks to my basic estimates, and to confirm that is the proportion of youngest students "in advance" in the classroom that has a positive effect on individual performance, i.e. among pupils "in advance", those having a higher impact on scores are the youngest.

Results from simulations on Normalized test scores and Rasch test scores are reported in Table 9 and 10 respectively. We can note that the effect of the proportion of students "in advance" on individual scores is decreasing with an increasing average age of pupils "in advance". For example, if I consider as students in advance those born either in March or in April of the year *t+1* (Simulation 1), the impact on Reading normalized scores is on average of 13.137 points, which reduces to 7.584 points considering as students in advance those born between January and April of the year *t+1* (Basic Estimates). The effect decreases again if I simulate a wider range of months that qualify pupils as *in advance*. In fact, assuming that early students are those born between November of the year *t* and April of the year *t+1* (Simulation 2), the coefficient reduces to 2.287. I also find a progressive reduction of parameters in Mathematics Normalized test scores. The proportion of students in advance has a parameter of 18.061 considering Simulation 1. It reduces to 8.999 in Basic Estimates and to 1.976 in Simulation 3.

I find same results estimating Simulations on Rasch test scores, both in Reading and in Mathematics. All coefficients associated to the proportion of students "in advance" are positive and statistically significant at level of 1% (See Table 9 and 10). The impact on scores reduces with a higher average age of pupils identified as “in advance”.

Through the simulation exercise, I first validate that the proportion of youngest pupils “in advance” in the classroom really affects individual score. Then, I confirm previous findings concening the absence of selection in classroom formation. Coefficients related to *Selection Effect* are also not statistically significant changing the threshold of limit birth date for first enrollment at school, both in Simulation 1 and 2 (see Figures 2 and 3 for Simulation strategy, and Table 9 and 10 for estimates).

3.8 Does Peer Age Effect differ by Students’ Age Group?

If previous estimates provide a robust evidence of a positive spillover on individual performance in sharing the classroom with youngest students “in advance”, an interesting question remains open.

The effect of the proportion of youngest pupils “in advance” in the classroom on scores could differ by age group. For example, the impact could be higher for the same youngest students than the oldest one, or vice versa. Moreover, only youngest pupils could benefit from sharing the classrooms with other youngest students “in advance”, while the effect could be inexistent for older classmates.

In this Section I explore the impact of the proportion of students born between January and April of the year *t+1*, who are youngest pupils “in advance” in the classroom, on test scores by students’ age group. In particular, focusing on pupils born in either year *t* or *t+1*, I run separate regressions by considering four age groups: *i)* students born between January and April of the year *t+1*, i.e. youngest students “in advance” in the classroom (age from 65 to 68 months at school entry); *iii)* students born between September and December *t* (age from 69 to 72 months); *iii)* students born between May and August *t* (age from 73 to 76 months); *iv)* students born between January and April of the year *t* (age from 77 to 80 months).

Table 11 presents findings of peer age effect on Normalized test scores and Table 12 reports results for Rasch test scores. The impact of the proportion of pupils born between January and April of the year *t+1* on scores is higher for other students born in same four months of *t+1*, both in Reading and in Mathematics. We can note a decreasing coefficient with increasing age. This suggests that the benefit of sharing the classroom with youngest peers “in advance” is higher for youngest pupils. The advantage decreases when student is older.

Figures 4 and 5 show the effect of the proportion of students born between January and April *t+1* on scores by students’ age group, indicating a decreasing effect with increasing age of sharing the classroom with youngest students “in advance”.

3.9 *Robustness Checks*

3.9.1 Is Selection Effect invalidated by parents' perception of wrong cutoff date?

In this chapter I focus on analyzing the impact of sharing the classroom with youngest students "in advance" on individual educational outcomes by using data on pupils attending both the 2nd and the 5th Grade in the year 2012/2013 and by implementing a novel identification strategy. Specifically, I exploit different rules of the corresponding 1st Grade enrollment for both grades to extract the effect on scores of the proportion of students born in year *t+1* (youngest pupils "in advance" in the classroom). As this effect could be bias by systematic assignment of students to classes by age at school entry, I follow the approach of Zimmerman (2003) according to which *"if peers are randomly assigned, the causal impact of peers is not invalidated by the hard bias that may occur when students' peers are endogenously determined"*. So, I verify there is not any selection bias by looking at the coefficient of students "in advance" for the 5th grade only, i.e. by checking the significance of the proportion of students born between September and December of the year *t* attending the 5th grade. In the absence of any selection in classroom formation, the correspondent coefficient will be not statistically significant. If so, peer effect estimated is the "true" peer effect.

Results from Table 8 show that selection in classroom formation seems do not take place and do not to be relevant in identifying pupils' performance in Italian primary schools. Hence, the impact of youngest students "in advance" in the classroom on individual scores is a causal impact of peers that does not suffer from selection bias.

A right question that may arise concerns what follows. Consider the enrollment rule at 1st grade for pupils attending the 5th grade in 2012/2013. It *impose* enrollment to children who will be aged 6 by August 31st of the year *t*, *allowing* for optional enrollment to pupils born from September *t* until April *t+1*.

As school start in September of the year *t*, parents' decision to enroll their children at school could, *de facto*, does not take into account the differences in cutoff date between *mandatory* and *optional* enrollment. This is to say that parents', independently from rules, could sent their offsprings to school if they are born in September *t*, October *t* or, however, by the end of the year *t*, considering as *students "in advance"* only those born in the year *t+1*. Indeed, having 72 months (corresponding to 6 years old) or 71, 70, 69 months at the beginning of the school may translate in very low differences in "readiness for school" so parents could decide to enroll children at school although they are considered as students "in advance" by law.

A sort of misperception of cutoff date for mandatory school could therefore take place. As a consequence, if I identify the absence of Selection Effect from pupils born between September and December of the year *t* attending the 5th grade, the estimated "true" peer age effect could be invalidated as Selection Effect on pupils could actually exists.

Looking at descriptive statistics of months and year of birth of students attending the 2nd and the 5th grade of primary schools (Table 13), we can note there is a substantial numerosity of pupils born between September and December of the year *t* in both grades, although for the 2nd grade students born in these months are "regular" while for the 5th grade pupils born in the last four months of the year *t* are "in advance”. This seems to mean that parents feel cutoff date for mandatory enrollment as “wrong” and enroll their children as if they were regular although the law defines them in advance. Frequencies of students enrolled at school in the year *t* reduce when pupils are born in the year *t+1*. This seems to confirm that parents have the perception that their children are in advance if born in year *t+1* only.

Starting from these relevant considerations, in this Section I provide further empirical evidence to confirm there is no systematic assignment of students to classroom by age at school entry – and, hence, that the peer age effect I find is an unbiased effect – by employing an alternative identification strategy (see Figure 6).

I focus on pupils attending the 8th grade – i.e. last year of lower secondary education – in 2012/2013 merging data with those of pupils attending the 2nd grade in the same scholastic year and I exploit once again the different enrollment rules for 1st grade.

As already said in Section 3.3, cutoff date for mandatory schooling for students attending the 2nd grade is December 31st of the year *t*. Limit birth date for optional enrollment is April 30th of the year *t+1*. For pupils attending the 8th grade, instead, cutoff date is fixed on August 31st while limit birth date for optional enrollment is March 31st of the year *t+1*[7]. It is easy to understand that students attending the 8th grade born between September and December of the year *t* are students “in advance” for the 8th grade only so I could extract the Selection Effect from these pupils. But, in doubt of a misperception of cutoff date by parents, I could find a “wrong” selection effect and estimates on peer age effect could be bias.

Indeed, also in Grade 8 students born between September and December of the year *t* are sizeable even if enrollment rule defines these months of birth as elegible for “optional” enrollment (see Table 13).

So, I move attention on students born in the year *t+1* as only these pupils could be considered the “right” students “in advance” and, hence, the systematic assignment of children to classroom could be take place with reference to youngest students “in advance” born in the year *t+1*. I can obtain information about Selection Effect checking the statistical significance of pupils attending the 8th grade and born between January and March *t+1*, i.e. looking at the coefficient of interaction variable *Grade8*Proportion of students born between* $January_{t+1}$-$March_{t+1}$. In the absence of systematic assignment of pupils to classes by early entry at school, I expect this parameter is not statistically significant.

I estimate peer age effect on Normalized and Rasch test score according equation (2) using data on grade 8 instead on grade 5. Beside to the interest in confirming the absence of Selection Effect, through these estimates I can also corroborate a positive spillover on individual score of the proportion of youngest students “in advance” in the classroom.

Results from Table 14 confirm that systematic assignment of pupils to classes by age seems do not take place in Italian first cycle of education[8]. Coefficients of interaction dummy *Grade8*Proportion of students born between* $January_{t+1}$-$March_{t+1}$ are not statistically significant both in Reading and in Mathematics.

[7] See Circular of the Ministry of Education n.90/2004.

[8] Italian Education System is structured into preprimary education following by two education cycle: *i) First cycle*, divided in primary education and lower secondary education; *ii) Second cycle*, including upper secondary school and vocational training. As previous findings provide evidence there is not selection in primary school (2nd and 5th grade) and results of alternative identification strategy implemented in this section provide evidence of the absence of selection by age in lower secondary school too, I can affirm that, overall, systematic assignment of pupils to classes seems do not take place in the first cycle of education.

Findings also confirm there is a positive impact on individual performance of sharing the classroom with youngest pupils “in advance”. All parameters of the proportion of students born between January and April of the year *t+1* are positive and statistically significant at 1% level. The effect is higher in Mathematics than in Reading. For example, one more youngest student “in advance” in the classroom generates a rise of on average 5.960 points more in Reading Normalized test scores and 8.096 points more in Mathematics Normalized test scores. Considering as educational outcome Rasch test scores, a pupil achieves on average 0.317 points more in Reading and 0.666 points more in Mathematics with an increase of one more youngest student “in advance” in the classroom.

I also present some graphs to confirm there is not systematic assignment of youngest pupils “in advance” to classrooms. Specifically, I focus on the proportion of students born between January and April *t+1* (for grade 2 and 5) and born between January and March *t+1* (for grade 8), representing distributional graphs with the aim to show how many observations in the respective samples have a certain proportion of youngest students “in advance” in the classroom. Figures from 7 to 12 indicate that between 30% and 40% of classrooms have not students born in the year *t+1*. Between 10% and 15% have around 5% of youngest pupils “in advance” in the classrooms and between 5% and 10% of classroom get to have 10% of students born in *t+1*. Only around 2%-3% of classes get to have 20% of youngest pupils “in advance”. Very low percentage of classes (near to zero) have a percentage of youngest students “in advance” ranging from 30% and 40%. Percentages higher than 50% represent outliers.

3.9.2 Does “cheating” bias results?

INVALSI data often suffer from prejudices about their reliability, although the effort that Institution takes every year to control the presence of anomalies. INVALSI, in fact, has always tried to ensure the correctness of tests by “sending” external supervisors in some sampled schools. In these schools do not emerge incorrect behaviours (see INVALSI Report, 2013). Moreover, in the last few years, INVALSI has realized, in partnership with MIUR (Ministry of Education, Universities and Research), an intensive formation/information campain in some regions of South Italy – Campania, Puglia, Calabria and Sicilia – where the cheating phenomenon is more prevalent. Also as a result of this campain, cheating behaviuor is considerably reduced over the last 2/3 years (see INVALSI Report, 2013).

Finally, in the year 2012/2013, INVALSI adopted further strategies to intensify methods and control measures. First, tests both in Reading and in Mathematics have been prepared in five different versions: for each question, the response options were arranged in a different order and, as regards Mathematics tests, were also rotate questions concering the different contents[9]. Second, the monitoring of the execution of tests has also been strengthened with the introduction of second-level controllers that, on a random basis, carried out checks on the processes taking place at different times of implementation and evaluation of tests.

[9] I remember that the content of Mathematics test consists of the following four areas (first three for the 2nd Grade): 1) Numbers, 2) Space and Figures; 3) Data and forecast; 4) Relations and functions.

Both measures have exerted a preventive and deterrent action of possible misconduct. On the basis of results of the sample classes such deterrent action seems to have been effective. In general, the distribution of the results is much more "regular" than it usually was (see INVALSI Report, 2013).

Given the changes made to the structure of the survey, the persistence of anomalies seems to be more connoted as *teacher cheating* in the sense of at least to having allowed to pupils to cheat because of a lack of supervision by teachers.

In light of these relevant considerations, in this Section I would like to check if my previous findings are biased by *cheating phenomenon*. Instead of replicating results for the subsample of *Sample Schools* in which results seems do not be invalidated by uncorrect behaviours, I provide additional evidence by using scores revised for "cheating". INVALSI datasets, in fact, contain for the year 2012/2013 a "correction factor"[10] allowing to compute "scores correct for the presence of cheating". In this way, I don't reduce my sample to observations of Sample Schools but I can obtain reliable estimates by entire universe of pupils, taking into account the precence of possible anomalies in scores.

I present results of classroom peer age effect on Rasch test scores in Tables 15, 16 and 17.

Specifically, in Table 15 I replicate results by using data on pupils attending both the 2nd and the 5th grade of primary school in Italy. Findings confirm the positive impact of the proportion of youngest students "in advance" on individual educational outcomes. The effect of sharing the classroom with a more youngest student "in advance" is higher in Mathematics than in Reading. Selection Effect results not statically significant, consistently with basic estimates.

When I consider the alternative identification strategy, using data of students attending the 2nd and the 8th grade and replicating regressions on "revised" test scores, I obtain similar results of previous estimates. Also in this case coefficients on the proportion of youngest students "in advance" are positive and statistically significant at 1% level. Moreover, the effect is wider in Mathematics than in Reading (see Table 16).

Finally, in Table 17 are reported estimates of peer age effect on Rasch test scores by students' age group, focusing on the proportion of youngest pupils "in advance" in the classroom. Taking into account pupils born in either year *t+1* or *t*, i.e. with an age at school entry ranging from 65 to 80 months, a higher effect with decreasing age is confirmed. The impact on individual performance of being placed in classrooms with youngest students "in advance" is greater in Mathematics with respect to Reading scores.

3.10 Concluding remarks

In his influential study, Coleman et al. (1966) assert that peer quality is one of the only factors that could influence student outcomes besides family background. Since then, a quite large empirical evidence has been

[10] For details on computing procedure see *Rapporto SNV PN 2013* at *www.invalsi.it*.

put forward to demonstrate that the quality of student's schoolmates is an important determinant of academic performance and, by extension, of other life outcomes.

If students are affected by characteristics of their schoolmates, that is if peer effect exists in education, then the school system that encourages an efficient distribution of peers will make human capital investiments more efficient and will, thus, increase economic growth. So, understanding the nature and importance of peer group effects in education becomes crucial for education policy.

Identifying and estimating peer effects raises some challenges. The main issue is that peer effects must be isolated from confounding factors. Especially, spurious correlation between students' outcome may arise from selection into groups and from common unobserved shocks. Spurious correlated effects may be important if the allocation of teachers and students to classes is not random (class-level selection biases).

In this study I face selection bias by exploiting changes in Italian enrollment rule occurred in the recent past. Results do not appear to be influenced by selection issues so that systematic assignment of students and teachers to classrooms does not seem to take place and to be relevant in determining students' performance of Italian primary school pupils. Through an identification strategy never used in previous studies, I show that peer age impact on academic performance may arise from a "true" spillover.

Specifically, results suggest that the proportion of youngest students "in advance" in the classrooms has a positive effect on Normalized and Rasch test scores both in Reading and in Mathematics.

Analyzing peer age effect on scores by students' age group, it appears that youngest children "in advance" have a higher benefit from being placed in a class with other youngest peers "in advance". The positive effect on individual educational outcomes of sharing the classroom with youngest pupils decreases when student is older.

All findings seem to be not invalidated by "cheating" phenomena.

The way through which the impact of the classmates' age interacts with individual performance remains an open issue. Other researches can be devoted to understand if the presence of youngest students "in advance" in the classrooms affects individual performance because of a better learning environment or through learning spillover. Spillover effect of youngest pupils "in advance" could be due to the teachers, who alter curriculum choices and redirect more attention towards students. Alternatively, the positive impact of youngest pupils "in advance" on cognitive achievements could be due to the ability of these children, who are ready for school despite their young age, or a more disciplinated school environment in which the process of teaching-learning can more easily take place.

References

Ammermueller A. and Pischke J.S. (2009), Peer effects in European primary schools: evidence from Pirls, *Journal of Labor Economics*, 27(3): 315-348.

Angrist J.D. and Lang K. (2004), Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program, *The American Economic Review*, 94(5): 1613-1634.

Arcidiacono P., Foster G., Goodpaster N. and Kinsler J. (2012), Estimating spillovers using panel data, with an application to the classroom, *Quantitative Economics*, 3: 421-470.

Boucher V., Bramoullé Y., Djebbari H. and Fortin B. (2010), Do peers affect student achievement? Evidence from Canada using group size variation, *IZA Discussion Papers* No. 4723.

Brunello G., De Paola M. and Scoppa V. (2010), Peer effects in higher education: does the field of study matter?, *Economic Inquiry*, 48(3): 621-634.

BurkeM.A. and Sass T.R. (2013), Classroom Peer Effects and Student Achievement, *Journal of Labor Economics*, 31(1): 51-82.

Calvó-Armengol A., Patacchini E. and Zenou Y. (2009), Peer effects and Social Networks in Education, *Review of Economic Studies*, 76: 1239-1267.

Carman K.G. and Zhang L. (2012), Classroom peer effects and academic achievement: Evidence from a Chinese middle school, *China Economic Review*, 23: 223-237.

Coleman J.S., Campbell E.Q., Hobson C.J., McPartlan J., Mood A.M., Weinfeld F.D. and York R.L. (1966), *Equality of Educational Opportunity*, Washington, DC: Government Printing Office.

Duflo E., Dupas P. and Kremer M. (2008), Peer effects, teacher incentives, and the impact of tracking: evidence from a randomized evaluation in Kenya, *NBER Working Paper* No. 14475.

Eisenkopf G. (2010), Peer effects, motivation, and learning, *Economics of Education Review*, 29: 364-374.

Elder T.E. and Lubotsky D.H. (2009), Kindergarten Entrance Age and Children.s Achievement: Impacts of State Policies, Family Background, and Peers, *Journal of Human Resources*, 44(3): 641-683.

Epple D. and Romano R.E. (1998), Competition between Private and Public Schools, Vouchers, and Peer-Group Effects, *The American Economic Review*, 88(1): 33-62.

Foster G. (2006), It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism, *Journal of Public Economics*, 90: 1455-1475.

Hanushek E.A., Kain J.F., Markman J.M. and Rivkin S.G. (2003), Does peer ability affect student achievement?, *Journal of Applied Econometrics*, 18: 527-544.

Hoxby C. (2000), Peer Effects in the Classroom: Learning from Gender and Race Variation, *NBER Working Paper* No. 7867.

Kang C. (2007), Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea, *Journal of Urban Economics*, 61: 458-495.

Kirk A.J. (2000), *The Peer Effect on Academic Achievement among Public Elementary School Students. A Report of the Heritage Centerfor Data Analysis*, Washington, DC: The Heritage Foundation.

Ladd G.W. (1990), Having Friends, Keeping Friends, Making Friends, and Being Liked by Peers in the Classroom: Predictors of Children's Early School Adjustment?, *Child Development*, 61: 1081-1100.

Lavy V., Silva O. and Weinhardt F. (2009), The good, the bad and the average: evidence on the scale and nature of ability peer effects in schools, *NBER Working Paper* No. 15600.

Lefgren L. (2004), Educational peer effects and the Chicago public schools, *Journal of Urban Economics*, 56: 169-191.

Leuven E. and Rønning M. (2011), Classroom grade composition and pupil achievement, *IZA Discussion paper series* No. 5922.

Levin J. (2001), For whom the reductions count: A quantile regression analysis of class size and peer effects on scholastic achievement, *Empirical Economics*, 26: 221-246.

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

McEwan P.J. (2003), Peer effects on student achievement: evidence from Chile, *Economics of Education Review*, 22: 131-141.

Ponzo M. and Scoppa V. (2014), The long-lasting effects of school entry age: Evidence from Italian students, *Journal of Policy Modeling*, 36: 578-599.

Rasch G. (1960), *Probabilistic models for some intelligence and attainment tests*, Oxford, England: Nielsen & Lydiche.

Sandgren S. and Strøm B. (2005), Peer effects in primary school: Evidence from age variation, *Annual Conference EALE*, Prague.

Stinebrickner R. and Stinebrickner T.R. (2006), What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds, *Journal of Public Economics*, 90: 1435-1454.

Vandenberghe V. (2002), Evaluating the magnitude and the stakes of peer effects analysing science and math achievement across OECD, *Applied Economics*, 34(10): 1283-1290.

Winston G.C. and Zimmerman D.J. (2003), Peer Effects in Higher Education, *NBER Working Paper* No. 9501.

Zimmer R.W. and Toma E.F. (2000), Peer Effects in Private and Public Schools across Countries, *Journal of Policy Analysis and Management*, 19(1): 75-92.

Zimmerman D.J. (2003), Peer effects in academic outcomes: Evidence from a natural experiment, *The Review of Economics and Statistics*, 85(1): 9-23.

Appendix of Tables and Figures

Table 1 – Peer Effects on Students' Performance Literature Review

Author	Country Object of Analysis	Data source	School Level	Research focus	Method	Main results
Ammermueller and Pischke (2009)	6 European Countries: Germany, France, Iceland, the Netherlands, Norway and Sweden.	PIRLS (Progress in International Reading Literacy Study)	Primary school (grade 4)	- Variation across classes within schools, which are formed roughly randomly - Peer effects on students test scores within schools	- WLS - OLS - IV	- Peer effect is modestly large - Measurement error is important in survey data - Selection plays little role in biasing peer effects estimates once measurement error issue is taken into account
Angrist and Lang (2004)	Boston (Massachusetts, USA)	- Data about Metco program - Massachusetts Comprehensive Assessment System (MCAS) testing program - Brookline Data: Test of Basic Skills (ITBS)	Primary school (grade 4) MCAS data Primary school (grades 3 and 5) ITBS data Middle school (grade 7) ITBS data	Racial school integration	- OLS - IV	Peer effects from Metco (Metropolitan Council for Educational Opportunity) program, i.e. effects on minority students in the host district, are modest and short-lived.
Arcidiacono et al. (2012)	Maryland (USA)	Administrative data from University of Maryland	Higher education	Spillover in education	Monte Carlo iterative algorithm	Small but significant peer effects are found, with evidence of heterogeneity by course type
Boucher et al. (2010)	Quebec (Canada)	Quebec Government MERS (Ministry of Education, Recreation and Sports)	Secondary school (grade 4 and 5)	Peer effects in students achievements	- OLS - Conditional Maximum Likelihood (CML) - IV - Monte Carlo simulation	- While a rise in own age is associated with a decline in own test score, peers' average age have a positive effect on own test scores in IV estimates but a negative effect in CML estimates - Peers' socioeconomic background has little effect on own schooling performance - Average test score of his peers increases a student's test score
Brunello et al. (2010)	South Italy	- Administrative data covering students who live on campus at the University of Calabria - PISA (Programme for International Student Assessment)	Higher education	- Rommate peer effects for freshman enrolled - Effort at college	OLS	- Roommate peer effects for freshmen enrolled are positive and significant for hard sciences students, and close to zero or negative in the humanities and social sciences - A theoretical model suggests that the uncovered differences between fields in the intensity of the peer effect could be generated by between-field variation in labor market returns, which affect optimal student effort
Burke and Sass (2013)	Florida (USA)	Florida Comprehensive Assessment Test-Norm Referenced Test (FCAT-NRT)	Elementary, middle and high schools (grades 3-10)	Peer effects on individual student performance	- Value-Added Model of Student Achievement - OLS - Quantile regression	- Peer effects only at the classroom level and not at the general grade level - Low-ability students appear to benefit significantly from having top-quality peers while highest-ability students benefit from mixing with students of middling ability

(Table 1 - continued on the next page)

Author	Country Object of Analysis	Data source	School Level	Research focus	Method	Main results
Calvó-Armengol et al. (2009)	USA	National Longitudinal Survey of Adolescent Health (Add Health)	Secondary school (grades 7-12)	Peer effects and Social Networks in Education	- Network fixed effects OLS model - Network fixed effects Maximum Likelihood model	Pupil school performance is affected by peer effects
Carman and Zhang (2012)	China	Data from a middle school in the capital city of a North China province	Middle schools (grades 7 to 9)	Peer effects on students achievements	- OLS - Quantile regression	- Peers have a positive and significant effect on math test scores, but no significant effect on Chinese and English test scores - Students at the middle quintile of the ability distribution tend to benefit from better peers, whereas students at both ends of the ability distribution do not
Duflo et al. (2008)	Kenya	Data from Extra-Teacher Program (ETP), a primary school class-size reduction experiment	Primary school	Peer effects, tracking and teacher incentives	- OLS - IV - RDD	- Students in tracking schools performed higher than those in non-tracking schools - Students in non-tracking schools scored higher if they were randomly assigned to peers with higher initial scores - Peers affect students both directly and indirectly by influencing teacher behavior, in particular teacher effort and choice of target teaching level
Eisenkopf (2010)	Switzerland	Data from a experiment conducted in some Cantones of Switzerland	High school (age 15-18)	Motivation and peer effects	- OLS - Poisson regression - Negative binomial regressions	Some of the "better" students improve the performance of their partner but they induce lower motivation
Elder and Lubotsky (2009)	USA	- Early Childhood Longitudinal Study-Kindergarten cohort (ECLS-K) - National Educational Longitudinal Study (NELS)	Kindergarten Middle school (grade 8)	Kindergarten entrance age and school achievement	- OLS - IV	- Being a year older at the beginning of kindergarten reduces the probability of repeating kindergarten, first, or second grade of primary school - Oldest children perform better than youngers in reading and maths scores, but gap tend to fade away as children progress through school - The entrance age effect is larger and more persistent among children from higher socioeconomic status families - Having older classmates tends to raise reading and math achievement but also increases the probabilities of repeating a grade
Epple and Romano (1998)	USA	Data from various sources	Not specified	- Peer group effects - Voucher system in education - Competition between private and public school	Cobb-Douglas specification	- Achievement depends on own and peers' ability - Because in private schools, high-ability low-income students receive tuition discounts, while low ability high-income students pay tuition premia, tuition vouchers increase the relative size of private sector and the premium on ability, benefiting high-ability students - Students remaining in the public sector are those with relatively low income and low ability, and those students experience losses

(Table 1 - continued on the next page)

Author	Country Object of Analysis	Data source	School Level	Research focus	Method	Main results
Foster (2006)	Maryland (USA)	Administrative data for undergraduates residing in University of Maryland housing	Higher education	Peer effects in education	- OLS - IV	Overall, no meaningful or statistically significant peer effects are found
Hanushek et al. (2003)	Texas (USA)	Dataset constructed by the UTD Texas Schools Project, which used Texas Assessment of Academic Skills (TAAS) data	Primary school (grades 3 to 5)	Peer effects on students achievements	- OLS - Value-added specification (achievement gain between current and previous grade)	- Peer achievement has a positive effect on individual achievement growth - Students throughout the school test score distribution appear to benefit from higher achieving schoolmates
Hoxby (2000)	Texas (USA)	- Texas Schools Microdata Panel - Texas Assessment of Academic Skills (TAAS) data	Primary school (grades 3 to 6)	Classroom peer effects	- WLS - IV	- Students are affected by peer achievements - Peer effect are stronger intra-race - Females' math performance is about the same as that of males, but both males and females perform better in math in classrooms that are more female
Kang (2007)	South Korea	TIMSS (Third International Mathematics and Science Study)	Middle schools (grades 7 and 9)	Classroom peer effects	- OLS - IV - Quantile regression	- Mean classroom achievement is positively correlated with a student's performance - Weak students interact more closely with other weak students than with strong students; hence their learning can be delayed by the presence of worst-performing peers. In contrast, strong students are found to interact more closely with other strong students; hence their learning can be improved by the presence of best-performing peers
Kirk (2000)	USA	National Assessment of Educational Progress (NAEP)	Primary, middle and high schools (grades 4, 8 and 12)	Peer effects in education	Jackknifed ordinary least squares model	- Peer effects have a strong influence on academic achievement, particularly in 4th grade; the significance of peer effect wanes by 8th grade - Peer effect is independent of other factors such as race, ethnicity, gender, income, and other background variables
Lavy et al. (2009)	England	- Standard National Tests (SATS) - Pupil Level Annual School Census (PLASC)	Secondary school (grade 9)	Peer ability effects	- OLS	- 'Bad' peers at school, as identified by students in the bottom 5% of the ability distribution, negatively and significantly affect the cognitive performance of schoolmates - Little evidence that the average peer quality and the share of 'very good' peers, as identified by students in the top 5% of the ability distribution, affect the educational outcomes of other pupils - Girls significantly benefit from the presence of very academically bright peers, while boys marginally losing out
Lefgren (2004)	Chicago (USA)	- Chicago Public Schools (ChiPS) - Iowa Test of Basic Skills (ITBS)	Primary school (grades 3 and 6)	Classroom peer ability effects in tracked and untracked schools	- OLS - IV	Peer effects are quite small, though generally positive and statistically significant

(Table 1 - continued on the next page)

Author	Country Object of Analysis	Data source	School Level	Research focus	Method	Main results
Leuven and Rønning (2011)	Norway	- Administrative enrollment data provided by Statistics Norway - School database GSI (Grunnskolens Informasjonssystem)	Junior High School (mixed grade classrooms)	Classroom grade (which equals to age) composition and pupil achievement	- OLS - IV	- Pupils in mixed grade classrooms outperform pupils in single grade classrooms - Pupils benefit from sharing the classroom with more mature peers from higher grades
Levin (2001)	The Netherlands	PRIMA	Primary school (grades 4, 6, 8)	Class size and peer effects	- IV - IV Quantile regression: 2SLAD (two stage least absolute deviation estimator)	Individuals in the lower proportion of the achievement distribution benefit more from being placed in classes with individuals of similar ability
McEwan (2003)	Chile	Sistema de Medición de la Calidad de Educación (SIMCE)	Primary school (grade 8)	Classroom peer effects	OLS	The classroom mean of mothers' education has the strongest link to individual achievement, though subject to diminishing marginal returns
Ponzo and Scoppa (2014)	Italy	- PIRLS (Progress in International Reading Literacy Study) - TIMSS (Third International Mathematics and Science Study) - PISA (Programme for International Student Assessment)	Primary school (grade 4, 8) Secondary school (Grade 10)	Absolute and relative age effects on students' performance	- IV - Discontinuity Sample Strategy	- Younger children score substantially lower than older peers - The advantage of older students does not dissipate as they grow - There is not any significant effect of the relative age of a child with respect to the classmates' age
Sandgren and Strøm (2005)	Norway	Administrative data and individual scores on national test in mathematics and reading in all public schools and a fraction of private schools	Primary school (grade 4)	Classmates age effects on individual student achievement	OLS	- Being in a class with older peers increases achievement in maths, but not in reading - Peer age effect is higher for the late born children than for the early born - Peer age effect found in mathematics seems to be most prevalent among the students with low educated parents
Stinebrickner and Stinebrickner (2006)	Kentucky (USA)	- Administrative data from Berea College - Data from the Berea Panel Study	Higher education	Roommate peer effects on freshmen student outcome	OLS	- No evidence of a relationship between college grades and roommate ACT score for females - Evidence in the relationship between college grades and both roommate high school grade point average and peer roommate family income for females - No peer effects on males performance
Vandenberghe (2002)	17 OECD Countries: Australia, Austria, Belgium (Flemish and French Speaking Community), Canada, France, Germany, Greece, South Korea, the Netherlands, New Zealand, Norway, Singapore, Switzerland, Spain, Scotland and the USA.	TIMSS (Third International Mathematics and Science Study)	Secondary school (grade 7 or 8)	Peer effects across OECD countries	OLS	- The higher the mean SES of the classmates, the higher the achievement level of the student - Low-SES pupils are more sensitive to peer effects than their more privileged mates - A student's achievement level is lower, the greater the underlying heterogeneity

(Table 1 - continued on the next page)

Author	Country Object of Analysis	Data source	School Level	Research focus	Method	Main results
Winston and Zimmerman (2003)	USA	College and Beyond Database	Higher education	Effect of roommates' academic characteristics on an individual's GPA (Grade Point Average)	OLS	- Students in the middle of the SAT distribution may do somewhat worse in terms of GPA if they share a room with a student who is in the bottom 15 percent of the SAT distribution - Students in the top of the SAT distribution are typically not affected by the SAT scores of their roommates
Zimmerman (2003)	Massachusetts (USA)	Data from Williams College	Higher education	Effect of roommates' academic characteristics on an individual's GPA (Grade Point Average)	OLS	- Peer effects are almost always linked more strongly with verbal SAT scores than with math SAT scores - Students in the middle of the SAT distribution may have somewhat worse grades if they share a room with a student who is in the bottom 15% of the verbal SAT distribution
Zimmer and Toma (2000)	5 countries: Belgium, France, Canada (Ontario), New Zealand, and the USA.	International Association for the Evaluation of Educational Achievement (IEA)	Secondary school (grade 8)	Peer effects in private and public schools	OLS	- Peer effects are a significant determinant of educational achievement - Peers play a larger role in the achievement levels of low-ability students than they do in high-ability student achievement: raising the average peer level increases individual student achievement levels in schools across countries

Table 2 – Peer Measures used in Previous Studies

Author	Peer Measures					Specification of Peer Measures and Notes
	Student's Gender	**Student's Age**	**Student's Race**	**Peer ability (performance)**	**Socioeconomic background**	
Ammermueller and Pischke (2009)	✓	✓			✓	Socioeconomic background: - Index of n. of books at home - Foreign parent - Foreign language spoken at home Class average of socioeconomic background variables, student's gender and age have been used for the decomposition of variance in class level means. For regressions, the only measure used is *Index of the number of books at home*, as follows: - Class average n. of books at home - Class average n. of books at home*individual level dummy variable for >100 books at home - Class average n. of books at home/Individual n. of books at home
Angrist and Lang (2004)			✓			- Fraction Metco on non-Metco students (OLS regressions) - Average number of Metco students per classroom (IV regressions) Metco is a desegregation program that sends students from Boston schools to more affluent suburbs, i.e. sends black students to schools that were previously all white and vice versa.
Arcidiacono et al. (2012)				✓		- Scholastic Aptitude Test (SAT) scores - High school Grade Point Average (GPA)
Boucher et al. (2010)	✓	✓		✓	✓	Socioeconomic background: - Foreign students=0 (whose language of instruction is the same as the mother tongue and the language spoken at home) - Index of SES Peer group of a student contains all other students in the same school.
Brunello et al. (2010)				✓		Academic ability is extracted by two components: marks at graduation from secondary school and standardized average test scores in each type of high school.
Burke and Sass (2013)				✓		- Average ability of the classroom or grade-level, not including students himself - Standard deviation of peer ability - Ability quantile* average ability of the classroom - Lowest/medium/highest quantile*Fraction of peer in lowest/highest quantile
Calvó-Armengol et al. (2009)	✓	✓	✓		✓	Peer group characteristics also include other individual socio-demographic measures, residential neighborhood variables, and protective factors such as parents at home or relationship with teachers. Peer effects are identified as average values of all control variables over the students' direct friends. A unique coefficient that includes peer effects is estimated.
Carman and Zhang (2012)				✓		- Average score of other students in the class
Duflo et al. (2008)				✓		- Average score of other students in the class
Eisenkopf (2010)				✓		- Score of the partner - Partner's marks in math Peer group variables also include the interest of the partner in logical puzzles.
Elder and Lubotsky (2009)		✓				- School average entrance age, except age of child *i*
Epple and Romano (1998)				✓		- Mean ability of the student body in the school attended
Foster (2006)				✓		- Student's peer group's mean/median SAT score - Student's peer group's mean/median high school GPA
Hanushek et al. (2003)				✓	✓	- Average math score and standard deviation of scores in grade G-2 - Proportion eligible for reduced price lunch Peer measures include all other students in the school and grade.

(Table 2 - continued on the next page)

Author	Peer Measeures					Specification of Peer Measures and Notes
	Student's Gender	**Student's Age**	**Student's Race**	**Peer ability (performance)**	**Socioeconomic background**	
Hoxby (2000)	✓		✓	✓		- Average achievements of males and females in a grade in a school in a cohort - Average achievements of students according their racial groups in a grade in a school in a cohort
Kang (2007)	✓			✓	✓	Gender: - Proportion of males Peer performance: - Average value of math scores of classroom peers excluding own score - Its square term - Its standard deviation - Proportion of weak peers (excluding oneself) within a classroom who are below the 25th percentile of the math score distribution - Proportion of strong peers (excluding oneself) who are above the 75th percentile Socioeconomic background: - Books over 200 - Computer at home - Father's and Mother's education
Kirk (2000)				✓		"Make Fun of Those Who Try to Do Well in School"
Lavy et al. (2009)				✓		- Average ability of peers at school, measured by test scores achieved by students at age 11 at the end of primary school (grade 6) - Fraction of very high-ability peers in one students' cohort (those who are above the 95th percentile) - Fraction of very low-ability peers in one students' cohort (those who are below the 5th percentile)
Lefgren (2004)				✓		- Average classmates ability, measured by the prior year's test scores
Leuven and Rønning (2011)		✓				- Average classroom grade composition (which equals to the average classroom age composition) Regressions also include Relative age, which equals 0 for the youngest pupil (born December 31st) and 1 for the relatively oldest one (born January 1st)
Levin (2001)				✓		- Number of classmates with similar IQ
McEwan (2003)			✓		✓	- Classroom mean of mother/father education - Classroom mean of family income - Classroom mean of student ethnicity (indigenous) - Squared terms of each peer variable
Ponzo and Scoppa (2014)		✓				- Average age of students in the class of *i* (excluding individual *i*)
Sandgren and Strøm (2005)		✓				- Average age (in months) of students' classmates in school
Stinebrickner and Stinebrickner (2006)				✓	✓	- Roommate ACT (American College Test) score - Roommate HSGPA (High School Grade Point Averages) - Roommate family income/10,000
Vandenberghe (2002)					✓	- Average SES of the pupil's classmates - Squared term of average SES of the pupil's classmates - SES of the student*average SES of the pupil's classmates - Average SES of the pupil's classmates**standard deviation* of the SES of the pupil's classmates
Winston and Zimmerman (2003)				✓		- Students' freshman roommate Scholastic Aptitude Test (SAT) scores - Students' freshman roommate SAT score range
Zimmerman (2003)				✓		- Students' first year roommate verbal SAT scores - Students' first year roommate math SAT scores - Students' first year roommate SAT total scores (verbal + math) - Students' first year roommate verbal/math SAT score range

(Table 2 - continued on the next page)

Author	Peer Measeures					Specification of Peer Measures and Notes
	Student's Gender	Student's Age	Student's Race	Peer ability (performance)	Socioeconomic background	
Zimmer and Toma (2000)				✓	✓	- Mean beginning-of-year test score of students in a classroom and its squared term - Mean beginning-of-year test score of students in a classroom*student score at the beginning-of-year - Mean beginning-of-year test score of students in a classroom*student score at the beginning-of-year by country variables and by school type variable (private school) - Standard deviation of the mean classroom scores - Mean beginning-of-year test score of students in a classroom*standard deviation of the mean classroom scores - Proportion of high/low ability students in the classroom*student score at the beginning-of-year - Proportion of the classmates' fathers whose occupation is professional or skilled - Proportion of the classmates' mothers who work outside the home (full or part-time) - Proportion of the classmates' fathers/mothers whose highest level of school is secondary or greater - Proportion of the classmates' fathers whose occupation is professional or skilled*student score at the beginning-of-year - Proportion of the classmates' mothers who work outside the home (full or part-time)*student score at the beginning-of-year - Proportion of the classmates' fathers/mothers whose highest level of school is secondary or greater*student score at the beginning-of-year - Private school*each of the four socioeconomic characteristics (father's/mother's occupation, father's/mother's education)

Table 3 – Description of Variables

OUTCOME VARIABLES:		*Description*	
Normalized Test Scores in Reading and Mathematics		Continuous variable	
Rasch Test Scores in Reading and Mathematics		Continuous variable	
COVARIATES:			
Group	*Dimensions*	*Description*	*Dummy variables*
Student-level variables (*Individual characteristics*)	Age at school entry (in months)	Discrete variable (range 65 to 116)	-
	Gender	Dummy variable	Male Female
	Country of birth	Dummy variable	Italy Foreign Country
	Pre-school attendance	Dummy variable	Daycare (yes/no) Kindergarten (yes/no)
Student-level variables (*Parents' background*)	Father's/Mother's country of birth	Dummy variable	Italy Foreign Country
	Father's/Mother's educational qualification	Dummy variable	'Low' if educational qualifications are: primary school certificate, lower secondary school certificate, vocational secondary school diploma (3 years of study) 'Medium' if educational qualifications are: upper secondary school diploma, another qualification higher than diploma (Fine Arts Academy, Conservatory, etc.) 'High' if educational qualifications are: university degree or postgraduate qualification
	Father's/Mother's employment status	Dummy variable	Unemployed Homemaker 'Low' if employment statuses are: Laborer, services personnel, member of cooperatives 'Medium' if employment statuses are: Self-employed worker (trader, farmer, craftsman, mechanic, etc.); Teacher, employee, military in career; Retired worker 'High' if employment statuses are: Entrepreneur, landowner; Manager, university lecturer, officer; Professional employee or freelancer (doctor, lawyer, psychologist, researcher, etc.)
School-level variables	School size (N. of classrooms in the school)	Discrete variable	-
	Index of Sample school	Dummy variable	Sample school School no sample
	School weekly hours	Dummy variable	Normal Time (Up to 30 hours in the 2nd Grade – Up to 39 hours in the 5th Grade) Full Time (40 hours)
Classroom-level variables (Peer age composition)	Proportion of delayed students (born in years t-n)	Continuous variable	-
	Proportion of students born between Jan_t-Apr_t	Continuous variable	-
	Proportion of students born between May_t-Aug_t	Continuous variable	-
	Proportion of students born between $Sept_t$-Dec_t	Continuous variable	-
	Proportion of students born between Jan_{t+1}-Apr_{t+1}	Continuous variable	-
Geographic-level variables	Province	Dummy variable	101 Dummy variables for Italian Provinces (Autonomous Provinces of Trento and Bolzano are excluded)

Table 4 – Descriptive Statistics of Outcome Variables

Otcome Variables	GRADE 2					GRADE 5				
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Normalized test score in reading	489,631	64,402	17,831	0	100	475,444	76.661	15.555	0	100
Rasch test score in reading	489,631	0,238	1,056	-4.471	3.801	475,444	0.176	1.119	-5.427	4.120
Normalized test score in maths	491,702	60,685	21,571	0	100	476,810	59.205	19.259	0	100
Rasch test score in maths	491,702	0,400	1.293	-4.728	4.713	476,810	0.242	1.067	-5.231	4.778

Table 5 – Pooled estimates of Classroom Peer Age Effect on Educational Achievements

	Y=Normalized test scores		Y=Rasch test scores	
	Reading	Mathematics	Reading	Mathematics
INDIVIDUAL AGE				
Student's Age at school entry (in months)	**0.240*****	**0.345*****	**0.016*****	**0.020*****
CLASSROOM PEER AGE COMPOSITION				
Proportion of delayed students	-7.856***	-9.224***	-0.497***	-0.475***
Proportion of students born between Jan_t-Apr_t	0.263	0.931***	0.006	0.050***
Proportion of students born between $Sept_t$-Dec_t	-0.906***	-1.665***	-0.059***	-0.097***
Proportion of students born between Jan_{t+1}-Apr_{t+1}	**1.139*****	**0.845****	**0.105*****	**0.083*****
Number of Obs	582.813	583.615	582.813	583.615

*Notes:*1) * p<0.1; ** p<0.05; *** p<0.01. 2) Coefficients are estimated with robust standard errors. 3) Estimates include student-level (individual characteristics and parents' background), school-level and territorial-level covariates; see Table 3 for details.

Table 6 – OLS estimates of the Classroom Peer Age Effect on Educational Achievements by Grade

	GRADE 2			
	Y=Normalized test scores		Y=Rasch test scores	
	Reading	Mathematics	Reading	Mathematics
INDIVIDUAL AGE				
Student's Age at school entry (in months)	**0.360*****	**0.496*****	**0.021*****	**0.028*****
CLASSROOM PEER AGE COMPOSITION				
Proportion of delayed students	-8.829***	-11.214***	-0.447***	-0.590***
Proportion of students born between Jan_t-Apr_t	0.496	0.729*	0.004	0.038
Proportion of students born between $Sept_t$-Dec_t	-0.357	-1.413***	-0.016	-0.078***
Proportion of students "in advance"	**3.199*****	**4.706*****	**0.233*****	**0.342*****
Number of Obs	294.207	294.550	294.207	294.550
	GRADE 5			
	Y= Normalized test scores		**Y=Rasch test scores**	
	Reading	**Mathematics**	**Reading**	**Mathematics**
INDIVIDUAL AGE				
Student's Age at school entry (in months)	**0.130*****	**0.202*****	**0.013*****	**0.011*****
CLASSROOM PEER AGE COMPOSITION				
Proportion of delayed students	-8.722***	-9.974***	-0.583***	-0.557***
Proportion of students born between Jan_t-Apr_t	0.188	1.354***	0.017	0.078***
Proportion of students "in advance"	**-0.998*****	**-1.712*****	**-0.066*****	**-0.098*****
Number of Obs	288.606	289.065	288.606	289.065

Notes: 1) Students "in advance" are those born between January and April of the year *t+1* for Grade 2 while are student born between September of the year *t* and April of the year *t+1* for Grade 5. 2) * p<0.1; ** p<0.05; *** p<0.01.
3) Coefficients are estimated with robust standard errors. 4) Estimates include student-level (individual characteristics and parents' background), school-level and territorial-level covariates; see Table 3 for details.

Table 7 – Comparison of Classroom Peer Age Effect between the 2nd and the 5th Grade

	Specification 1			
	Y=Normalized test scores		Y=Rasch test scores	
	Reading	Mathematics	Reading	Mathematics
CLASSROOM PEER AGE COMPOSITION				
Proportion of students born between Jan$_{t+1}$-Apr$_{t+1}$ (2nd Grade)	**3.199*** **	**4.706*** **	**0.233*** **	**0.342*** **
Proportion of students born between Jan$_{t+1}$-Apr$_{t+1}$ (5th Grade)	**-0.274**	**-1.965*** **	**0.006**	**-0.100*** **
	Specification 2			
	Y= Normalized test scores		**Y=Rasch test scores**	
	Reading	**Mathematics**	**Reading**	**Mathematics**
CLASSROOM PEER AGE COMPOSITION				
Proportion of students born between Jan$_{t+1}$-Feb$_{t+1}$ (2nd Grade)	**1.771*** **	**1.763** **	**0.154*** **	**0.137*** **
Proportion of students born between Jan$_{t+1}$-Feb$_{t+1}$ (5th Grade)	**-0.979** **	**-2.426*** **	**-0.059***	**-0.129*** **
Proportion of students born between Mar$_{t+1}$-Apr$_{t+1}$ (2nd Grade)	**6.373*** **	**11.281*** **	**0.410*** **	**0.801*** **
Proportion of students born between Mar$_{t+1}$-Apr$_{t+1}$ (5th Grade)	**1.272***	**-0.951**	**0.148*** **	**-0.038**

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Coefficients are estimated with robust standard errors. 3) Estimates include student-level (individual characteristics and parents' background), school-level, territorial-level and other classroom peer age composition covariates; see Table 3 for details.

Table 8 – Classroom Peer Age Effect and Identification of Selection Effect: 2nd and 5th Grade

	Y=Normalized test scores		Y=Rasch test scores	
	Reading	Mathematics	Reading	Mathematics
INDIVIDUAL AGE				
Student's Age at school entry (in months)	**0.240*** **	**0.345*** **	**0.016*** **	**0.020*** **
CLASSROOM PEER AGE COMPOSITION				
Proportion of delayed students	-7.896***	-10.776***	-0.334***	-0.571***
Proportion of students born between Jan$_t$-Apr$_t$	-1.130***	-0.723*	-0.082***	-0.086***
Proportion of students born between Sept$_t$-Dec$_t$	-0.925***	-2.053***	-0.040**	-0.116
Proportion of students born between Jan$_{t+1}$-Apr$_{t+1}$	**7.584*** **	**8.999*** **	**0.443*** **	**0.700*** **
Grade 5*Proportion of delayed students	1.126	5.583	-0.113	0.398**
Grade 5*Proportion of students born between Jan$_t$-Apr$_t$	4.849**	7.554**	0.380**	0.605***
Grade 5*Proportion of students born between May$_t$-Aug$_t$	1.864	4.023	0.193	0.316*
Grade 5*Proportion of students born between Sept$_t$-Dec$_t$	**2.058**	**4.096**	**0.164**	**0.367**
Grade 5*Proportion of students born between Jan$_{t+1}$-Apr$_{t+1}$	-10.127***	-11.114***	-0.438**	-0.828***
Number of Obs	582.813	583.615	582.813	583.615

*Notes:*1) * p<0.1; ** p<0.05; *** p<0.01. 2) Coefficients are estimated with robust standard errors. 3) Estimates include student-level (individual characteristics and parents' background), school-level and territorial-level covariates; see Table 3 for details.

Table 9 – Classroom Peer Age Effect: Simulations on Normalized Test Scores

	Reading			Mathematics		
	Simulation 1	**Basic Estimates**	**Simulation 2**	**Simulation 1**	**Basic Estimates**	**Simulation 2**
INDIVIDUAL AGE						
Student's Age at school entry (in months)	**0.240*** **	**0.240*** **	**0.238*** **	**0.346*** **	**0.345*** **	**0.343*** **
CLASSROOM PEER AGE COMPOSITION						
Proportion of students "in advance"	**13.137*** **	**7.584*** **	**2.287*** **	**18.061*** **	**8.999*** **	**1.976*** **
Selection Effect	**0.382**	**2.058**	**1.064**	**2.950**	**4.096**	**4.017**
Number of Obs	582.813	582.813	582.813	583.615	583.615	583.615

Notes: 1) Proportion of students "in advance" refers to those "in advance" for both the 2nd and the 5th Grade, i.e. to students born in March or April *t+1* in Simulation 1, pupils born between January *t+1* and April *t+1* (Basic Estimates) and students born between November *t* and April *t+1* in Simulation 2. For a clearer interpretation of *Proportion of students "in advance"* and *Selection Effect* parameters see Figures 1, 2 and 3. 2) * p<0.1; ** p<0.05; *** p<0.01. 3) Coefficients are estimated with robust standard errors. 4) Estimates include student-level (individual characteristics and parents' background), school-level, territorial-level and other classroom peer age composition covariates; see Table 3 for details.

Table 10 – Classroom Peer Age Effect: Simulations on Rasch Test Scores

	Reading			Mathematics		
	Simulation 1	**Basic Estimates**	**Simulation 2**	**Simulation 1**	**Basic Estimates**	**Simulation 2**
INDIVIDUAL AGE						
Student's Age at school entry (in months)	**0.016*** **	**0.016*** **	**0.016*** **	**0.020*** **	**0.020*** **	**0.019*** **
CLASSROOM PEER AGE COMPOSITION						
Proportion of students "in advance"	**0.739*** **	**0.443*** **	**0.126*** **	**1.361*** **	**0.700*** **	**0.181*** **
Selection Effect	**0.072**	**0.164**	**0.088**	**0.219**	**0.367**	**0.295**
Number of Obs	582.813	582.813	582.813	583.615	583.615	583.615

Notes: 1) Proportion of students "in advance" refers to those "in advance" for both the 2nd and the 5th Grade, i.e. to students born in March or April *t+1* in Simulation 1, pupils born between January *t+1* and April *t+1* (Basic Estimates) and students born between November *t* and April *t+1* in Simulation 2. For a clearer interpretation of *Proportion of students "in advance"* and *Selection Effect* parameters see Figures 1, 2 and 3. 2) * p<0.1; ** p<0.05; *** p<0.01. 3) Coefficients are estimated with robust standard errors. 4) Estimates include student-level (individual characteristics and parents' background), school-level, territorial-level and other classroom peer age composition covariates; see Table 3 for details.

Table 11 – Classroom Peer Age Effect on Normalized Test Scores by Students' Age Group

	Reading			
	65 ≤ Age ≤ 68 (Students born between Jan_{t+1}-Apr_{t+1})	**69 ≤ Age ≤ 72 (Students born between $Sept_t$-Dec_t)**	**73 ≤ Age ≤ 76 (Students born between May_t-Aug_t)**	**77 ≤ Age ≤ 80 (Students born between Jan_t-Apr_t)**
CLASSROOM PEER AGE COMPOSITION				
Proportion of students born between Jan_{t+1}-Apr_{t+1}	**10.018*** **	**10.090*** **	**7.437*** **	**5.179*** **
Number of Obs	47.078	191.035	195.848	138.286
	Mathematics			
	65 ≤ Age ≤ 68 (Students born between Jan_{t+1}-Apr_{t+1})	**69 ≤ Age ≤ 72 (Students born between $Sept_t$-Dec_t)**	**73 ≤ Age ≤ 76 (Students born between May_t-Aug_t)**	**77 ≤ Age ≤ 80 (Students born between Jan_t-Apr_t)**
CLASSROOM PEER AGE COMPOSITION				
Proportion of students born between Jan_{t+1}-Apr_{t+1}	**14.731*** **	**10.999*** **	**6.979*** **	**4.938*** **
Number of Obs	47.027	191.318	196.001	138.586

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Coefficients are estimated with robust standard errors. 3) Estimates include student-level (individual characteristics and parents' background), school-level, territorial-level and other classroom peer age composition covariates; see Table 3 for details.

Table 12 – Classroom Peer Age Effect on Rasch Test Scores by Students' Age Group

	Reading			
	65 ≤ Age ≤ 68 (Students born between Jan_{t+1}-Apr_{t+1})	**69 ≤ Age ≤ 72 (Students born between $Sept_t$-Dec_t)**	**73 ≤ Age ≤ 76 (Students born between May_t-Aug_t)**	**77 ≤ Age ≤ 80 (Students born between Jan_t-Apr_t)**
CLASSROOM PEER AGE COMPOSITION				
Proportion of students born between Jan_{t+1}-Apr_{t+1}	**0.628*** **	**0.529*** **	**0.415*** **	**0.363*** **
Number of Obs	47.078	191.035	195.848	138.286
	Mathematics			
	65 ≤ Age ≤ 68 (Students born between Jan_{t+1}-Apr_{t+1})	**69 ≤ Age ≤ 72 (Students born between $Sept_t$-Dec_t)**	**73 ≤ Age ≤ 76 (Students born between May_t-Aug_t)**	**77 ≤ Age ≤ 80 (Students born between Jan_t-Apr_t)**
CLASSROOM PEER AGE COMPOSITION				
Proportion of students born between Jan_{t+1}-Apr_{t+1}	**1.060*** **	**0.814*** **	**0.587*** **	**0.439*** **
Number of Obs	47.027	191.318	196.001	138.586

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Coefficients are estimated with robust standard errors. 3) Estimates include student-level (individual characteristics and parents' background), school-level, territorial-level and other classroom peer age composition covariates; see Table 3 for details.

Table 13 – Frequencies of students by month and year of birth

Month and year of birth	Grade 2		Grade 5		Grade 8	
	Reading	**Mathematics**	**Reading**	**Mathematics**	**Reading**	**Mathematics**
$January_t$	21,637	21,764	20,438	20,553	20,052	20,053
$February_t$	26,137	26,253	23,291	23,355	22,908	22,909
$March_t$	34,125	34,327	30,663	30,711	34,194	34,194
$April_t$	35,006	35,214	33,325	33,549	33,260	33,263
May_t	42,778	42,890	39,900	39,881	36,993	36,993
$June_t$	39,889	39,987	38,342	38,433	37,794	37,794
$July_t$	42,377	42,471	41,195	41,333	42,236	42,237
$August_t$	40,983	41,152	39,930	40,039	39,843	39,842
$September_t$	**42,621**	**42,791**	**42,104**	**42,125**	**40,866**	**40,867**
$October_t$	**41,980**	**42,192**	**41,397**	**41,546**	**37,875**	**37,878**
$November_t$	**38,167**	**38,362**	**36,190**	**36,303**	**34,443**	**34,446**
$December_t$	**39,440**	**39,594**	**36,506**	**36,623**	**35,322**	**35,325**
$January_{t+1}$	19,374	19,370	18,921	18,966	19,830	19,830
$February_{t+1}$	9,291	9,304	11,622	11,598	12,050	12,049
$March_{t+1}$	6,499	6,466	7,240	7,220	8,718	8,718
$April_{t+1}$	4,295	4,325	4,631	4,604	-	-

Notes: 1) Cutoff date for first enrollment is August 31st of the year *t* for students attending both the 5th and 8th Grade in 2012/2013 while is December 31st of the year *t* for those attending the 2nd Grade in the same year. 2) Limit birth date for optional enrollment is April 30th of the year *t+1* for pupils attending both the 2nd and the 5th Grade while is March 31st of the year *t+1* for those attending the 8th Grade in 2012/2013.

Table 14 – Classroom Peer Age Effect on Normalized Test Scores and Identification of Selection Effect: 2nd and 8th Grade

	Reading	**Mathematics**
CLASSROOM PEER AGE COMPOSITION		
Proportion of delayed students	-5.518***	-7.857***
Proportion of students born between Jan_t-Apr_t	-0.017	-0.176
Proportion of students born between $Sept_t$-Dec_t	-1.833***	-3.488***
Proportion of students born between Jan_{t+1}-Apr_{t+1}	**5.960*** **	**8.096*** **
Grade 8*Proportion of delayed students	3.294	8.533**
Grade 8*Proportion of students born between Jan_t-Apr_t	10.052***	15.336***
Grade 8*Proportion of students born between May_t-Aug_t	8.488***	11.344***
Grade 8*Proportion of students born between $Sept_t$-Dec_t	8.666***	14.052***
Grade 8*Proportion of students born between Jan_{t+1}-Mar_{t+1}	**-1.826**	**-2.702**
Number of Obs	557.734	558.058

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Coefficients are estimated with robust standard errors. 3) Estimates include student-level (individual characteristics and parents' background), school-level and territorial-level covariates; see Table 3 for details.

Table 15 – Classroom Peer Age Effect on Rasch Test Scores Revised for "Cheating" and Identification of Selection Effect: 2nd and 5th Grade

	Reading	Mathematics
CLASSROOM PEER AGE COMPOSITION		
Proportion of delayed students	-0.385***	-0.562***
Proportion of students born between Jan_t-Apr_t	-0.055***	-0.039*
Proportion of students born between $Sept_t$-Dec_t	-0.046***	-0.123***
Proportion of students born between Jan_{t+1}-Apr_{t+1}	**0.291*** **	**0.373*** **
Grade 5*Proportion of delayed students	-0.119	0.319*
Grade 5*Proportion of students born between Jan_t-Apr_t	0.210	0.405**
Grade 5*Proportion of students born between May_t-Aug_t	0.049	0.199
Grade 5*Proportion of students born between $Sept_t$-Dec_t	**0.041**	**0.277**
Grade 5*Proportion of students born between Jan_{t+1}-Apr_{t+1}	-0.426**	-0.572***
Number of Obs	582.813	583.615

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Coefficients are estimated with robust standard errors. 3) Estimates include student-level (individual characteristics and parents' background), school-level and territorial-level covariates; see Table 3 for details.

Table 16 – Classroom Peer Age Effect on Rasch Test Scores Revised for "Cheating" and Identification of Selection Effect: 2nd and 8th Grade

	Reading	Mathematics
CLASSROOM PEER AGE COMPOSITION		
Proportion of delayed students	-0.220***	-0.413***
Proportion of students born between Jan_t-Apr_t	0.048***	0.017
Proportion of students born between $Sept_t$-Dec_t	-0.107***	-0.200***
Proportion of students born between Jan_{t+1}-Apr_{t+1}	**0.108*** **	**0.255*** **
Grade 8*Proportion of delayed students	-0.118	0.355**
Grade 8*Proportion of students born between Jan_t-Apr_t	0.381**	0.661***
Grade 8*Proportion of students born between May_t-Aug_t	0.409**	0.480***
Grade 8*Proportion of students born between $Sept_t$-Dec_t	0.387**	0.644***
Grade 8*Proportion of students born between Jan_{t+1}-Mar_{t+1}	**0.168**	**-0.079**
Number of Obs	557.734	558.058

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Coefficients are estimated with robust standard errors. 3) Estimates include student-level (individual characteristics and parents' background), school-level and territorial-level covariates; see Table 3 for details.

Table 17 – Classroom Peer Age Effect on Rasch Test Scores Revised for "Cheating" by Students' Age Group

	Reading			
	$65 \leq$ Age ≤ 68 (Students born between Jan_{t+1}-Apr_{t+1})	**$69 \leq$ Age ≤ 72 (Students born between $Sept_t$-Dec_t)**	**$73 \leq$ Age ≤ 76 (Students born between May_t-Aug_t)**	**$77 \leq$ Age ≤ 80 (Students born between Jan_t-Apr_t)**
CLASSROOM PEER AGE COMPOSITION				
Proportion of students born between Jan_{t+1}-Apr_{t+1}	**0.396*** **	**0.389*** **	**0.315*** **	**0.231*** **
Number of Obs	47.078	191.035	195.848	138.286
	Mathematics			
	$65 \leq$ Age ≤ 68 (Students born between Jan_{t+1}-Apr_{t+1})	**$69 \leq$ Age ≤ 72 (Students born between $Sept_t$-Dec_t)**	**$73 \leq$ Age ≤ 76 (Students born between May_t-Aug_t)**	**$77 \leq$ Age ≤ 80 (Students born between Jan_t-Apr_t)**
CLASSROOM PEER AGE COMPOSITION				
Proportion of students born between Jan_{t+1}-Apr_{t+1}	**0.522*** **	**0.482*** **	**0.321*** **	**0.221*** **
Number of Obs	47.027	191.318	196.001	138.586

Notes: 1) * p<0.1; ** p<0.05; *** p<0.01. 2) Coefficients are estimated with robust standard errors. 3) Estimates include student-level (individual characteristics and parents' background), school-level, territorial-level and other classroom peer age composition covariates; see Table 3 for details.

Figure 1 – Conceptual Framework of Identification Strategy

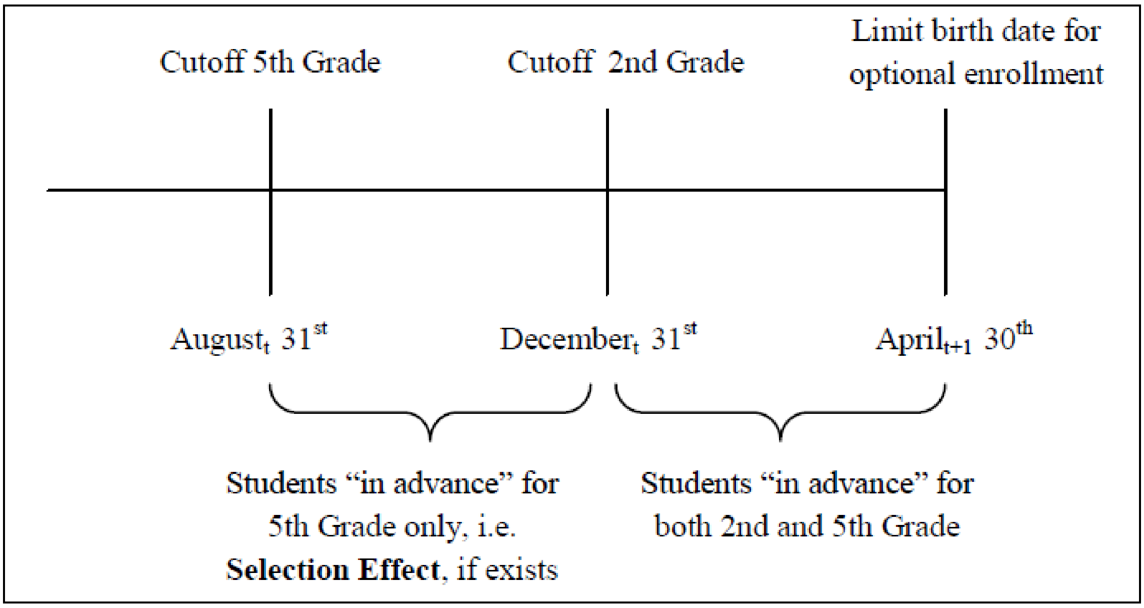


Figure 2 – Simulation 1

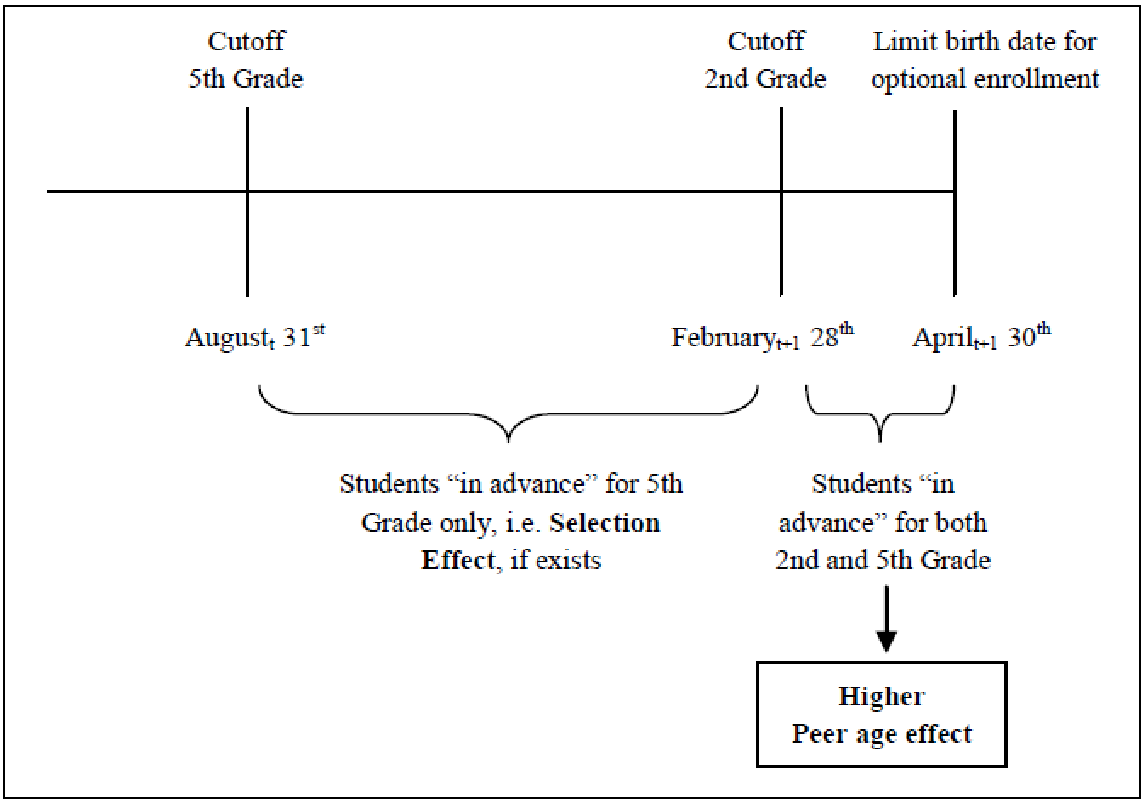


Figure 3 – Simulation 2

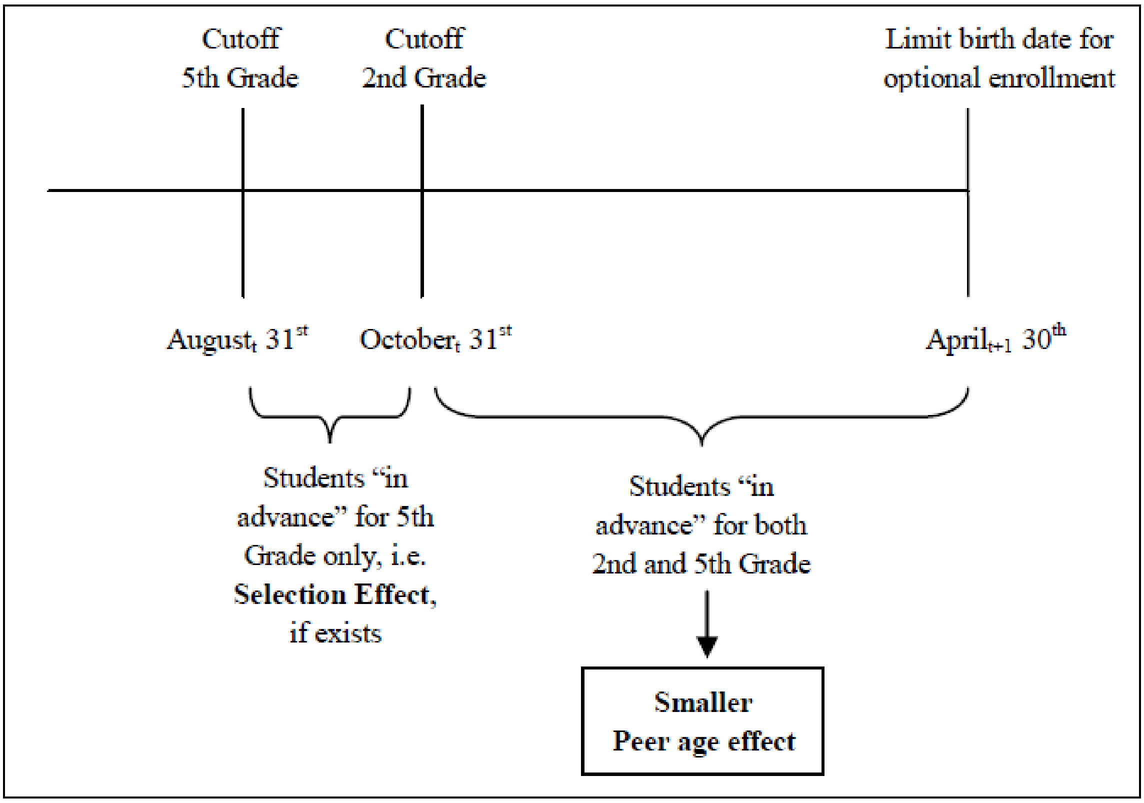
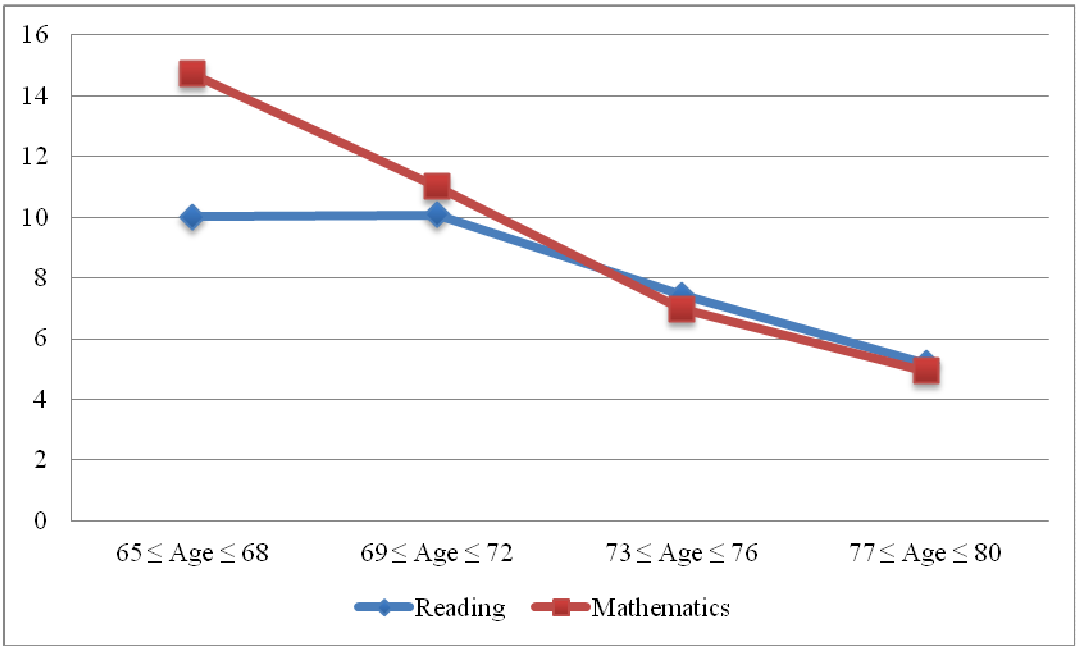
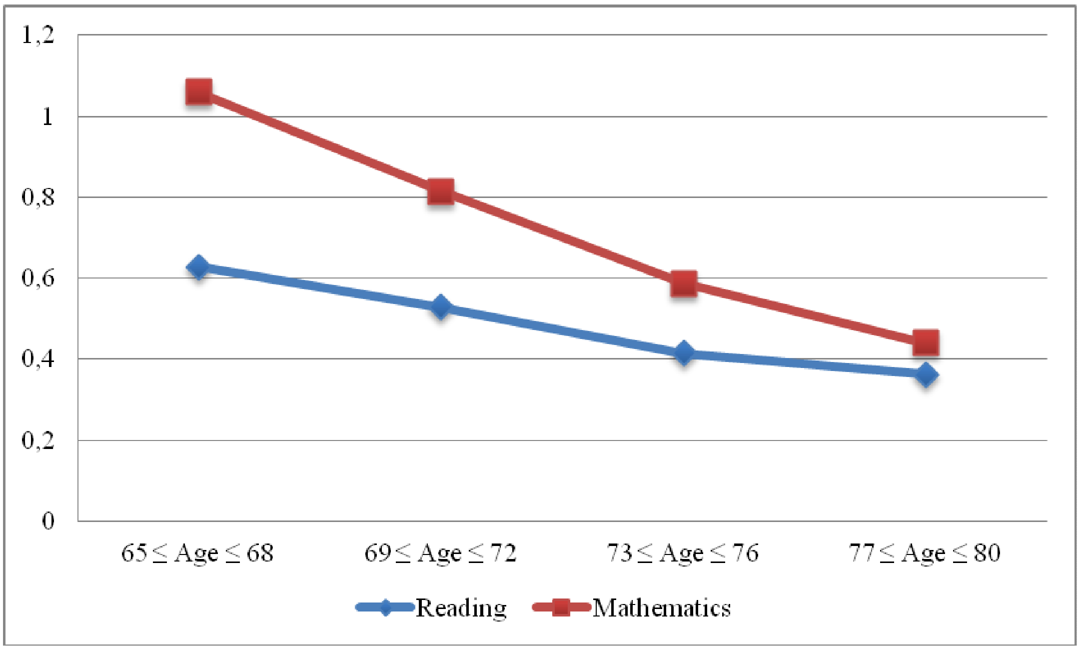


Figure 4 – Effect of the Proportion of students born between January *t+1* and April *t+1* on Normalized test scores by Students' Age Group



The horizontal axis reports four students' age groups. The first one identifies pupils born between January *t+1* and April *t+1*, who are youngest pupils in the classroom (age from 65 to 68 months at school entry). Others groups include students born in the others four months of the year *t*, identifying an increasing age. The vertical axis reports the effects of the Proportion of students born between January *t+1* and April *t+1*, i.e. of youngest students "in advance" on Normalized test scores in Reading and Mathematics.

Figure 5 – Effect of the Proportion of students born between January *t+1* and April *t+1* on Rasch test scores by Students' Age Group



The horizontal axis reports four students' age groups. The first one identifies pupils born between January *t+1* and April *t+1*, who are youngest pupils in the classroom (age from 65 to 68 months at school entry). Others groups include students born in the others four months of the year *t*, identifying an increasing age. The vertical axis reports the effects of the Proportion of students born between January *t+1* and April *t+1*, i.e. of youngest students "in advance" on Rasch test scores in Reading and Mathematics.

Figure 6 – Alternative Identification Strategy

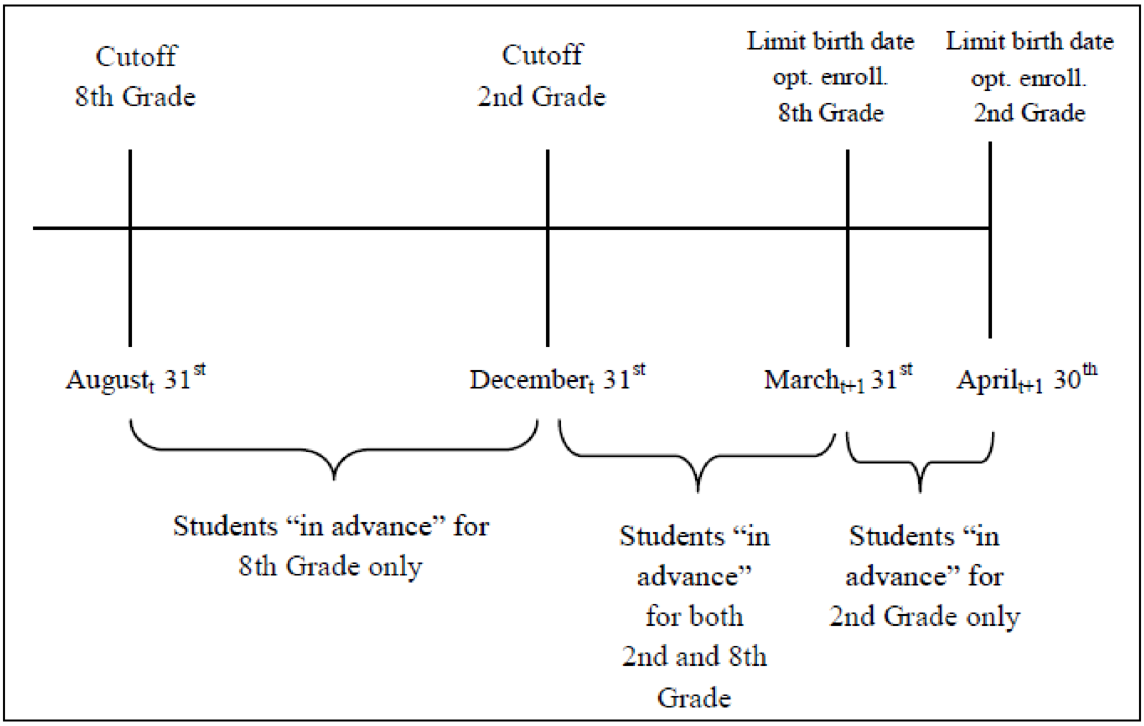
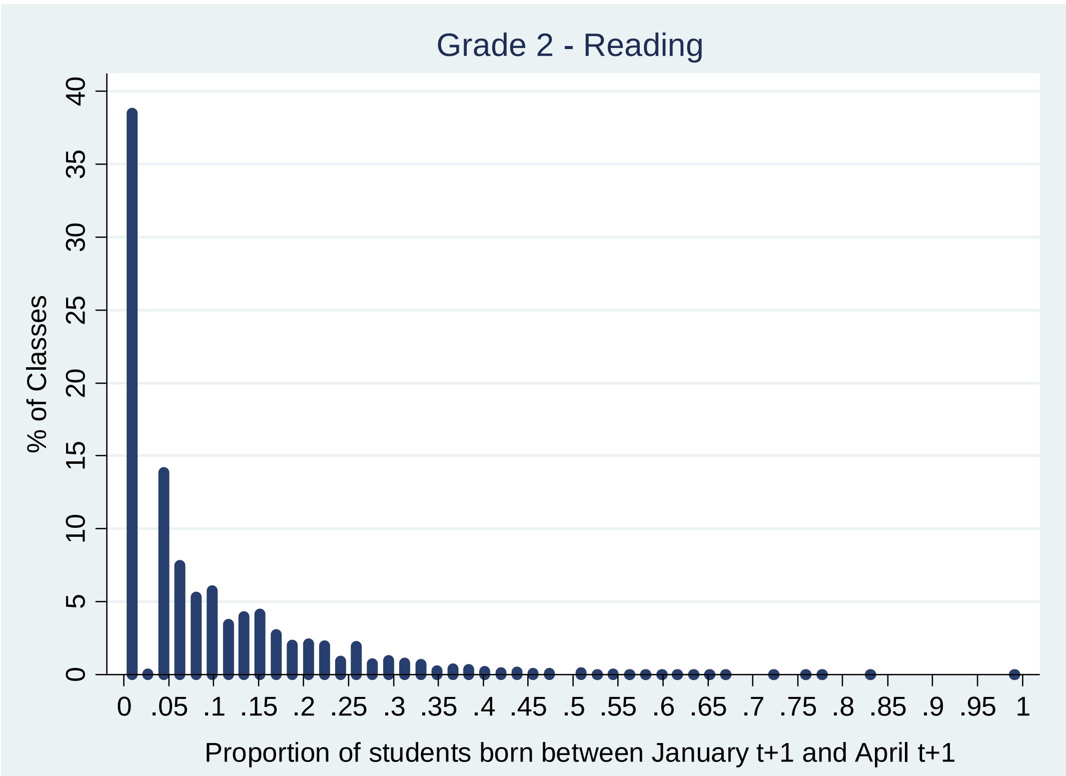
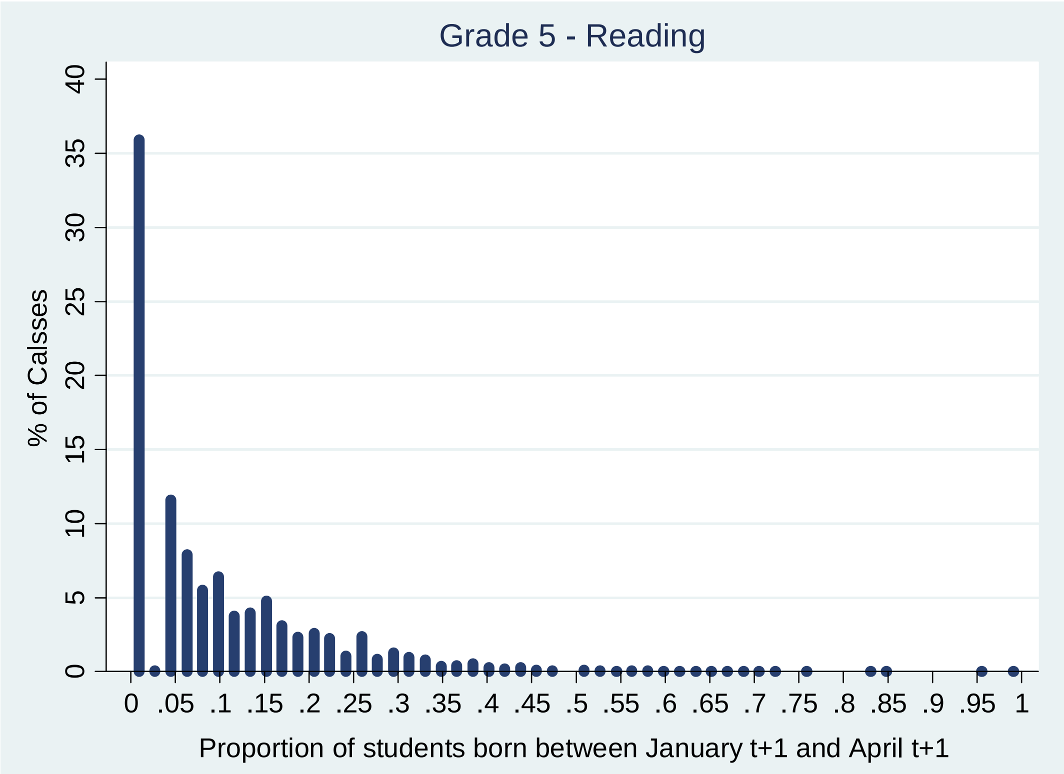


Figure 7 – Share of youngest students "in advance" in the classroom



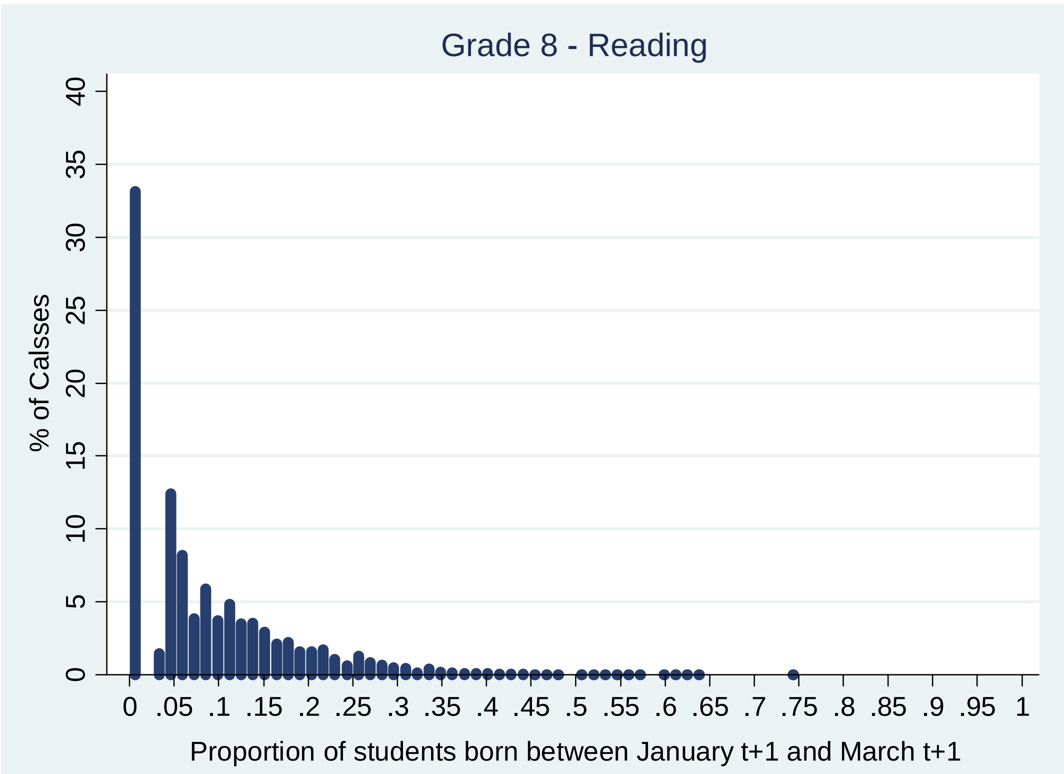
The horizontal axis reports the proportion of youngest pupils "in advance" in the classroom, i.e. of students born between January and April of the year *t+1*. The vertical axis reports the percentage of classes in the sample with the correspondent proportion of students.

Figure 8 – Share of youngest students "in advance" in the classroom



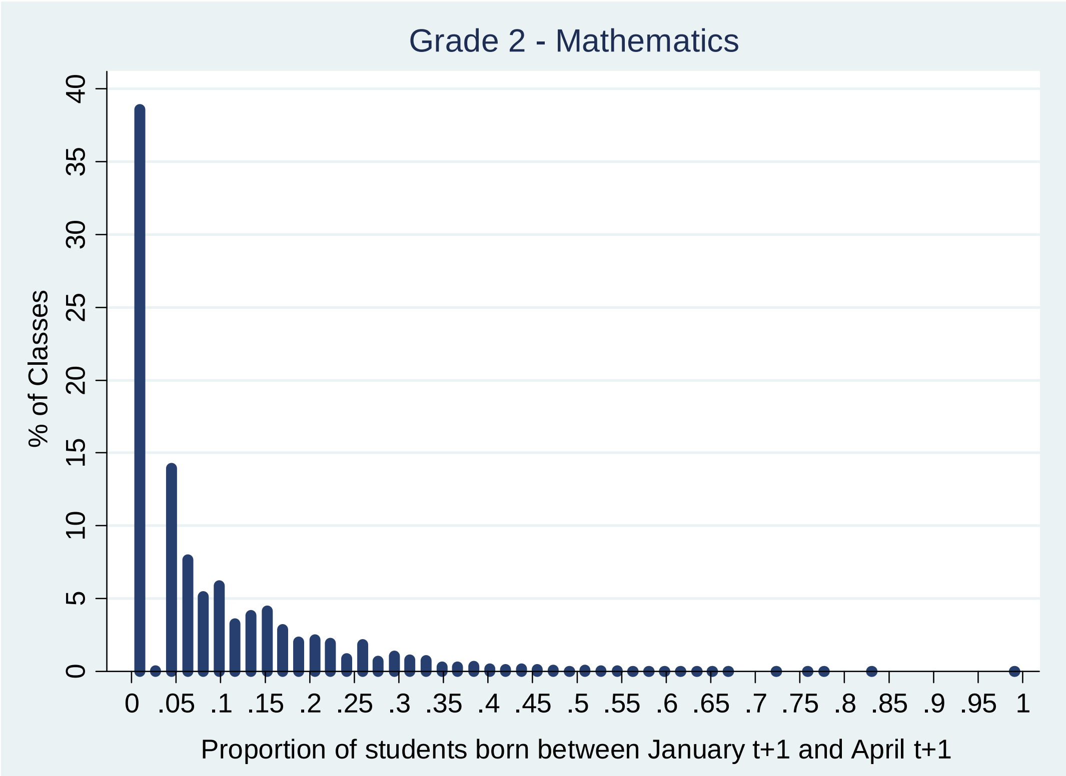
The horizontal axis reports the proportion of youngest pupils "in advance" in the classroom, i.e. of students born between January and April of the year *t+1*. The vertical axis reports the percentage of classes in the sample with the correspondent proportion of students.

Figure 9 – Share of youngest students "in advance" in the classroom



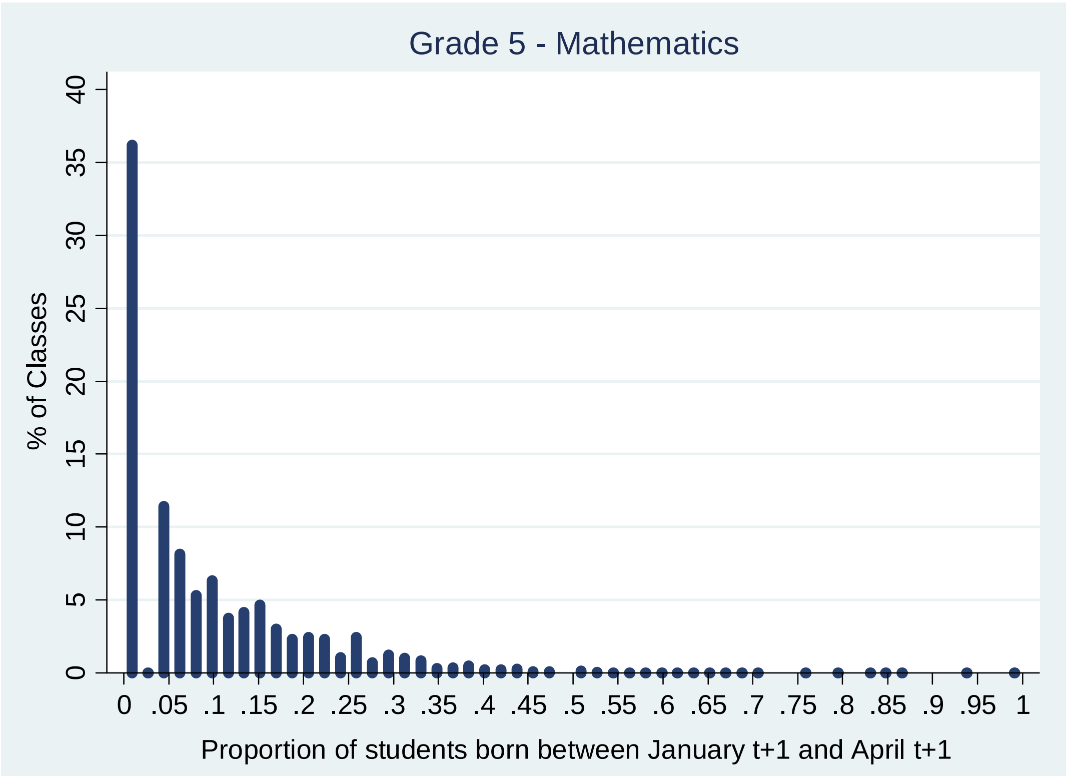
The horizontal axis reports the proportion of youngest pupils "in advance" in the classroom, i.e. of students born between January and March of the year *t+1*. The vertical axis reports the percentage of classes in the sample with the correspondent proportion of students.

Figure 10 – Share of youngest students “in advance” in the classroom



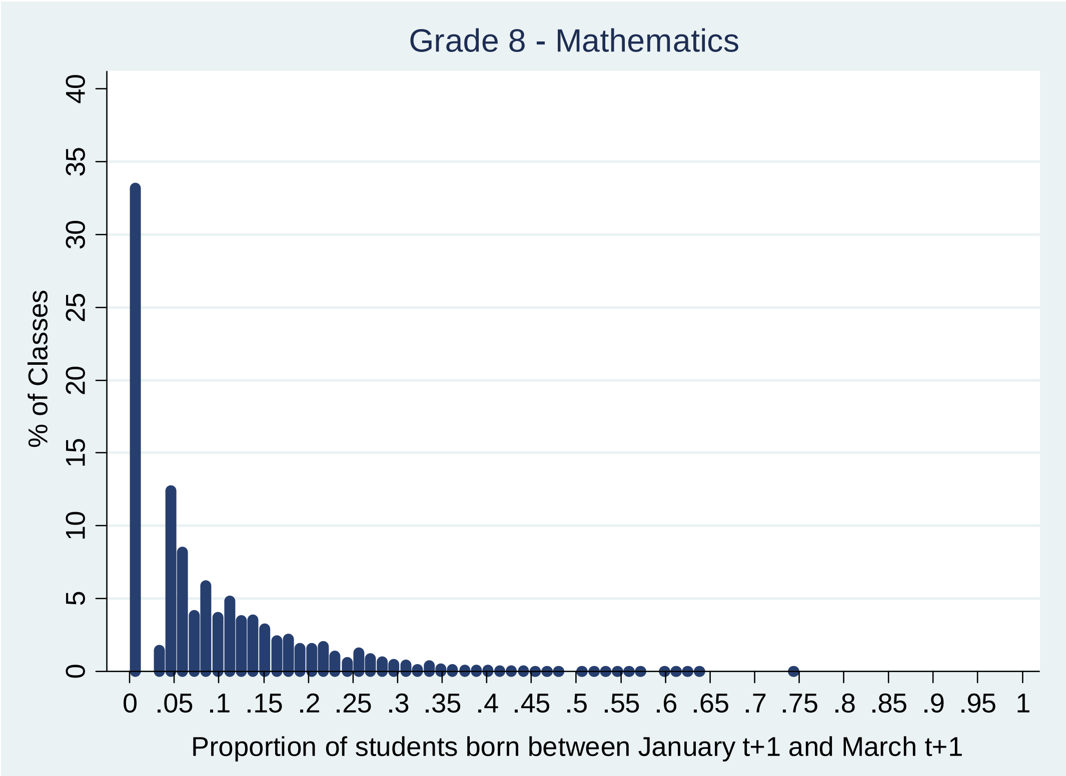
The horizontal axis reports the proportion of youngest pupils “in advance” in the classroom, i.e. of students born between January and April of the year *t+1*. The vertical axis reports the percentage of classes in the sample with the correspondent proportion of students.

Figure 11 – Share of youngest students “in advance” in the classroom



The horizontal axis reports the proportion of youngest pupils “in advance” in the classroom, i.e. of students born between January and April of the year *t+1*. The vertical axis reports the percentage of classes in the sample with the correspondent proportion of students.

Figure 12 – Share of youngest students "in advance" in the classroom



The horizontal axis reports the proportion of youngest pupils "in advance" in the classroom, i.e. of students born between January and March of the year *t+1*. The vertical axis reports the percentage of classes in the sample with the correspondent proportion of students.

Concluding remarks

In the last years, international surveys (PISA, PIRLS, TIMSS) have tested the level of students' skills through standardized tests across several countries, with the aim to evaluate the effectiveness and efficiency of the different educational systems. In most of the classifications drawn up based on results of these assessments, Italy is in rather low positions.

At national level, the INVALSI (National Institute for the Educational Evaluation of Instruction and Training) yearly assesses Italian pupils' competences in Reading and Mathematics through the National Service for the Evaluation of Education and Training (SNV). Results in test scores show a significant delay of Italians whether compared with other countries.

In this framework, understanding the possible determinants of educational outcomes becomes interesting.

First of all, we might attribute the poor performance of students to the presence of a lacking cultural environment at home. If pupils have parents with poor education, they do not receive enough support and family pressures to achieve good results at school. Moreover, parents with a high occupational status have more resources to provide a better environment for their children to do well in school.

Another channel through which parents can operate on educational performance of their children is the decision to delay or expedite their offsprings' school entry. The choice, indeed, is certainly related to both families' and pupils' characteristics which influence educational outcomes.

Finally, effects on pupils achievements can be exert in classrooms where pupils interact all together, generating the so-called peer effect.

The importance of family background as well as the individual and peer age effects occurs from the first year of schooling. Better or lower school performance translates into social inequalities along upper secondary school and then in labor market. Hence, I support the importance of acting in the early stages of schooling process to mitigate inequalities and improve equality of opportunity and social mobility. For these reasons, in the thesis I focus on primary school analyzing some possible determinants of achievements of Italian students through three chapter.

The first one aims at investigating if parents socio-economic background has a strong impact on pupils' outcomes in Italian primary school. The sample consists of pupils attending both the 2nd and the 5th grade of primary education. To handle missing data I generate multiple imputations preferring this method respect to others according to literature which considers it the best (see, for example, Graham et al., 2003). Additional specifications without imputations and including school, classroom and provincial fixed effects lead to very similar results. The additional evidence, which is consistent in sign and magnitude to basic estimates, lends strong support to the causal impact of parents' background on individual achievements both in Reading and in Mathematics. Results show that parents' educational qualification is the most fundamental factor in explaining the child's success in school. The effect of the parents' occupational status is also strong. The impact of parents' background on students' achievements does not fade away but persists during primary education.

In Chapter 2 I examine the effect of early school entry on Italian Normalized test scores achieved by pupils attending primary school in 2011/2012. Unlike other studies, I deal with selection on unobservable by estimating the potential selection bias comparing pupils who should start school in year *t* and pupils who have the opportunity to start school in that year. I point out that a severe distortion in the evaluation of early entry arises when neglecting the effect of unobserved characteristics driving school entry decisions. In particular, in the presence of a positive *Family Gift* leading best pupils to enter school in advance, the penalty imposed by early school entry is substantially underestimated. After neutralizing this "schooling ability" effect, I find that pupils in advance perform worse than regular ones. This gap does not fade away during primary school.

Chapter 3 examines peer age effects on educational outcomes (both Normatized and Rasch test scores) in academic year 2012/2013. Findings suggest that the proportion of youngest students "in advance" in the classrooms has a positive effect on test scores both in Reading and in Mathematics. This impact differ by students' age group.

I face the selection bias by exploiting changes in Italian enrollment rule occurred in the recent past. Results do not appear to be influenced by selection issues so systematic assignment of students and teachers to classrooms does not seem to take place and to be relevant in determining performance of Italian primary school pupils. Through an identification strategy never used in previous studies, I show that peer age impact on academic performance may arise from a "true" spillover.

Given the central role of students' educational performance for the future economic prospects of societies, the evidence presented in this thesis may reveal interesting aspects for educational and social policies in Italy. Results from the first chapter allow to understand how much of inequality in educational achievements is due to socioeconomic status of students' family. The analysis shows that intergenerational educational persistence and social immobility originates in the early stages of the schooling process. Pupils with an advantaged family background perform better in primary school. This persists during school career and it obviously translates into social inequalities along upper secondary school and then in labor market. This has policy implications both for schools and for governments. These authorities, in fact, should reduce social disparities in terms of education opportunities and improve social mobility.

Important policy implications also arise from the issue concerning the age at which a child should start school. On the one hand, parents struggle with the question of whether they should send their children to school as soon as they are eligible. Then, governments could change cutoff birth date for first enrollment into school, weighting penalties of being younger at school entry against the costs for parents in terms of child care and delayed entrance in the labor market. The research presented in this thesis contributes to this debate providing evidence of a skill gap and its persistence during primary school. Further researches, however, should be devoted to understand if this gap is actually bridged in the long run.

Finally, understanding the nature and importance of peer group effects in education becomes crucial for education policy. Focusing on peer age effects, I find that youngest students "in advance" in the classroom has a positive impact on individual achievements. The way through which the impact of the classmates' age interacts with individual performance remains an open issue. Other researches should be devoted to understand if the presence of youngest students "in advance" in the classrooms affects individual performance through a learning spillover between classmates – i.e. youngest pupils are ready for formal education and particulary able to perform better in school despite their young age, or because of a better learning environment: *i)* teachers alter curriculum choices and redirect more attention towards students; *ii)* in a more disciplinated school environment, the process of teaching-learning can more easily take place with positive effects on the entire classroom.

Acknowledgments

I am very grateful to Professors Patrizia Ordine and Giuseppe Rose (University of Calabria) for their willingness and for precious suggestions for thesis. A heartfelt thanks also goes to Professor Orazio Attanasio (University College London) for his suggestions and his availability as supervisor during my visiting research stay in London. Finally, a special thanks for moral support goes to my husband, my parents and my brother.