POLITECNICO DI MILANO

Scuola di Ingegneria Industriale e dell'Informazione
Corso di Laurea Magistrale in Ingegneria Gestionale

School factors helping disadvantaged students to succeed: empirical evidence from five Italian cities

Relatore: Prof. Tommaso AGASISTI

Tesi di Laurea di:

Mara SONCIN 796922
Riccardo VALENTI 796578

Anno Accademico 2013 - 2014

Acknowledgments

We would like to thank professor T. Agasisti for his continued support and interest.
We had interesting discussions about preliminary results of this thesis with J. Scott-Clayton, H. Levin and L. Stiefel who provided interesting comments and suggestions. We acknowledge the essential help by INVALSI, which sent us the original dataset, and gave useful statistical assistance (especially, we are indebted on this with P. Falzetti). We are also grateful to Teachers College, Columbia University, in the person of B. Bowden, P. Bergman, Y. Pan which hosted us during the writing of this research (Fall 2014). All eventual errors are our only responsibility.

Table of Contents

Index of Figures

Index of Tables

Index of Graphs

Abstract

This research investigates a particular category of disadvantaged students, namely those who are able to overcome their disadvantaged condition obtaining good academic results (here named “resilient students”). We use micro-data provided by the Italian National Evaluation Committee for Education (INVALSI) to focus on class and school-level characteristics that help disadvantaged students to become resilient when they switch from primary (grade 5) to lower secondary school (grade 6). We focus our attention on five big cities of Northern and Central Italy: Bologna, Milan, Padua, Rome and Turin. The preliminary study on these cities has been conducted implementing an educational production function aiming at investigating class and school-level factors that affect students’ achievement (overall considered) in grade 6. Similarly, we implement value added models in order to verify how much the *variation* in class and school-level features affect the *variation* in test scores between grade 5 and 6. We find that individual characteristics matter, especially when considering the socioeconomic background and prior achievement (in grade 5). However, school’s features play an important role, with particular relevance to factors related to school environment, namely the school average socioeconomic background and the average test score.

Through a descriptive analysis, we observe that resilient students are not those marginally wealthier, but those who are relatively closer to the cut-off point of the scores’ distribution. Employing a probit regression model, we investigate which factors have a correlation with the probability to be resilient. In this analysis, we find that class and school factors do matter. In particular, we estimate the positive impact that peers’ outcomes have on the probability of disadvantaged students to become resilient. When considering this finding as a treatment through a propensity score matching, results are significant only at class level. Thus, we find that peers’ influence not only fosters children’s outcomes, but also affects the performance of a particular group of students, i.e. disadvantaged students, raising their probability to “beat the odds” and to become resilient.

Keywords: Educational equality; students’ resilience; school factors; educational production fucntion, propensity score matching

Abstract (italian)

Lo scopo della presente ricerca è quello di studiare una particolare categoria di studenti svantaggiati, ovvero coloro in grado di superare un'iniziale condizione di svantaggio ottenendo buoni risultati scolastici (e qui definiti "studenti resilienti"). I dati utilizzati sono forniti dall'Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di Formazione (INVALSI). Lo scopo dell'analisi è quello di indagare quali caratteristiche a livello classe e scuola siano maggiormente correlate alla capacità di uno studente di diventare resiliente nel passaggio dalla scuola primaria (classe quinta) alla scuola secondaria di primo grado (classe prima). La nostra attenzione è rivolta a cinque grandi città del Nord e Centro Italia: Bologna, Milano, Padova, Roma e Torino. Lo studio preliminare effettuato sul campione complessivo di studenti delle cinque città oggetto di analisi è stato condotto attraverso l'utilizzo di una *educational production function* e di *value added model*. Nel primo caso, lo scopo è quello di individuare fattori a livello scuola che influenzino il risultato degli studenti nel primo anno di scuola secondaria. Nel secondo, l'obiettivo è verificare quanto la variazione delle caratteristiche di classe e scuola impattino sulla variazione del risultato ottenuto nel test INVALSI tra la classe quinta primaria e prima secondaria. I risultati mostrano come le caratteristiche individuali contino, specialmente per quanto riguarda il contesto socioeconomico e il risultato ottenuto nel test in quinta primaria. Tuttavia, anche i fattori scuola giocano un ruolo importante, con particolare rilevanza del contesto scolastico rappresentato dalla condizione socioeconomica media e dal punteggio medio ottenuto nel test a livello classe/scuola. Mediante un'analisi statistica emerge come gli studenti resilienti non siano caratterizzati da un relativo vantaggio socioeconomico rispetto al gruppo di controllo, ma che piuttosto mostrino una distribuzione dei punteggi (già nella classe quinta) relativamente più elevata e più vicina all'estremo superiore della definizione di studente resiliente. Successivamente, utilizzando un modello di regressione *probit* e un *propensity score matching* si dimostra quali fattori siano maggiormente correlati alla probabilità di diventare uno studente resiliente nel passaggio dalla classe quinta primaria alla prima secondaria. Da quest'analisi risulta come le caratteristiche classe e scuola abbiano una forte rilevanza. In particolare, emerge l'impatto positivo che il risultato scolastico dei propri pari (i compagni di classe) ha sulla probabilità di diventare uno studente resiliente. In conclusione, la presente ricerca mostra l'influenza che un ambiente scolastico positivo, rappresentato da studenti che ottengono buoni risultati scolastici, non solo influenzi l'andamento dei propri pari, ma abbia una particolare rilevanza per quella categoria di studenti che, nonostante un background socioeconomico svantaggiato, riesce a "superare le barriere", diventando uno studente resiliente.

Parole chiave: Equità dell'istruzione; resilienza; fattori scolastici; educational production fucntion, propensity score matching

Sommario

La realizzazione di un sistema educativo in grado di garantire equità di accesso e un livello minimo di istruzione per tutti è un obiettivo di primaria importanza nella politica pubblica, grazie alla capacità del sistema di istruzione di ridurre le più ampie disparità economiche e sociali (Field *et al.*, 2007, OECD, 2012). L'assenza di equità si traduce in più elevati tassi di ripetenza e di abbandono scolastico determinando, di conseguenza, un maggior numero di individui senza le competenze minime per essere parte integrante del tessuto sociale ed economico del Paese. Tra le principali motivazioni che influenzano il livello di istruzione raggiunto da un individuo, il background socioeconomico riveste un ruolo di primo piano. Studenti provenienti da famiglie economicamente svantaggiate dimostrano una probabilità doppia di ottenere scarsi risultati scolastici (OECD, 2012). L'impatto che il contesto socioeconomico dello studente può avere sui suoi risultati accademici è stato ampiamente investigato fin dalla pubblicazione del Rapporto Coleman (1966). Il filone di ricerca che studia l'effetto del background socioeconomico sulla performance degli studenti può essere classificato in quella corrente di letteratura che individua le determinanti dell'andamento scolastico nelle caratteristiche personali dello studente. Una seconda, in qualche modo parallela, corrente di ricerca si concentra invece sull'influenza delle caratteristiche della scuola frequentata.

Nella primo filone di studi, i fattori generalmente considerati come rilevanti sono correlati alla condizione socioeconomica, espressa da (I) il livello di educazione dei genitori (*e.g.* Hermisch & Francesconi, 2000, Lauer, 2003), e (II) il reddito familiare (*e.g.* Shea, 2000, Chevalier *et al.*, 2005), oppure legati a caratteristiche individuali come (I) lo status di immigrato (*e.g.* Schnepf, 2007, Meunier, 2011) e (II) le differenze di genere (Lubienski *et al*., 2013). Al contrario, il ruolo rivestito dalle caratteristiche della scuola frequentata è oggetto di maggiore dibattito. In generale, è condivisa l'importanza del ruolo svolto dalla qualità dell'insegnamento (Rockoff, 2004, Rivkin, Hanushek & Kain, 2005, Hanushek & Woessman, 2010) e – in minor misura – dalla dimensione della classe (Krueger, 1999, Rivkin, Hanushek & Kain, 2005) e dalle risorse economiche impiegate (Greenwald *et al*., 1996, 2013, Tajalli & Opheim, 2004, Hanushek & Woessman, 2012).

La letteratura riguardante, nello specifico, il ruolo che la scuola può avere nell'influenzare il risultato scolastico degli studenti svantaggiati è invece piuttosto scarna. Tuttavia, fornire evidenza del fatto che le politiche scolastiche possano effettivamente fare la differenza nel sostenere i risultati degli studenti svantaggiati, può avere enormi implicazioni in termini di *policy*. Per questo motivo, la nostra attenzione è rivolta, in questa ricerca, a quella particolare categoria di studenti che, provenendo da un contesto familiare socialmente ed economicamente svantaggiato, riesce ad ottenere

buoni risultati scolastici. Definiremo questi studenti come "studenti resilienti" (OECD, 2012). Gli studi rivolti a questa categoria di studenti si focalizzano, spesso, sulle caratteristiche individuali legate alla resilienza (come, tra le altre, una maggiore motivazione e sicurezza di sé), e solo marginalmente si concentrano su quelle caratteristiche che, a livello scolastico, stimolano questo fenomeno. Generalmente, la letteratura concorda sul ruolo svolto dal supporto genitoriale e della comunità (Benard, 1991, Borman & Rachuba, 2001) o del corpo insegnate (Borman & Rachuba, 2001, Walsh & Black, 2009). Altri interventi a livello scolastico (come la riduzione della dimensione delle classi) sono state considerate rilevanti solo quando accompagnate da interventi più ampi, come piani di supporto formativo per gli insegnanti (Faubert, 2012).

Il presente studio si pone l'obiettivo di estendere quest'ultimo filone di ricerca, utilizzando i dati forniti dall'Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di Formazione (INVALSI) per gli anni scolastici 2011/12 e 2012/13, e rispondendo alle seguenti domande di ricerca:

Quali sono gli elementi principali che determinano il risultato scolastico degli studenti e come la variazione delle caratteristiche di scuola e classe ne influenza l'andamento?

Quali caratteristiche sono legate alla resilienza e quanto la variazione dei fattori scuola e classe influenza la probabilità di diventare resiliente?

La seguente ricerca fornisce un contributo innovativo principalmente perché studia il tema della resilienza focalizzandosi sul *cambiamento* delle caratteristiche di classe e scuola nel passaggio dall'ultimo anno di scuola primaria al primo anno di scuola secondaria di primo grado. Inoltre, i dati a disposizione permettono di definire lo studente resiliente analizzandone la performance tra un anno e l'altro, coerentemente con la definizione di resilienza come fenomeno longitudinale – e non statico.

Il contenuto della ricerca è sviluppato come segue: il primo capitolo contiene una revisione della letteratura sull'argomento ed indaga il tema dell'equità nel sistema scolastico italiano, mettendo in luce i principali provvedimenti attuati a questo scopo. Inoltre, il capitolo introduce il *dataset* utilizzato, indagando anche i principali limiti legati ai dati disponibili. Il secondo capitolo contiene una serie di analisi preliminari svolte sull'intero campione di studenti residenti nelle città analizzate, allo scopo di indagare quali caratteristiche di classe e scuola siano maggiormente correlate al risultato scolastico dello studente. Il terzo capitolo rappresenta il fulcro della ricerca e si concentra sull'analisi empirica degli studenti resilienti. In apertura è fornita la definizione di studente resiliente e di gruppo di controllo e una serie di statistiche descrittive e di confronti. Successivamente, è presentata un'analisi descrittiva delle caratteristiche delle scuole e classi dove la concentrazione di studenti resilienti è particolarmente elevata, che vengono definite "scuole e classi resilienti". I paragrafi successivi presentano l'analisi empirica svolta sul campione di studenti resilienti allo scopo di individuare quali fattori a livello classe e scuola siano maggiormente correlati con la probabilità di diventare uno studente resiliente. Infine, il quarto capitolo presenta le implicazioni manageriali e di *policy*, e conclude.

Come precedentemente accennato, i dati utilizzati si riferiscono ai risultati ottenuti nei test per la valutazione del sistema scolastico annualmente effettuati da INVALSI a livello nazionale. In particolare, il *dataset* si riferisce ai risultati dei test di matematica ed italiano degli ultimi due anni disponibili al momento in cui la nostra ricerca è cominciata: 2011/2012 per la classe quinta primaria e 2012/2013 per la classe prima secondaria. Volendo concentrare la nostra attenzione sul passaggio da un ciclo di studi all'altro, abbiamo identificato gli studenti che hanno sostenuto la prova sia in quinta elementare che in prima media. Nonostante, in questo processo, circa il 40-50% del campione iniziale venga perso a causa di un problema di *matching* nei codici studenti forniti dal Ministero, il campione a nostra disposizione risulta essere statisticamente rappresentativo. La nostra analisi si concentra su cinque grandi citta del Nord e Centro Italia (Bologna, Milano, Padova, Roma e Torino) per due motivi principali:
la disparità socioeconomiche sono più evidenti in città ad elevata concentrazione di popolazione (Wang *et al*., 1997),
una maggior libertà di scelta comporta maggiori differenze nelle caratteristiche delle scuole, e rende l'analisi più significativa.
Il fatto che, nelle città, queste due caratteristiche si verifichino contemporaneamente, rende la resilienza un fenomeno maggiormente osservabile. Il *dataset* così selezionato è composto da 17.961 studenti residenti nelle cinque città sopra elencate. Le variabili a nostra disposizione permettono di effettuare una ricerca a tre livelli: a livello individuale disponiamo di informazioni relative alla condizione socioeconomica (espressa dall'indice ESCS), al genere, allo status di immigrato e agli anni di scuola frequentati (nel caso si tratti di uno studente anticipatario o posticipatario); a livello classe le informazioni riguardano il numero di studenti per classe, il punteggio medio, la condizione socioeconomica media e la composizione della classe (proporzione di femmine o studenti immigrati); infine, a livello scuola, disponiamo di informazioni relative alla dimensione della scuola, al numero delle classi, al fatto che la scuola sia pubblica/privata od un Istituto comprensivo).
Lo studio preliminare svolto sul campione complessivo (formato dalle cinque città menzionate) ha lo scopo di indagare quali caratteristiche siano maggiormente correlate al risultato ottenuto nel test, attraverso due regressioni lineari. Il modello seguito è quello dell*'educational production function*, dove gli input a livello individuale, di classe e di scuola vengono combinati per determinare il risultato scolastico. Nel primo caso, la regressione ha permesso di individuare i fattori correlati al risultato ottenuto nel test in prima secondaria. Nel secondo caso, un *value added model* è stato utilizzato per verificare come la variazione nelle caratteristiche di classe e scuola frequentata (tra quinta primaria e prima secondaria) influenzasse la variazione del risultato scolastico. Dal primo modello emerge come un più elevato contesto socioeconomico familiare influenzi la performance, dimostrando l'effettiva correlazione tra i due fattori. Inoltre, i risultati mostrano la positiva correlazione con la performance registrata nel test svolto l'anno precedente (nella classe quinta) coerentemente con l'ipotesi sostenuta riguardo alle abilità innate e agli effetti dei precedenti anni di educazione espressi dal punteggio

ottenuto in quinta. Considerando le variabili riguardanti la scuola frequentata, i fattori più significativi sono legati al contesto scolastico, in particolare il background socioeconomico e il punteggio medio registrati nella scuola. In particolare, un più elevato contesto socioeconomico è negativamente correlato all'output, fornendo un risultato controverso (rispetto alla letteratura sul tema) ma persistente nella nostra ricerca. Al contrario, il punteggio medio nella scuola è positivamente correlato al risultato del singolo studente, sottolineando l'effetto positivo dell'influenza dei propri pari.

I modelli *value added* mostrano risultati simili a livello individuale, ma forniscono informazioni di maggior interesse a livello classe e scuola. Infatti, i risultati suggeriscono che il background socioeconomico e la performance dei compagni non solo influenzano il punteggio dello studente in un determinato istante temporale ma, variando tra un anno e l'altro, impattano fortemente sulla variazione del punteggio. Infine questi modelli sottolineano due interessanti risultati: il primo è legato al fatto che minore è il numero di studenti per classe (in prima media), maggiore è la variazione del punteggio; il secondo indica invece che maggiore è il numero di studenti immigrati (in prima media), maggiore è la variazione del risultato. Il primo dato può essere interpretato nella corrente di letteratura che dibatte l'effetto della riduzione del numero di studenti per classe, mentre il secondo risulta particolarmente significativo nel contesto italiano, dove generalmente gli immigrati ottengono minori risultati scolastici rispetto ai nativi.

Il capitolo tre introduce l'analisi degli studenti resilienti. Abbiamo così definito gli studenti il cui punteggio e background socioeconomico in quinta primaria compaiono nel primo terzile della distribuzione di ogni città, e che in prima secondaria ottengono un punteggio superiore alla media della città. Al contrario, il gruppo di controllo è selezionato tra gli stessi studenti svantaggiati che, in prima secondaria, ottengono una *performance* inferiore alla media della città. I due campioni così definiti costituiscono il gruppo di riferimento per l'applicazione di un modello *probit*, applicato per ogni città allo scopo di identificare i fattori maggiormente correlati alla probabilità di diventare resiliente. L'output del modello è infatti una variabile binaria pari ad 1 quando lo studente è resiliente e 0 altrimenti. Dai risultati emerge come aver frequentato una classe o una scuola (in quinta primaria) dove i compagni ottengono buone *performance* abbia una correlazione positiva con la probabilità di diventare resiliente. Inoltre, frequentare una classe o una scuola dove i compagni ottengono risultati mediamente superiori rispetto a compagni della classe/scuola frequentata in quinta ha una correlazione positiva con il manifestarsi della resilienza. Tuttavia, considerando questo risultato come trattamento ed implementando un *propensity score matching,* i risultati sono significativi solo a livello classe. Il trattamento considerato può essere definito come "frequentare una classe in prima secondaria dove i compagni ottengono un punteggio mediamente superiore rispetto ai compagni della classe frequentata l'anno precedente". Essendo l'output del modello una variabile binaria pari ad 1 quando lo studente è resiliente e 0 altrimenti, l'effetto stimato può essere interpretato come una

variazione nella probabilità di diventare resiliente dato il trattamento subito. I risultati dimostrano un incremento nella probabilità di diventare resiliente compreso tra il 3% (a Torino) e il 27% (a Roma), risultati che permangono consistenti anche dopo l'analisi di robustezza. Fintanto che rimane valida l'ipotesi di allocazione casuale degli studenti nelle classi, questo risultato può essere considerato come un rapporto di causalità.
I risultati di questa ricerca possono essere interpretati nella corrente di letteratura che sostiene l'influenza positiva delle performance dei compagni sul raggiungimento di buoni risultati scolastici, focalizzandosi però su un particolare gruppo di studenti, definiti studenti resilienti. Da un punto di vista manageriale i nostri risultati sostengono l'importanza di favorire la creazione di classi caratterizzate da un'elevata diversificazione, che influirebbe positivamente sulla capacità degli studenti svantaggiati di ottenere migliori risultati scolastici.

Chapter 1

Introduction and Data

1.1 Motivation and research question

The ability of an educational system to guarantee equity of access to education is a key challenge for policymakers. An equitable education system can reduce the effect of broader social and economic inequalities that, otherwise, would result in higher costs in terms of health, income support, child welfare and security (Field *et al.*, 2007, OECD, 2012). As Field *et al.* (2007) affirm, equity in education has two dimensions: fairness, interpreted as the equality of opportunity in access to education, independently from gender, ethnicity or socioeconomic status; inclusion, in terms of ensuring a minimum level of education for all. The two dimensions are closely related: lack of fairness or inclusion stokes school failure and drop-out, resulting in a higher portion of people without the minimum skills to be integral part of social and economic system.

Among circumstances that can induce low attainment and high levels of school failure, the socioeconomic background plays a fundamental role. In fact, students from low socioeconomic background are twice as likely to be low performers (OECD, 2012). Even when they are able to achieve high standards, they continue to perform worse than high achieving peers from higher socioeconomic background and are less likely to apply for selective colleges (Bromberg & Theokas, 2014). Moreover, the increasing disparity between wealthy and disadvantaged families makes the necessity to work for greater equity in education even more urgent. Reardon (2011) starts from this point to investigate whether the achievement gap between high and low-income families has increased as well. Focusing on Northern American society, he uses data from nineteen national representative studies since 1950's. He compares the size of the achievement gap across studies through test-scores differences between groups in standard deviation units, finding that the achievement gap between the bottom 10th and top 90th percentile of the score distribution has not increased from 1950's to 1970's. On the contrary, It raised by 30-40% from late 1970's to 2001. The relevance of the income achievement gap (defined as the comparison between the bottom 10th and top 90th) has increased so much to become twice as large as the black-white achievement gap, which has been one of the most serious issues in US education so far. Students from low-SES background tend to experience lower family support and to live in communities with fewer resources, and this impacts negatively on their attainment. Furthermore, income inequalities tend to be passed on from one generation to the other. The more the intergenerational mobility is low, the more the society fails to provide equality of opportunity for all. Corak (2013) demonstrates that countries with grater inequality of incomes tend to be countries where a greater fraction of any economic advantage and disadvantage is passed on between parents and their children. Measuring the relationship between generational earnings elasticity and income inequality (proxied by the Gini coefficient), he finds that in Italy, the United Kingdom and the United States roughly 50 percent of any advantage or disadvantage is passed on between one generation and another. In this context, inequality affects children's future not only because parents with more human capital have more capacity to invest, but also because incentives to do so are greater if intergenerational mobility is low. Nevertheless, a proportion of disadvantaged students is able to overcome the initial condition, obtaining good academic results. This group of students has been defined as resilient students (OECD, 2012). The determinants of students' resiliency (discussed in the section below) are heterogeneous. Anyway, providing evidence about the ability of an educational system to foster this phenomenon would have great implications in terms of policy designed to promote equality. In addition, Agasisti & Longobardi (2014a) demonstrate a relationship between the percentage of resilient students in OECD countries and the average OECD-PISA 2009 score, suggesting that a higher proportion of resilient students is associated with higher (average) students' achievement and a more performing educational system. For all these reasons, our attention has been caught by this particular category of students, and by the role that schools have in determining their ability to "beat the odds". Furthermore, we want to verify if the determinants of students' achievement - overall considered - are the same that foster resiliency, and in which measure they differ. The present study extends this line of research, investigating resilient students in the Italian context using INVALSI data for

2011/2012 and 2012/2013, focusing on the impact of school-level factors' switch between grade 5 and 6.

The research questions can be summarized as follows:

a) *Which are the determinants of students' achievement and how much the variation in class and school's factors influences academic outcomes?*
b) *Which characteristics are most related to resiliency and how much the variation in class and school's factors affect the probability to become resilient?*

This research is innovative mainly because it faces the issue of resilient students focusing on the *changes* at class and school-level between the primary and lower secondary school (grade 5 and 6). This switching between the two cycles enables us to identify a variation that would not be investigated otherwise. Moreover, it allows us to define resilient students analyzing their changes over time, consistently with the definition of resilience as a longitudinal phenomenon – and not a static one based on "expected performance" given the observable characteristics.

The remainder of the research is organized as follows. Subsequent sections in this chapter present a review of the literature on this topic and the dataset used in the study. In chapter 2 are investigated the determinants of students' achievement, while chapter 3 contains the analysis on resilient students. Section 4 discusses managerial and policy implication, and concludes.

1.2 Background

The effect of the socio-economic (SES) condition on students' achievement has been largely investigated since the *Coleman's Report* (1966) opened the way. In one of the largest social science research project in history (involving more than 600,000 students) universally considered as a milestone of education research, James Coleman and his colleagues find that poor and minority students' performances in US schools are more related to their personal status and background than to school resources. These finding, still largely discussed in educational research, do not affirm that schools' features do not make a difference (teachers qualities and student body characteristics are described as the most predictive features), but that "*variation between schools* in their resource levels mattered little for *variation among individual students*" (Gamoran & Long, 2006, p.3). In other words, as Coleman's objective is to focus on the achievement gap between white and black students, he finds that schools have limited impact on the *difference* between black and white's outcomes but not on students' achievement overall considered (Rothstein, 2004). Subsequent researches have deeper studied how the family background affects student's abilities from both a descriptive and a quantitative point of view. Rothstein (2004) investigates major family factors affecting students' ability. Among them, the genetic potential of children plays a role. The intergenerational impact of innate ability has been highlighted by many "adoption studies" that show how adopted children's attainment is more similar to that of biological parents than adoptive ones. Moreover, he underlines the importance of childrearing practices in pre-school years: children's achievement difference during school years can be partially due to differences in how ready they are to learn when they enter school. Higher SES families tend – on average – to read aloud more and differently to their children, use less "baby talk" (enriching offspring's vocabulary) and give more indirect guidance to children instead of precise instructions about what to do or not. Also, he describes cultural and immigrant status influences, with some ethnic groups that foster children's achievement more than others (reflected, for example, by different results achieved by Asian children, whose families have more academic expectations, and other immigrant groups like Latin Americans). Finally, low SES families show poorer health (reflected by worse nutrition, medical care *etc.*) and higher mobility rates, which could arguably result in lower attainment.

Many other studies investigate the same topic from a quantitative standpoint, trying to quantify the impact of family and socio-economic background on children academic results. Among them, Rasbash *et al.* (2010) implement a multilevel cross-classified model on UK data in order to identify how much shared environmental factors beyond families impact on students' progress in secondary school. In particular, they study the environmental background at three levels: school (considering both secondary and carry-over effects from primary school), neighborhood and local education authority (LEA). Besides, when considering the family influence, they separate genetic factors from environmental effects of the immediate family identifying a group of twins and non-twins children. They find that, for twins, the family effect (composed by genes and shared environment) is the most relevant, accounting for 40% of the unexplained variation in learning progress; school, neighborhood and LEA account for 22%, while the remaining 38% can be considered as variation at pupil level. This study well summarizes the two main research lines followed when investigating factors affecting students' achievement: a first stream of literature can be classified among researches that investigate the determinants of educational outcomes related to individual-level characteristics within the family. A second, somehow parallel, stream of literature investigate school-level characteristics and environmental elements beyond the family.

In the first research line, the determinants of children's attainment are related to the socioeconomic status, expressed by (I) parental education (*e.g.* Hermisch & Francesconi, 2000, Lauer, 2003) and (II) family income (*e.g.* Shea, 2000, Chevalier *et al.*, 2005), or related to individual characteristics such as (I) the immigrant status (*e.g.* Schnepf, 2007, Meunier, 2011) or (II) gender (Lubienski *et al.*, 2013).

Empirical studies on the determinants of economic success have a long history, dating back to 1920s. Early studies paid attention especially to social mobility, demonstrating the close relationship between parents and children's earnings or between the family background and children's occupation (Haveman & Wolfe, 1995, provide a review of the literature on this topic). In more recent years, literature investigating children's socioeconomic status, has focused on parental education and income as main proxies of the family background. Chevalier *et al.* (2005) investigate how early school leaving (at age 16) is related to variations in permanent income and parental education level in UK. Through an OLS regression, they find a stronger impact of maternal education on sons' probability of post-16 participation in education (increased by 2.9%) than paternal education on daughters' attainment (increased by 2.4%). Adding the household income, results about parental education are still significant. The impact of current income is then reduced when a proxy for permanent income (the paternal occupation) is considered. When using paternal union status and paternal occupation as instrumental variables, the positive effect of paternal education is visible only for daughters (and quantified as an increase of remaining in education by seven percent per additional year of father's education), while the effect of maternal education disappears. Using a similar methodology, Shea (2000) focuses on father's permanent income variation due to factors (like the job loss due to plant closing) that are likely to represent luck, analyzing their impact on children's ability. The human capital of children is measured by the amount of years of schooling, wages, earnings and total family income, using data from the Panel Study of Income Dynamics that collect information on 5,000 families in the United States since 1968. Shea's hypothesis is that income's positive correlation across generation does not necessarily mean that family money do matter, as the income is presumably correlated with ability. In order to study the effect of exogenous income variation, he measures the impact of external variation in permanent income on the human capital of children. Interestingly, he finds that parents' money does matter for families whose father has low educational attainment (less than 12 years), but not in families with low income per se. This poses the interesting question of why intergenerational earnings' are so closely related if parents' money do not have a strong impact. It could be the case that other factors, such as inheritable ability or parental education do have a stronger causal effect than parents' income. Regarding inheritable ability, part of this stream of literature also account for the endogeneity of parental education, considering genetic effects through the comparison between adopted and natural children, or considering the children of twins. Ermish and Francesconi (2002) focus on the effect of parental employment on children's attainment as young adults (when they are in their early 20's). They use data from the British Household Panel Survey collected in the period spanning 1991-97 to show that the strongest negative impact on children's attainment is linked to mother's full-time employment when the child was aged 0-5. Nevertheless, a higher full family income increases the child's educational achievement through the increase of parents' time allocated to human capital investment in children.

Indirect effects of parental income and education related to parents' behavior and expectations are also deeply studied. Sewell & Shah (1968) demonstrate the influence of parents' educational aspirations on the achievements of their children (in terms of perceived parental encouragement resulting in higher college plans, attendance or graduation). Observing a court of students from Wisconsin for a 7-year period after their graduation from high school, they highlight that father's education has a stronger impact than mother's education on perceived encouragement for males and that the discrepancy in parents' educational achievement is far less important than high-level achievement of both parents. In more recent years, Davis-Kean (2005) investigates, through structural equation modeling techniques, how parents' education and income are indirectly related to children's academic outcomes through parents' beliefs and behaviors. The author analyses how the socioeconomic background can indirectly influence children's achievement through higher educational expectations, more reading, play and affective behaviors. Using a national US sample of 868 8-12 years old students, half African American and half non-Hispanic European American, she finds a structural difference by racial group: African American families show a strong and positive indirect effect of the socioeconomic background, while European American families tend to have a moderate relation. Also the years of schooling appear to do influence the home environment, as well as how parents interact with their children in promoting academic success. Similarly, Corak (2013) underlines that parents' investment on children's education is not only monetary (due to high household income that allows them to develop their children's skills and attitudes) but also nonmonetary (in terms of motivation and aspiration). Summarizing, Haveman & Wolfe (1995) describe home investments on children as based on both parents' abilities (partially transferred genetically) and education (that influence family's income and the quantity and quality of time and goods invested as inputs). Finally, children's ability, parental investments and income determine their final schooling level and future earnings (figure 1).

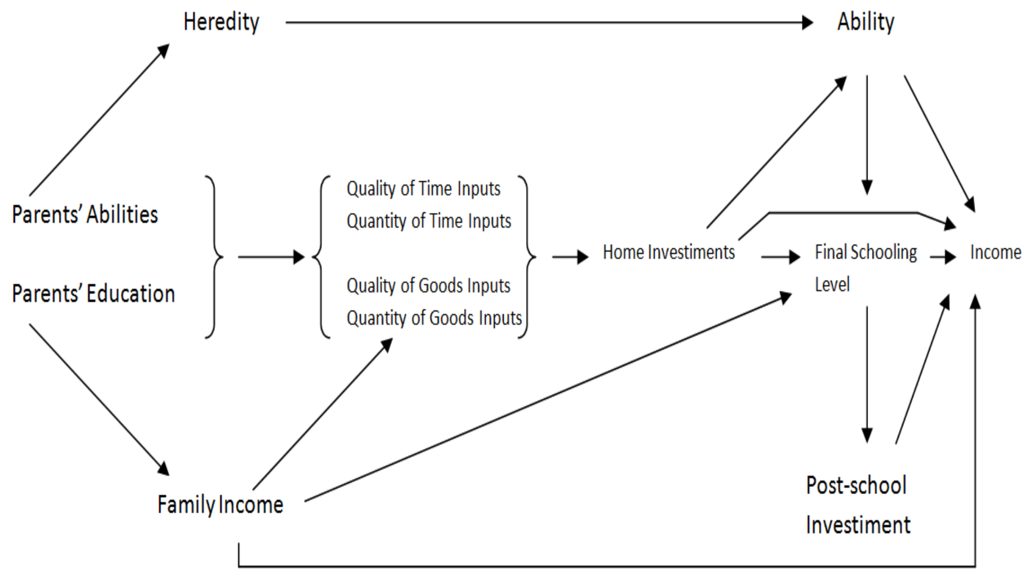


Figure 1. Home investments in children (adapted from Leibowitz, 1974)

When considering individual characteristics, the immigrant status of the student is generally the most investigated achievement's determinant. Schnepf (2007) considers immigrants' educational disadvantage analyzing ten OECD countries (where the immigration rate is similar or higher than 10%) and three OECD dataset: PISA, TIMSS and PIRLS. Implementing three regression models, she controls for the three main characteristics linked to immigrants' poor outcomes: the socioeconomic background, language skills and the characteristics of the school attended. She finds that the immigrant disadvantage is much more evident in Continental European than in English-speaking countries. Analyzing the main determinants, the author demonstrates that in the first group of countries (Continental Europe), the socioeconomic background and school segregation appear to be much more relevant than in English-speaking countries, where language skills remain the most significant factor. At country level, Meunier (2010) provides similar explanations when considering the Swiss context. Using PISA 2000 data, the author demonstrates a negative relationship between the immigrant status and achievement, even when controlling for a set of personal and school characteristics. When analyzing the determinants, a decomposition model is used to show how second generation immigrants' gap is mainly explained by lower endowment. Considering first generation immigrants, part of the gap is explained by lower endowment as well, but nearly one quarter of the achievement difference can be explained by lower return, which raises the issue of educational segregation in schools. Also Azzolini *et al.* (2012) use more recent data (PISA 2009) to obtain similar results for two "new" immigration countries, e.g. Italy and Spain. In particular, as adult immigrants tend to obtain poor labor market attainment in both countries, they investigate how much of the achievement gap of their children is due to the socioeconomic background. They find that immigrant students systematically underperform natives in both countries, especially when they belong to the first generation of immigrants. Having one native-born parent makes the gap achievement disappear, acting like a sort of buffer against low educational performance. Moreover, they show how the socioeconomic background of the family is less relevant in Italy, due to the stronger school segregation by the socioeconomic status that mediate the family effect. Finally, results report that the language spoken at home and the socioeconomic background are the two family factors that explain nearly a third of the achievement gap between first and second generation immigrant. To report an example of a different tendency, immigrant students tend to outperform natives in New York City public schools, as showed by Schwartz & Stiefel (2006) through three sets of regression models. They provide three possible explanations to their findings: the first regards New York schools, which could be able to avoid the achievement gap; the second involve the greater will of immigrant families to “make it” in America; the third regards the better schooling background that immigrant students have when they enroll in the US. Summarizing, they suggest that nativity itself explains a little part of the disparity in children’s performance, while more predictive are the same variables which explain for variability in native-born performances (like the prior performance or the gender).

Researches have also shown that student’s performance is influenced by both his/her own socio-economic background and by the SES background and achievement of his/her peers (the so-called peer effect, see van Ewijk & Sleegers, 2012, for a summary), entailing important policy implications about school environment. The main factors considered in studies about the effect of peers on students’ achievement regard the achievement of peers themselves, their socioeconomic background and the proportion of immigrants students in the class/school. Ammermueller & Pischke (2009) analyze six OECD countries in Center and Northern Europe to investigate the effect of peers’ achievement on primary school students using PIRLSS data. In order to account for peers’ background, they create a vector of variables at class-level that include: number of books at home, student’s sex and age, whether at least one parent is born abroad and whether a foreign language is spoken at home. Through an OLS regression, they find that a one-standard-deviation change in peers’ characteristics leads to 0.17-standard-deviation change in reading test scores in the countries considered. As they initially estimate peer effects at class level (under the assumption that students are randomly allocated across classes), they also implement IV model to control for both across-schools and within-school variation. Results suggest a more important role of measurement error than a self-selection issue of students across school that could influence the peer effects. Similarly, Zimmer & Toma (2000) implement a value added model comparing test scores at the beginning and the end of the school year when students are 13/14 years old. Using data from the International Association for the

Evaluation of Educational Achievement (IEA) and comparing five countries (Belgium, New Zealand, France, United States and Ontario), they show that the effect of peers is greater on low-ability students than on high-ability students. In this study, they define the peer variable as composed by: (I) the mean test scores of students in the classroom; (II) the proportion of high-ability students in the class (upper 20th percentile of test scores distribution among all students across the nation); (III) the proportion of low-ability students in the class (bottom 20th percentile of test scores distribution). Finally, Hoxby & Weingarth (2005) study the reassignment of a large number of students in the Wake County school district, North Carolina (US), as natural experiments on the interaction of a student with his/her new peers in the class. Implementing regression models, they support the Boutique and Focus models of peer effects. The first refers to the fact that a student's learning would benefit from being part of a group where peers have characteristics which are similar to his/hers; the second regards the positive impact that peers' homogeneity have, even though the student is not part of the homogeneous group. Moreover, when peers' achievement is taken into account, other characteristics (such as peers' race, income or parental education) have slight effects.

On the other hand, many studies focus on the integration of immigrant (or minority) students at class/school level. The main factors considered in studies regarding the effect of immigrant students on their peers are generally related to the achievement effect that can be both negative (due to the amount of time that teachers could be forced to address to non-native children or to the lower expectations that teachers could have for all students because of the worse results of immigrant students) or positive (like the positive environment created by immigrant families that have high expectations for their children). Angrist & Lang (2004) consider another kind of reassignment project (Metco) that, aiming at desegregation, send disadvantaged (mostly black) students to more affluent suburb schools in Boston. Through both OLS and 2SLS, they find that Metco students generally pull results down at class-level, mainly because of the worse achievement of Metco students themselves. Nevertheless, once this factor is taken into account, they find little evidence of positive statistically significant effect of Metco students on non-Metco peers. On this topics, differences between the United States and Europe appear to be remarkable. To mention a European study, Jensen & Rasmussen (2011) study the effect of immigrant concentration in Danish schools using PISA-2000 data. They obtain (implementing OLS and 2SLS) that immigrant concentration in schools affect negatively the educational outcomes of both immigrant and native students, also when instrumenting for immigrant concentration across neighborhoods. Nevertheless, they state that since immigrants have lower socioeconomic status than native Danes, the estimated effect of immigrant concentration could be biased by any effects that come from low socioeconomic status of the child's peers in the school. For this reason, they suggest that educational policy should provide support to all disadvantaged students, without focusing only on the immigrant status or immigrant concentration in schools. To underline the close relationship between the effect of the immigrant status and that of the socioeconomic background, we cite an example from the Italian context. Contini (2013) uses national data to investigate the effect of immigrant concentration on test scores in primary and lower secondary schools. The author finds that the negative impact of immigrant students on peers' outcomes is larger for children from disadvantaged family background (immigrant or low socioeconomic background), but is negligible or even positive for high socioeconomic background students. Thus, once again, the concentration of immigrant students is not a fundamental issue *per se*, but it is their relative disadvantage that needs to be addressed. In their investigation, Ballatore *et al.* (2013) start from the assumption that immigrant concentration cannot abstract from the class size in which they are. Therefore, they use Italian data (from INVALSI) to identify the effect on students' performance of increasing the number of immigrant students holding the class size constant. As the class size is an endogenous factor chosen by the school (between a minimum of 10 students and a maximum of 25), it is interesting to note that they actually find that principals tend to increase the class size to compensate the effect of a higher number of immigrant students, and *vice versa*. Despite this finding, the effect of a higher concentration of immigrant students is negative for children in grade 2 (-12% in language test scores and -7% in mathematics) but it vanishes when children are in grade 5. This is remarkable as "it indicates that the school system is somehow capable of implementing educational strategies aimed at neutralizing the negative effects of immigrant inflows in classrooms" (p. 18).

Analyzing the second stream of literature, the debate about school-factors influencing students' performance is multi-faceted (see Vignoles *et al.*, 2000, for a summary).

Most of studies generally agree about the importance of the quality of the teaching force (a review of the main factors is provided by Beteille & Loeb, 2009) and - to a smaller extent - of class size (Krueger, 1999, Rivkin *et al.*, 2005) and school resources (Greenwald *et al*., 1996, 2013, Tajalli & Opheim, 2004, Hanushek & Woessman, 2010). In particular, Rockoff (2004) investigates the issue of teacher quality measuring the impact of teachers' experience on achievement of two cohorts of students in New Jersey. He also accounts for the variance of teaching fixed effects to catch the impact of fixed teacher quality, finding that raising teacher quality could be a key element in improving student outcomes. The debate centered on how increasing teacher quality and effectiveness is a major point of interest. About effectiveness, Lavy (2002) shows that providing teachers with monetary incentives in secondary schools is positively related to students' achievement and less expensive than providing teachers with conventional resources (like on-the-job teacher training) in Israeli context. Rivkin *et al.* (2005) investigate schools' role in influencing students' achievement in Texas between grade 3 and 7 through a value added model. They find that teachers and therefore schools do matter, but achievement gain due to observable characteristics generally explains little of the variance in teacher quality. In particular, they find no evidence of the impact of teacher's qualification (e.g. having a bachelor's or master's degree) on achievement, while the years of experience seem to play an important role. When looking at the other school input characteristic closely related to achievement, namely the class size, they find a small but significant effect of class size reduction on both reading and mathematics scores. Nevertheless, it is estimated that a ten student reduction in class size has smaller effect than increasing the teacher quality distribution of one standard deviation, highlighting the importance of the teaching force. Moreover, they find that class size reduction have marginally smaller but positive effect on test scores until grade 6. This result is compared to the most famous random assignment experiment on class size reduction conducted in Tennessee: the STAR project (Krueger, 1999). Though It involved pupils from kindergarten to grade 3, the main beneficial effect from class size reduction were recorded in the first year of treatment and not in subsequent classes. The STAR project also accounted for the teacher effect, designing three groups: small classes (13-17 students), regular-size classes (22-25 students) and regular/aide classes (22-25) with a full time teacher's aide. Students assigned to small classes score about five-seven percent better than those assigned to regular classes, who in turn perform as well as students provided with teacher aide. Despite the debate between Dr. Krueger and Dr. Hanushek on these topic is still open (see Krueger, Hanushek and Rice, 2002) they agree on the fact that school resources have a larger impact on disadvantaged students than on other students. Mentioning similar studies in the European context, Browning & Heinesen (2003) measure the effect of class size and the number of pupils per weekly teacher hour on educational attainment - represented by years of education - in Denmark. Employing both OLS and 2SLS, they find a positive correlation with both factors when not controlling for background variables, but insignificant estimates when they are included. However, the small effect of reducing class size (an increase of 0.07 years of education when reducing the size by one pupil), could be related to the rather small average class size in Denmark (nearly 20 children), while one may expect to find stronger effect when reduction start from a higher level. In UK, Bradley & Taylor (1998) focus on the effect of the school size on the probability to obtain a high score in the General Certificate of Secondary Education (GCSE). They find that exam performances increase with school size but at a decreasing rate. They highlight that policy implication about school closures or school mergers cannot rely only on these data, but should also take into account aspects related to non-cognitive educational skills provided by schools or aspects related to all-round education. Focusing on both family background and school quality in the Italian context, Brunello & Checchi (2005) find through a two-step regression that a better school quality is particularly significant for individuals from poor family background and provide evidences that school is a technical substitute to parental education, helping children to overcome the social gap. Nevertheless, they highlight how the family background is still, in Italy, the main predictor of children’s educational attainment, with a strong intergenerational persistence. Moreover, as Agasisti et al. (2014) show, strong differences occur between Northern, Central and Southern Italy, with school effects much more consistent in the South (indicating a more diversified quality of schools). In their research, they not only claim that the quality of the educational system varies, but also that “the interplay between individual and school characteristics are not uniform across the country” (p. 24) claiming that differences across schools tend to increase, instead of reducing, the gap between the more advantaged and disadvantaged students.

Despite the evident relationship between school achievement and the socioeconomic background on the one hand, and the important role that school factors can play in overcoming (or sharpening) the social and economic gap on the other hand, the literature concerning specifically how schools may foster academic achievement of disadvantaged students is quite poor. Nonetheless, providing evidence that school policies can make the difference in disadvantaged students' outcomes would have great implications for policymakers when designing interventions for promoting equality. Focusing on a particular category of students, those who obtain good academic results despite a disadvantaged background – namely "resilient students" (OECD, 2012) – it can be argued that the proportion of resilient students as a proxy of the equality of an educational system; thus a higher number of resilient students would imply a more equal (and efficient) educational system.

Research about students resiliency often focuses on behavioral and personal aspects of students able to succeed despite adversity. Individual characteristics of resilient students generally include an internal locus of control, self-esteem, higher engagement in school and strong interpersonal skills (Wang *et al.*, 1997, Borman & Rachuba, 2001, OECD, 2012). Walsh & Black (2009) analyze the Australian context recognizing the major role of disengagement from school in determining poor achievement, especially among disadvantaged students. For this reason, they describe two programs (Cityscape and ruMAD) aimed at a student-centred design of school, in order to overcome the social background and geographic location, which is the other important factor influencing achievement of Australian students. In their discussion, motivation and educational engagement are described as the main factors that allow this process. Generally, researchers also agree about the importance of greater engagement by parents or family (Benard, 1991, Borman & Rachuba, 2001, Australian Government Productivity Commission, 2012) through high expectations and participation.

Less quantitative support is provided about school features fostering resiliency. Generally, researchers agree about the importance of caring and supportive teachers (Benard, 1991, Borman & Rachuba, 2001, Walsh & Black, 2009), healthy peers group (Benard, 1991, Johnson, 1997), strong school leadership (Muijs *et al*., 2004) and positive school climate (Wang *et al.*, 1997, Agasisti *et al.*, 2014). Finally, the role of community and the need for a close relationship with school are important factors in the discussion about resilience (Wang *et al.*, 1997, Borman & Rachuba, 2001). Bromberg & Theokas (2014), in their research about the characteristics of high achieving students in American high schools, highlight the important role of school as a community that foster students' achievement. Also Borman & Rachuba (2001) make a similar statement focusing on poor and minority students in US from three cohorts of students from grade 1, 3 and 7. Testing four models through two-way MANOVAs (accounting for both SES and race), they find that the school community model, which includes elements that actively shield children from adversity, is the most powerful predictor of children's resiliency. However, they do not find significant difference when accounting for students' race and ethnicity, suggesting the applicability of uniform individual and school-level models of academic resiliency to all low-SES students. To summarize, Figure 2 reports the model presented by Benard (1991) in her review about resiliency. She claims that resilient students tend to show a series of individual characteristics in early childhood that will help them to become resilient: problem solving skills, that include the ability to think abstractly and flexibly, autonomy, which is often referred to as "internal locus of control", sense of purpose and future, that includes achievement motivation and educational aspiration, and social competence, in terms of being responsive and caring. When analyzing the protective factors within the family, she highlights the importance of a close relationship with at least one member of the family who provide the child with caring and supportive behavior, have high expectations for his/her education and life, and also encourage the children's active role as participant of life and work of the family. The school has also an important role in fostering resiliency, being a "protective shield" that may help students who are least likely to obtain support elsewhere. The main factors refer to a caring and supportive environment, both from teachers and peers, that establish high expectations for all kids and give them the opportunity to participate with roles of responsibility within the school environment. Finally, the community plays a fundamental role when providing support mainly through the availability of resources necessary for healthy human development (like child care or education), high expectations (in terms of "cultural norms" like the tendency to consider youth as resources instead of sources of problems), and opportunities for participation as a contributing member of the community itself.

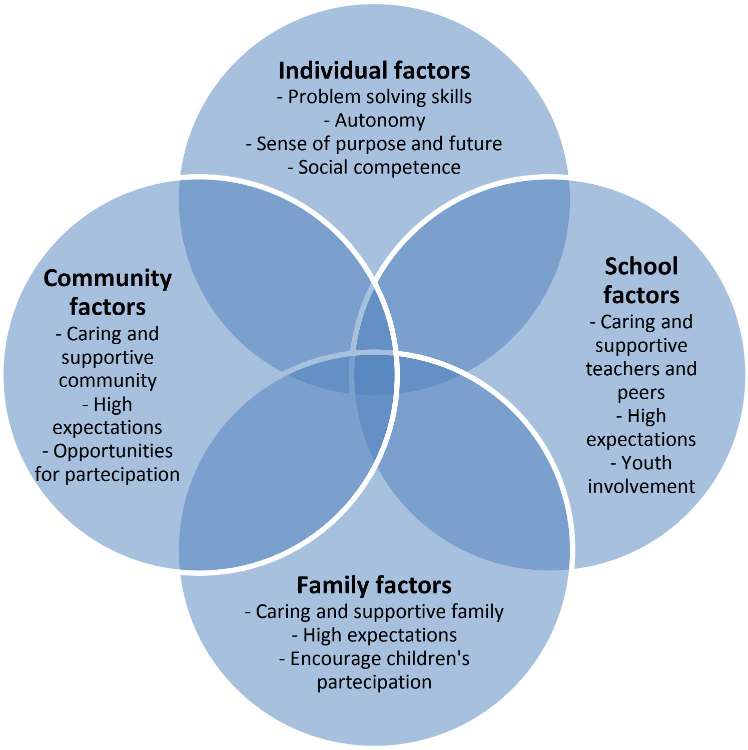


Figure 2. Factors fostering resiliency. Authors' elaboration from Benard (1991)

Other interventions on school resources, (like reducing school and class size), are considered to have small impact if applied on their own. Indeed such actions need to be accompanied by changes in school and classroom practices such as establishing a clear educational mission in the school or providing ongoing professional learning for teachers (Faubert, 2012). On the contrary, researchers generally highlight the importance of the human relationship at school (such as supportive school personnel or healthy peer group) to provide the support that often is missing from the family when children come from a disadvantaged condition (Johnson, 1997, Wang *et al.*, 1997). OECD (2011) stresses the importance of "classroom practices and teaching methods that encourage learning and foster motivation and self-confidence" (p. 4) like high-quality mentoring programs in order to fill the self-confidence gap. Moreover, the literature show how the gap between advantaged and disadvantaged arises early in children's life, making early intervention particularly important. Heckman (2008) argue how much later interventions are ineffective, and even when they show some benefits, the performance of disadvantaged students is still behind that of children who experienced earlier interventions (in preschool years), as shown in figure 3.

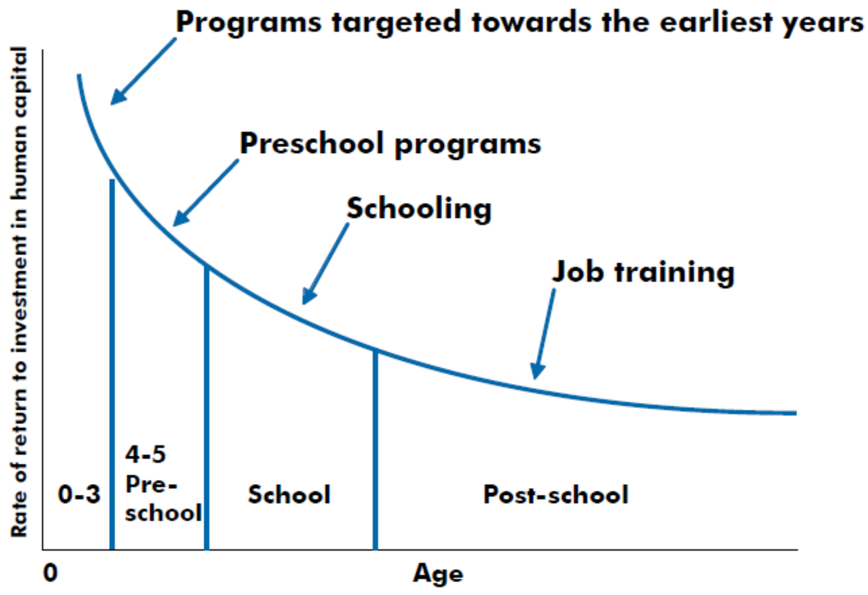


Figure 3. Returns to a unit dollar invested, Heckman (2008).

The most powerful programs are those that actively involve parents and family aiming at changing home environment or influencing children's character and motivation as well as their cognitive abilities. Aiming at early intervention towards disadvantaged students, Levin (1988) presents the model of the Accelerated School as a "transitional elementary school that is designed to bring disadvantaged children up to grade level by the completion of the sixth grade" (p. 218). This model has been implemented in several schools across the US with strongly positive results. School's resources and environment are not the only factors investigated when focusing on disadvantaged students. The role played by institutions in fostering resiliency has also been object of discussion. As an example, Agasisti & Longobardi (2014a) analyze OECD-PISA 2009 data for EU-15 countries to determine, through a logit regression, which characteristics are more related to the resilient status - defined as the ability to obtain good academic results despite the disadvantaged condition -. Their results suggest that, together with parents engagement and school climate, school's autonomy and the provision of extracurricular activities are school-level characteristics positively related to resilience. These findings highlight the positive role that school's independency and other educational systems' "institutions" can have in influencing student's achievement and resiliency.

As disadvantaged students are often geographically concentrated, especially in inner-cities and rural areas, part of the literature on this topic is focused on disadvantaged schools, i.e. schools characterized by a particularly high concentration of low-SES students that for this reason have to face the greatest challenges. Tajalli & Opheim (2004) consider schools characterized by the same low socioeconomic background in order to investigate how school-level characteristics affect performances variation. Working on a dataset of Texas schools at grade 4, 8 and 10, they implement a logistic regression where the output is a dichotomous variable identifying low and high-performing schools. Thirteen independent variables are used as predictors, providing information about school resources (mainly expenditure ratios), size, SES and race composition, teachers' experience and salary. The major finding refers to the importance of teachers' characteristics, both in terms of salary and experience, that are positively related to better outcomes at school level. The SES and race composition of schools are found to be significant in grade 4 and 8. In particular, an increase of one percent on the proportion of disadvantaged students at school makes the probability to be a high-performing school drop by 6.3% for grade 4 and 8.4% for grade 8. On the other hand, school expenditure and size are not found to be generally related to the average achievement. Conducting a review of the literature on this topic, Mujis *et al.* (2004) report three theoretical frameworks regarding school improvement in disadvantaged areas: the *contingency* theory is based on the statement that what makes an organization effective is linked to situational factors that are both internal and external to the organization itself; in this context, effective schools in disadvantaged areas would be characterized by particular configuration or policies that make them different from the others. The *compensatory* model deals with the idea of school as an institution that needs to compensate for the lack of resources in the children's homes, concentrating on students' basic needs. Finally, the hypothesis of *additivity* of schools and background factor effects is based on the observation that schools in disadvantaged areas still perform worse than others even when students' background has been controlled, suggesting the idea that these schools tend to increase the socioeconomic gap instead of reducing it. From the literature, authors find more evidences supporting the contingency theory than other models, claiming that effective low-SES schools could actually differ from high-SES campuses in the sphere of teaching and learning, providing different instructional strategies. They also show a lists of overall factors that may help schools in disadvantaged areas to improve, which include creating an information-rich environment, focusing on teaching and learning, having a strong school leadership, a positive school culture and a learning community, providing professional development and involving parents. Agasisti & Longobardi (2014b) implement a multilevel logistic approach on OECD-PISA 2009 data focusing on the Italian context and selecting a particular category of low-SES students, namely those who come from disadvantaged families and attend disadvantaged schools. Their aim is to find which school-level characteristics are more related to resilience, defined as the ability to overcome the disadvantaged background, obtaining good level of educational performance. They define disadvantaged schools as those in the bottom 33rd percentile of the whole distribution of the socioeconomic index and disadvantaged students (within disadvantaged schools) as those whose socioeconomic indicator is in the bottom quartile of the new distribution (that of low-SES schools). Then, they define resilient students regressing their performance of the square of the socioeconomic index, allowing for non-linearity in this relationship. They find that not only the socioeconomic background of the family and context factors (like the geographic location) are important, but also that school factors matter. Relevant school features are not only related to resources but also to “soft” managerial skills of school’s organization. The most significant factors are (I) better relationship between students and teachers (II) the provision of (good) extracurricular activities and (III) school resources aimed to avoid teachers’ shortage and increase the quality of teaching activities.

1.3 The Italian educational system and policy measures towards equity

The Italian educational system includes 7.8 million of students attending 366,000 classes in 41,483 schools (year 2013/2014). Among them, 32.5% are kindergartens (preschool years), 37% are primary schools (grade 1-5), 17.5% are lower secondary schools (grade 6-8) and 13% are upper-secondary schools (grade 9-13). The latter are then divided into *licei* (designed to give students the skills to progress to any higher educational institution), technical and vocational schools.

The issue of equity in the Italian educational system has become particularly important in the last decade. Many national and international studies show how, despite a theoretically high grade of homogeneity of the educational system, strong regional differences are evident (Falzetti & Ricci). For example, in OECD PISA 2012 scores, there is a difference between Trentino – Northern region – and Calabria – Southern region – of 94 points, which means approximately two years of school. Among national policies implemented to increase equity in the educational system, many efforts aim at limit school drop-out and school failure, which are very serious issues in Italian schools. According to a report of the Educational Commission of Italian Parliament, 3 millions of high school students have dropped out before the 5th year since 1999. Considering that the total number of students was 9 millions, this means that one out of three haven't completed the upper secondary school. Focusing on last 5 years, 597,915 students enrolled in a public high school in 2009/2010. Five years later (2013/2014), 167,083 have dropped-out, with a drop-out rate of 27.9% (dossier Tuttoscuola, "Dispersione scolastica nella scuola secondaria superiore statale" from Ministry of Public Education data). About one out of four of that students completed schooling in private schools, whereas the remaining part has become Neet (Not in Employment, Education or Training), who are 22.9% of Italian people between 15 and 29 (ISTAT 2013). Most of students drop out at the end of compulsory education, after the second year of high school. Last year (2013/2014) drop-out rate from second to third year of upper secondary school was 14.8% (90,866 students). Moreover, drop-out rate is not uniform neither along Italy nor among different types of school. From a geographical point of view, 46% of students who dropped out live in the south, where drop-out rate is 27.5%. Nevertheless, the highest rate is in the North-West (29.1%), while the lowest is in the North-East (24.5%). Two out of three students who dropped out attend technical or vocational institutes (precisely 28.1 and 38%), types of school that are often attended by students with lower socio-economic condition or cultural background. Policy addressed to limit drop-out rate are rather slow to bring results. Generally, these type of policies refer to the inclusion objective, starting from the increase of compulsory school until 16 years old (law n.296, 2006). In October 2013 was presented an amendment to increase the compulsory education from 16 to 18 years old, pointing to the creation of higher cultural level and the possibility to attend a higher education, supporting educational progression.

Projects aimed at reducing drop-out rate can be recognized in particular in PON (National Operative Programmes) projects, both in 2000/2006 programme and in 2007/2013 one. In 2007/2013 programme, for the specific objective F – Promoting school success, equal opportunities and social inclusion - € 270millions were allocated, financing 5700 projects. It turns out that change in drop-out rate is slow to happen, even though there has been a decrease - of about 10% - in the number of students who dropped out (who were about 37% in 2000). Considering that one of the goal of the programme *Europa 2020* is to keep drop-out rate under 10%, It is clear that Italy has to solve a long standing issue. The mentioned National Operative Programmes (PON) are structural funds allocated by the European Union in order to narrow the gap between more advanced and less developed areas. In Italy, they are allocated in Southern regions, characterized by lower socio-economic condition than Northern areas. Funds have been allocated in the period spanning 1994/99, 2000/06 and 2007/13. Since 2000, two National Operative Programmes has been carried through: PON "Competenze per lo sviluppo", co-funded by the European Social Fund (ESF), which promotes employment through actions on human capital, and PON "Ambienti per l'apprendimento", co-funded by the European Regional Development Fund (ERDF) which promotes regional development. The common objective is to foster "convergence", i.e. the growth of underdeveloped regions (whose per capita Gross Domestic Product is under 75% of the EU average). In Italy, four regions have been involved: Calabria, Campania, Puglia and Sicily (in the 2000/2006 programme, Sardegna and Basilicata were also involved). Programmes have been 50% co-funded by the Italian National Government and completely managed by the Ministry of Public Education. Even though performance's variance between Northern and Southern Italy is still evident, some advances have appeared. Referring to reading literacy, in the period spanning 2000/2009 the number of students with low proficiency has decreased from 28.5% to 27.5% in Southern regions. The result is much more evident considering that the percentage of low proficient students reached 35% in 2003. About Mathematics

literacy, the number of low proficient students has decreased from 47.5% to 33.5% in the period 2003/2006.
Other policies aimed at creating a more equitable education system and labor market entry regard the vocational training. In 2011, 62.4% of vocational training providers began at least one educational activity financed with public money. Students are, in prevalence, female and young between 18 years and 34 years old. Unemployed are twofold of employed, with a peak in the South. Most of members have junior high school degree or high school degree. Only a marginal part of members are foreigners. Vocational education and development of human resources simplify employment, contribute to economic development and guide social inclusion. In addition, thanks to vocational training, we register a decreasing of drop-out rate, for example helping the students through modular courses.
Finally, some interventions have also been directed towards the integration of foreign students, who represent nearly 10% of students population (year 2013/2014). Foreign students integration is often complicated because of cultural and language difficulties on the one side, and because of poor socio-economic condition on the other. Inequality is observable from data: 38.2% of foreign students are late-enrolled versus 11.6% of natives. Moreover, the difference rises with the school level: late-enrolled foreign students are 16.3% in primary school (2% among Italian students) but they are 67.1% in upper secondary school. In order to make integration easier, Italian government has established a national plan for teaching Italian as a second language (*Piano Nazionale Italiano L2*) targeted to students living in Italy for less than two years and attending lower or upper secondary school. Moreover, a national fund for schools placed in areas of high immigrant concentration allows to finance projects to limit foreigners marginalization and drop-out rate. Funds for 2013/2014 school year were € 29,730,000.00. Money allocated by the National Government are mediated by Regional Governments and Municipalities, and then assigned to schools, according to the project interest. Finally, school integration is also supported by the *European Fund for the integration of third country nationals*. Among actions implemented through these resources, action number 3 is targeted to *“school integration and social inclusion of young immigrants”*, which is partially co-funded by the national government. Data demonstrate that policies and funds have been supplied both at national and European

level, but much more efforts are needed to increase equity in a substantial measure in the Italian educational system.

1.4 Data

1.4.1 The Dataset

In this research we use data from the standardized test administered by the Italian National Evaluation Committee for Education (INVALSI), an organization subjected to the control of the Ministry of Public Education. The aim of this organization is to periodically evaluate abilities of Italian students and to control the quality of the educational system.

Evaluation of students' ability is realized through five tests taken at grade 2, 5, 6, 8 and 10 (since 2013/14, test at grade 6 has been suppressed) . This means that the first two tests are taken at the primary school, two others at the lower secondary school and the last at the upper secondary school. Standardized tests assess both language (Italian) and mathematical skills of students. Scores are corrected for *cheating propensity* through a procedure that accounts for average scores, scores variation within classes, answers homogeneity and repeated missing answers. Results are then adjusted according to the Rasch model, which takes into account the complexity of questions and the level attended, enabling direct comparisons. Scores are also standardized to have a mean of 200 and a standard deviation of 40.

In order to investigate how the school system affects students' abilities, we have decided to concentrate our attention on tests taken at grade 5 and 6. In Italian educational system, grade 5 corresponds to the last year of the primary school and grade 6 corresponds to the first year of the lower secondary school. This change is particularly significant in our research as enables us to observe a change in school's characteristics. We use data taken from tests of the last two years available when the research started: 2011/2012 for grade 5 and 2012/2013 for grade 6.

In addition to test scores, dataset contains a large number of individual socioeconomic characteristics as well as personal information collected through a questionnaire. Questions regard, for example, the number of books at home or how students spend their free time. The answers to these questions are largely used to compute the Index of Economic, Social and Cultural Status (ESCS) which is then provided for each student.

We are also provided with some variables that regard the classroom and the school attended by the student, which give us precious information for our analysis. The complete list of variables used in the study is in table 1. The dataset provided by INVALSI permits to conduct a three-levels investigation. At the first level we analyze students' characteristics in terms of test achievement (both in Reading and Mathematics test), gender, immigrant status (divided into first and second generation immigrants), age of schooling (students who enrolled one year before the standard age of 6 are defined "early-enrolled"; students who enrolled after the age of 6 or repeated one or more years are defined "late-enrolled"), information about economic, social and cultural status of the family (calculated through the ESCS index) and about the family size (through the number of siblings for each child). At the second level we take into consideration classroom-level characteristics. We have information about the average socioeconomic background of the classroom, the proportion of female, immigrants, early and late-enrolled students, the number of students in the classroom and the time spent at school in a week. With the expression "*tempo pieno*", we mean a class where students spend at school 36 or 40 hours a week instead of 24, 27 or 30 hours. The choice between the first and the second option is made by the family depending on the offer of the school. Every characteristic is available for both 5 and 6 grade. We also investigate school-level features, creating nine variables related to the dimension of the school at each grade (number of classrooms and students) and to the average number of students per class in each school. A dummy variable indicates if the school is private or public. We also have information about the number of schools that are an *Istituto comprensivo*, which is a school that gathers in the same organization kindergarten, primary and lower secondary school, and about students who attended the same *Istituto comprensivo* between grade 5 and 6. Finally, we compute – for each student - *delta variables* to take into account how school and classroom-level characteristics change between grade 5 and 6.

Table 1. Variables used in the research (from INVALSI 2011/12 and 2012/13).

Category	Variable	Description
Student-level characteristics	Achievement	Test score in Mathematics corrected by cheating propensity and Rasch model (Mean=200, Stdev=40).
	Gender	Gender of student (0=male, 1=female).
	Immigration Status	Status (1=native, 2=first generation immigrant, 3=second generation immigrant).
	Age of schooling	Status (1=regular, 2=early-enrolled student, 3=late-enrolled students.
	ESCS (Index of Economic, Social and Cultural Status)	The OECD index of economic, social and cultural status (ESCS) is based on information from students on parental occupations, parental education, and home possessions. It is standardized to have a mean of 0 and a standard deviation of 1.
	Students who have siblings	Number of siblings of a student.
Classroom-level characteristics	Class average socio-economic background	Average ESCS of the classroom.
	Class average score	Class average score in mathematics test.
	Proportion of female students	Number of female students with respect to the total number of students in the class.
	Proportion of immigrant students	Number of immigrant students (first or second generation) with respect to the total number of students in the class.
	Proportion of early-enrolled students	Number of early-enrolled students with respect to the total number of students in the class.
	Proportion of late-enrolled students	Number of late-enrolled students with respect to the total number of students in the class.
	Number of students in the class	Number of students enrolled in the class (independently from the number of students who took the test).
	Class with "tempo pieno"	Number of classroom offering at least 36 school hours per week.
School-level characteristics	School average socio-economic background	Average ESCS in the school.
	School average score	School average score in the mathematics test.
	Number of classrooms in the school	Number of classrooms in the school per each grade.
	Number of students in the school	Number of students enrolled in the school per each grade.
	Average number of students per class in the school	Number of students enrolled in each class computed at school-level.
	Private school	Status (0=public school, 1=private school).
	Istituto Comprensivo	Status (0=generic school, 1=*Istituto Comprensivo*).
	Same *Istituto Comprensivo*	Proportion of students who attended the same *Istituto* from primary to lower secondary school.
Variation in classroom-level characteristics	Delta class average score	Difference in class average mathematics test scores between grade 6 and 5.
	Delta class average socio-economic background	Difference in class average ESCS between grade 6 and 5.
	Delta number of students in the class	Difference in the number of students per class between grade 6 and 5.
	Delta number of immigrant students in the class	Difference in the number of immigrant students between grade 6 and 5.
Variation in school-level characteristics	Delta school average score	Difference in school average mathematics test scores between grade 6 and 5.
	Delta school average socio-economic background	Difference in school average ESCS between grade 6 and 5.
	Delta number of classrooms in the school	Difference in the number of classes of each grade between grade 6 and 5.
	Delta number of students in the school	Difference in the number of students of each grade between grade 6 and 5.
	Delta number of immigrant students in the school	Difference in the number of immigrant students between grade 6 and 5.

Notes. Authors' elaboration on INVALSI data 2011/12 and 2012/13.

INVALSI tests are taken at national level, so initial data regard the entire population of students that attended grade 5 in 2011/2012. However, students take the reading and mathematics tests in two different days. For this reason we haven't a complete correspondence between reading and mathematics' datasets and we lose 2% of observations of grade 5 and 1% of grade 6. The first step is to identify students that took both tests, obtaining a sample of 473,869 students at grade 5 and 481,119 students at grade 6 (table 2).

Table 2. Sample size at national level.

	BEGINNING SAMPLE		STUDENTS WHO TOOK BOTH ITA AND MATHS
	READING	**MATH**	
Grade 5	489,581	489,279	473,869
Grade 6	484,033	484,441	481,119

One of the main points in our research is to study the switch between grade 5 and 6, in order to investigate a variation in school's characteristics that would not be caught otherwise. This is why we have merged the two dataset obtaining a complete sample that contains all the information about students who took the test both in 5^{th} and 6^{th} grade (285,066 students, which correspond to nearly 60% of the entire population). This loss of data is due to a matching problem in student's codes from data provided by the Ministry. The procedure is still problematic because the data about tests are collected by INVALSI, while those about individual students' characteristics are archived in the Ministry's registry. The impact of this loss of information on the representativeness of the subsample is discussed below in this section. Finally, we select students living in five big cities in Northern and Central Italy: Milan, Bologna, Turin, Padua and Rome. The decision to concentrate our attention on big cities is driven by the following factors: (I) inequality is more evident in cities with high population (Wang *et al*., 1997), (II) greater freedom of choice involves more variation in the characteristics of school and make investigation of school features more meaningful. The fact that, in big urban areas, these two characteristics occur jointly makes resilience a more probable phenomenon. Furthermore, considering big cities we can investigate the immigrants issue - generally related to poor socio-economic background and low educational outcomes (OECD, 2012) - from a closer point of view. In fact, as Azzolini *et al.* (2012) show analyzing PISA 2009, immigrant students in Italy tend to systematically underperform natives, especially when they are first-generation immigrants. Making this selection, we obtain a sample of 90,598 students (table 3).

Table 3. Final sample size.

STUDENTS WHO TOOK THE TEST IN 5th & 6th GRADE (at national level)	FINAL SAMPLE (5 cities)
285,066	90,598

Finally, we decide to exclude Southern cities for two reasons: (I) smaller number of immigrants live in Southern regions or Islands, (II) cheating is a minor problem in the North, but it seems to be a relevant problem in the South. Bertoni, Brunello & Rocco (2013) analyze how the presence of external examiners impact on INVALSI test scores. They find that external monitoring reduces test scores due to less cheating possibility. Cheating propensity is much more evident in Southern regions – where the total effect is a 8.9% reduction with respect to the mean scores in untreated schools – than in Northern Italy (2.6%) .

We have also made the decision to remove from the sample students living in the province of each city. The decision to focus on inner-cities is supported by the analysis made on the difference between province and inner-city (tables 4).

Tables 4. Comparison between province and inner-city.

Bologna

Population size in the sample	**N**		
Province	6927		
Inner-city	2153		
Immigrant concentration in the sample	**%**		
Province	8.4		
Inner-city	17		
Number of schools in the sample	**N**	**Private schools (%)**	**Schools per student (‰)**
Province (Bologna excluded)	64	2.3	9.24
Inner-city	37	18.4	17.19
Private schools per student	**‰**		
Province	2.8		
Inner-city	6.97		

Milan

Population size in the sample	N		
Province	30193		
Inner-city	8987		
Immigrant concentration in the sample	**%**		
Province	0.132		
Inner-city	0.19		
Number of schools in the sample	**N**	**% Private**	**Schools per student (‰)**
Province (Milan excluded)	281	7	9.31
Inner-city	141	19.7	15.69
Private schools per student	**‰**		
Province	3.61		
Inner-city	6.9		

Padua

Population size in the sample	N		
Province	7320		
Inner-city	1621		
Immigrant concentration in the sample	**%**		
Province	0.127		
Inner-city	0.154		
Number of schools in the sample	**N**	**% Private**	**Schools per student (‰)**
Province (Padua excluded)	70	8.57	12.28
Inner-city	27	48.1	16.66
Private schools per student	**‰**		
Province	2.60		
Inner-city	8.02		

Rome

Population size in the sample	N		
Province	29952		
Inner-city	19242		
Immigrant concentration in the sample	**%**		
Province	0.101		
Inner-city	0.097		
Number of school in the sample	**N**	**% Private**	**Schools per student (‰)**
Province (Rome excluded)	168	7.9	5.61
Inner-city	342	17.7	17.77
Private schools per student	**‰**		
Province	5.91		
Inner-city	7.8		

Turin

Population size in the sample	N		
Province	16206		
Inner-city	5705		
Immigrant concentration in the sample	**%**		
Province	0.126		
Inner-city	0.177		
Number of schools in the sample	**N**	**% Private**	**Schools per student (‰)**
Province (Turin excluded)	141	13.48	13.43
Inner-city	78	30.77	13.67
Private schools per student	**‰**		
Province	2.65		
Inner-city	4.21		

Notes: all data refer to grade 5, year 2011/2012

Despite the higher number of students in province, the percentage of schools per thousand students in inner-cities is higher than in provinces. This information support our idea about the greater freedom of choice in big cities. The same can be stated for the number of private schools, much more concentrated in towns. Moreover, the percentage of immigrant students is higher in inner-cities, allowing us to better understand this phenomenon. Aiming at finding out more about these differences, we compare characteristics of the province and the city through a three-levels investigation.

We also compare these data with the characteristics of students who took the test at both grade 5 and 6 (after the matching procedure). Tables 5 report the results obtained.

Tables 5. Comparison between province and inner-city (pre and after matching).

Bologna

	Province		Inner-city		Inner-city after matching	
	N=6,770		N=2,153		N=1,296	
Student-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Achievement in Reading	205.54	41.77	203.48	38.51	207.20	47.23
Achievement in Math	167.66	37.02	178.63	24.69	178.86	42.23
Female Student	0.49		0.501		0.51	
1st generation immigrant	0.06		0.085		0.079	
2nd generation immigrant	0.07		0.079		0.076	
Early-enrolled student	0.008		0.013		0.008	
Late-enrolled student	0.04		0.057		0.053	
Socioeconomic background	0.59	1.05	0.76	1.053	0.724	1.03
Student who has siblings	0.74		0.71		0.70	
Classroom- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Class average socioeconomic background (ESCS)	0.57	0.85	0.74	0.78	0.67	0.70
Proportion of female in the classroom	0.49		0.49		0.49	
Proportion of 1st generation immigrants in the classroom	0.06		0.08		0.09	
Proportion of 2nd generation immigrants in the classroom	0.08		0.08		0.08	
Proportion of Early-enrolled students in the classroom	0.007		0.01		0.01	
Proportion of Late-enrolled students in the classroom	0.04		0.05		0.06	
Number of student in the classroom	21.74	3.23	22.55	2.49	22.37	2.55
Class with "tempo pieno"	0.54		0.55		0.54	
School-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Number of classrooms in the school	4.85	1.77	4.55	1.939	4.59	1.93
Number of students in the school	84.73	43.32	76.43	47.92	80.74	49.03
Average numbers of students per class, in the school	21.75	3.23	22.55	2.5	22.38	2.55
Private school	0.07		0.18		0.18	
Istituto comprensivo	0.68		0.47		0.5	

Milan

	Province		Inner-city		Inner-city after matching	
	N=29,865		N=8,987		N=5,608	
Student-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Achievement in Reading	206.21	38.68	204.99	35.78	207.94	38.3
Achievement in Mathematics	205.02	37.15	204.43	32.66	206.62	36.54
Female Student	0.5		0.505		0.509	
1st generation immigrant	0.05		0.06		0.052	
2nd generation immigrant	0.08		0.13		0.13	
Early-enrolled student	0.005		0.007		0.009	
Late-enrolled student	0.03		0.041		0.038	
Socioeconomic background	0.3	1	0.49	1.104	0.595	1.121
Student who has siblings	0.7800		0.78		0.77	
Classroom- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Class average socioeconomic background (ESCS)	0.28	0.81	0.44	0.79	0.529	0.78
Proportion of female in the classroom	0.50		0.50		0.50	
Proportion of 1st generation immigrants in the classroom	0.05		0.06		0.056	
Proportion of 2nd generation immigrants in the classroom	0.08		0.13		0.14	
Proportion of Early-enrolled students in the classroom	0.005		0.007		0.007	
Proportion of Late-enrolled students in the classroom	0.03		0.04		0.04	
Number of student in the classroom	21.65	3.09	21.62	3.19	21.83	3.18
Class with "tempo pieno"	0.82		0.77		0.78	
School-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Number of classrooms in the school	4.69	1.81	5.07	2.205	4.93	2.13
Number of students in the school	85.14	43.55	76.57	50.57	77.27	50.07
Average numbers of students per class, in the school	21.65	3.09	21.62	3.195	21.84	3.19
Private school	0.10		0.197		0.220	
Istituto comprensivo	0.51		0.43		0.42	

Padua

	Province		Inner-city		Inner-city after matching	
	N=7,320		N=1,621		N=828	
Student-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Achievement in Reading	206.20	39.77	209.00	39.36	212.67	37.67
Achievement in Mathematics	169.82	37.00	168.35	38.47	172.58	41.88
Female Student	0.51		0.51		0.53	
1st generation immigrant	0.065		0.086		0.093	
2nd generation immigrant	0.062		0.068		0.058	
Early-enrolled student	0.005		0.009		0.007	
Late-enrolled student	0.041		0.056		0.052	
Socioeconomic background	0.2723	1.0189	0.7031	1.0773	0.7711	1.0543
Student who has siblings	0.79		0.78		0.78	
Classroom- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Class average socioeconomic background (ESCS)	0.2474	0.8830	0.6494	0.7953	0.7280	0.7922
Proportion of female in the classroom	0.51		0.51		0.53	
Proportion of 1st generation immigrants in the classroom	0.065		0.093		0.107	
Proportion of 2nd generation immigrants in the classroom	0.063		0.071		0.059	
Proportion of Early-enrolled students in the classroom	0.005		0.009		0.007	
Proportion of Late-enrolled students in the classroom	0.041		0.056		0.060	
Number of student in the classroom	20.11	3.86	21.40	3.42	22.03	3.17
Class with "tempo pieno"	0.20		0.69		0.35	
School-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Number of classrooms in the school	5.71	2.46	4.70	2.30	4.39	2.11
Number of students in the school	88.46	48.01	72.88	44.24	72.76	42.38
Average numbers of students per class, in the school	20.32	3.01	22.06	2.79	21.80	3.31
Private school	0.07		0.25		0.28	
Istituto comprensivo	0.72		0.75		0.72	

Rome

	Province		Inner-city		Inner-city after matching	
	N=27,872		N=18,360		N=8,081	
Student-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Achievement in Reading	205.12	38.97	207.18	39.53	208.89	37.53
Achievement in Mathematics	183.83	40.74	186.86	41.82	186.15	41.53
Female Student	0.49		0.493		0.50	
1st generation immigrant	0.04		0.04		0.038	
2nd generation immigrant	0.05		0.057		0.056	
Early-enrolled student	0.01		0.022		0.024	
Late-enrolled student	0.03		0.034		0.031	
Socioeconomic background	0.37	0.94	0.4588	0.9498	0.587	1
Student who has siblings	0.78		0.77		0.77	
Classroom- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Class average socioeconomic background (ESCS)	0.34	0.73	0.41	0.69	0.53	0.66
Proportion of female in the classroom	0.49		0.49		0.50	
Proportion of 1st generation immigrants in the classroom	0.05		0.04		0.04	
Proportion of 2nd generation immigrants in the classroom	0.05		0.06		0.05	
Proportion of Early-enrolled students in the classroom	0.01		0.02		0.02	
Proportion of Late-enrolled students in the classroom	0.03		0.03		0.03	
Number of student in the classroom	21.44	3.66	21.31	3.47	21.45	3.42
Class with "tempo pieno"	0.52		0.58		0.58	
School-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Number of classrooms in the school	5.14	2.47	5.13	2.62	5.14	2.77
Number of students in the school	76.56	51.52	70.96	52.52	74.36	53.52
Average numbers of students per class, in the school	21.45	3.66	21.32	3.47	21.45	3.42
Private school	0.13		0.177		0.20	
Istituto comprensivo	0.36		0.29		0.31	

Turin

	Province		Inner-city		Inner-city after matching	
	N=16,206		N=5,705		N=2,562	
Student-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Achievement in Reading	203.42	38.08	204.42	40.52	207.94	38.34
Achievement in Mathematics	205.58	39.64	201.41	39.48	203.62	38.14
Female Student	0.496		0.50		0.51	
1st generation immigrant	0.065		0.084		0.086	
2nd generation immigrant	0.061		0.093		0.098	
Early-enrolled student	0.005		0.005		0.005	
Late-enrolled student	0.037		0.05		0.041	
Socioeconomic background	0.2032	0.9951	0.2693	1.0615	0.3660	1.0748
Student who has siblings	0.79		0.78		0.78	
Classroom- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Class average socioeconomic background (ESCS)	0.2504	0.8252	0.2504	0.8252	0.2949	0.8484
Proportion of female in the classroom	0.493		0.50		0.50	
Proportion of 1st generation immigrants in the classroom	0.064		0.087		0.088	
Proportion of 2nd generation immigrants in the classroom	0.060		0.097		0.107	
Proportion of Early-enrolled students in the classroom	0.005		0.004		0.005	
Proportion of Late-enrolled students in the classroom	0.038		0.042		0.042	
Number of student in the classroom	20.93	4.11	22.17	2.90	22.11	2.98
Class with "tempo pieno"	0.62		0.69		0.74	
School-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Number of classrooms in the school	5.49	1.89	5.32	2.10	5.34	2.47
Number of students in the school	91.04	45.26	89.24	51.05	84.80	53.5
Average numbers of students per class, in the school	20.32	3.95	21.59	3.23	21.53	2.9
Private school	0.07		0.12		0.18	
Istituto comprensivo	0.39		0.29		0.24	

Notes:

Province: Students at grade 5 (last year of primary school) who took the INVALSI test in 2011/12 and live in the area around the city considered.

Inner-city: Students of grade 5 who took the INVALSI test in 2011/12 and live in the city.

Inner-city after matching: Students who took both the INVALSI test in grade 5 in 2011/12 and in grade 6 in 2012/13 (whose data were available) and live in the city.

Apart from enhancing the previous consideration about the number of immigrant students and private schools in inner-cities, It stands out a better socioeconomic condition in towns than in provinces. The existing different condition between provinces and inner-cities is demonstrated by a Student's t-test between means (table 6).

Table 6. t-test between province and inner-city

Student-level characteristics	BOLOGNA	MILAN	PADUA	ROME	TURIN
Achievement in Reading	0.9033	0.3745	0.0097	0	0.0621
Achievement in Mathematics	0	0.9345	0.1631	0	0.0009
Female Student	0	0	0	0	0
1st generation immigrant	0	0	0	0	0
2nd generation immigrant	0	0	0	0	0
Early-enrolled student	0.0095	0	0	0	0
Late-enrolled student	0	0	0	0	0
Socioeconomic background	0	0	0	0	0
Classroom- level characteristics					
Class average socioeconomic background (ESCS)	0	0	0	0	0
Proportion of female in the classroom	0	0.2381	0	0.0014	0
Proportion of 1st generation immigrants in the classroom	0	0	0	0.9303	0.4894
Proportion of 2nd generation immigrants in the classroom	0	0	0	0.9303	0.4894
Proportion of Early-enrolled students in the classroom	0	0	0	0	0
Proportion of Late-enrolled students in the classroom	0	0	0	0	0
Number of student in the classroom	0	0.4566	0	0.0008	0
School-level characteristics					
Number of classrooms in the school	0	0	0	0.2689	0
Number of students in the school	0	0	0	0.7155	0

Notes:
p-values reported. α=0.05

On the contrary, we would like to demonstrate the representativeness of the subsample that we obtain after matching dataset of inner-cities in grade 5 and 6 (named inner-cities after matching). In this case, we want to accept H_0 hypothesis to verify that, despite all the lost data, our subsample is statistically similar to the original dataset (table 7). The t-test show some differences in test scores (averagely higher in the subsample after matching) and in some classroom-level characteristics.

Table 7. t-test between inner-cities pre and after-matching.

Student-level characteristics	BOLOGNA	MILAN	PADUA	ROME	TURIN
Achievement in Reading	0.4245	0.0002	0.0249	0.004	0
Achievement in Mathematics	0.8898	0.0022	0.0152	0.32	0.0164
Female Student	0.694	0.609	0.462	0.194	0.603
1st generation immigrant	0.812	0.036	0.585	0.378	0.793
2nd generation immigrant	0.891	0.865	0.318	0.781	0.498
Early-enrolled student	0.288	0.407	0.61	0.427	0.823
Late-enrolled student	0.793	0.377	0.665	0.223	0.06
Socioeconomic background	0.3586	0	0.1412	0	0
Classroom- level characteristics					
Class average socioeconomic background (ESCS)	0.1314	0	0.0305	0	0
Proportion of female in the classroom	0.425	0	0	0.0069	0.7089
Proportion of 1st generation immigrants in the classroom	0	0	0	0	0
Proportion of 2nd generation immigrants in the classroom	0	0	0	0	0
Proportion of Early-enrolled students in the classroom	0	0.0004	0.0355	0.0081	0.0114
Proportion of Late-enrolled students in the classroom	0.7384	0.5495	0.055	0.8052	0
Number of student in the classroom	0.0103	0	0	0.0068	0.4267
School-level characteristics					
Number of classrooms in the school	0.5714	0	0.001	0.4113	0.8251
Number of students in the school	0.0009	0.4473	0.145	0.6349	0.625

Notes:
p-values reported. α=0.05

As mentioned above, in the matching procedure between grade 5 and 6, we lose 40-60% of original data because of mismatching problem in data provided by the Ministry. Table 8 contains the sample size for each city, comparing provinces and inner-cities. The last column contains the dimension of the sample of students who took the test in both grade 5 and 6 (after the matching procedure) and live in the same city for both years, who are our starting sample.

Table 8. Comparison between sample sizes.

	Province	Inner-city	Inner-city after matching
Bologna	6,927	2,153	1,229 (60.2%)
Milan	30,193	8,987	5,390 (62.4%)
Padua	7,320	1,621	828 (51.1%)
Rome	29,952	19,242	7,952 (42%)
Turin	16,206	5,705	2,562 (44.9%)

Notes:
Province: Students at grade 5 (last year of primary school) who took the INVALSI test in 2011/12 and live in the area around the city considered.
Inner-city: Students of grade 5 who took the INVALSI test in 2011/12 and live in the city.
Inner-city after matching: Students who took both the INVALSI test in grade 5 in 2011/12 and in grade 6 in 2012/13 (whose data were available) and live in the city both 2011/2012 and 2012/2103. Percentage within brackets refer to the proportion of data available in respect to the total population.

1.4.2 Preliminary observations about data

1.4.2.1 Self-Selection

In statistics, self-selection is commonly used to describe situations where individuals select themselves into a group. This means that observed relationship are more likely to be endogenous outcomes of an optimization problem rather than an exogenous causal relationship. Families, making decisions about the school for their children, affect the dislocation of students among schools (with potential self-selection). For this reason, we cannot treat any of the estimates as causal, but only as correlational. In the implementation of Propensity Score Matching approach (see section 3.4.2), we try to make our best to control for as many observable factors as possible, but we cannot assure that we completely identified the determinants of self-sorting. Therefore, albeit we cannot eliminate the endogenous component, we can partially deal with it. In fact, observing students' distribution in grade 5 and 6, we are able to catch the observable part of self-selection, such as what is related to school's characteristics or socioeconomic segmentation across schools or classes. Thus, despite unobservable part of self-selection lasts, we can partly reduce the impact of this phenomenon controlling for the observable part of it.

1.4.2.2 Missing Data

The second issue regards missing data, which means that some variables don't have a measurement. High proportion of missing data can be a serious problem in terms of loss of information and statistical power. Reading and mathematics tests are taken in two different days, so that some students took only one of them, generating a loss of data corresponding to 2% of the sample of grade 5 and 1% of the sample of grade 6. Moreover, having to deal with students who took the test both in grade 5 and 6 we lose nearly 40-60% of students' sample because of mismatch in data provided by INVALSI, which is the most serious problem of missing data we have. The administrative process wants the data to be sent by schools to the Ministry, which then matches student's codes between one year and another. Unfortunately, this procedure encounters some technical problems that cause a significant loss of information. Nevertheless, comparing variables before and after matching through a t-test on means, we can consider the subsample as representative.

Chapter 2

The determinants of students' achievement

Before focusing on the analysis of resilient students, we want to preliminary investigate factors that affect students' achievement overall. We concentrate our attention on the sample of students living in the five cities mentioned above. Implementing regression models, we study which characteristics (at individual, class and school level) are correlated with the performance attained in INVALSI test. Literature show a lack of consensus about schooling inputs that affect students' performance (see Hanushek, 1998, Krueger 1998, 2000). We follow the stream of literature that studies the determinants of cognitive achievement using an educational production function (EPF), examining the productivity relationship between schooling inputs and test score outcomes for school-age children (Todd & Wolpin, 2001). Following this approach, we want to understand how the technology, represented by school inputs, can be combined in order to obtain achievement outcomes.

2.1 The Educational Production Function

Having to deal with unobservable variables such as family and school inputs until grade 5 and student's endowment, we make an essential assumption: let t=0 correspond to the period of school attended prior to grade 5 and t=1 correspond to grade 5.

Student's achievement at grade 5 (t=1 time interval), represented by Y_1, depends on school inputs received until t=1 (A_0 and A_1), personal and family's characteristics -like the socioeconomic condition- until t=1 (X_o and X_1) and student's endowment (μ), which is innate:

$$\boldsymbol{Y_1 = g_1(A_0, A_1, X_0, X_1, \mu)} \quad (1)$$

In other words, we are assuming that the achievement in grade 5 contains all the historical information needed to estimate student's achievement in grade 6. Similarly, let t=2 be the period of time corresponding to grade 6. Thus student's outcome in grade 6 depends on student's historical inputs, held in his/her prior achievement (Y_1) and on school and family inputs at t=2 (respectively A_2 and X_2). We use a vector of individual characteristics and the ESCS index as a proxy for family inputs:

$$\boldsymbol{Y_2 = g_2(Y_1, A_2, X_2, \mu)} \qquad (2)$$

Specifically, we use the following formulation of the EPF:

$$\boldsymbol{y_{iw(t)} = \alpha_0 + \alpha_1 y_{iw(t-1)} + \propto_2 \bar{X}_{2iw(t)} + \alpha_3 \bar{X}_{3w(t)} + \varepsilon_{ijw}} \qquad (3)$$

Where y_{ijw} is the performance of the ith student, in the wth school at time t; $y_{iw(t-1)}$ is the pupil's prior achievement that depends on all the variables listed in (1); $\bar{X}_2$ is the vector of student's characteristics; $\bar{X}_3$ describe school-level features at time t. ε_{ijw} is the robust standard error clustered at school level.

Results in table 9 are reported for each city, distinguishing the dependent variable for reading and math score. The last column contains results for the sample of cities considered globally. Among student-level characteristics, there is a strong and positive relationship between the outcome in grade 6 and the one in grade 5, as well as with the socioeconomic background of the student. This is consistent with the assumption that prior achievement includes the endowment and mental capacity of the student (which is innate and so related to the outcome in grade 6 as well). Similarly, the correlation with the socioeconomic background is consistent with most of the literature on this topic. When accounting for gender, being female has a positive correlation with reading test score and a negative one with mathematics, suggesting a natural propensity of females for language skills. Moreover, immigrant students are less likely to obtain good academic results, especially in reading test, highlighting the greater difficulties immigrant students have in language proficiency. Finally, being late-enrolled is not always significant, but generally negatively related to achievement. Turning to school-level characteristics, both the school average socioeconomic condition and test score are strongly related to the outcome, but in different ways. In fact, we can observe a positive relationship with the average test score, suggesting that a peer effect may occur, but a negative correlation with the average socioeconomic background. Furthermore, even though immigrant students are less likely to be high achieving, the number of immigrant students at school is positively related to achievement when it's significant, except for the city of Padua. Finally, attending a private school or an *Istituto comprensivo* have a negative correlation with student's academic outcome.

Table 9. Results from the Educational Production Function.

Dependent variable: TEST GRADE 6	**BOLOGNA**		**MILAN**		**PADUA**		**ROME**		**TURIN**		**ENTIRE SAMPLE**	
Indipendent variables	**ITA**	**MAT**	**ITA**	**MAT**	**ITA**	**MAT**	**ITA**	**MAT**	**ITA**	**MAT**	**ITA**	**MAT**
Student-level characteristics												
Prior achievement (grade 5)	**0.491*****	**0.457*****	**0.625*****	**0.682*****	**0.589*****	**0.356*****	**0.516*****	**0.315*****	**0.565*****	**0.622*****	**0.553*****	**0.458*****
	(0.058)	(0.054)	(0.015)	(0.014)	(0.054)	(0.040)	(0.018)	(0.017)	(0.025)	(0.042)	(0.013)	(0.017)
Female Student	**0.043*****	**-0.123*****	**0.019****	-0.012	**0.070*****	-0.059	**0.027*****	**-0.074*****	**0.030*****	**-0.048*****	**0.028*****	**-0.065*****
	(1.595)	(2.072)	(0.758)	(0.748)	(1.704)	(2.817)	(0.641)	(0.813)	(0.861)	(0.947)	(0.409)	(0.536)
1st generation immigrant	**-0.059*****	-0.028	**-0.029*****	**-0.021*****	-0.031	-0.057	-0.005	-0.007	**-0.039*****	0.016	**-0.026*****	**-0.018*****
	(2.964)	(4.197)	(2.047)	(1.802)	(4.455)	(9.754)	(2.080)	(2.192)	(1.443)	(1.900)	(1.102)	(1.297)
2nd generation immigrant	**-0.044*****	-0.026	**-0.036*****	-0.012	-0.013	-0.031	**-0.028*****	-0.015	-0.029	-0.009	**-0.034*****	**-0.023*****
	(2.422)	(3.151)	(1.160)	(0.987)	(3.875)	(7.232)	(1.466)	(1.652)	(2.353)	(1.713)	(0.809)	(0.882)
Early-enrolled student	0.020	0.000	-0.01	0.0002	-0.022	-0.033	0.006	0.005	-0.002	0.012	0.002	0.002
	(6.216)	(8.838)	(4.292)	(2.852)	(6.975)	(11.709)	(2.024)	(2.217)	(3.880)	(7.449)	(1.720)	(1.864)
Late-enrolled student	-0.043	-0.013	-0.002	-0.0009	-0.014	-0.023	**-0.021****	-0.018	**-0.037*****	**-0.046*****	**-0.018*****	**-0.021*****
	(4.430)	(5.948)	(2.094)	(2.347)	(8.071)	(7.592)	(2.287)	(2.448)	(3.060)	(3.200)	(1.345)	(1.515)
Socioeconomic background	**0.195*****	**0.228*****	**0.123*****	**0.086*****	**0.132*****	**0.221*****	**0.142*****	**0.180*****	**0.145*****	**0.131*****	**0.145*****	**0.169*****
	(1.289)	(1.254)	(0.537)	(0.505)	(0.901)	(0.911)	(0.475)	(0.630)	(0.680)	(0.777)	(0.309)	(0.363)
Student who has siblings	0.041	**0.058*****	0.002	**0.020*****	**0.085*****	**0.070***	0.005	0.002	-0.003	0.004	0.008	**0.012****
	(2.033)	(1.694)	(0.760)	(0.781)	(2.635)	(3.692)	(0.742)	(0.968)	(1.237)	(1.259)	(0.484)	(0.620)
School-level characteristics												
School average socio-economic background (ESCS)	**-0.165*****	**-0.114*****	**-0.062*****	**-0.074*****	**-0.134*****	**-0.102*****	**-0.109*****	**-0.109*****	**-0.078*****	-0.044	**-0.102*****	**-0.116*****
	(2.550)	(2.937)	(1.344)	(1.195)	(3.760)	(2.329)	(1.097)	(1.579)	(2.223)	(2.669)	(0.705)	(1.105)
Average test score in the school	**0.278*****	**0.362*****	**0.168*****	**0.155*****	**0.111*****	**0.108*****	**0.229*****	**0.234*****	**0.251*****	**0.131*****	**0.204*****	**0.192*****
	(0.095)	(0.163)	(0.073)	(0.074)	(0.219)	(0.116)	(0.044)	(0.064)	(0.179)	(0.136)	(0.039)	(0.050)
Immigrant students per School	**0.096*****	**0.215*****	**0.041****	0.018	-0.002	**-0.040*****	0.015	0.004	**0.084*****	-0.004	**0.021****	-0.01
	(0.095)	(0.198)	(0.039)	(0.050)	(0.086)	(0.061)	(0.038)	(0.059)	(0.062)	(0.044)	(0.024)	(0.030)
Number of classrooms in the school	**-0.152*****	**-0.479*****	0.018	-0.067	0.001	-0.035	-0.003	-0.066	0.153	0.012	0.032	-0.026
	(1.309)	(2.993)	(1.008)	(1.195)	(2.529)	(3.241)	(0.557)	(0.747)	(1.043)	(1.108)	(0.392)	(0.631)

Number of students in the school	0.07	0.262	-0.056	0.028	-0.025	-0.012	0.007	0.028	-0.144	0.016	-0.026	0.023
	(0.052)	(0.124)	(0.034)	(0.041)	(0.073)	(0.110)	(0.021)	(0.026)	(0.049)	(0.046)	(0.015)	(0.023)
Private school	**-0.110***	**-0.184***	-0.006	-0.014		**-0.237***	**-0.054***	**-0.124***	0.041	0.012	0.015	**-0.060***
	(3.708)	(8.227)	(2.351)	(2.318)		(3.698)	(1.166)	(1.512)	(3.575)	(3.842)	(1.463)	(2.186)
Istituto Comprensivo	**-0.115***	0.0223	0.014	-0.001	**-0.124***		**-0.066***	**-0.038***	0.022	-0.019	0.01	0.018
	(2.755)	(8.466)	(1.359)	(1.552)	(4.340)		(1.427)	(2.292)	(2.042)	(2.664)	(0.950)	(1.315)
Cons.	-141.23	-216.99	-96.55	-104.47	-93.54	-106.16	-140.61	-139.75	-149.30	-84.86	-124.81	-119.51
	(18.54)	(30.36)	(15.21)	(14.93)	(47.79)	(26.89)	(9.29)	(13.85)	(36.17)	(29.09)	(8.40)	(10.60)
N	1.229	1.229	5.390	5.390	828	828	7.952	7.952	2562	2562	17.961	17.961
Adjusted R^2	0.45	0.32	0.52	0.55	0.45	0.2	0.39	0.21	0.49	0.51	0.45	0.32

NOTES: *** is statistically significant at 1%; ** 5%, * 10%. Robust standard errors in parentheses are clustered at school level.
Omitted variables for the city of Padua are due to collinearitity.

2.2 Value Added Specifications

Value-added models are a class of statistical procedures that use longitudinal test score data to measure the extent to which a student has improved in a specific period of time. As Doran and Izumi (2004) state: "VAMs are an attempt to determine *how much value has a school added to a student's learning*?" The most influential value-added model is the Tennessee Value Added Assessment System (TVAAS) developed in late 1980's by Dr. William L. Sanders as a tool to measure teacher's effectiveness. Since then It has been implemented by many school districts across the USA and investigated in many researches (Corcoran, 2101, Rothstein, 2010, Todd & Wolpin, 2006) despite some caveats (Kuppermintz, 2003). One of the main issues related to this model concerns the real ability to isolate the "effect" of the school from other non-school factors (Doran & Izumi, 2004). In our model, we maintain student-level characteristics fixed, and consider how the variation in class and school-level factors affect student's performance. In the first model (4), we only consider individual characteristics in order to create a baseline model to which time-variant variables are added subsequently. Thus, we consider individual characteristics as time invariant. Let *t* correspond to grade 6 and *t – 1* correspond to grade 5.

The baseline equation is the following:

$$\Delta \boldsymbol{y}_{\boldsymbol{i}(\boldsymbol{t},\boldsymbol{t}-\mathbf{1})} = \boldsymbol{\alpha}_{\mathbf{0}} + \boldsymbol{\alpha}_{\mathbf{1}} \overline{\boldsymbol{X}}_{\mathbf{1}\boldsymbol{i}(\boldsymbol{t}-\mathbf{1})} + \boldsymbol{\varepsilon}_{\boldsymbol{i}} \tag{4}$$

Where $\Delta \boldsymbol{y}_{\boldsymbol{i}(\boldsymbol{t},\boldsymbol{t}-\mathbf{1})}$ is the variation in the performance of ith student between grade 5 and 6, $\bar{X}_{1i}$ is the vector of student's characteristics in grade 5 and ε_i is the robust standard error.

Secondly, we enrich the model with two more vectors related to classroom's characteristics:

$$\Delta \boldsymbol{y}_{\boldsymbol{ij}(\boldsymbol{t},\boldsymbol{t}-\mathbf{1})} = \boldsymbol{\alpha}_{\mathbf{0}} + \boldsymbol{\alpha}_{\mathbf{1}} \overline{\boldsymbol{X}}_{\mathbf{1}\boldsymbol{ij}(\boldsymbol{t}-\mathbf{1})} + \boldsymbol{\alpha}_{\mathbf{2}} \overline{\boldsymbol{X}}_{\mathbf{2}\boldsymbol{ij}(\boldsymbol{t}-\mathbf{1})} + \boldsymbol{\alpha}_{\mathbf{3}} \Delta \overline{\mathbf{X}}_{\mathbf{3ij(t,t-1)}} + \boldsymbol{\alpha}_{\mathbf{4}} \Delta \mathbf{ESCS}_{\mathbf{ij(t,t-1)}} + \boldsymbol{\varepsilon}_{\boldsymbol{ij}} \tag{5}$$

Where $\Delta y_{ij(t,t-1)}$ is the variation in the achievement of ith student in jth classroom, $\bar{X}_1$ refers to individual characteristics, $\bar{X}_2$ is a vector of variables describing initial classroom's factors (grade 5) and $\Delta\overline{\mathrm{X}}_3$ refers to the variation in class-level characteristics. $\Delta\mathrm{ESCS}_{\mathrm{ij}}$ catches the variation in student's socio-economic background (the main time variant individual factor) and ε_{ij} is robust standard error clustered at school level, in order to control for the within-school correlation between school factors. Moreover, we use school fixed effects to catch the impact of missing school-level data on the output, under the assumption that students in the same school receive the same level of inputs.

Finally, we consider school-level variations through the following equation:

$$\Delta \boldsymbol{y}_{\boldsymbol{iw}(\boldsymbol{t},\boldsymbol{t}-\mathbf{1})} = \boldsymbol{\alpha}_{\mathbf{0}} + \boldsymbol{\alpha}_{\mathbf{1}}\overline{\boldsymbol{X}}_{\mathbf{1}\boldsymbol{iw}(\boldsymbol{t}-\mathbf{1})} + \boldsymbol{\alpha}_{\mathbf{2}}\overline{\boldsymbol{X}}_{\mathbf{2}\boldsymbol{iw}(\boldsymbol{t}-\mathbf{1})} + \boldsymbol{\alpha}_{\mathbf{3}}\Delta\overline{\boldsymbol{X}}_{\mathbf{3}\boldsymbol{iw}(\boldsymbol{t},\boldsymbol{t}-\mathbf{1})} + \boldsymbol{\alpha}_{\mathbf{4}}\Delta\mathbf{ESCS}_{\mathbf{iw(t,t-1)}} + \boldsymbol{\varepsilon}_{\boldsymbol{iw}} \tag{6}$$

Where $\Delta y_{iw(t,t-1)}$ is the variation in the achievement of ith student in wth school, $\bar{X}_1$ refers to individual characteristics, $\bar{X}_2$ is a vector of variables describing initial school's factors (grade 5), $\Delta\overline{\mathrm{X}}_3$ refers to the variation in school-level characteristics. When controlling for these variations, we also take into account if the student is switching from a private to a public school (and *vice versa*) or from/towards an *Istituto Comprensivo.* Finally, $\Delta\mathrm{ESCS}_{\mathrm{iw}}$ catches the variation in student's socioeconomic background and ε_{ij} is the robust standard error clustered at school-level. We describe 6 results: one for each city and the last concerning the five cities jointly.

Results from model (4) reported in table 10 show that individual characteristics do matter, even though the proportion of variance explained varies from 14.5% of Rome to 53% of Milan. Generally, higher socio-economic condition is related to positive score variations, while higher score in grade 5 is related to smaller increase the next year (as It is rational to think since delta score are obtained subtracting 5 grade scores from 6 grade ones). Moreover, consistently with international studies, female students are more likely to obtain better results in Italian and worse outcomes in mathematics. Similarly, being immigrant student has a negative correlation with score variation, independently from the status (first or second generation immigrant). Finally, being a late-enrolled student is generally related to a decrease in student's outcome.

Analyzing the second model (5), whose results are reported in table 11, we find similarities with the previous model when considering the significance of the socioeconomic background and the prior achievement, which are strongly related to the output, and of the female status (positively related to reading score's variation and negatively to mathematics'). Among classroom-level characteristics in grade 5, the most predictive are the class average socioeconomic background (with a negative impact on scores variation) and the class average score (with a positive one). Likewise are *delta* predictors; among them, the variation in the class average score recurs as the most significant in terms of both statistical relevance and absolute value of the coefficient. The positive impact the variable has on the outcome can be only compared (in its absolute value) with the estimate of prior achievement, and it has a greater impact on the mathematics' score. Nonetheless, each delta variable has a high level of significance showing how the variation in these characteristics can affect the outcome's variation. More in detail, the variation in the class average socioeconomic background impact negatively on the output, as well as the increase in the number of students per class (consistently with part of literature, e.g. Krueger, 1999). Finally, a higher number of immigrant students in grade 6 is related to better outcomes. Despite the literature generally agrees about the negative impact of immigrant students on peers' attainment in Italy, some researches provide evidences of a different tendency. Ballatore, Fort & Ichino (2013) analyze, in the Italian context, the effect of increasing the number of immigrant students keeping class size constant and find that the negative impact on peers' performance vanish when students attend grade 5. It could actually be the case that the effect has disappeared by the time the observation was conducted. The adjusted R squared increases (when compared to results from model 4) variably from city to city. The increase in the portion of variance explained is particularly significant for the city of Bologna (from 21 to 55%) and Rome (from 14.5 to 58%).

When analyzing the correlation between school-level characteristics and the variation in students' outcomes, results in table 12 are consistent with what we described at class-level. In particular, among individual variables, the previous score (grade 5) and the socioeconomic background (in terms of both initial condition in grade 5 and variation between the two years) are the most predictive factors. Among *delta* variables, the variation in the socioeconomic background of school peers is negatively related to higher achievement in grade 6, whereas an opposite effect is recorded for an average improvement, in grade 6, in peers achievement. Once again, Bologna and Rome register the highest gain in adjusted R squared (respectively 50 and 54%). In conclusion, *delta* variables have a relevant impact on score variation, especially when considering the impact of peers both at class and school-level.

Table 10. Value Added Model, Student-level characteristics.
NOTES: *** is statistically significant at 1%; ** 5%, * 10%. Robust standard errors in parentheses.

	BOLOGNA		MILANO		PADOVA		ROMA		TORINO		TOTALE	
student level characteristic	ITA	MAT	ITA	MAT	ITA	MAT	ITA	MAT	ITA	MAT	ITA	MAT
Socioeconomic background	**0.152*** **	**0.165*** **	**0.158*** **	**0.130*** **	**0.140*** **	**0.093*** **	**0.156*** **	**0.119*** **	**0.206*** **	**0.181*** **	**0.1647*** **	**0.132*** **
	(1.042)	(1.798)	(0.506)	(0.478)	(1.326)	(1.413)	(0.505)	(0.659)	(0.916)	(0.048)	(0.308)	(0.401)
Score (grade 5)	**-0.709*** **	**-0.639*** **	**-0.489*** **	**-0.380*** **	**-0.537*** **	**-0.696*** **	**-0.545*** **	**-0.670*** **	**-0.493*** **	**-0.427*** **	**-0.537*** **	**-0.579*** **
	(0.0584)	(0.0443)	(0.0167)	(0.0153)	(0.059)	(0.023)	(0.017)	(0.019)	(0.0249)	(0.9150)	(0.013)	(0.0178)
Female student	**0.0453** **	**-0.115*** **	**0.0242** **	-0.018	**0.098*** **	-0.042	**0.027*** **	**-0.063*** **	**0.0366** **	**-0.0630*** **	**0.0326*** **	**-0.064*** **
	(1.750)	(2.154)	(0.772)	(0.730)	(1.975)	(2.411)	(0.657)	(0.799)	(1.087)	(1.147)	(0.422)	(0.527)
Fist generation immigrant	**-0.080*** **	-0.042	**-0.051*** **	**-0.041*** **	-0.056	-0.048	**-0.019***	**-0.015***	**-0.0564*** **	0.006	**0.0424*** **	**-0.0263*** **
	(3.300)	(4.39)	(1.942)	(1.683)	(4.221)	(8.213)	(2.334)	(2.510)	(1.893)	(2.010)	(1.131)	(1.330)
Second generation immigrant	**-0.0496** **	-0.022	**-0.0593*** **	**-0.027** **	-0.040	**-0.052***	**-0.038*** **	**-0.0149***	**-0.052** **	**-0.030***	**-0.048*** **	**-0.0278*** **
	(2.062)	(3.513)	(1.194)	(0.946)	(4.277)	(5.595)	(1.712)	(1.767)	(2.177)	(1.865)	(0.865)	(0.869)
Early enrolled student	0.0245	-0.0118	-0.006	0.004	**-0.0162***	-0.010	0.010	0.002	-0.011	0.014	0.003	-0.003
	(7.641)	(6.718)	(4.320)	(3.007)	(3.439)	(8.877)	(1.982)	(2.219)	(5.295)	(6.400)	(1.699)	(1.931)
Late enrolled students	**-0.044***	-0.024	-0.011	-0.009	0.002	-0.026	**-0.022** **	**-0.018***	**-0.038** **	**-0.0625** **	**-0.023*** **	**-0.026*** **
	(4.861485)	(7.077016)	(2.075)	(2.373)	(6.121)	(7.132)	(2.425)	(2.659)	(3.051)	(3.661)	(1.368)	(1.585)
Student who has siblings	0.0294	**0.0483** **	0.003	**0.022***	**0.086*** **	**0.061*** **	-0.002	-0.007	**-0.028***	-0.009	0.004	0.007
	(2.314)	(1.949)	(0.768)	(0.801)	(2.260)	(2.387)	(0.814)	(1.128)	(1.299)	(1.311)	(0.510)	(0.684)
Cons.	-5.94	-2.36	-2.24	-1.98	-10.48	-4.77	-3.31	0.72	-0.33	1.21	-3.25	-0.013
	(2.38)	(2.98)	(0.92)	(1.05)	(1.96)	(3.28)	(1.01)	(1.38)	(1.50)	(1.36)	(0.61)	(0.86)
adj. R^2	0.402	0.210	0.507	0.534	0.271	0.431	0.343	0.145	0.217	0.171	0.430	0.294

Table 11. Value Added Model, Student and Classroom-level characteristics.

	BOLOGNA		MILAN		PADUA		ROME		TURIN		ENTIRE SAMPLE	
	N=1,199		N=5,373		N=825		N=7,898		N=2,559		N=17,854	
Variable												
student level characteristic												
Socioeconomic background	**0.138*** **	**0.224*** **	**0.156*** **	**0.103*** **	**0.099*** **	**0.202*** **	**0.148*** **	**0.137*** **	**0.179*** **	**0.147*** **	**0.161*** **	**0.162*** **
	(1.182)	(1.22)	(0.500)	(0.479)	(1.298)	(1.60)	(0.476)	(0.516)	(0.745)	(0.706)	(0.294)	(0.310)
Score (grade 5)	**-0.641*** **	**-0.323*** **	**-0.491*** **	**-0.314*** **	**-0.503***	**-0.445*** **	**-0.523*** **	**-0.373*** **	**-0.493*** **	**-0.296*** **	**-0.519*** **	**-0.389*** **
	(0.021)	(0.034)	(0.017)	(0.010)	(0.027)	(0.057)	(0.009)	(0.012)	(0.015)	(0.016)	(0.005)	(0.007)
Female student	**0.037****	**-0.095*** **	**0.020***	-0.015	**0.074*** **	**-0.060****	**0.024*** **	**-0.052*** **	**0.034*** **	**-0.043*** **	**0.027*** **	**-0.050*** **
	(1.682)	(1.79)	(0.688)	(0.676)	(1.83)	(2.064)	(0.628)	(0.720)	(0.851)	(0.857)	(0.381)	(0.464)
Fist generation immigrant	**-0.055*** **	-0.0005	**-0.039*** **	-0.017	-0.054	**-0.064****	-0.003	-0.003	-0.031	**0.037*** **	**-0.024*** **	**-0.011*** **
	(3.648)	(3.91)	(1.860)	(1.744)	(4.718)	(9.156)	(2.225)	(2.015)	(1.820)	(1.914)	(1.093)	(1.095)
Second generation immigrant	**-0.043*** **	-0.002	**-0.037*** **	-0.015	-0.013	-0.029	**-0.021*** **	-0.001	-0.024	0.003	**-0.027*** **	-0.005
	(3.247)	(3.51)	(1.16)	(1.11)	(3.763)	(5.465)	(1.40)	(1.53)	(2.218)	(1.725)	(0.766)	(0.820)
Early enrolled student	0.02	-0.005	-0.004	0.008	-0.015	-0.029	0.008	**0.012***	-0.004	0.01	0.004	0.008
	(8.67)	(3.91)	(3.73)	(3.058)	(10.86)	(13.70)	(1.965)	(2.21)	(4.900)	(7.49)	(1.637)	(1.646)
Late enrolled students	-0.031	-0.01	-0.012	-0.014	-0.018	-0.026	-0.016	**-0.015***	**-0.036***	**-0.048*** **	**-0.018*** **	**-0.018*** **
	(4.307)	(4.66)	(2.139)	(2.072)	(5.56)	(7.240)	(2.156)	(2.339)	(3.392)	(2.949)	(1.259)	(1.31)
Student who has siblings	0.024	**0.031***	-0.0003	0.021	**0.080*** **	**0.043***	-0.002	0.003	-0.014	0.012	0.002	**0.013****
	(1.804)	(1.689)	(0.747)	(0.796)	(2.16)	(3.228)	(0.778)	(1.62)	(1.28)	(1.22)	(0.473)	(0.497)
Delta student's ESCS	**0.123*** **	**0.088*** **	**0.085*** **	**0.026*** **	**0.075*** **	**0.114*** **	**0.093*** **	**0.078*** **	**0.080*** **	**0.077*** **	**0.093*** **	**0.078*** **
	(1.33)	(1.43)	(0.60)	(0.584)	(1.529)	(1.92)	(0.515)	(0.560)	(0.928)	(0.883)	(0.339)	(0.355)
Classroom- level characteristics												
Class-average socioeconomic background	**-0.079*** **	**-0.11*** **	**-0.089*** **	**-0.027*** **	**-0.103*** **	**-0.218*** **	-0.011	**-0.038*** **	-0.049	0.009	**-0.052*** **	**-0.123*** **
	(4.062)	(4.39)	(0.047)	(1.712)	(4.83)	(5.83)	(1.50)	(1.62)	(2.52)	(2.42)	(0.964)	(0.559)

Class-average score	**0.177*****	**0.288*****	**0.263*****	**0.195*****	**0.201*****	**0.426*****	**0.199*****	**0.231*****	**0.137*****	-0.015	**0.193*****	**0.256*****
	(0.108)	(0.104)	(0.047)	(0.044)	(0.144)	(0.149)	(0.037)	(0.038)	(0.069)	(0.062)	(0.025)	(0.017)
Immigrants in the classroom	**-0.071*****	-0.049	-0.0002	-0.003	-0.016	0.031	0.01	0.007	-0.055	**-0.079****	-0.015	0.013
	(0.750)	(0.736)	(0.260)	(0.246)	(0.822)	(1.056)	(0.315)	(0.341)	(0.465)	(0.435)	(0.122)	(0.110)
Females in the classroom	-0.016	**-0.046***	0.004	-0.01	-0.007	0.077	**0.023***	0.008	-0.004	-0.005	0.012	0.002
	(0.400)	(0.426)	(0.195)	(0.187)	(0.614)	(0.751)	(0.166)	(0.181)	(0.324)	(0.308)	(0.100)	(0.130)
Early enrolled in the classroom	**0.029***	0.017	**-0.02***	-0.019	-0.002	0.024	-0.002	-0.008	-0.01	**-0.034*****	-0.009	**-0.002*****
	(2.44)	(2.841)	(0.959)	(0.925)	(2.70)	(3.897)	(0.467)	(0.509)	(1.88)	(1.801)	(0.392)	(0.334)
Late enrolled in the classroom	-0.02	0.03	**0.034*****	**0.037*****	0.049	0.034	-0.003	0.012	-0.004	0.006	0.008	0.01
	(0.817)	(0.662)	(0.219)	(0.245)	(0.564)	(1.323)	(0.507)	(0.536)	(0.808)	(0.781)	(0.223)	(0.220)
Number of students in the classrooms	**0.103***	0.052	-0.06	**-0.068***	-0.104	-0.032	-0.026	**-0.042*****	-0.061	0.042	-0.024	-0.01
	(0.942)	(0.690)	(0.380)	(0.342)	(0.977)	(1.23)	(0.256)	(0.252)	(0.540)	(0.519)	(0.182)	(0.167)
Delta classroom level characteristics												
Delta class-average socioeconomic background	-0.036	-0.046	**-0.09*****	-0.014	-0.092	**-0.122*****	**-0.004*****	**-0.029***	**-0.086*****	**-0.073*****	**-0.073*****	**-0.084*****
	(3.24)	(3.58)	(1.56)	(1.51)	(4.19)	(4.93)	(1.24)	(1.34)	(1.89)	(1.799)	(0.811)	(0.510)
Delta class-average score	**0.385*****	**0.659*****	**0.420*****	**0.470*****	**0.428*****	**0.692*****	**0.441*****	**0.647*****	**0.333*****	**0.396*****	**0.416*****	**0.643*****
	(0.095)	(0.085)	(0.037)	(0.037)	(0.107)	(0.068)	(0.029)	(0.026)	(0.052)	(0.049)	(0.020)	(0.015)
Delta number of students in the classroom	-0.002	0.012	**-0.113*****	**-0.073*****	**-0.140*****	**-0.091*****	**-0.05*****	**-0.047*****	-0.074	-0.004	**-0.0740*****	**-0.039*****
	(0.676)	(0.504)	(0.259)	(0.191)	(0.513)	(0.641)	(0.172)	(0.187)	(0.411)	(0.397)	(0.127)	(0.071)
Delta number of immigrants in the classroom	**0.08****	**0.072***	**0.097*****	**0.066*****	0.046	0.02	**0.046*****	**0.035*****	**0.063*****	**0.047***	**0.069*****	**0.059*****
	(0.538)	(0.568)	(0.206)	(0.195)	(0.472)	(0.746)	(0.212)	(0.203)	(0.306)	(0.242)	(0.128)	(0.149)
Constant	-99.28	-121.16	-93.25	-65.06	-65.68	-142.87	-79.55	-57.11	-25.30	7.73	-71.58	-72.11
	(33.09)	(29.58)	(13.23)	(12.81)	(35.92)	(38.26)	(15.80)	(16.62)	(23.91)	(21.52)	(4.41)	(3.87)
adjusted R²	0.49	0.55	0.28	0.26	0.37	0.56	0.39	0.58	0.31	0.33	0.36	0.47

NOTES: *** is statistically significant at 1%; ** 5%, * 10%. Robust standard errors in parentheses are clustered at school level.

Table 12. Value Added Model, Student and School-level characteristics.

	BOLOGNA		MILAN		PADUA		ROME		TURIN		ENTIRE SAMPLE	
	N=1,199		N=5,373		N=825		N=7,898		N=2,559		N=17,854	
Variable												
student level characteristic												
Socioeconomic background	**0.216*** **	**0.222*** **	**0.173*** **	**0.140*** **	**0.134** **	**0.203*** **	**0.173*** **	**0.159*** **	**0.183*** **	**0.175*** **	**0.182*** **	**0.176*** **
	(1.152)	(1.231)	(0.490)	(0.46)	(1.26)	(1.217)	(0.465)	(0.687)	(0.682)	(0.690)	(0.325)	(0.304)
Score (grade 5)	**-0.719*** **	**-0.321*** **	**-0.501*** **	**-0.344*** **	**-0.522*** **	**-0.447*** **	**-0.550*** **	**-0.419*** **	**-0.496*** **	**-0.374*** **	**-0.546*** **	**-0.401*** **
	(0.020)	(0.040)	(0.015)	(0.013)	(0.026)	(0.048)	(0.018)	(0.0188)	(0.022)	(0.042)	(0.013)	(0.006)
Female student	**0.043** **	**-0.096*** **	**0.022* **	-0.013	**0.080** **	**-0.048** **	**0.026*** **	**-0.055*** **	**0.036** **	**-0.051*** **	**0.030*** **	**-0.052*** **
	(1.675)	(1.878)	(0.757)	(0.745)	(1.811)	(2.956)	(0.638)	(0.714)	(1.053)	(1.004)	(0.407)	(0.481)
Fist generation immigrant	**-0.053** **	-0.026	**-0.032** **	-0.021	-0.039	**-0.056** **	-0.002	-0.008	**-0.036** **	0.024	**-0.025*** **	**-0.017*** **
	(3.726)	(4.097)	(1.960)	(1.782)	(3.858)	(7.470)	(2.117)	(2.125)	(1.639)	(1.933)	(1.081)	(1.266)
Second generation immigrant	**-0.036** **	-0.015	**-0.04*** **	-0.016	-0.024	-0.028	**-0.024*** **	-0.006	-0.020	-0.003	**-0.033*** **	**-0.008** **
	(3.330)	(2.638)	(1.122)	(0.889)	(4.571)	(5.072)	(1.479)	(1.497)	(2.208)	(1.88)	(0.797)	(0.177)
Early enrolled student	0.020	-0.001	-0.016	-0.001	-0.017	-0.024	0.005	0.007	0.001	0.013	0.002	0.006
	(8.70)	(9.47)	(3.74)	(2.645)	(10.72)	(11.39)	(2.011)	(2.126)	(4.587)	(6.085)	(1.701)	(1.737)
Late enrolled students	-0.036	0.009	-0.001	-0.004	-0.010	-0.015	**-0.022* **	-0.010	**-0.044** **	**-0.055** **	**-0.019*** **	**-0.015** **
	(4.378)	(4.74)	(2.122)	(2.353)	(5.42)	(6.880)	(2.220)	(2.413)	(3.087)	(3.021)	(1.340)	(1.466)
Student who has siblings	**0.043* **	**0.051*** **	0.002	**0.023** **	**0.095** **	**0.049** **	0.001	0.001	-0.003	0.014	0.007	**0.013** **
	(1.861)	(1.492)	(0.797)	(0.752)	(2.20)	(2.611)	(0.773)	(0.819)	(1.396)	(1.27)	(0.499)	(0.517)
Delta student's ESCS	**0.126*** **	**0.111*** **	**0.088*** **	**0.035** **	**0.106*** **	**0.111*** **	**0.101*** **	**0.079*** **	**0.068*** **	**0.084*** **	**0.099*** **	**0.080*** **
	(1.874)	(1.391)	(0.613)	(0.556)	(1.393)	(1.91)	(0.507)	(0.687)	(0.899)	(0.880)	(0.37)	(0.357)
School-level characteristics												
School-average socioeconomic background	**-0.182*** **	**-0.124** **	**-0.071*** **	**-0.106*** **	**-0.237*** **	**-0.171** **	**-0.076*** **	**-0.063*** **	**-0.077*** **	-0.054	**-0.091*** **	**-0.122*** **
	(3.54)	(3.92)	(1.08)	(0.992)	(4.86)	(5.68)	(0.984)	(1.041)	(2.144)	(1.99)	(0.627)	(0.645)

School-average score	**0.238*** **	**0.399** **	**0.112*** **	**0.115*** **	**0.113* **	**0.322* **	**0.148*** **	**0.258*** **	**0.102*** **	-0.016	**0.120*** **	**0.203*** **
	(0.139)	(0.200)	(0.056)	(0.049)	(0.229)	(0.263)	(0.040)	(0.045)	(0.104)	(0.082)	(0.028)	(0.024)
Number of classrooms in the school	-0.198	0.036	-0.006	-0.012	0.042	-0.087	0.034	0.004	**0.255** **	0.016	0.063	**0.082* **
	(3.349)	(3.68)	(0.774)	(0.747)	(3.012)	(3.64)	(0.63)	(0.661)	(1.284)	(1.165)	(0.407)	(0.435)
Number of students in the school	0.117	-0.104	0.027	0.022	-0.037	0.145	-0.049	-0.025	-0.123	0.092	-0.026	-0.036
	(0.128)	(0.137)	(0.030)	(0.028)	(0.099)	(0.123)	(0.024)	(0.025)	(0.049)	(0.046)	(0.019)	(0.025)
Private school	-0.080	-0.072	0.036	0.010	**0.270*** **	0.083	**-0.071*** **	**-0.107*** **	0.059	0.015	0.003	-0.037***
	(9.159)	(10.30)	(2.19)	(2.12)	(5.595)	(omitted)	(4.87)	(5.36)	(3.603)	3.55)	(1.28)	(2.141)
Istituto comprensivo	**-0.169* **	0.137	0.042	0.005	**-0.270*** **	-0.130	**-0.068**	-0.001	-0.030	-0.003	-0.016	0.003
	(7.44)	(1.078)	(1.58)	(1.51)	(omitted)	(12.561)	(4.93)	(5.65)	(3.66)	(0.068)	(0.921)	(1.486)
Delta school-level characteristics												
Delta school-average socioeconomic background	**-0.118*** **	**-0.059* **	**-0.067*** **	**-0.069*** **	**-0.227*** **	**-0.164*** **	**-0.096*** **	**-0.055*** **	**-0.042* **	**-0.058* **	**-0.090*** **	**-0.088*** **
	(2.933)	(3.267)	(1.05)	(0.974)	(4.76)	(5.98)	(0.923)	(1.006)	(1.870)	(1.73)	(0.595)	(0.626)
Delta school-average score	**0.286*** **	**0.629*** **	**0.217*** **	**0.341*** **	**0.267*** **	**0.637*** **	**0.298*** **	**0.533*** **	**0.291*** **	**0.288*** **	**0.258*** **	**0.491*** **
	(0.119)	(0.112)	(0.05)	(0.047)	(0.201)	(0.251)	(0.039)	(0.038)	(0.076)	(0.075)	(0.025)	(0.024)
Delta number of students in the school	0.120	-0.026	-0.104	0.008	-0.107	0.009	0.011	0.030	**-0.233* **	0.015	-0.043	-0.036
	(0.075)	(0.098)	(0.022)	(0.021)	(0.082)	(0.107)	(0.019)	(0.017)	(0.035)	(0.035)	(0.011)	(0.017)
Delta number of classrooms in the school	**-0.200** **	-0.067	0.008	-0.131	0.094	-0.024	-0.037	-0.052	**0.183* **	-0.038	0.021	0.035
	(2.193)	(2.38)	(0.613)	(0.047)	(2.925)	(3.739)	(0.529)	(0.48)	(0.912)	(0.702)	(0.31)	(0.461)
Delta number of immigrants in the school	0.064	0.058	**0.084*** **	**0.097*** **	-0.045	-0.006	0.019	-0.004	**0.138*** **	**0.072** **	**0.038*** **	**0.026*** **
	(0.139)	(0.139)	(0.038)	(0.035)	(0.135)	(0.223)	(0.041)	(0.043)	(0.098)	(0.048)	(0.028)	(0.031)
Switch from private to public school	**0.059**	0.049	0.014	0.008	-0.015	0.026	0.015	**0.038*** **	-0.016	-0.036	-0.004	0.009
	(8.51)	(9.16)	(2.38)	(2.310)	(6.08)	(6.558)	(4.82)	(5.31)	(3.95)	(3.85)	(1.33)	(1.425)
Switch from public to private school	**-0.113*** **	-0.060	**-0.030** **	**-0.041** **	0.016	-0.001	**-0.024**	**-0.035*** **	0.009	0.022	-0.016	-0.014
	(8.94)	(9.59)	(2.62)	(2.684)	(6.266)	(11.53)	(5.87)	(8.11)	(4.93)	(4.85)	(1.769)	(2.484)
Switch from an *Istituto comprensivo* to another school	0.035	0.003	0.009	0.026	0.016	0.005	0.005	0.013	0.001	-0.006	**-0.013* **	-0.004
	(9.020)	(9.63)	(1.613)	(1.587)	(omitted)	(omitted)	(5.760)	(8.11)	(4.557)	(4.688)	(1.264)	(1.587)
Switch from a school to an *Istituto comprensivo*	-0.079	-0.004	0.011	-0.009	-0.029	0.013	**-0.083**	**-0.107*** **	-0.012	0.008	-0.010	**-0.095*** **

	(6.17)	(7.686)	(1.640)	(1.62)	(omitted)	(omitted)	(1.376)	(1.456)	(6.132)	(2.630)	(1.077)	(1.218)
Same *Istituto comprensivo*	0.035	-0.042	-0.005	0.022	0.057	0.010	0.005	0.001	0.072	-0.009	0.014	-0.002
	(2.838)	(2.561)	(1.292)	(1.259)	(3.079)	(2.923)	(1.194)	(1.488)	(3.754)	(3.418)	(0.838)	(1.025)
Constant	-104.25	-120.47	-65.71	-54.00	-77.43	-127.28	-87.97	-77.30	-75.05	0.08	-73.79	-65.83
	(29.50)	(37.83)	(12.14)	(10.49)	(49.05)	(57.32)	(10.05)	(10.54)	(22.29)	(17.76)	(6.23)	(5.27)
adjusted R^2	0.447	0.503	0.230	0.192	0.310	0.497	0.336	0.539	0.278	0.250	0.302	0.428

NOTES: *** is statistically significant at 1%; ** 5%, * 10%. Robust standard errors in parentheses are clustered at school level.
Omitted variables for the city of Padua are due to collinearitity.

Chapter 3

Resilient students: the empirical analysis

3.1 Defining Resilient Students

After describing our reference group, we must define the subsample of students to be examined i.e. the "resilient students"; this concept, following OECD (2012), deals with the ability of disadvantaged students to "beat the odds" and – despite their disadvantaged background – obtain good academic results.

In order to identify resilient students, we start selecting disadvantaged students according to their socio-economic condition (ESCS index) in 5^{th} grade; more specifically, we analyze the ESCS index distribution for each city and define "Disadvantaged students" those whose index is below the 33th percentile of the city distribution. Once defined this group, we start again from the initial sample and, for each city, we take into consideration test scores at grade 5. In this case, we select students whose score is below the 33th percentile of the city's distribution in reading or mathematics test and define them "Low Achievers".

Finally we merge the two dataset obtaining the "Disadvantaged Low Achievers" (DLA) sample in 5^{th} grade, formed by those students characterized by jointly poor socioeconomic background and low performance (table 13).

Table 13. Disadvantaged Low Achiever students, sample size.

	Disadvantaged Low Achiever students
Bologna	252
Milan	1024
Padua	177
Rome	1833
Turin	512

Next step permits to select resilient students observing the same sample one year later, when students attend grade 6. According to previous explanations, we consider as "resilient" only students that, at grade 6, are still disadvantaged. This means that the

ESCS index has still to be in the bottom 33th percentile of city's distribution of the year 2012/2013. Therefore, resilient students are those who overtake the average score of the city's distribution in the Reading or Mathematics test and were low achievers in the same subject in grade 5 (table 14). We control for jointly these conditions in order to avoid considering resilient in mathematics those students who were selected among low achievers for their reading score. It is important to highlight that the distribution of test scores has a mean of 200 and a standard deviation of 40, but our cut point partially changes from city to city, according to the average score of the city. This way, we can take into account structural differences across cities.

Table 14. Resilient students, sample size.

	Resilient students in Reading	**Resilient students in Mathematics**
Bologna	18	20
Milan	50	49
Padua	2	19
Rome	93	174
Turin	22	16

After having selected and labelled resilient students, we create the control group (table 15). We consider the same group of "Disadvantaged low achiever students" (at grade 5) and we select those who perform under the average of the city in grade 6 both in reading and mathematics tests. Moreover, we select low achiever students whose ESCS index is in the bottom 33th percentile of the distribution, in order to consider the same socioeconomic background of resilient students and to have a control group corrected by the social class.

Table 15. Control group, sample size.

	Control Group
Bologna	143
Milan	603
Padua	79
Rome	656
Turin	277

3.2 Further Investigation About Data

Having defined resilient students and control group, we have to face two main issues:

- Regression to the mean
- *Antiresilient* students

3.2.1 Disadvantaged Low Achievers distribution at grade 6: Regression to the Mean

In order to compare students' results between the two years, we have to consider the difference between test scores. In order to make results more easily interpretable, we have created *relative scores*, calculated as follow:

$$R_{ij} = X_{ij} - E(X)_j \tag{7}$$

Where R_{ij} is the relative score of the ith student in the jth city, X_{ij} is the initial score and $E(X)_j$ is the average score of the city. Every student has two relative scores: one for the Reading test and one for the Mathematics test, considered separately.

Our analysis starts from students who were low achievers in grade 5 (bottom third of each city scores distribution). Taking into account the possibility that regression to the mean may occur, we study how low achievers perform in grade 6, in order to understand if there is a regress towards the mean of the overall population.

Sample size

In each city, we take into consideration relative scores in grade 5 and we select students whose scores belong to the 1st, 2nd or 3rd decile of the city distribution. The first column of table 16 and 17 shows the sample we obtain. Subsequently, we observe the same group in grade 6 to understand changes in scores' distribution (columns 2 to 4 of table 16 and 17).

We create three groups:

- **Decile ≤ 3rd** contains students whose Reading/Mathematics test score remains in the lowest third of the distribution. These students could have worsened their performance or be unvaried, belonging to the same decile or to a lower one.

- **$4^{th} \leq$ Decile $\leq 5^{th}$** refers to students who have improved their relative performance, overtaking the third decile but performing still under the average score of the city.
- **Decile $\geq 6^{th}$** includes students who have improved their performance in an absolute way. These students have obtained a relative score higher than the average of the city.

Table 16. Distribution of disadvantaged low achievers in grade 6, Reading test.

	Decile $\leq 3^{rd}$ Grade 5	Decile $\leq 3^{rd}$ Grade 6	$4^{th} \leq$ Decile $\leq 5^{th}$ Grade 6	Decile $\geq 6^{th}$ Grade 6
Bologna	367	236	93	38
Milan	1617	1052	395	170
Padua	248	157	50	41
Rome	2385	1431	542	412
Turin	779	493	179	107

Table 17. Distribution of disadvantaged low achievers in grade 6, Mathematics test.

	Decile $\leq 3^{rd}$ Grade 5	Decile $\leq 3^{rd}$ Grade 6	$4^{th} \leq$ Decile $\leq 5^{th}$ Grade 6	Decile $\geq 6^{th}$ Grade 6
Bologna	370	176	69	125
Milan	1617	1156	228	233
Padua	248	115	45	88
Rome	2167	1066	318	783
Turin	768	523	111	134

Observing the tables above, we notice that students' distribution among the three groups is quite similar in Reading test scores: 63% of students selected in grade 5 remain (in grade 6) in the bottom third of distribution of the city; 23% improve in a relative way; 14% overtake the GPA. This pattern is repeated for each city. As regard to Mathematics test, three groups are differently populated among cities. Nevertheless, we point out a similarity between Milan and Turin on one side, and Bologna, Padua and Rome on the other side.

Regression to the mean

Regression to the mean is a statistical phenomenon that occurs when repeated measures of a nonrandom sample are imperfectly correlated. It is a group phenomenon that happens because of random errors in values observed. Random error can be caused by (I) the measurement instrument (the test itself in this case) or (II) subjects themselves (within-student variation from test to test).

We want to investigate the impact of this phenomenon on our data as It could entail two main issues: on the one hand It makes difficult to distinguish a real change from this expected change due to natural variation; on the other hand, It reduces the variation itself from one year and another, making the identification of resilience-related characteristics harder. Considering the same sample of disadvantaged low achiever students, we calculate the GPA in grade 5 and 6 using relative scores (table 18 and 19).
Tables underline the presence of a general regression of the mean, much more clear in Mathematics than in Reading test. In addition, we find the same similarity between Milan and Turin on one side, and Bologna, Padua and Rome on the other side.

Table 18. Sample size and GPA of disadvantaged low achievers, Reading test.

	Sample size	GPA Grade 5	GPA Grade 6
Bologna	367	-51.60	-34.27
Milan	1617	-44.19	-33.29
Padua	248	-44.24	-29.61
Rome	2385	-42.68	-26.49
Turin	779	-44.75	-31.62

Table 19. Sample size and GPA of disadvantaged low achievers, Mathematics test.

	Sample size	GPA Grade 5	GPA Grade 6
Bologna	370	-41.04	-16.74
Milan	1617	-41.83	-31.32
Padua	248	-37.23	-15.89
Rome	2167	-46.58	-13.58
Turin	768	-42.53	-30.58

A similar regression happens when considering “high achiever students”, namely those students whose test score was on the top 33^{th} percentile of the distribution in grade 5.
As we can see from the tables 20 and 21, also high achievers’ score move towards the mean in grade 6; this reinforce the hypothesis that a regression to the mean happens.
Thus, commenting our results we bear in mind that the variation between grade 5 and 6 tends to reduce as a consequence of smaller differences due to regression. For this reason, some characteristics related to resilience could be difficult to observe because we don’t have sufficient variation to detect them clearly.

Table 20. Sample size and GPA of high achiever students, Reading test.

	Sample size	**GPA Grade 5**	**GPA Grade 6**
Bologna	369	47.65	29.32
Milan	1717	40.73	28.47
Padua	255	41.01	24.36
Rome	2551	39.63	23.43
Turin	843	40.26	28.50

Table 21. Sample size and GPA of high achiever students, Mathematics test.

	Sample size	**GPA Grade 5**	**GPA Grade 6**
Bologna	369	54.37	17.82
Milan	1671	41.74	32.56
Padua	257	49.75	11.30
Rome	2389	51.27	17.78
Turin	768	44.58	35.02

3.2.2 *Antiresilient* Students

The second issue regards *antiresilient* students, namely those students whose score is above the average of the city in 5th grade but who decreased their results in 6th grade, performing below the average. The matter is: *It would be good to have a school that helps low performers, but what happens if the same school compromises high achievers?* A school that helps low performers while disadvantaging high achievers might not be so desirable.

In order to investigate this matter, we classify the *antiresilient* sample. As made for resilient students, we define as “advantaged” those students whose ESCS is in the top 33% of the ESCS distribution in grade 5. In the same way we name “high achievers” those students whose test score is above the 70th percentile of the test scores distribution in grade 5. Finally we merge two dataset obtaining the “Advantaged high achievers” sample (table 22).

Table 22. Advantaged high achievers sample.

	Advantaged High Achievers
Bologna	278
Milan	1128
Padua	206
Rome	1616
Turin	533

Once obtained the sample, we look for students who decrease their performance in grade 6. We consider *antiresilient* those students whose test score is under the mean of the test scores distribution in grade 6, which is exactly the reverse of our definition of resilient student (table 23).

Table 23. Antiresilient students, sample size.

	***Antiresilient* in Reading**	***Antiresilient* in Mathematics**	**Overall Percentage**
Bologna	15	75	28
Milan	69	120	16
Padua	20	51	34
Rome	146	286	26
Turin	38	47	15

Notes: The percentage refers to the proportion of Advantaged high achievers who has become *Antiresilient* in Reading or Mathematics.

With regard to the reading test, Bologna, Milan, Rome and Turin do not show a high concentration of *antiresilient* students in particular schools. In Bologna they are distributed across 10 schools with a mean of 1.5 students per school and a standard deviation of 0.70. A quite similar distribution arises in Milan and Rome, with a mean of 1.70 with a standard deviation of 1.12 for Milan and 1.89 for Rome, even though *antiresilient* students are distributed across 82 in Rome schools and 41 in Milan. In Turin they are allocated among 30 schools with a mean of 1.26 and standard deviation of 0.52. If the situation is similar for this four cities, Padua show a very different pattern. In this city there is a greater concentration of this kind of students, with a mean of 2.5 students per school.

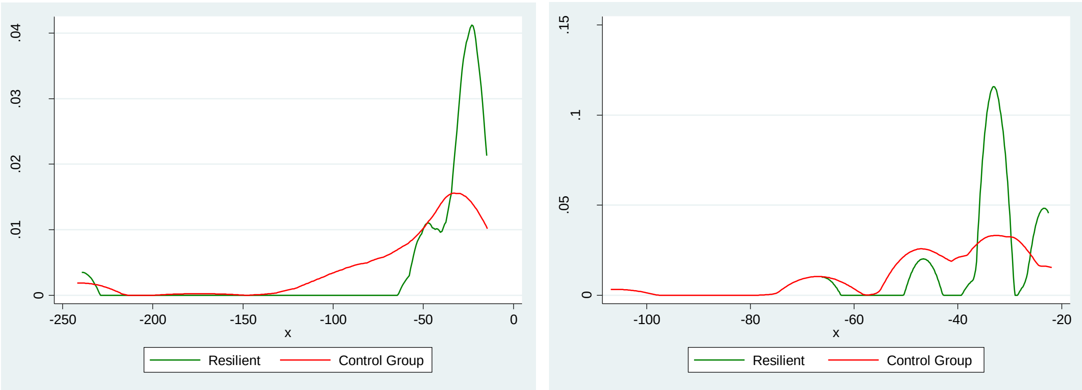
When considering the Mathematics test, in Milan *antiresilient* are distributed across 50 schools with a mean of 2.4 students per school and a standard deviation of 1.56. In Rome, they are allocated across 101 schools with a mean of 2.8 students per school (similarly to Milan) but a higher standard deviation of 3.15. Turin follows closely behind with a mean of 2.13 and a standard deviation of 1.2, but the number of schools is only 22. Nevertheless, their distribution across schools is not so relevant to suggest that their worse outcomes could be related to the school attended. In Bologna and Padua the situation changes. In both cities students are distributed across 12 schools with a mean of 5.76 (and a standard deviation of 4.51) for Bologna and 4.25 (with a standard deviation of 4.39) for Padua. Focusing on these two cities, we define "antiresilient" schools those with a concentration of "antiresilient" students higher than the mean (of "antiresilient" students per school) plus a standard deviation. We aim at understanding if "antiresilient" schools are the same that we identify as "resilient schools", because of the high concentration of resilient students (see paragraph 3.3.3). In this comparison, none of the schools with a high concentration of resilient students has also shown a high concentration of "antiresilient" students. For this reason, we can say that schools that we consider as "resilient" do not compromise the performance of high achiever students when arguably help disadvantaged students to succeed.

3.3 Descriptive Statistics

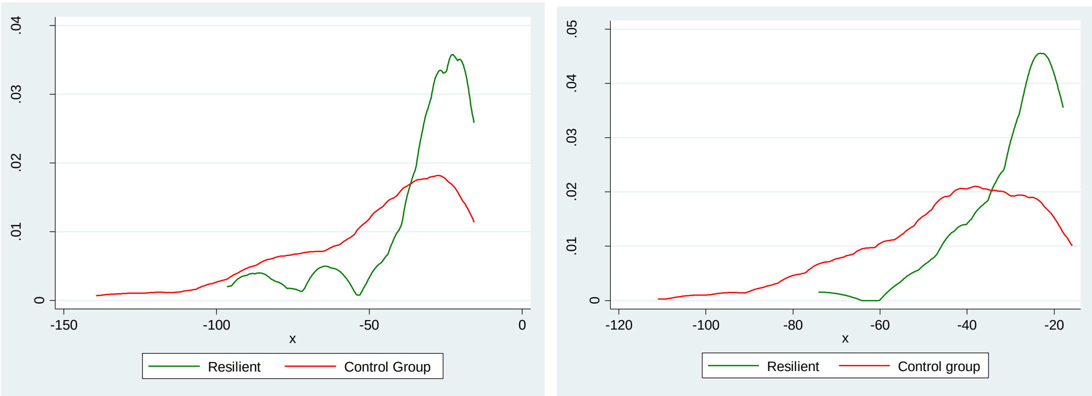
3.3.1 Describing Resilient Students

Having defined resilient students, the first step is to analyze their characteristics in order to better understand the potential factors associated with this phenomenon. Specifically, the investigation of baseline descriptive statistics enables a double comparison: on one side it is possible to compare the characteristics of resilient and control group, on the other it allows investigating the role of the variables of the same group among the cities considered. For this purpose, we start comparing the distribution of the groups we defined. Figures from 4 to 8 show the distribution of test scores in grade 5 in the cities considered. Distributions are quite similar for both subjects. Exceptions are represented by the Mathematics test score distribution in the city of Padua, where resilient students perform worse on average, and Rome, where the distributions are similar. Figures show how resilient students tend to perform better than control group not only in grade 6, but also in grade 5.

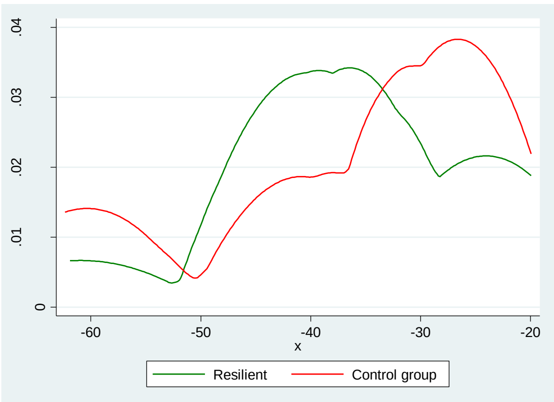
Graphs 1a and 1b. Reading (on the left) and mathematics (on the right) test scores distribution in grade 5 (2011/2012), Bologna – kernel density distribution.



Graphs 2a and 2b. Reading (on the left) and mathematics (on the right) test scores distribution in grade 5 (2011/2012), Milan – kernel density distribution.

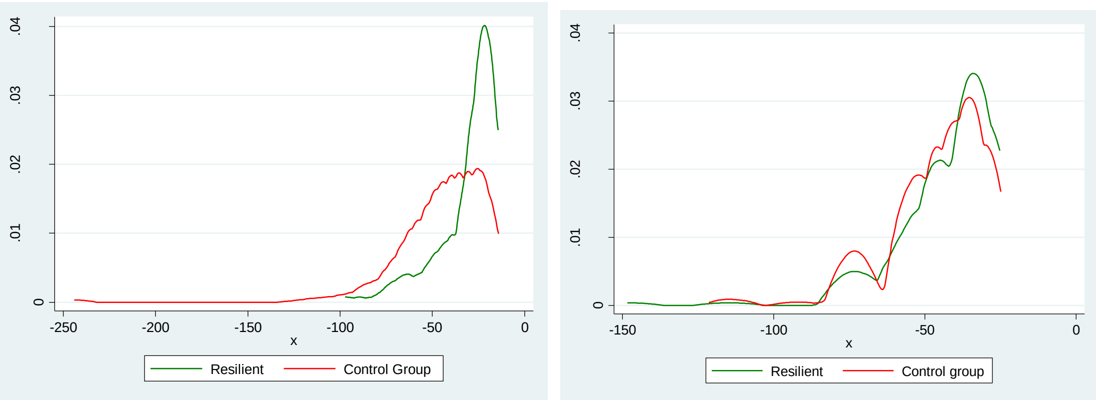


Graph 3. Mathematics test scores distribution in grade 5 (2011/2012), Padua – kernel density distribution.

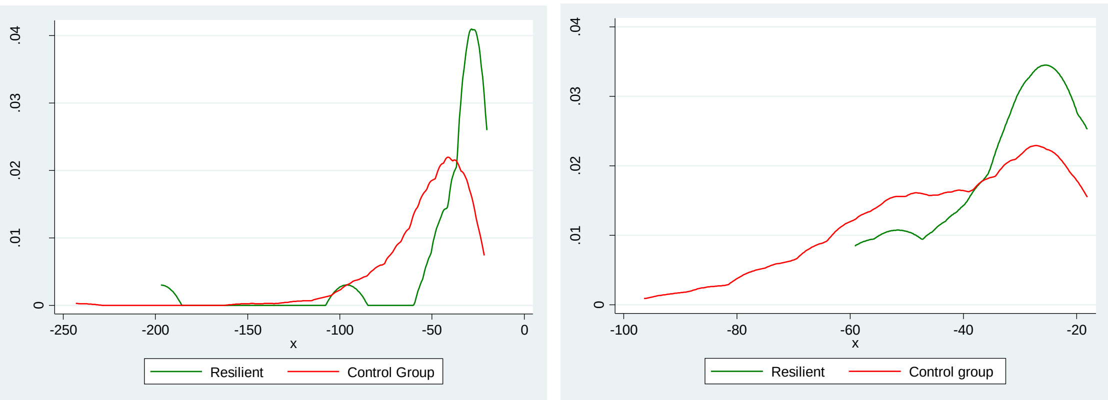


Note: The missing graph of reading's distribution is due to lack of observations (two only resilient students).

Graphs 4a and 4b. Reading (on the left) and mathematics (on the right) test scores distribution in grade 5 (2011/2012), Rome – kernel density distribution.

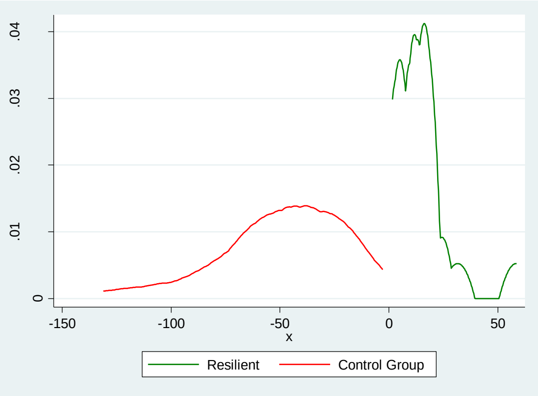


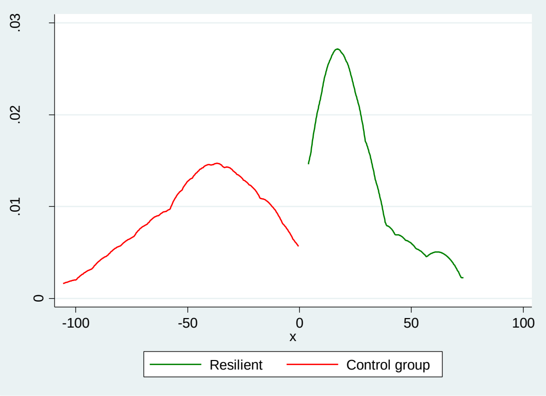
Graphs 5a and 5b. Reading (on the left) and mathematics (on the right) test scores distribution in grade 5 (2011/2012), Turin – kernel density distribution.



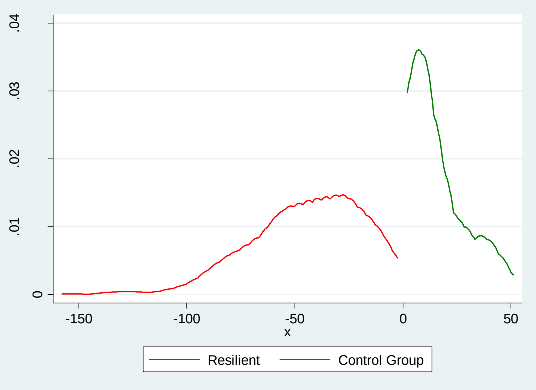
Similarly, figures from 9 to 13 illustrate scores distribution in grade 6, revealing a higher density of resilient students close to the cut-off point, a pattern repeated for each city. When investigating the distribution of the socioeconomic index, figures from 14 to 18 show very similar results across the cities. Resilient students and control group have similar trend in their socioeconomic background. The only isolated case is represented by the distribution of mathematics resilient students in Bologna (figure 14b) who have a SES background significantly higher than the control group. This suggests that resilient students are more likely to be identified among disadvantaged students who achieve relatively higher scores than among wealthier students.

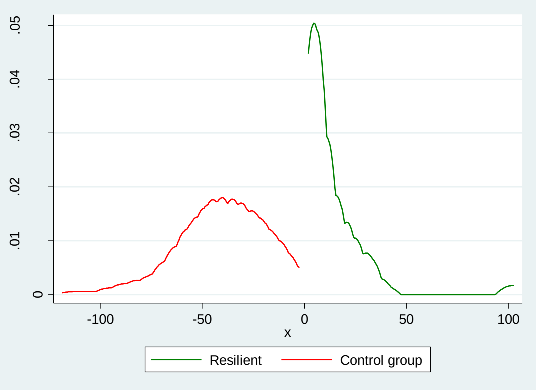
Graphs 6a and 6b. Reading (on the left) and mathematics (on the right) test scores distribution in grade 6 (2012/2013), Bologna – kernel density distribution.



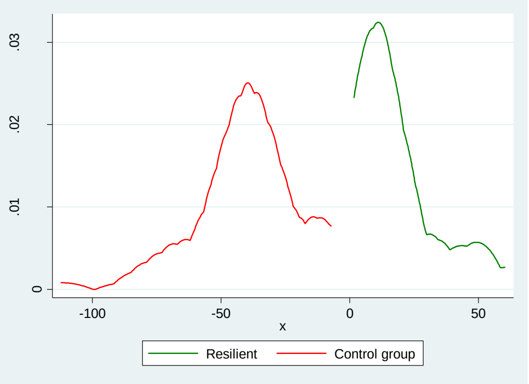


Graphs 7a and 7b. Reading (on the left) and mathematics (on the right) test scores distribution in grade 6 (2012/2013), Milan – kernel density distribution.



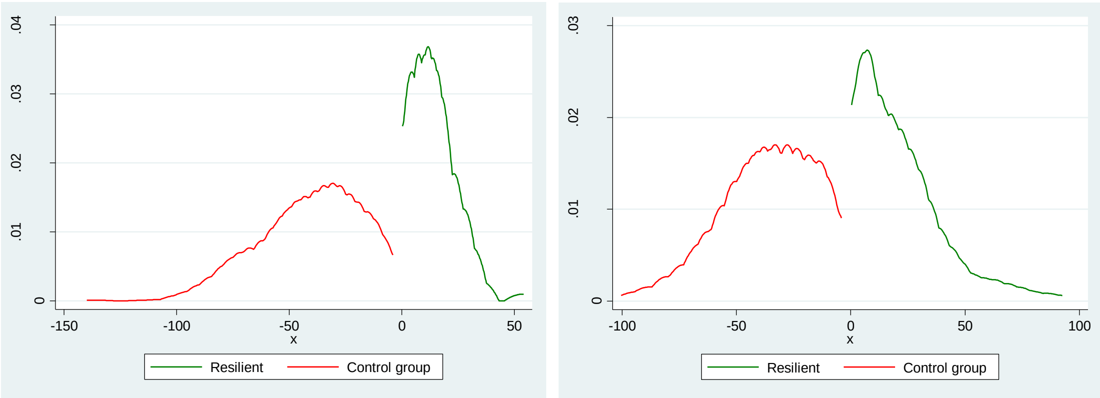


Graphs 8. Mathematics test scores distribution in grade 6 (2012/2013), Padua – kernel density distribution.

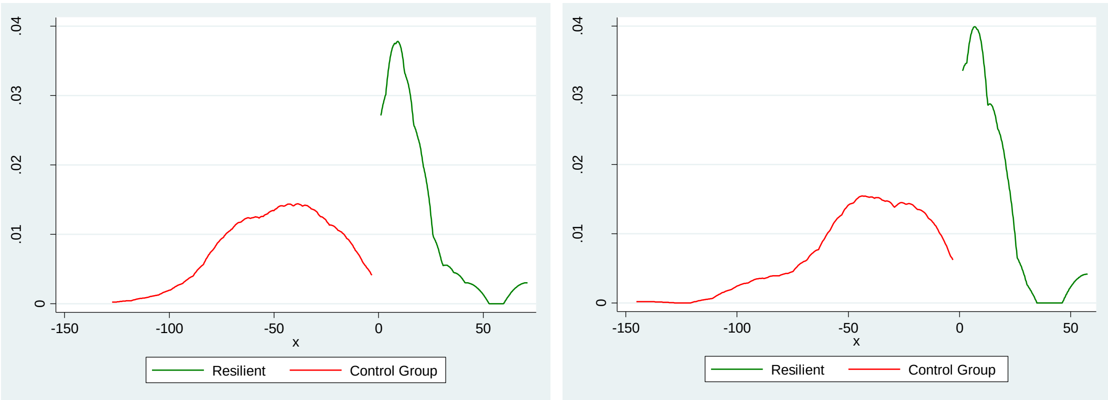


Note: The missing graph of reading's distribution is due to lack of observations (two only resilient students).

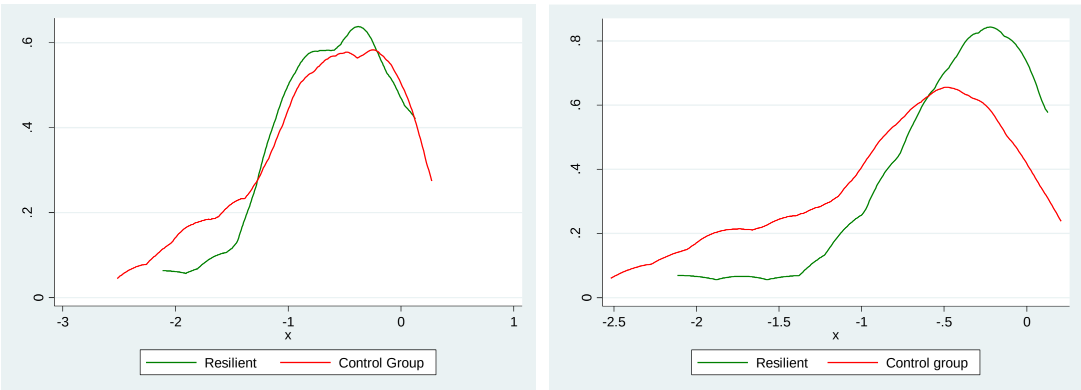
Graphs 9a and 9b. Reading (on the left) and mathematics (on the right) test scores distribution in grade 6 (2012/2013), Rome – kernel density distribution.



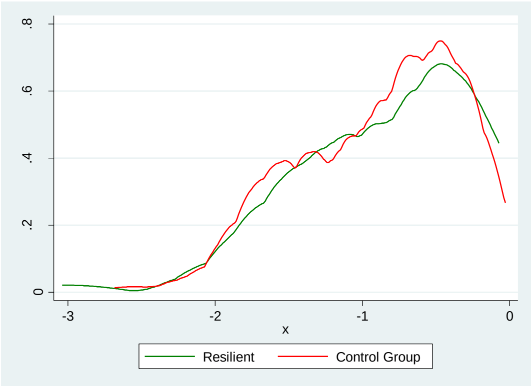
Graphs 10a and 10b. Reading (on the left) and mathematics (on the right) test scores distribution in grade 6 (2012/2013), Turin – kernel density distribution.

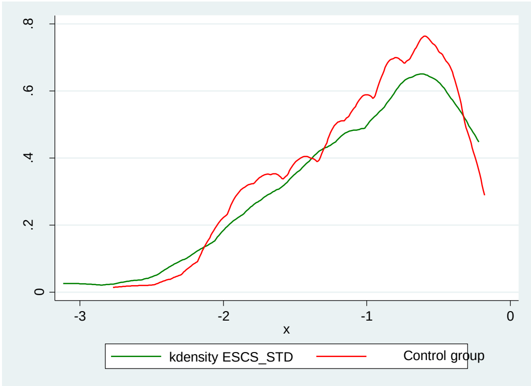


Graphs 11a and 11b. ESCS index distribution for reading (on the left) and mathematics (on the right) sample of students. Data refer to grade 5 (2011/2012), Bologna – kernel density distribution.

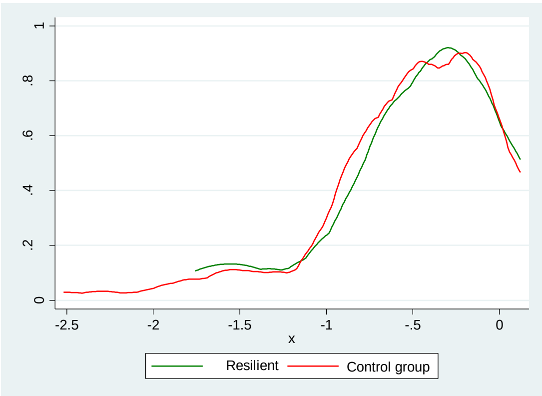


Graphs 12a and 12b. ESCS index distribution for Reading (on the left) and mathematics (on the right) sample of students. Data refer to grade 5 (2011/2012), Milan – kernel density distribution.



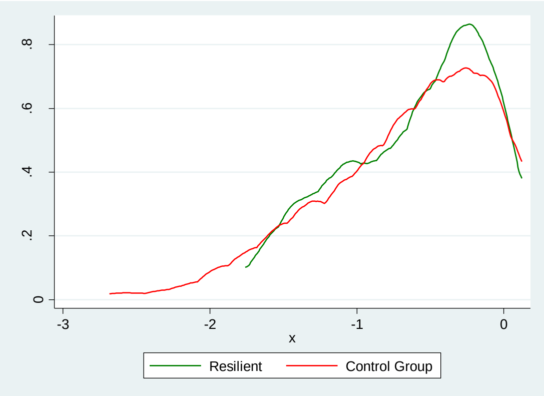


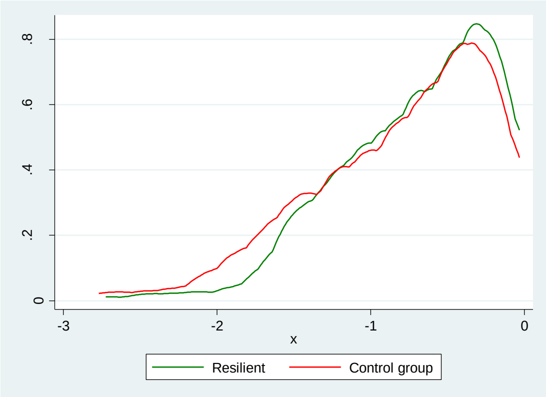
Graphs 13. ESCS index distribution for mathematics sample of students. Data refer to grade 5 (2011/2012), Padua – kernel density distribution.



Note: The missing graph of reading's distribution is due to lack of observations (two only resilient students).

Graphs 14a and 14b. ESCS index distribution for Reading (on the left) and mathematics (on the right) sample of students. Data refer to grade 5 (2011/2012), Rome – kernel density distribution.





Graphs 15a and 15b. ESCS index distribution for Reading (on the left) and mathematics (on the right) sample of students. Data refer to grade 5 (2011/2012), Turin – kernel density distribution.

Besides, we analyze the characteristics of resilient students and control group implementing a three-level analysis (individual, classroom and school-level characteristics). Each table (one for each city, see tables 24-28) presents three columns of descriptive indicators: the first contains data about resilient students in reading, the second refers to resilient students in mathematics and the last show statistics computed on the control group. We want to focus on the switch from grade 5 to 6, so every time variant variable is compared between one year (2011/2012) and another (2012/2013).

Comparing resilient students and control groups among the cities considered, some aspects can be noted. At student-level, the average score of the three groups (Reading, Mathematics resilient students and control group) is clearly affected by our definition of resilient student. Nevertheless, as we observed from the distributions above, resilient students tend to perform on average better than the control group in grade 5 as well. The second variable involved in the definition of resilient student, the socio-economic background measured by the ESCS index, show that all students (independently from the group or the city) tend to have an ESCS index in 2012/2013 lower than the previous year, which means an average worse socioeconomic condition. Focusing on the immigrant status, It can be noted that first generation immigrants are averagely less numerous among resilient students than in control group, except for the city of Rome, where the proportion is similar for all three groups. On the contrary, such a clear pattern cannot be noticed for second generation immigrant. This could suggest that first generation immigrants have greater difficulties in overtaking their disadvantaged condition than second generation immigrants, at least for the cities mentioned. Finally, descriptive analysis shows a lower rate of late-enrolled students in the resilient group than in the control. This scheme is repeated in every city, except for Rome.

Investigating data at classroom level, It is again to be stressed the fact that classrooms attended by students are on average poorer than the year before, both for resilient students and control group. Furthermore, the class average score increases between one year and the other, much more clearly for resilient students: they are concentrated in classrooms characterized by lower performances in grade 5 and higher in grade 6.

In this case, the only exception is Padua. About this city, It is important to say that the small number of reading resilient students (with just two people forming the sample) could heavily affect results.

Another important comment to classroom-level characteristics regards the proportion of students who attend class with *"tempo pieno".* In fact, It is much more common among primary than secondary school, which explains why the proportion of students attending this kind of class is so small at grade 6. In addition, the overall number of junior secondary schools is lower than primary schools (from INVALSI data we have counted up a number of 7,150 primary schools and 5,863 lower secondary schools across Italy), so that the average number of classrooms and students in a school is higher switching from grade 5 to 6, and this is reflected in our analysis. Nonetheless, we can notice that resilient students not only attend bigger school at grade 6, but they also come from bigger schools at grade 5, in terms of both number of students and classrooms in a school.

Table 24. Descriptive analysis, Bologna.

	Reading		Mathematics		Control	
	N=18		N=20		N=143	
Variable						
Studen- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Achievement in reading (grade 5)	168.91	51.24	194.32	62.2	154.13	55.91
Achievement in reading (grade 6)	229.42	13.51	223.66	24.29	168.31	28.45
Achievement in Math (grade 5)	165.29	32.36	145.45	10.11	147.73	23.06
Achievement in Math (grade 6)	216.17	31.52	238.94	18.26	170.33	25.1
Female Student	0.39		0.35		0.52	
1st generation immigrant	0.11		0.1		0.21	
2nd generation immigrant	0.28		0.25		0.21	
Early-enrolled student	0		0		0	
Late-enrolled student	0.05		0.05		0.11	
Socioeconomic background (grade 5)	-0.56	0.58	-0.43	0.55	-0.71	0.65
Socioeconomic background (grade 6)	-0.55	0.53	-0.49	0.35	-0.85	0.61
Student who has siblings	0.72		0.8		0.75	
Classroom- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Class average socioeconomic background (ESCS) (grade 5)	0.16	0.52	0.12	0.53	0.37	0.55
Class average socioeconomic background (ESCS) (grade 6)	0.15	0.49	0.004	0.53	0.24	0.62
Classroom average score in reading (grade 5)	185.74	28.2	189.83	32.06	196.35	39.81
Classroom average score in reading (grade 6)	202.58	16.39	198.04	15.79	198.94	17.75
Classroom average score in mathematics (grade 5)	164.53	27.16	149.14	5.69	165.26	27.17
Classroom average score in mathematics (grade 6)	204.06	11.2	204.67	16.61	200.04	21.49
Proportion of female in the classroom (grade 6)	0.37		0.38		0.41	
Proportion of female in the classroom (grade 6)	0.44		0.39		0.41	
Proportion of 1st generation immigrants in the classroom (grade 5)	0.1		0.11		0.1	
Proportion of 1st generation immigrants in the classroom (grade 6)	0.1		0.17		0.15	
Proportion of 2nd generation immigrants in the classroom (grade 5)	0.15		0.17		0.11	
Proportion of 2nd generation immigrants in the classroom (grade 6)	0.15		0.18		0.13	
Proportion of Early-enrolled students in the classroom (grade 5)	0.002		0.006		0.005	
Proportion of Early-enrolled students in the classroom (grade 6)	0.002		0.007		0.006	
Proportion of Late-enrolled students in the classroom (grade 5)	0.04		0.06		0.05	
Proportion of Late-enrolled students in the classroom (grade 6)	0.11		0.11		0.09	
Number of student in the classroom (grade 5)	23.41	2	22.73	2.05	22.5	2.17
Number of student in the classroom (grade 6)	24.33	1.91	24.73	1.79	24.48	2.4
Class with "tempo pieno" (grade 5)	0.7		0.85		0.72	
Class with "tempo pieno" (grade 6)	0		0		0.02	

School-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
School average socioeconomic background (ESCS) (grade 5)	0.28	0.33	0.31	0.41	0.54	0.45
School average socioeconomic background (ESCS) (grade 6)	0.27	0.59	0.19	0.52	0.28	0.63
School average score in reading (grade 5)	195.30	13.86	190.87	11.21	203.10	15.55
School average score in reading (grade 6)	203.53	14.39	201.92	13.46	201.73	19.09
School average score in mathematics (grade 5)	167.36	27.89	149.79	2.43	170.58	28.54
School average score in mathematics (grade 6)	204.33	10.26	205.42	10.76	201.96	21.98
Number of classrooms in the school (grade 5)	4.63	1.68	4.44	1.42	4.26	1.97
Number of classrooms in the school (grade 6)	4.6	1.34	5.1	1.52	5	2.09
Number of students in the school (grade 5)	107.93	38.13	99.51	35.05	96.8	45.87
Number of students in the school (grade 6)	113.07	27.19	126.85	34.84	123.88	54.12
Average numbers of students per class, in the school (grade 5)	23.45	2.08	22.17	1.91	22.39	2.01
Average numbers of students per class, in the school (grade 6)	24.97	2.12	22.15	2	24.61	2.56
Private school (grade 5)	0.06		0		0.03	
Private school (grade 6)	0.06		0		0.03	
Istituto comprensivo (grade 5)	0.77		1		0.86	
Istituto comprensivo (grade 6)	0.94		1		0.95	
Same *Istituto comprensivo* grade 5 - 6	0.63		0.65		0.56	

Notes. Authors' elaboration from INVALSI data, 2011/12 and 2012/13.

Table 25. Descriptive analysis, Milan.

	Reading		Mathematics		Control	
	N=50		N=50		N=603	
Variable						
Studen- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Achievement in reading (grade 5)	181.5	10.32	189.02	35.12	166.75	31.82
Achievement in reading (grade 6)	229	13.32	206.22	24.92	170.79	25.74
Achievement in Math (grade 5)	194.76	31.23	178.15	11.34	172.11	26.06
Achievement in Math (grade 6)	216.6	30.95	227.2	16.27	177.22	22.02
Female Student	0.54		0.55		0.5	
1st generation immigrant	0.12		0.16		0.17	
2nd generation immigrant	0.22		0.24		0.33	
Early-enrolled student	0		0		0.003	
Late-enrolled student	0.1		0.06		0.11	
Socioeconomic background (grade 5)	-0.76	0.52	-0.85	0.65	-0.87	0.54
Socioeconomic background (grade 6)	-0.77	0.48	-0.88	0.59	-0.87	0.51
Student who has siblings	0.86		0.83		0.78	
Classroom- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Class average socioeconomic background (ESCS) (grade 5)	0.19	0.6	-0.02	0.67	0.26	0.6
Class average socioeconomic background (ESCS) (grade 6)	0.06	0.46	0.06	0.61	0.06	0.52
Classroom average score in reading (grade 5)	199.87	15.47	194.61	18.74	199.41	16.45
Classroom average score in reading (grade 6)	207.77	12.8	204.67	19.65	200.28	15.86
Classroom average score in mathematics (grade 5)	198.54	15.21	191.72	17.99	200.35	18.39
Classroom average score in mathematics (grade 6)	206.94	12.6	211.77	15.8	204.11	13.76
Proportion of female in the classroom (grade 6)	0.43		0.4		0.42	
Proportion of female in the classroom (grade 6)	0.42		0.42		0.4	
Proportion of 1st generation immigrants in the classroom (grade 5)	0.06		0.1		0.08	
Proportion of 1st generation immigrants in the classroom (grade 6)	0.09		0.11		0.11	
Proportion of 2nd generation immigrants in the classroom (grade 5)	0.16		0.19		0.18	
Proportion of 2nd generation immigrants in the classroom (grade 6)	0.16		0.18		0.18	
Proportion of Early-enrolled students in the classroom (grade 5)	0.01		0.009		0.005	
Proportion of Early-enrolled students in the classroom (grade 6)	0.005		0.006		0.008	
Proportion of Late-enrolled students in the classroom (grade 5)	0.04		0.05		0.04	
Proportion of Late-enrolled students in the classroom (grade 6)	0.07		0.1		0.08	
Number of student in the classroom (grade 5)	21.6	3.54	20.07	3.45	21.06	3.19
Number of student in the classroom (grade 6)	22.76	2.57	22.08	2.25	22.67	2.64
Class with "tempo pieno" (grade 5)	0.84		0.82		0.94	
Class with "tempo pieno" (grade 6)	0.02		0.02		0.03	
School-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
School average socioeconomic background (ESCS) (grade 5)	0.19	0.61	0.21	0.61	0.41	0.62
School average socioeconomic background (ESCS) (grade 6)	0.23	0.54	0.14	0.59	0.18	0.55
School average score in reading (grade 5)	202.14	12.85	202.21	12.67	203.82	11.52

School average score in reading (grade 6)	209.64	12.11	205.94	15.95	204.46	15.67
School average score in mathematics (grade 5)	200.49	12.87	199.08	13.37	203.82	14.13
School average score in mathematics (grade 6)	210.53	8.97	211.32	13.36	207.60	12.51
Number of classrooms in the school (grade 5)	4.75	2.01	5.21	2.1	4.61	2.22
Number of classrooms in the school (grade 6)	5.93	3.04	5.96	2.66	5.28	2.53
Number of students in the school (grade 5)	102.5	47.46	106.12	44.5	97.99	48.26
Number of students in the school (grade 6)	135.89	76.71	136.48	66.83	120.04	64.41
Average numbers of students per class, in the school (grade 5)	21.29	3.65	20.39	3.53	21.19	3.12
Average numbers of students per class, in the school (grade 6)	22.45	2.49	22.69	2.42	22.85	2.82
Private school (grade 5)	0.16		0.1		0.04	
Private school (grade 6)	0.1		0.08		0.02	
Istituto comprensivo (grade 5)	0.7		0.78		0.78	
Istituto comprensivo (grade 6)	0.62		0.78		0.79	
Same *Istituto comprensivo* grade 5 - 6	0.44		0.47		0.53	

Notes. Authors' elaboration from INVALSI data, 2011/12 and 2012/13.

Table 26. Descriptive analysis, Padua.

	Reading		Mathematics		Control	
	N=2		N=19		N=79	
Variable						
Studen- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Achievement in reading (grade 5)	191.914 6	5.99034	221.62	33.68	170.13	35.48
Achievement in reading (grade 6)	235.490 4	7.31126	225.54	25.24	174.71	26.14
Achievement in Math (grade 5)	156.515 5	14.3263 5	136.09	11.8	147.58	20.34
Achievement in Math (grade 6)	205.712 8	28.6555 6	237.03	16.15	180.7	20.35
Female Student	1		0.32		0.53	
1st generation immigrant	0		0.16		0.28	
2nd generation immigrant	0		0.16		0.18	
Early-enrolled student	0		0		0.01	
Late-enrolled student	0		0.11		0.15	
Socioeconomic background (grade 5)	-0.31	0.042	- 0.4591	0.4824	- 0.5237	0.5186
Socioeconomic background (grade 6)	-0.322	0.31	- 0.5693	0.6499	- 0.6672	0.5356
Student who has siblings	0.72		0.74		0.72	
Classroom- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Class average socioeconomic background (ESCS) (grade 5)	0.663	0.417	0.2083	0.7449	0.453	0.549
Class average socioeconomic background (ESCS) (grade 6)	0.657	0.458	0.3335	0.7744	0.5	0.41
Classroom average score in reading (grade 5)			217.34	31.72	209.45	12.82
Classroom average score in reading (grade 6)			211.61	12.9	208.08	14.22
Classroom average score in mathematics (grade 5)			134.96	11.52	156.33	21.28
Classroom average score in mathematics (grade 6)			216.83	12.61	213.01	14.44
Proportion of female in the classroom (grade 6)	0.61		0.51		0.51	
Proportion of female in the classroom (grade 6)	0.62		0.49		0.47	
Proportion of 1st generation immigrants in the classroom (grade 5)	0.09		0.18		0.13	
Proportion of 1st generation immigrants in the classroom (grade 6)	0.1		0.12		0.11	
Proportion of 2nd generation immigrants in the classroom (grade 5)	0.07		0.07		0.07	
Proportion of 2nd generation immigrants in the classroom (grade 6)	0.08		0.04		0.06	
Proportion of Early-enrolled students in the classroom (grade 5)	0		0.009		0.005	
Proportion of Early-enrolled students in the classroom (grade 6)	0		0.005		0.006	
Proportion of Late-enrolled students in the classroom (grade 5)	0.02		0.1		0.07	
Proportion of Late-enrolled students in the classroom (grade 6)	0.05		0.12		0.12	
Number of student in the classroom (grade 5)	23.5	0.71	21.18	3.15	20.87	3.26
Number of student in the classroom (grade 6)	22	1.41	24.19	3.33	23.71	3.06
Class with "tempo pieno" (grade 5)	0.5		0.24		0.37	
Class with "tempo pieno" (grade 6)	0		0		0	
School-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
School average socioeconomic background (ESCS) (grade 5)	0.74	0.14	0.45	0.58	0.67	0.61

School average socioeconomic background (ESCS) (grade 6)	0.52	0.12	0.55	0.33	0.63	0.38
School average score in reading (grade 5)	213.40	5.57	209.73	7.93	211.16	7.22
School average score in reading (grade 6)	212.88	5.13	210.50	7.96	211.92	8.03
School average score in mathematics (grade 5)	152.33	3.04	150.44	2.77	168.58	32.65
School average score in mathematics (grade 6)	215.52	2.74	214.83	8.53	215.74	7.49
Number of classrooms in the school (grade 5)	4	0	5.52	1.77	4.51	2.23
Number of classrooms in the school (grade 6)	4	0	4.67	1.32	4.54	1.45
Number of students in the school (grade 5)	94	2.83	112.23	45.97	87.55	37.41
Number of students in the school (grade 6)	88	5.66	111.56	94.86	108.82	32.26
Average numbers of students per class, in the school (grade 5)	23.5	0.71	21.11	3.27	20.98	3.66
Average numbers of students per class, in the school (grade 6)	22	1.41	24.07	3.74	24.23	2.56
Private school (grade 5)	0		0		0.05	
Private school (grade 6)	0		0		0.03	
Istituto comprensivo (grade 5)	1		1		0.95	
Istituto comprensivo (grade 6)	1		1		0.97	
Same *Istituto comprensivo* grade 5 - 6	1		1		0.78	

Notes. Authors' elaboration from INVALSI data, 2011/12 and 2012/13.

Table 27. Descriptive analysis, Rome.

	Reading		Mathematics		Control	
	N=93		N=174		N=656	
Variable						
Studen- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Achievement in reading (grade 5)	178.83	16.53	209.68	33.53	173.13	30.65
Achievement in reading (grade 6)	224.61	10.35	216.55	28.62	171.17	22.85
Achievement in Math (grade 5)	178.39	34.24	144.14	16.46	158.76	30.21
Achievement in Math (grade 6)	202.91	26.46	227.83	18.83	171.94	21.2
Female Student	0.57		0.47		0.5	
1st generation immigrant	0.1		0.12		0.1	
2nd generation immigrant	0.08		0.1		0.12	
Early-enrolled student	0.01		0.01		0	
Late-enrolled student	0.09		0.07		0.06	
Socioeconomic background (grade 5)	-0.59	0.49	-0.54	0.52	-0.66	0.58
Socioeconomic background (grade 6)	-0.7	0.48	-0.66	0.47	-0.75	0.52
Student who has siblings	0.8		0.75		0.8	
Classroom- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Class average socioeconomic background (ESCS) (grade 5)	0.26	0.59	0.08	0.5	0.2	0.5
Class average socioeconomic background (ESCS) (grade 6)	0.11	0.54	0.06	0.4	0.07	0.49
Classroom average score in reading (grade 5)	198.37	15.01	203.11	18.19	202.89	16.73
Classroom average score in reading (grade 6)	209.64	15.02	205.3	14.42	200.27	15.48
Classroom average score in mathematics (grade 5)	177.74	30.5	151.16	12.27	175.88	31.9
Classroom average score in mathematics (grade 6)	204.88	15.32	205.34	16.59	198.46	15.69
Proportion of female in the classroom (grade 6)	0.42		0.4		0.41	
Proportion of female in the classroom (grade 6)	0.42		0.39		0.4	
Proportion of 1st generation immigrants in the classroom (grade 5)	0.06		0.06		0.05	
Proportion of 1st generation immigrants in the classroom (grade 6)	0.06		0.06		0.07	
Proportion of 2nd generation immigrants in the classroom (grade 5)	0.06		0.07		0.08	
Proportion of 2nd generation immigrants in the classroom (grade 6)	0.06		0.07		0.08	
Proportion of Early-enrolled students in the classroom (grade 5)	0.01		0.008		0.01	
Proportion of Early-enrolled students in the classroom (grade 6)	0.01		0.007		0.01	
Proportion of Late-enrolled students in the classroom (grade 5)	0.04		0.04		0.03	
Proportion of Late-enrolled students in the classroom (grade 6)	0.05		0.05		0.06	
Number of student in the classroom (grade 5)	21.04	3.28	20.54	3.54	20.49	3.47
Number of student in the classroom (grade 6)	22.78	2.87	22.52	2.38	21.94	3.3
Class with "tempo pieno" (grade 5)	0.6		0.7		0.66	
Class with "tempo pieno" (grade 6)	0.01		0.02		0.03	
School-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
School average socioeconomic background (ESCS) (grade 5)	0.33	0.56	0.17	0.41	0.36	0.48
School average socioeconomic background (ESCS) (grade 6)	0.19	0.50	0.13	0.36	0.19	0.51
School average score in reading (grade 5)	202.56	12.46	203.48	12.48	205.18	12.39

School average score in reading (grade 6)	207.16	10.64	204.34	10.56	203.89	11.95
School average score in mathematics (grade 5)	180.47	29.12	155.49	16.37	181.28	30.63
School average score in mathematics (grade 6)	203.16	11.70	202.97	12.70	201.39	13.48
Number of classrooms in the school (grade 5)	4.64	2.28	4.88	2.22	4.34	2.43
Number of classrooms in the school (grade 6)	5.6	2.69	5.77	2.23	5.29	2.87
Number of students in the school (grade 5)	97.14	47.92	101.75	48.43	89.49	51.61
Number of students in the school (grade 6)	130.17	70.09	129.74	53.61	117.45	69.51
Average numbers of students per class, in the school (grade 5)	20.79	3.37	20.81	3.11	20.58	3.17
Average numbers of students per class, in the school (grade 6)	22.84	2.67	22.36	2.17	21.88	3.28
Private school (grade 5)	0.14		0.01		0.07	
Private school (grade 6)	0.06		0.01		0.03	
Istituto comprensivo (grade 5)	0.43		0.94		0.55	
Istituto comprensivo (grade 6)	0.94		0.99		0.97	
Same *Istituto comprensivo* grade 5 - 6	0.37		0.68		0.41	

Notes. Authors' elaboration from INVALSI data, 2011/12 and 2012/13.

Table 28. Descriptive analysis, Turin.

	Reading		Mathematics		Control	
	N=22		N=16		N=277	
Variable						
Studen- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Achievement in reading (grade 5)	166.15	38.16	192.09	36.28	167.85	32.42
Achievement in reading (grade 6)	225.85	26.85	215	27.48	167.65	25.66
Achievement in Math (grade 5)	196.89	26.85	171.55	13.06	169.75	24.32
Achievement in Math (grade 6)	213.97	29.53	223.17	14.22	171.49	24.96
Female Student	0.36		0.5		0.549	
1st generation immigrant	0.09		0.13		0.2	
2nd generation immigrant	0.27		0.06		0.22	
Early-enrolled student	0		0		0.004	
Late-enrolled student	0.09		0.13		0.1	
Socioeconomic background (grade 5)	-0.9443	0.5211	-0.9041	0.435	-0.9867	0.507438
Socioeconomic background (grade 6)	-1.1591	0.4547	-1.0953	0.4102	-1.1446	0.485375
Student who has siblings	0.68		0.81		0.79	
Classroom- level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Class average socioeconomic background (ESCS) (grade 5)	-0.1804	0.8066	-0.1116	0.9019	0.0144	0.5294
Class average socioeconomic background (ESCS) (grade 6)	-0.2276	0.8428	-0.0243	0.8129	-0.1229	0.5548
Classroom average score in reading (grade 5)	193.37	13.22	197.38	17.65	201.44	15.5
Classroom average score in reading (grade 6)	199.07	16.86	203.93	16.3	199.42	18.38
Classroom average score in mathematics (grade 5)	186.25	22.55	179.95	25.64	198.52	20.6
Classroom average score in mathematics (grade 6)	199.79	17.14	205.06	17.8	200.92	16.52
Proportion of female in the classroom (grade 6)	0.52		0.49		0.5	
Proportion of female in the classroom (grade 6)	0.5		0.54		0.51	
Proportion of 1st generation immigrants in the classroom (grade 5)	0.11		0.12		0.13	
Proportion of 1st generation immigrants in the classroom (grade 6)	0.13		0.11		0.13	
Proportion of 2nd generation immigrants in the classroom (grade 5)	0.2		0.16		0.14	
Proportion of 2nd generation immigrants in the classroom (grade 6)	0.13		0.1		0.11	
Proportion of Early-enrolled students in the classroom (grade 5)	0.01		0.004		0.005	
Proportion of Early-enrolled students in the classroom (grade 6)	0.003		0.008		0.003	
Proportion of Late-enrolled students in the classroom (grade 5)	0.06		0.07		0.07	
Proportion of Late-enrolled students in the classroom (grade 6)	0.13		0.11		0.12	
Number of student in the classroom (grade 5)	21.2	2.91	21.19	2.95	21.61	2.69
Number of student in the classroom (grade 6)	23	2.36	23.19	2.01	23.15	2.34
Class with "tempo pieno" (grade 5)	0.89		0.81		0.88	
Class with "tempo pieno" (grade 6)	0		0		0.007	
School-level characteristics	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
School average socioeconomic background (ESCS) (grade 5)	0.02	0.60	0.03	0.60	0.19	0.68
School average socioeconomic background (ESCS) (grade 6)	-0.17	0.55	0.05	0.55	-0.06	0.57
School average score in reading (grade 5)	200.66	12.29	199.93	14.00	203.57	12.78

School average score in reading (grade 6)	200.16	15.54	204.41	14.28	200.86	15.58
School average score in mathematics (grade 5)	194.37	21.39	191.68	22.08	198.81	17.37
School average score in mathematics (grade 6)	201.42	13.10	205.08	14.26	200.88	14.00
Number of classrooms in the school (grade 5)	5.88	2.33	5.36	2.21	4.79	2.4
Number of classrooms in the school (grade 6)	8.18	3.8	8.07	3.77	6.78	3.82
Number of students in the school (grade 5)	123.74	44.54	114.46	44.71	105.05	54.63
Number of students in the school (grade 6)	189.82	94.64	191.13	95.45	159.03	94.64
Average numbers of students per class, in the school (grade 5)	21.48	3	21.39	3.06	21.75	2.32
Average numbers of students per class, in the school (grade 6)	22.99	2.25	23.2	2.08	22.91	2.57
Private school (grade 5)	0		0.06		0.06	
Private school (grade 6)	0		0.06		0.05	
Istituto comprensivo (grade 5)	0.45		0.38		0.3	
Istituto comprensivo (grade 6)	0.59		0.31		0.3	
Same *Istituto comprensivo* grade 5 - 6	0.36		0.19		0.21	

Notes. Authors' elaboration from INVALSI data, 2011/12 and 2012/13.

3.3.2 Sensitivity of the Definition

In this section we modify our definition of resilient student, in order to understand how much flexible it is. More in detail, we change the threshold on the score and the socioeconomic background and we study the variation in the sample size. As described above, our definition of resilient student start from students who are below the 33th percentile of the city distribution of scores and socioeconomic background (in grade 5). We then consider as resilient those who, despite the disadvantage condition, improve his/her score over the average of the city distribution in grade 6 .

We want to investigate how our samples change when we replace constraints both upward and backward. We consider four variations of the threshold:

1) Socioeconomic background: ESCS index under the 50^{th} percentile of the distribution of each city, both in grade 5 and 6 (upward change), *caeteris paribus*.
2) Socioeconomic background: ESCS index under the 25^{th} percentile of the distribution of each city, both in grade 5 and 6 (backward change), *caeteris paribus*.
3) Score: test score in grade 6 over 200 points, which is the national mean (upward change), *caeteris paribus*.

4) Score: test score in grade 6 in the top of 33^{th} percentile of the distribution of each city (backward change), *caeteris paribus*.

This allow us to understand which characteristics affect more our definition of resiliency, whether the variation in the threshold of the socioeconomic condition or the score. Results are reported in tables 24-27, where the first column of each table refers to the number of resilient students currently used in this research and the second column shows the sample size obtained once considered the upward constraint. In the third column of each table we compute the absolute increase or decrease in sample size due to the variation in the threshold, while the fourth column show the same results in percentage. The second part of each table (columns 5, 6 and 7) regards the backward variation.

Table 29. Sample size of resilient students in reading changing the threshold of the ESCS index.

Reading							
	Current	50^{th} percentile	Increase	%	25^{th} percentile	Decrease	%
BOLOGNA	18	33	15	83%	10	-8	-44%
MILAN	50	102	52	104%	33	-17	-34%
PADUA	2	13	11	550%	2	0	0%
ROME	93	133	40	43%	91	-2	-2%
TURIN	22	31	9	41%	17	-5	-23%

Notes:
Current: number of resilient students in our original sample.
50^{th} percentile: number of resilient students after considering the ESCS index upward constraint.
Increase/decrease: growth or drop in the number of resilient students.
%: percentage of increase or decrease.
25^{th} percentile: number of resilient students after considering the ESCS index backward constraint.

Table 30. Sample size of resilient students in reading changing the threshold of the score in grade 6.

Reading							
	Current	Score > 200	Increase	%	33^{th} percentile	Decrease	%
BOLOGNA	18	85	67	372%	10	-8	-44%
MILAN	50	153	103	206%	19	-31	-62%
PADUA	2	15	13	650%	1	-1	-50%
ROME	93	179	86	92%	50	-43	-46%
TURIN	22	57	35	159%	6	-16	-73%

Notes:
Current: number of resilient students in our original sample.
Score > 200: number of resilient students after considering upward constraint of the score.
33^{th} percentage: number or resilient students after considering backward constraint of the score.

Table 31. Sample size of resilient students in mathematics changing the threshold of the ESCS index.

Mathematics							
	Current	50^{th} percentile	Increase	%	25^{th} percentile	Decrease	%
BOLOGNA	20	63	43	215%	11	-9	-45%
MILAN	49	85	36	73%	36	-13	-27%
PADUA	19	50	31	163%	15	-4	-21%
ROME	174	394	220	126%	133	-41	-24%
TURIN	16	43	27	169%	13	-3	-19%

Table 32. Sample size of resilient students in mathematics changing the threshold of the score in grade 6.

Mathematics							
	Current	Score > 200	Increase	%	33^{th} percentile	Decrease	%
BOLOGNA	20	83	63	315%	13	-7	-35%
MILAN	49	167	118	241%	15	-34	-69%
PADUA	19	84	65	342%	6	-13	-68%
ROME	174	647	473	272%	102	-72	-41%
TURIN	16	72	56	350%	7	-9	-56%

When analyzing the sample of resilient students in reading (tables 29-30) we note that, enlarging the definition through a variation of the ESCS index constraint, our sample of resilient students raises about 40 percent for Rome and Turin. This increase is relatively little if we consider that in Bologna the raise is double, whereas in Milan the sample increases of 100 percent. On the contrary when we pull downwards the ESCS index threshold restricting the sample, the smallest decrease is recorded for Rome, (2 percent), while all other cities have a reduction of 20 – 40 percent. Focusing on score's threshold, the effect on the number of resilient students is bigger than the impact of the ESCS index. In fact, samples increase from 100 to 372 percent when the definition is enlarged, while the drop is about 40 – 70 percent when it is restricted. Summarizing we can say that, considering the reading test, our definition of resilient student is more sensitive to a change in the score threshold than in the socioeconomic background.

The same results stand out from tables 31 and 32 that refer to the mathematics test. Considering upward constraint for the ESCS index (broadening the definition), samples' increase is greater than in reading, except for the cities of Milan and Padua (especially in Padua, the small dimension of the reading sample amplifies any percentage increase or decrease and make it less comparable). On the contrary, when considering backward constraint of the ESCS index, the decrease in the number of resilient students in mathematics is similar to the one recorded in reading.

Table 32 shows results obtained when changing the score threshold. Also in this case we note that, for mathematics, the number of resilient students increases more than for reading (especially in Rome and Turin) while our definition appears to be much less sensitive to the backward variation, highlighting a decrease between 35 and 70 percent.

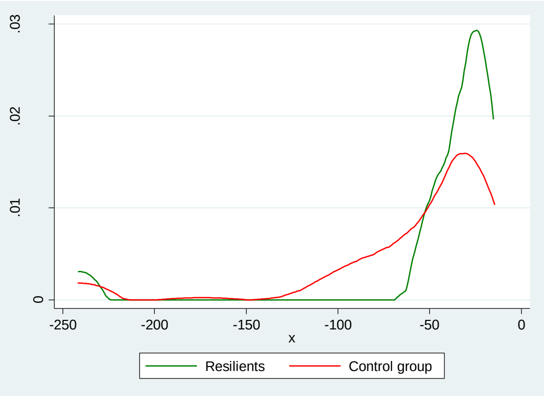
Summarizing, when comparing the two groups, the definition of mathematics resilient students is more sensitive to an upward variation of the threshold both in the socioeconomic condition and the score. On the contrary, the reduction among samples is quite similar for all cities. More in general, we record greater variation in the sample size when we consider score's constraint, suggesting a greater sensitive of the definition to score variation.

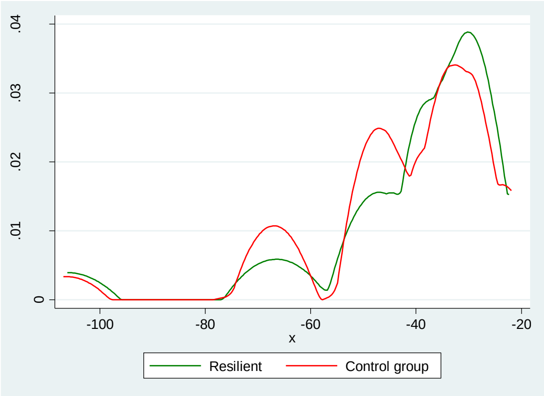
Moreover, we analyze the test scores distribution of resilient students and control group in order to investigate whether the variation in the definition (both upward and backward) has affected the distribution curve. We focus on test score distribution in grade 5 to understand how the difference in the distribution of the two groups (which was marked in the original definition, see section 3.3.1) would change when the definition of resilient student changes. Figures 19-23 show that, broadening the definition (defining resilient students those who, in grade 6, obtain a score higher than

200 points), resilient students and control group tend to show more similar distribution than in the original definition. This could be due to the fact that, defining a lower threshold, we include in the sample of resilient students more children who, also in grade 5, performed on average very close to the control group. In fact, on the basis of this new definition, the discrimination between being resilient or not, is really slight.

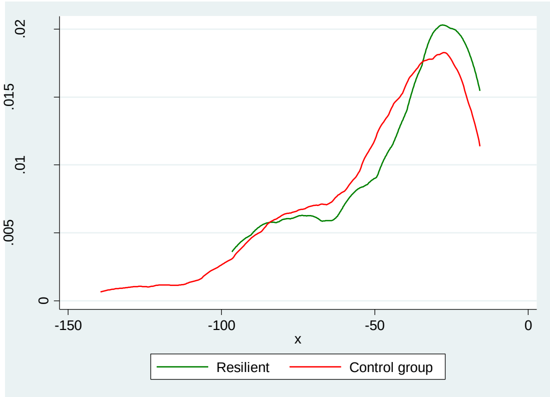
On the contrary, figures 24-28 about backward constraints show different patterns for reading and mathematics test score. Considering the reading distribution we observe that resilient students perform on average closer to the control group with respect to the original definition. On the contrary, mathematics distribution show a greater difference in the performance of the groups, suggesting that becoming resilient in mathematics can be more related to the ability endowed (which would explain a relative higher performance of resilient students also in grade 5).

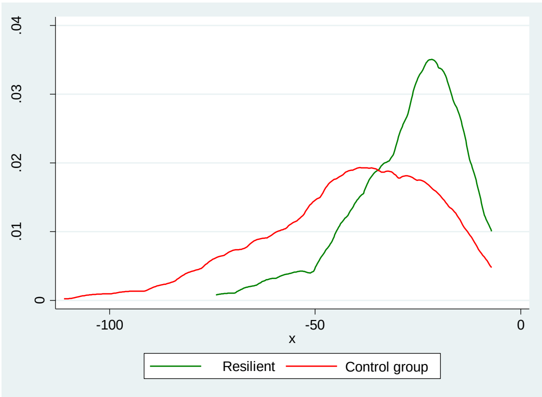
Graphs 16a and 16b. Broadening the definition of resilient student. Reading (on the left) and mathematics (on the right) test scores distribution in grade 5 (2011/2012), Bologna – kernel density distribution.



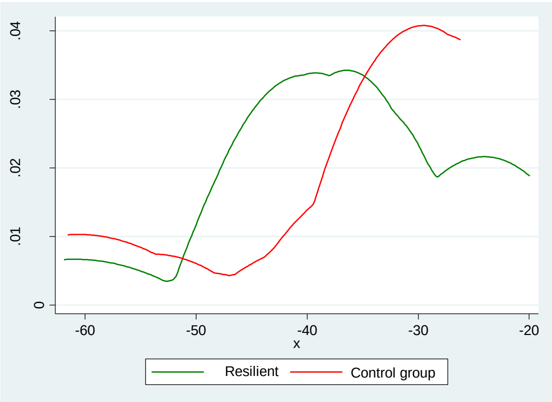


Graphs 17a and 17b. Broadening the definition of resilient student. Reading (on the left) and mathematics (on the right) test scores distribution in grade 5 (2011/2012), Milan – kernel density distribution.



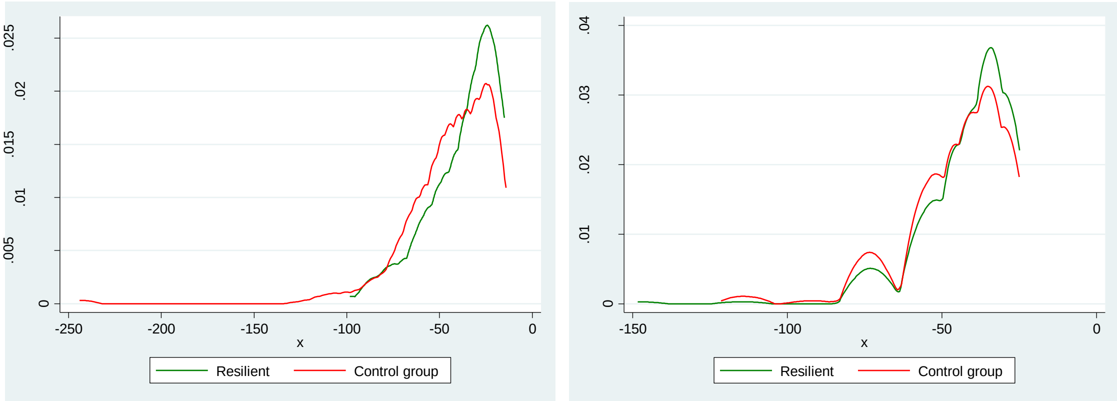


Graph 18. Broadening the definition of resilient student. Mathematics test scores distribution in grade 5 (2011/2012), Padua – kernel density distribution.

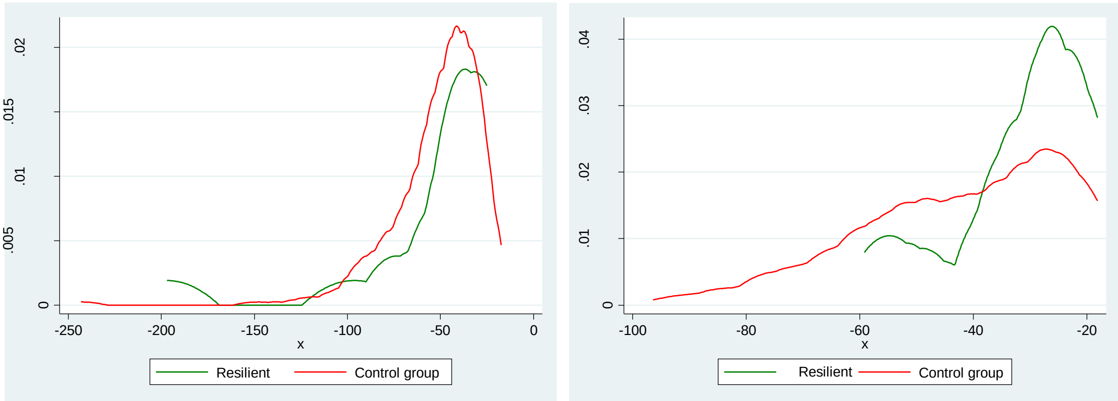


Note: The missing graph on reading's distribution is due to lack of observations in the original definition (two only resilient students).

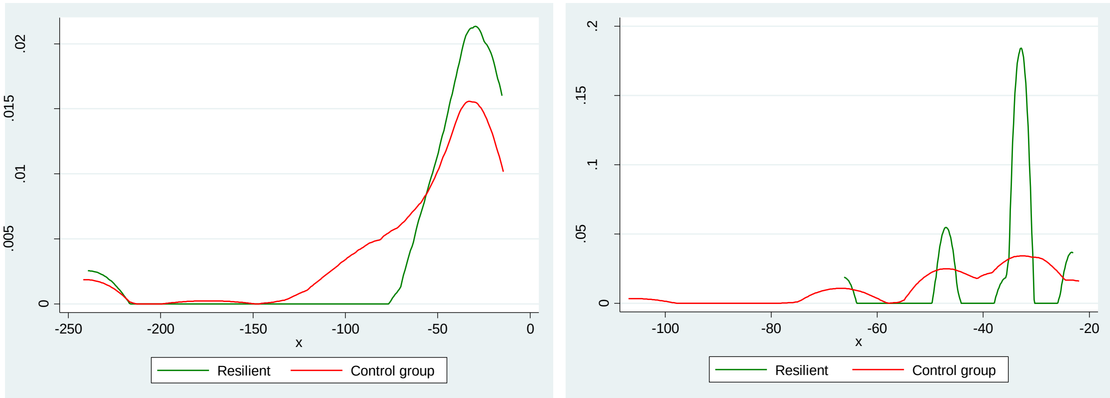
Graphs 19a and 19b. Broadening the definition of resilient student. Reading (on the left) and mathematics (on the right) test scores distribution in grade 5 (2011/2012), Rome – kernel density distribution.



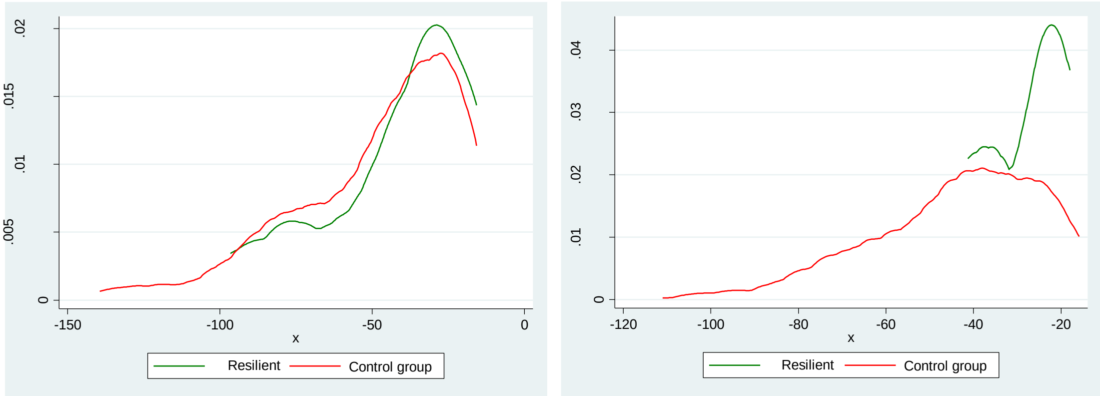
Graphs 20a and 20b. Broadening the definition of resilient student. Reading (on the left) and mathematics (on the right) test scores distribution in grade 5 (2011/2012), Turin – kernel density distribution.



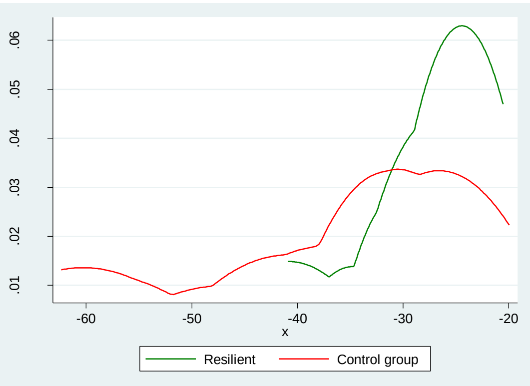
Graphs 21a and 21b. Restricting the definition of resilient student. Reading (on the left) and mathematics (on the right) test scores distribution in grade 5 (2011/2012), Bologna – kernel density distribution.



Graphs 22a and 22b. Restricting the definition of resilient student. Reading (on the left) and mathematics (on the right) test scores distribution in grade 5 (2011/2012), Milan – kernel density distribution.

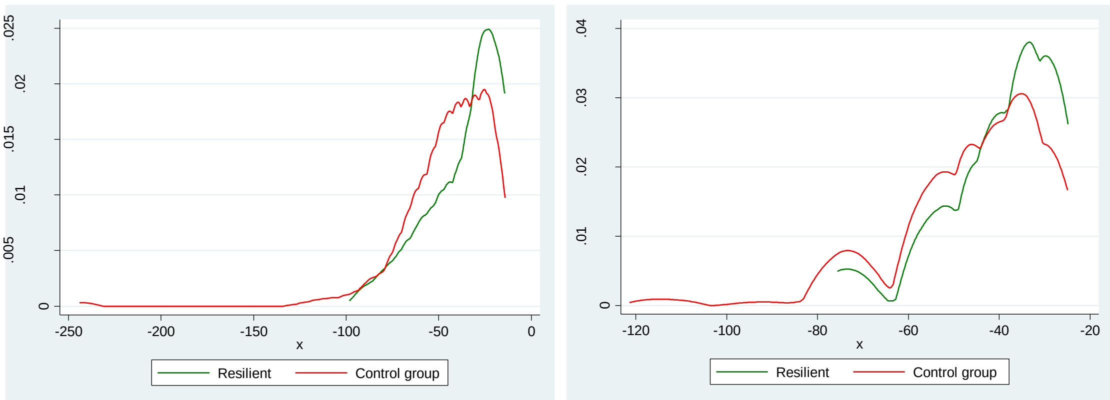


Graphs 23. Restricting the definition of resilient student. Mathematics test scores distribution in grade 5 (2011/2012), Padua – kernel density distribution.

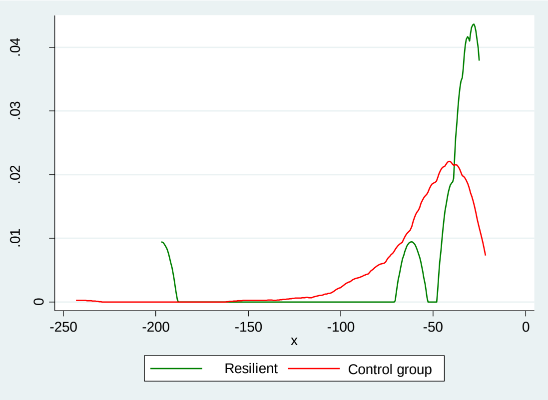


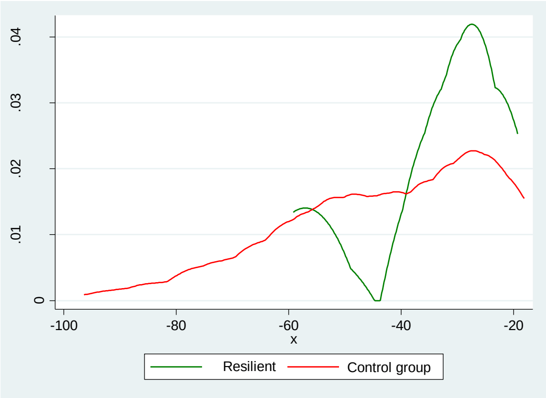
Note: The missing graph on reading's distribution is due to lack of observations in the original definition (two only resilient students).

Graphs 24a and 24b. Restricting the definition of resilient student. Reading (on the left) and mathematics (on the right) test scores distribution in grade 5 (2011/2012), Rome – kernel density distribution.



Graphs 25a and 25b. Restricting the definition of resilient student. Reading (on the left) and mathematics (on the right) test scores distribution in grade 5 (2011/2012), Turin – kernel density distribution.





3.3.3 Resilient schools and classes

3.3.3.1 Resilient schools

Having defined and described the group of resilient students, we want to investigate how they are distributed among schools. Table 33 show their concentration across schools with respect to the total number of schools per city.

Table 33. Distribution of resilient students across schools.

	Number of RES students (ita)	Number of schools where RES students are (ita)	Number of RES students (mat)	Number of schools where RES students are (mat)	Number of schools in the city
Bologna	18	10	20	10	27
Milan	50	33	49	27	110
Padua	2	2	19	9	18
Rome	93	56	174	57	234
Turin	22	17	16	15	56

Notes. Data from INVALSI 2012/2013.

Some schools have a higher concentration of resilient students than others, so we have selected and investigated a group of schools that we can define “resilient schools”. This group is formed by schools with a concentration of resilient students higher than $\mu + \sigma$, where μ is the mean of resilient students per school and σ is the standard deviation from the mean. The level of investigation is always represented by the city. Resilient students are, as always, considered separately for the reading and mathematics test. We lose information about resilient students in reading for Padua as we have just two resilient students who attend different school, so no further investigation can be done about these schools. Analyzing resilient schools across cities we can identify some recurring characteristics, shown in tables 34-38.

In general, mathematics-resilient schools tend to be attended by poorer students (measured by the school average ESCS index), except for the city of Padua. These schools have, on average, a higher test score, especially in the reading test. This is particularly clear when comparing results with the control group, but not with the city average. This finding suggests that resilient students are more concentrated in schools with better outcomes, whereas control group's students tend to attend lower-performing schools than the average of each city. Furthermore, resilient schools are bigger (both in term of number of students and number of classes per school) especially when selected upon reading score. The better outcome and the larger dimension could be both linked to unobservable factors like a better capacity of bigger schools to attract effective teachers, that results in better outcomes, even though a further investigation on this topic can't be carried on. Finally, private schools are generally under-represented, both in resilient and control group.

Table 34. Descriptive analysis of resilient schools, Bologna.

	RESILIENT SCHOOLS	RESILIENT SCHOOLS	CONTROL GROUP		CITY AVERAGE	
	ITA	MAT	ITA	MAT	ITA	MAT
School Average ESCS	0.40	-0.64	0.11		0.51	
School Average score	220.92	193.08	197.46	200.14	206.67	205.33
Proportion of immigrant students in the school	0.26	0.45	0.29		0.20	
Number of students in the school	152.00	120.00	128.01		97.46	
Number of classes in the school	6.00	5.00	5.27		4.00	
Proportion of private schools	0.00	0.00	0.12		0.33	
Proportion of *Istituto comprensivo*	1.00	1.00	0.82		0.63	
Number of students per class, in the school	25.40	24.00	24.58		23.43	
N	1	1	17		27	

Notes. Data from INVALSI 2012/2013.

Table 35. Descriptive analysis of resilient schools, Milan

	RESILIENT SCHOOLS	RESILIENT SCHOOLS	CONTROL GROUP		CITY AVERAGE	
	ITA	MAT	ITA	MAT	ITA	MAT
School Average ESCS	-0.01	-0.25	-0.04		0.43	
School Average score	207.70	207.33	200.01	204.21	209.35	210.54
Proportion of immigrant students in the school	0.28	0.35	0.29		0.23	
Number of students in the school	238.00	153.91	139.28		96.16	
Number of classes in the school	9.30	6.81	6.08		4.18	
Proportion of private schools	0.00	0.00	0.10		0.35	
Proportion of *Istituto comprensivo*	0.33	0.83	0.79		0.55	
Number of students per class, in the school	25.67	22.55	22.68		22.46	
N	3	6	71		110	

Notes. Data from INVALSI 2012/2013.

Table 36. Descriptive analysis of resilient schools, Padua

	RESILIENT SCHOOLS	CONTROL GROUP	CITY AVERAGE
	MAT	MAT	MAT
School Average ESCS	0.35	0.63	0.69
School Average score	215.81	215.74	214.82
Proportion of immigrant students in the school	0.22	0.15	0.15
Number of students in the school	177.79	108.82	88.93
Number of classes in the school	4.50	4.54	3.72
Proportion of private schools	0.00	0.15	0.33
Proportion of *Istituto comprensivo*	1.00	0.85	0.67
Number of students per class, in the school	25.50	24.23	23.80
N	2	13	18

Notes. Data from INVALSI 2012/2013.

Table 37. Descriptive analysis of resilient schools, Rome

	RESILIENT SCHOOLS	RESILIENT SCHOOLS	CONTROL GROUP		CITY AVERAGE	
	ITA	MAT	ITA	MAT	ITA	MAT
School Average ESCS	-0.05	-0.02	0.04		0.44	
School Average score	204.40	200.53	199.82	197.11	206.98	203.61
Proportion of immigrant students in the school	0.16	0.14	0.16		0.12	
Number of students in the school	157.05	140.07	136.90		94.45	
Number of classes in the school	6.68	6.16	6.18		4.21	
Proportion of private schools	0.00	0.00	0.14		0.30	
Proportion of *Istituto comprensivo*	1.00	1.00	0.86		0.70	
Number of students per class, in the school	23.46	22.77	21.90		21.92	
N	10	9	127		110	

Notes. Data from INVALSI 2012/2013.

Table 38. Descriptive analysis of resilient schools, Turin

	RESILIENT SCHOOLS	RESILIENT SCHOOLS	CONTROL GROUP		CITY AVERAGE	
	ITA	MAT	ITA	MAT	ITA	MAT
School Average ESCS	0.23	0.24	-0.06		0.25	
School Average score	203.25	208.81	200.86	200.88	204.95	205.22
Proportion of immigrant students in the school	0.29	0.08	0.21		0.20	
Number of students in the school	219.75	230.00	158.56		124.08	
Number of classes in the school	9.50	10.00	6.68		5.34	
Proportion of private schools	0.00	0.00	0.11		0.34	
Proportion of *Istituto comprensivo*	0.50	0.00	0.44		0.38	
Number of students per class, in the school	19.75	23.00	22.84		22.13	
N	4	1	36		56	

Notes. Data from INVALSI 2012/2013.

3.3.3.2 Resilient classes

Having defined as resilient those schools where the concentration of resilient students is greater than the mean (of resilient students per school) plus a standard deviation from the mean, we use the same cut point to define resilient classes in resilient schools. This means that, starting from the sample of resilient schools, we select classes where the number of resilient students is higher than the mean plus a standard deviation. In this process we lose the observation of resilient classes in Turin in Mathematics as none concentration of students can be noticed in the only resilient school. The sample of resilient classes is then compared with the average results of control group and of the city. Analyzing the tables 39-43, we notice that resilient classes tend to have an average ESCS index lower than both the control group and the city average. As this pattern is repeated for both resilient schools and classes (except for Turin, as mentioned above and partially for Bologna), we could deduce a remarkable aspect: Disadvantaged Low Achievers (defined in grade 5) are more likely to concentrate, in grade 6, not only in schools but also in classrooms characterized by lower socioeconomic condition. This occurs independently from the set which they belong to (resilient students or control group). Nevertheless, resilient students are concentrated in schools and classrooms whose average ESCS index is frequently lower than the control group. Moreover, class average scores are higher in resilient classes than in control group especially when considering the reading test, suggesting that a possible peer effect could help students to improve reading attainment more than mathematics'. These classes are also attended by a higher proportion of female students than resilient classes selected on the mathematics score, which could be another aspect to take into consideration when referring to a possible peer effect.

Table 39. Descriptive analysis of resilient classes, Bologna

	RESILIENT CLASSES	RESILIENT CLASSES	CONTROL GROUP		CITY AVERAGE	
	ITA	MAT	ITA	MAT	ITA	MAT
Class Average ESCS	0.087	-0.616	0.091		0.489	
Class Average score	211.66	194.84	196.97	199.10	207.75	207.94
Proportion of immigrant students in the classroom	0.22	0.55	0.31		0.198	
Proportion of female students in the classroom	0.43	0.39	0.47		0.498	
Proportion of early-enrolled students in the classroom	0	0	0.007		0.015	
Proportion of late-enrolled students in the classroom	0.09	0.11	0.13		0.085	
Number of students per classroom	25	24	24.58		24.21	
N	1	1	58		104	

Notes. Data from INVALSI 2012/2013.

Table 40. Descriptive analysis of resilient classes, Milan

	RESILIENT CLASSES	RESILIENT CLASSES	CONTROL GROUP		CITY AVERAGE	
	ITA	MAT	ITA	MAT	ITA	MAT
Class Average ESCS	-0.049	-0.195	-0.119		0.392	
Class Average score	223.66	220.07	196.90	201.85	209.04	211.12
Proportion of immigrant students in the classroom	0	0.38	0.33		0.220	
Proportion of female students in the classroom	0.57	0.38	0.47		0.497	
Proportion of early-enrolled students in the classroom	0	0	0.008		0.013	
Proportion of late-enrolled students in the classroom	0	0.13	0.12		0.079	
Number of students per classroom	24	22	22.68		22.934	
N	1	2	244		442	

Notes. Data from INVALSI 2012/2013.

Table 41. Descriptive analysis of resilient classes, Padua

	RESILIENT CLASSES	CONTROL GROUP	CITY AVERAGE
	MAT	MAT	MAT
Class Average ESCS	0.169	0.500	0.623
Class Average score	208.672	213.018	214.993
Proportion of immigrant students in the classroom	0.167	0.171	0.147
Proportion of female students in the classroom	0.571	0.473	0.508
Proportion of early-enrolled students in the classroom	0	0.007	0.008
Proportion of late-enrolled students in the classroom	0.143	0.114	0.091
Number of students per classroom	26	23.714	23.633
N	2	35	60

Notes. Data from INVALSI 2012/2013.

Table 42. Descriptive analysis of resilient classes, Rome

	RESILIENT CLASSES	RESILIENT CLASSES	CONTROL GROUP		CITY AVERAGE	
	ITA	MAT	ITA	MAT	ITA	MAT
Class Average ESCS	-0.240	-0.312	-0.028		0.367	
Class Average score	209.66	203.41	196.605	194.421	207.538	204.679
Proportion of immigrant students in the classroom	0.14	0.12	0.17		0.122	
Proportion of female students in the classroom	0.52	0.42	0.49		0.495	
Proportion of early-enrolled students in the classroom	0.008	0.02	0.01		0.025	
Proportion of late-enrolled students in the classroom	0.10	0.01	0.09		0.061	
Number of students per classroom	24.36	22.2	21.90		22.237	
N	5	14	342		895	

Notes. Data from INVALSI 2012/2013.

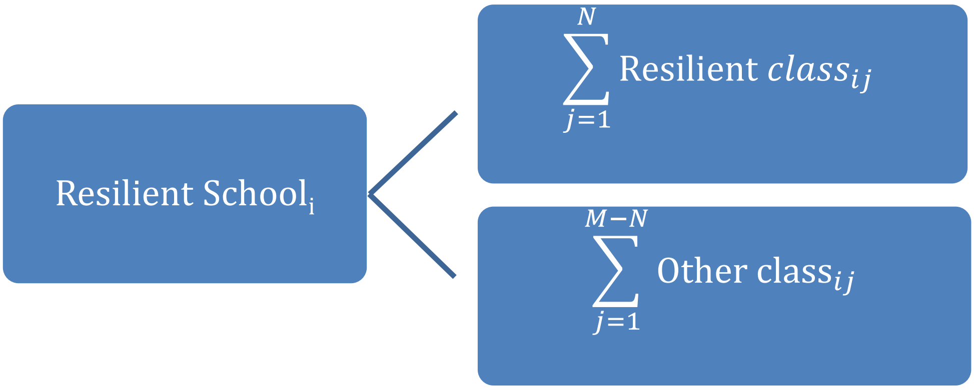
Table 43. Descriptive analysis of resilient classes, Turin

	RESILIENT CLASSES	CONTROL GROUP	CITY AVERAGE
	ITA	ITA	ITA
Class Average ESCS	-0.093	-0.123	0.142
Class Average score	214.388	199.429	205.356
Proportion of immigrant students in the classroom	0.212	0.238	0.194
Proportion of female students in the classroom	0.576	0.510	0.497
Proportion of early-enrolled students in the classroom	0.030	0.005	0.006
Proportion of late-enrolled students in the classroom	0.242	0.115	0.099
Number of students per classroom	20	23.153	23.20
N	2	137	275

Notes. Data from INVALSI 2012/2013.

In order to verify last considerations, we make a comparison between classes in resilient schools. Considered the group of resilient schools, we compare the characteristics of resilient classes (the same selected above) with all the other classes of the same school. The figure below represents this process of selection. The index i refers to the resilient school, j refers to the class, M is the number of classes in the school and N the number of resilient classes in the school (figure 29).

Figure 4. Within school comparison, the selection process.



This enables us to demonstrate a further consideration: classes with a higher concentration of resilient students have a lower average ESCS index than the other classes of the same school.

Results are reported in tables 44-48. The first column contains the variables investigated, the second contains the average characteristics of resilient classes and the third refers to the average features of the remaining classes of resilient schools. Actually, resilient students are more concentrated in classes with lower socioeconomic condition, with the exception of Milan. Moreover, reading average test scores are generally higher in resilient classes, except for the city of Bologna.

Table 44. Within school comparison, Bologna.

	RESILIENT CLASSES	OTHER CLASSES IN RESILIENT SCHOOLS	RESILIENT CLASSES	OTHER CLASSES IN RESILIENT SCHOOLS
	ITA	ITA	MAT	MAT
Class Average ESCS	0.087	0.445	-0.616	-0.599
Class Average score	211.66	222.430	194.84	194.505
Proportion of immigrant students in the classroom	0.22	0.21	0.55	0.55
Proportion of female students in the classroom	0.43	0.55	0.39	0.45
Proportion of early-enrolled students in the classroom	0	0.008	0	0.016
Proportion of late-enrolled students in the classroom	0.09	0.07	0.11	0.18
Number of students per classroom	25	25.38	24	24.43
N	1	5	1	4

Notes. Data from INVALSI 2012/2013.

Table 45. Within school comparison, Milan.

	RESILIENT CLASSES	OTHER CLASSES IN RESILIENT SCHOOLS	RESILIENT CLASSES	OTHER CLASSES IN RESILIENT SCHOOLS
	ITA	ITA	MAT	MAT
Class Average ESCS	-0.049	-0.257	-0.195	-0.331
Class Average score	223.66	211.27	220.07	207.994
Proportion of immigrant students in the classroom	0	0.23	0.38	0.40
Proportion of female students in the classroom	0.57	0.56	0.38	0.42
Proportion of early-enrolled students in the classroom	0	0.03	0	0
Proportion of late-enrolled students in the classroom	0	0.05	0.13	0.13
Number of students per classroom	24	23.92	22	21.65
N	1	9	2	12

Notes. Data from INVALSI 2012/2013.

Table 46. Within school comparison, Padua.

	RESILIENT CLASSES	OTHER CLASSES IN RESILIENT SCHOOLS
	MAT	MAT
Class Average ESCS	0.169	0.345
Class Average score	208.672	214.009
Proportion of immigrant students in the classroom	0.167	0.250
Proportion of female students in the classroom	0.571	0.467
Proportion of early-enrolled students in the classroom	0	0
Proportion of late-enrolled students in the classroom	0.143	0.158
Number of students per classroom	26	24.86
N	2	7

Notes. Data from INVALSI 2012/2013.

Table 47. Within school comparison, Rome.

	RESILIENT CLASSES	OTHER CLASSES IN RESILIENT SCHOOLS	RESILIENT CLASSES	OTHER CLASSES IN RESILIENT SCHOOLS
	ITA	ITA	MAT	MAT
Class Average ESCS	-0.240	-0.122	-0.312	-0.060
Class Average score	209.66	200.925	203.41	197.268
Proportion of immigrant students in the classroom	0.14	0.17	0.12	0.16
Proportion of female students in the classroom	0.52	0.47	0.42	0.49
Proportion of early-enrolled students in the classroom	0.008	0.009	0.02	0.005
Proportion of late-enrolled students in the classroom	0.10	0.09	0.01	0.09
Number of students per classroom	24.36	22.35	22.2	23.72
N	5	29	14	14

Notes. Data from INVALSI 2012/2013.

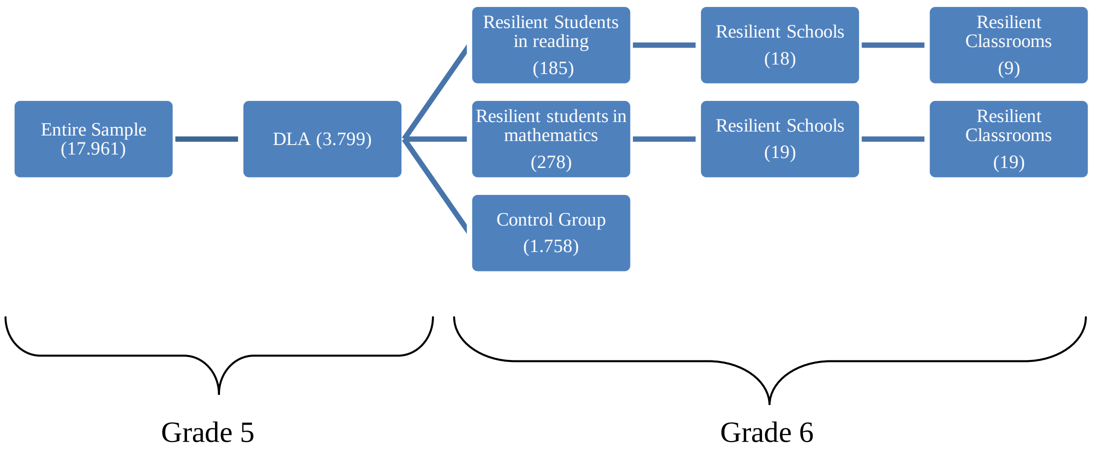
Table 48. Within school comparison, Turin.

	RESILIENT CLASSES	OTHER CLASSES IN RESILIENT SCHOOLS
	ITA	ITA
Class Average ESCS	-0.093	0.469
Class Average score	214.388	209.612
Proportion of immigrant students in the classroom	0.212	0.289
Proportion of female students in the classroom	0.576	0.511
Proportion of early-enrolled students in the classroom	0.030	0.006
Proportion of late-enrolled students in the classroom	0.242	0.108
Number of students per classroom	20	23.606
N	2	33

Notes. Data from INVALSI 2012/2013.

Finally, the figure below represents the entire process of selection made. In parenthesis is reported the sample size. Numbers refer to the five cities considered globally (figure 30). It is noteworthy to mention that reading resilient students are more likely to concentrate in a smaller number of classes, starting from a similar number of schools.

Figure 5. The selection process and sample size.



Notes. The entire sample refers to the five cities jointly considered. DLA is the acronym for Disadvantaged low achiever students.

3.4 Methodology

In this section we investigate which class and school-level factors are more related to resiliency. Implementing a probit regression and a propensity score matching, we find that school and class features actually have an impact on student's ability to succeed.

3.4.1 Probit Regression Models

In order to investigate which school and class factors are more related to resiliency, we consider an educational production function where the estimated outcome is a dichotomous variable which identify resilient students, being equal to 1 when the student is resilient and 0 otherwise (the definition of resilient student coming from section 3.1). Aiming at investigating how the variation in class/school-level characteristics affects the probability to be resilient, we consider two different cluster of explanatory variables: (I) characteristics in grade 5, those of the school/classroom where the student spent the first level of education, (II) the variation in the same characteristics between grade 5 and 6 (being grade 6 the new class/school in which the student has been tested). School and class-level variables used in the model to predict resiliency are: school/classroom average socioeconomic background (ESCS index), school/classroom average test score, number of students in the school/classroom, number of immigrants in the school/classroom and the number of classes in the school.

Class and school-level features are investigated separately because of collinearity and with the aim to check if the characteristics that are more predictive are class or school-level factors. We also include individual-level characteristics (always considered as time-invariant) in order to include all explanatory variables.
Analytically, the equation of the probit regression can be expressed as follows:

$$y_{ij}(Y_i = 1|Y_i = 0) = f\left(X_{1_{ij}}, X_{2_j}, \Delta X_{3_j}\right) + \varepsilon_{ij} \tag{8}$$

where y_{ij} is the outcome of the ith student in the jth class/school, X_1 is the vector containing his/her own individual characteristics, X_2 refers to the characteristics of the class/school attended and ΔX_3 contains the variation in those characteristics switching from grade 5 to 6. More specifically, let *t* correspond to grade 6 and *t – 1* correspond to grade 5. We estimate the following:

$$y_{ij(t)} = \alpha_0 + \alpha_1 \bar{X}_{1_{ij(t-1)}} + \alpha_2 \bar{X}_{2_{j(t-1)}} + \alpha_3 \Delta \bar{X}_{3_{j(t,t-1)}} + \varepsilon_{ij} \tag{9}$$

where $y_{ij(t)}$ is a dummy variable for the ith student being resilient (when It is equal to 1) or control group student (when It is equal to 0) attending the jth classroom identified at time t (grade 6); $\bar{X}_{1_{ij(t-1)}}$ refers to individual characteristics at time *t – 1* (also controlling for prior achievement, grade 5); $\bar{X}_{2_{j(t-1)}}$ contains the features of the jth classroom attended by the student at grade 5 and $\Delta\bar{X}_{3_{j(t,t-1)}}$ corresponds to the variation in class-level characteristics between period *t* and *t – 1*. Our reference sample is composed by resilient students and control group from each city separately, but the model is also applied on the cities considered jointly; in this latter case, a dummy variable for the city permits to control for structural differences among the cities. In this analysis, we do not consider results for reading resilient students in the city of Padua because of the small number in the sample (only two resilient students) that would have made any further investigation not statistically significant. Standard errors are robust and clustered at school-level, in order to catch the within-school correlation between school factors.

In the alternative model, we substitute class-level with school-level characteristics (represented by the index w) and estimate the following equation:

$$y_{iw(t)} = \alpha_0 + \alpha_1 \bar{X}_{1_{iw(t-1)}} + \alpha_2 \bar{X}_{2_{w(t-1)}} + \alpha_3 \Delta \bar{X}_{3_{w(t,t-1)}} + \varepsilon_{iw} \tag{10}$$

where variables considered are the same descripted above except for $\bar{X}_{2_{w(t-1)}}$ and $\Delta \bar{X}_{3_{w(t,t-1)}}$, that describe respectively school's features in grade 5 and the variation in school features between grade 5 and 6.

Results from the probit regression are reported in tables 49 and 50. Every table has two columns for each city, one related to reading test and one to mathematics test. The last column refers to the entire sample.

Table 49 recaps the results of equation (9) and shows that the most relevant individual characteristics are the socioeconomic background and the prior achievement (test score in grade 5). Both are positively related with the probability of becoming resilient and are statistically significant at 1% level. This finding is reasonable when considering that among disadvantaged low achievers, students who better perform and have relatively higher SES background are more likely to become resilient. In fact, we check for these two variables in order to catch the relative position of the student along the city distribution. Among other variables, being a second generation immigrant has a negative correlation with resilience, though results for first generation immigrant status are not statistically significant. Considering classroom's characteristics, a relevant aspect is that the class average socioeconomic background, in grade 5, has a negative correlation with the outcome, even though the most predictive variable is the class average score that it is positively related to the probability of being resilient. These outcomes are in line with those expressed by Hanushek *et al.* (2003), who found that peers' achievement has a positive effect on the outcomes' growth. Finally, studying the change in classroom characteristics between grade 5 and 6, the most relevant variable is the variation in the classroom average score. We find that attending a class, at grade 6, where peers achieve better results than those in grade 5 is positively related with resiliency. Another characteristic that has a positive influence on the resilience is the variation in the number of immigrant students in the classroom. From the table stands out that attending a class with a higher number of immigrant students in grade 6 have a small but positive relation with resiliency. The literature generally agrees about the negative impact of immigrant students on peers' attainment, but some researches provide evidences of a different tendency. Ballatore, Fort & Ichino (2013) analyze, in the Italian context, the effect of increasing the number of immigrant students keeping class size constant and find that the negative impact on peers' performance vanish when students attend grade 5.

So, summarizing results at class-level, we can highlight that the most significant result is linked to the average class score, considered both as initial condition (at grade 5) and as positive variation between grade 5 and 6. This entails two main issues: on the one hand, having attended (in grade 5) a class with peers with relatively higher achievement scores increase the probability to be resilient; on the other hand, attending a class (in grade 6) where the average score is higher than the previous year has a positive correlation with the resilient status.

Analyzing the correlation between school-level characteristics and the probability to be resilient, results are consistent with what we described at class-level (and are presented in table 50). In particular, among individual variables, the previous score (grade 5) and the socio-economic background (in terms of both initial condition in grade 5 and variation between the two years) are the most predictive factors. Among other aspects, being a female student is not highly predictive, but generally related to a higher probability to become resilient in reading than in mathematics, consistently with the existing literature e.g. Agasisti & Longobardi (2014). Moreover, being a first generation immigrant has a negative correlation with resiliency. Among school characteristics in grade 5, the school average score stands out as the most predictive factor and it shows a positive correlation with the probability of being resilient. Similar results can be noted among delta variables, as the higher is the average score of school-peers in grade 6 (with respect to the previous year), the higher is the probability to become resilient. Moreover, attending a school with a fewer number of classes in grade 6 than in grade 5, is negatively related to resiliency, suggesting a negative effect of attending a bigger school.

In conclusion, results show that having attended a class or a school (in grade 5) where peers perform averagely high has a positive correlation with the probability to become a resilient student. More importantly, attending in grade 6 a class or a school where peers achieve better performances than previous year has a positive relationship with resilience.

Table 49. Results from the probit regression (class-level characteristics).

	BOLOGNA		MILAN		PADUA	ROME		TURIN		ENTIRE SAMPLE	
	ITA	MAT	ITA	MAT	MAT	ITA	MAT	ITA	MAT	ITA	MAT
	N=150	N=131	N=632	N=569	N=63	N=714	N=660	N=238	N=247	N=1,734	N=1,629
Variable											
Stident-level characteristics											
Milan (dummy)										-0.011	**0.063***
										(0.025)	(0.038)
Turin (dummy)										**-0.047****	**-0.072*****
										(0.028)	(0.035)
Rome (dummy)										0.04	**0.123*****
										(0.031)	(0.036)
ESCS (grade 5)	**0.06*****	**0.17*****	0.046	0.004	0.164	0.029	**0.099*****	-0.003	0.006	0.019	**0.046*****
	(0.032)	(0.054)	(0.021)	*(0.012)*	*(0.119)*	(0.025)	*(0.031)*	(0.006)	(0.007)	(0.013)	(0.0163)
Prior achievement (grade 5)	0.001	**0.004****	**0.002*****	**0.002*****	0.002	**0.003*****	0.001	**0.0008*****	0.0001	**0.02*****	**0.002*****
	(0.0007)	(0.002)	(0.0004)	*(0.0003)*	*(0.001)*	(0.0009)	*(0.001)*	(0.0004)	(0.0001)	(0.013)	(0.0007)
Females	0.0625	-0.006	0.014	0.005	-0.220	-0.001	-0.044	-0.0005	0.001	0.007	0.005
	(0.056)	(0.073)	(0.018)	*(0.011)*	*(0.153)*	(0.024)	*(0.030)*	(0.006)	(0.002)	(0.013)	(0.017)
First generation immigrant	0.016	-0.045	0.002	-0.002	**-0.235*****	0.072	0.112	0.025	0.025	0.028	0.008
	(0.062)	(0.070)	(0.027)	*(0.019)*	*(0.108)*	(0.006)	*(0.046)*	(0.020)	(0.015)	(0.021)	(0.027)
Second generation immigrant	-0.020	0.049	**-0.040***	**-0.019****	-0.101	-0.002	-0.028	0.068	0.014	-0.006	-0.022
	(0.058)	(0.056)	(0.002)	*(0.013)*	*(0.120)*	(0.042)	*(0.028)*	(0.035)	(0.014)	(0.019)	(0.021)
Characteristics of classroom attended at grade 5											
Class average socio-economic background (grade 5)	-0.037	**-0.188*****	-0.024	-0.020	-0.125	-0.024	**-0.116****	0.007	-0.0001	-0.021	-0.030
	(0.087)	(0.090)	(0.030)	*(0.026)*	*(0.128)*	(0.037)	*(0.065)*	(0.014)	(0.001)	(0.020)	(0.039)

Class average score (grade 5)	0.001	-0.0007	0.001	**0.002****	0.002	**0.004*****	**0.004*****	0.00	-0.0001	**0.002*****	**0.002*****
	(0.002)	(0.0003)	(0.001)	*(0.001)*	*(0.002)*	(0.001)	*(0.001)*	(0.00)	(0.0001)	(0.0007)	(0.001)
Number of immigrants in the classroom (grade 5)	-0.001	-0.021	0.001	0.003	**0.071*****	-0.001	0.001	-0.003	-0.0004	-0.0007	-0.0001
	(0.014)	(0.013)	(0.004)	*(0.005)*	*(0.032)*	(0.009)	*(0.008)*	(0.001)	(0.0002)	(0.003)	(0.0003)
Number of students in the classroom (grade 5)	**0.028*****	0.001	-0.003	-0.006	0.03	0.007	0.005	-0.0006	-0.0002	0.002	0.0001
	(0.012)	(0.001)	(0.004)	*(0.005)*	*(0.025)*	(0.005)	*(0.008)*	(0.002)	(0.0002)	(0.003)	(0.0003)
Change in the classroom characteristics between grade 5 and 6											
Delta class average score	0.003	**0.005*****	**0.003*****	**0.003*****	0.004	**0.007*****	**0.009*****	0.0002	0.0001	**0.004*****	**0.005*****
	(0.002)	(0.002)	(0.001)	*(0.001)*	*(0.003)*	(0.0009)	*(0.001)*	(0.0002)	(0.0001)	(0.0006)	(0.0008)
Delta number of students in the classroom	0.004	0.009	**-0.001*****	-0.002	-0.007	0.004	0.003	0.0003	-0.0002	-0.0002	-0.0001
	(0.001)	(0.012)	(0.003)	*(0.005)*	*(0.004)*	(0.004)	*(0.006)*	(0.001)	(0.0001)	(0.002)	(0.003)
Delta class average socio-economic background	0.054	-0.100	-0.001	-0.004	-0.063	-0.014	-0.032	0.003	0.001	-0.009	-0.001
	(0.081)	(0.096)	(0.002)	*(0.033)*	*(0.170)*	(0.028)	*(0.058)*	(0.006)	(0.002)	(0.014)	(0.029)
Delta number of immigrants in the classroom	0.008	-0.001	**0.014*****	**0.005*****	0.031	-0.002	0.0002	0.001	-0.0002	**0.009*****	**0.007*****
	(0.009)	(0.013)	(0.003)	*(0.003)*	*(0.025)*	(0.007)	*(0.005)*	(0.001)	(0.0001)	(0.002)	(0.003)
Constant	-5.95	0.2	-2.12	-3.9	-4.04	-4.92	-4.51	1.27	7.09	-3.40	-3.94
	(3.58)	(2.96)	(1.61)	*(2.50)*	*(4.98)*	(1.33)	*(1.40)*	(4.24)	(3.28)	(0.97)	(1.04)
Pseudo R^2	0.16	0.22	0.12	0.20	0.29	0.13	0.12	0.27	0.31	0.13	0.17

NOTES: *** is statistically significant at 1%; ** 5%, * 10%. Robust standard errors in parentheses are clustered at school level.

Table 50. Results from the probit regression (school-level characteristics).

	BOLOGNA		MILAN		PADUA	ROME		TURIN		ENTIRE SAMPLE	
	ITA	MAT	ITA	MAT	MAT	ITA	MAT	ITA	MAT	ITA	MAT
	N=150	N=131	N=632	N=569	N= 63	N=714	N=660	N=238	N=247	N=1,734	N=1,629
Variable											
Student-level characteristics											
Milan (dummy)										-0.013	**0.113*** **
										(0.033)	(0.05)
Turin (dummy)										0.0320	0.108
										(0.004)	(0.07)
Rome (dummy)										0.056	**0.162*** **
										(0.003)	(0.048)
Socioeconomic background (ESCS) (grade 5)	**0.053*** **	**0.068*** **	-0.001	0.0020	0.116	0.024	**0.008*** **	0.006	0.014	0.018	**0.041*** **
	(0.031)	(0.032)	(0.002)	*(0.011)*	*(0.078)*	(0.025)	*(0.029)*	(0.003)	(0.021)	(0.001)	**(0.016)**
Prior achievement (grade 5)	0.0006	**0.002*** **	**0.002*** **	**0.002*** **	0.001	**0.003*** **	0.002	**0.001***	**0.002*** **	**0.002*** **	**0.003*** **
	(0.0001)	(0.0001)	(0.0004)	*(0.0005)*	*(0.001)*	(0.0008)	*(0.001)*	(0.0009)	(0.0006)	(0.0005)	(0.0007)
Females students	0.0690	-0.015	0.013	0.007	-0.231	-0.014	**-0.054***	-0.042	-0.013	-0.001	-0.0003
	(0.005)	(0.003)	(0.001)	*(0.009)*	*(0.174)*	(0.025)	*(0.032)*	(0.032)	(0.02)	(0.013)	(0.007)
First generation immigrant	-0.002	-0.032	0.0003	0.002	**-0.204*** **	0.073	0.089	**-0.073** **	**-0.034*** **	0.005	-0.015
	(0.051)	(0.025)	(0.002)	*(0.001)*	*(0.103)*	(0.048)	*(0.06)*	(0.026)	(0.018)	(0.02)	(0.025)
Second generation immigrant	-0.016	0.027	-0.043	-0.012	-0.22	-0.006	-0.032	-0.019	-0.034	-0.028	-0.039
	(0.060)	(0.003)	(0.023)	*(0.010)*	*(0.119)*	(0.043)	*(0.058)*	(0.034)	(0.021)	(0.018)	(0.022)
Characteristics of school attended at grade 5											
School average socio-economic background (grade 5)	**-0.281*** **	-0.001	-0.053	-0.028	**0.462*** **	-0.07	**-0.140** **	0.070	**0.060*** **	**-0.045** **	-0.024
	(0.095)	(0.005)	(0.004)	*(0.025)*	*(0.166)*	(0.004)	*(0.07)*	(0.005)	(0.034)	(0.025)	(0.038)

School average score	**0.011*** **	**0.012*** **	**0.002* **	**0.003*** **	0.010	**0.006*** **	0.003	-0.003	-0.001	**0.003*** **	0.002
	(0.004)	(0.002)	(0.001)	*(0.001)*	*(0.008)*	(0.002)	*(0.002)*	(0.003)	(0.001)	(0.001)	(0.001)
Number of immigrant students in the school (grade 5)	0.0080	-0.003	0.001	0.0010	**0.026*** **	0.0002	0.0010	-0.0009	0.00	0.0001	0.0003
	(0.005)	(0.003)	(0.001)	*(0.0008)*	*(0.005)*	(0.0001)	*(0.003)*	(0.001)	(0.00)	(0.0009)	(0.001)
Number of classes in the school (grade 5)	-0.134	0.0022	-0.002	0.017	**-0.318* **	**-0.059*** **	-0.018	**0.051*** **	0.005	-0.019	-0.008
	(0.123)	(0.067)	(0.016)	*(0.014)*	*(0.184)*	(0.002)	*(0.031)*	(0.027)	(0.002)	(0.013)	(0.018)
Number of students in the school (grade 5)	0.0020	-0.003	-0.0006	**-0.001*** **	**0.019*** **	**0.002*** **	0.0006	-0.001	-0.0003	0.0004	-0.0004
	(0.005)	(0.003)	(0.0001)	*(0.0006)*	*(0.004)*	(0.002)	*(0.001)*	(0.001)	(0.0008)	(0.0005)	(0.0007)
Change in the school characteristics between grade 5 and 6											
Delta school average socio-economic background	-0.230	**-0.134*** **	-0.030	-0.024	0.709	-0.037	-0.106	-0.016	0.005	-0.031	-0.016
	(0.171)	(0.006)	(0.04)	*(0.024)*	*(0.444)*	(0.003)	*(0.07)*	(0.071)	(0.043)	(0.024)	(0.038)
Delta number of immigrants in the school	-0.0003	-0.001	**0.003*** **	**0.001*** **	**0.012*** **	0.001	0.001	**0.002*** **	-0.0006	**0.028*** **	**0.002*** **
	(0.002)	(0.001)	(0.0008)	*(0.0009)*	*(0.005)*	(0.001)	*(0.001)*	(0.0001)	(0.0001)	(0.0008)	(0.0008)
Delta school average score	**0.012*** **	**0.003*** **	**0.005*** **	**0.004*** **	0.002	**0.008*** **	**0.008*** **	0.0010	**0.0009* **	**0.059*** **	**0.006*** **
	(0.004)	(0.001)	(0.001)	*(0.001)*	*(0.004)*	(0.002)	*(0.002)*	(0.001)	(0.001)	(0.001)	(0.001)
Delta number of classes in the school	-0.007	-0.019	-0.007	-0.005	-0.133	**-0.048*** **	-0.031	0.0080	-0.019	**-0.021** **	-0.019
	(0.054)	(0.004)	(0.018)	*(0.017)*	*(0.129)*	(0.001)	*(0.023)*	(0.02)	(0.013)	(0.001)	(0.014)
Delta number of students in the school	-0.001	0.00	-0.0007	-0.0003	-0.001	**0.001*** **	0.0008	-0.0006	0.00	0.0002	0.0001
	(0.002)	(0.001)	(0.0006)	*(0.0005)*	*(0.001)*	(0.0007)	*(0.001)*	(0.0001)	(0.0005)	(0.0004)	(0.000)
Constant	-13.27	22.50	4.02	-8.18	8.50	5.47	-2.78	4.19	3.87	-4.14	-3.16
	(4.90)	(10.71)	(2.56)	*(2.61)*	*(35.38)*	(1.76)	*(1.73)*	(5.53)	(3.83)	(1.32)	(1.48)
Pseudo R^2	0.18	0.24	0.11	0.22	0.35	0.1	0.08	0.15	0.19	0.09	0.13

NOTES: *** is statistically significant at 1%; ** 5%, * 10%. Robust standard errors in parentheses are clustered at school level.

3.4.2 Propensity Score Matching

Having found that the average score of peers (both at classroom and school level) has a particularly high statistical significance, we design an empirical analysis to investigate whether being in a class with that peers, in grade 6, is able to influence the probability to become a resilient student. In so doing, we define two groups of students, assuming that the treatment is *"to be assigned to a class/school where peers have higher achievement scores than previous peers (grade 5)"*. Procedurally, we define: (I) treated students, who attend a class/school (in grade 6) where the variation of the average test score is positive; (II) untreated students, who attend a class/school with a negative variation. Thus we want to estimate the impact, on the probability to be resilient, of attending a classroom/school where peers obtain averagely higher outcomes than the year before. Having defined such a treatment, we implement a Propensity Score Matching (PSM) to match similar treated and untreated students. Propensity score matching model was firstly proposed by Rosenbaum and Rubin (1983) in order to reduce the bias in the estimation of treatment effects in observational studies. As Becker & Ichino (2002) claims "propensity score matching is a way to *correct* the estimation of treatment effects controlling for the existence of confounding factors based on the idea that the bias is reduced when the comparison of outcomes is performed using treated and control subjects who are as similar as possible" (p.360). The cited "confounding factors" are related to the fact that, in observational researches, the assignment of individuals to treatment or control group is not random, causing bias in the estimation. For this reason, PSM constructs a statistical comparison group based on the propensity score using observed characteristics. Individuals in the treatment group are matched to untreated individuals on the basis of similar probabilities of receiving treatment, while the average treatment effect of the program is estimated by calculating the mean difference in outcomes across two groups. Using the definition provided by Rosenbaum and Rubin (1983), propensity score is the conditional probability of receiving a treatment given pre-treatment characteristics. Analytically, this means the following:

$$p(X) = Pr\{D = 1|X\} = E\{D|X\} \quad (11)$$

where D = {0, 1} is the indicator of exposure to treatment and X is the multidimensional vector of pre-treatment characteristics.

In our model, the outcome is a dichotomous variable that identifies resilient students. As applied above, the variable is equal to 1 when the student is resilient and 0 otherwise. Thus the Average effect of Treatment on the Treated (ATT) can be estimated as follows:

$$ATT = E(Y_{1i} - Y_{0i} | D = 1) \tag{12}$$

where and Y_{1i} and Y_{0i} are the potential outcomes, once that individuals are exposed to treatment (D).

We obtain propensity scores through a probit regression that controls for individual and classroom/school-level characteristics at grade 5. We also consider the variation in the classroom/school average ESCS index in order to check for the feature most related to peer effect (Ammermueller & Pischke, 2009) along with peers' performance (our treatment). Propensity score models are implemented in Stata 12 using *psmatch2* (Leuven & Sianesi, 2003).

The equation of the probit regression is estimated as follows:

$$y_{ij(t)} = \alpha_0 + \alpha_1 \bar{X}_{1_{ij(t-1)}} + \alpha_2 \bar{X}_{2_{j(t-1)}} + \alpha_3 \Delta\bar{X}_{3_{j(t,t-1)}} + \varepsilon_{ij} \tag{13}$$

where the outcome is the dummy variable for resilient students, $\bar{X}_1$ and $\bar{X}_2$ refer to individual and class/school-level characteristics, and $\Delta\bar{X}_3$ controls for the variation in class/school average socio-economic background. As the probability to be resilient cannot abstract from the characteristics of the classroom/school attended in grade 5, we control for this set of variables and for the variation in the most predictive feature related to peer effect (ΔESCS index). Tables 50-55a summarize results from the probit regression. The first column refers to outcomes for reading, the second column contains the coefficient for mathematics. R^2 is around 30% for reading and from 30% to 70% for mathematics, so we consider that our probit regression as sufficiently predictive.

Considering these findings, we present two alternative model of PSM using two different treatments. The first defines the treatment as "attending a *class* where peers have higher achievement scores than previous peers" and considers class-level characteristics among predictors. The second assumes that the treatment is "to be assigned to a *school* where peers have higher achievement scores than previous peers" and considers school-level characteristics among predictors. The outcome is kept unvaried and is the probability to become resilient, expressed as a dummy variable, so that resilient students=1. As regards mathematics' regression we perform a one-to-one matching with replacement (where each observation in the untreated group can be matched with more than one observation in the treated one) because the number of treated students exceeds the number of control group students. For the reading test, we implement a one-to-one matching both with and without replacement having a sufficient number of observations in both groups (treated and not). Results in tables below present the model without replacement, which show a higher level of statistical significance.

Results for PSM are presented in tables 50-55b. We interpret the outcome as a marginal increase or decrease in the probability to be resilient of being exposed to the treatment. Looking at the Average Treatment effect on the Treated (ATT), we observe a significant and strong treatment effect at class-level. In particular, attending a class where the average students' score in grade 6 is higher than in grade 5 increases the probability to be resilient. The size of this effect changes from city to city. In Bologna, attending a class in grade 6 with higher achieving peers increases the probability to become a resilient student by 13% in reading and 19% in mathematics, with a higher T-test for the latter. Considering the reading model with replacement, the effect is lower than the first one, 12%, but is not statistically significant. We find a quite similar situation in Milan, where the increase in the probability to become resilient is higher in mathematics, 9.8%, than reading, 6.7%. T-test is around 2 points for both, so the effect is positive even though it is slightly significant. For the reading model, the estimation with replacement is not statistically significant despite the effect is positive (4.6%). Considering the city of Rome we observe that the difference, in percentage, between reading and mathematics is particularly large. The probability to become resilient in reading increases by 9.7%, while considering the same treatment for mathematics, the treatment makes the probability raise by 27.3%. Moreover, Rome is the only city which preserves the statistical significance also in the model with replacement, with a treatment effect by 16.5%. The city of Turin is the only case where the effect of the treatment is not significant in reading, but it is positive as well. For mathematics we record a positive impact of 3.6% with a statistically significant t-test of 2.27.

Finally, considering the entire sample formed by the four cities globally, we find that the average difference between treated and untreated groups in the probability to be resilient is 8.2% for reading and 22% for mathematics model, that somehow represent an average of the estimate of each city. Results for Padua are omitted because of the small sample size in the reading test and because of the lack of untreated classes in mathematics.

On the contrary, none of the results obtained at school-level is significant, except for cities considered jointly (tables 56-60). This happens despite the high significance level that the variation in the school average score has in the probit regression discussed above. This result could be related to the fact that the high significance level previously observed at school-level is actually a consequence of the overlap of classroom's effects, which are the only significant indeed.

Summarizing the results of PSM, attending in grade 6 a class or a school where peers achieve better performances than previous year has a positive relationship with resilience. Nevertheless, when considering this finding as a treatment, results are significant only at class-level. Thus, the effect of a positive variation in class average score between grade 5 and 6 is to increase the probability to become resilient. To the extent that class assignment is random, this effect is causal. On the contrary, families, making decisions about the school attended by their children, affect the probability to be part of the treatment group at school-level, making the effect not recognizable. Therefore, the only effect we are able to observe is a within-school phenomenon that happens at class-level. Our results can be interpreted in the stream of literature that affirms how peer ability affects students' achievement (Zimmer & Toma, 2000, Hanushek *et al.*, 2003). Moreover, Zimmer & Toma (2000) argue that this effect appears to be greater for low-ability students than for high-ability students.

Table 51a. Results from the probit regression to estimate propensity score – class-level, Bologna

	Reading N=150	Mathematics N=131
Variable		
Student-level characteristics	coeff.	coeff.
ESCS (grade 5)	-0.178	-1.03
	(0.241)	*(0.844)*
Female student	0.06	-0.585
	(0.282)	*(1.083)*
First generation immigrant	0.127	-0.935
	(0.393)	*(1.33)*
Second generation immigrant	-0.262	-0.725
	(0.355)	*(1.18)*
Characteristics of the classroom attended at grade 5		
Classroom average socio-economic background (grade 5)	**1.33*** **	1.83
	(0.533)	*(1.4)*
Classroom average score (grade 5)	**-0.091*** **	**-0.065*** **
	(0.015)	*(0.032)*
Number of immigrant students in the classroom (grade 5)	**-0.229*** **	0.253
	(0.075)	*(0.289)*
Number of female students in the classroom (grade 5)	0.094	0.288
	(0.069)	*(0.264)*
Change in the classroom characteristics between grade 5 and 6		
Delta classroom socio-economic background	**1.87*** **	**3.23*** **
	(0.355)	*(1.33)*
Constant	18.01	9.99
	(3.277)	*(4.71)*
Pseudo R^2	0.47	0.65

NOTES: *** is statistically significant at 1%; ** 5%, * 10%. Robust standard errors in parentheses.

Table 51b. Results of propensity score matching – class-level, Bologna

Bologna	**Output**	**#Untreated**	**#Treated**	**Sample**	**Treated**	**Controls**	**Difference**	**S.E**	**T-stat**
Reading	DUM_RESI	61	89	Unmatched	0.191	0.066	0.125	0.057	2.2
				ATT	0.197	0.066	0.131	0.060	2.17
Mathematics	DUM_RESI	6	125	Unmatched	0.192	0.000	0.192	0.162	1.18
				ATT	0.192	0.000	0.192	0.035	5.43

Table 52a. Results from the probit regression to estimate propensity score – class-level, Milan

	Reading N=633	Mathematics N=569
Variable		
Student-level characteristics	coeff.	coeff.
ESCS (grade 5)	0.121	**-0.274***
	(0.115)	*(0.129)*
Female student	0.115	0.08
	(.0126)	*(0.143)*
First generation immigrant	-0.105	-0.289
	(0.182)	*(0.204)*
Second generation immigrant	-0.048	0.024
	(0.155)	*(0.171)*
Characteristics of the classroom attended at grade 5		
Classroom average socio-economic background (grade 5)	**1.53*****	**0.945*****
	(0.182)	*(0.199)*
Classroom average score (grade 5)	**-0.078*****	**-0.078*****
	(0.006)	*(0.005)*
Number of immigrant students in the classroom (grade 5)	**-0.065*****	**-0.098*****
	(0.023)	*(0.022)*
Number of female students in the classroom (grade 5)	0.027	0.046
	(0.024)	*(0.028)*
Change in the classroom characteristics between grade 5 and 6		
Delta classroom socio-economic background	**1.38*****	**1.48*****
	(0.164)	*(0.197)*
Constant	16	15.18
	(1.30)	*(1.36)*
Pseudo R^2	0.35	0.37

NOTES: *** is statistically significant at 1%; ** 5%, * 10%. Robust standard errors in parentheses.

Table 52b. Results of propensity score matching – class-level, Milan

Milan	**Output**	**#Untreated**	**#Treated**	**Sample**	**Treated**	**Controls**	**Difference**	**S.E**	**T-stat**
Reading	DUM_RESI	270	363	Unmatched	0.132	0.078	0.054	0.025	2.18
				ATT	0.144	0.078	0.067	0.027	2.47
Mathematics	DUM_RESI	172	396	Unmatched	0.101	0.052	0.049	0.026	1.90
				ATT	0.101	0.003	0.098	0.044	2.23

Table 53a. Results from the probit regression to estimate propensity score – class-level, Rome

	Reading N=725	Mathematics N=660
Variable		
Student-level characteristics	coeff.	coeff.
ESCS (grade 5)	0.151	**-0.193*** **
	(0.103)	*(0.354)*
Female student	-0.065	-0.033
	(0.112)	*(0.351)*
First generation immigrant	-0.114	-1.52
	(0.172)	*(0.550)*
Second generation immigrant	-0.09	-0.39
	(0.177)	*(0.580)*
Characteristics of the classroom attended at grade 5		
Classroom average socio-economic background (grade 5)	**1.093*** **	0.240
	(0.156)	*(0.608)*
Classroom average score (grade 5)	**-0.072*** **	**-0.091*** **
	(0.005)	*(0.015)*
Number of immigrant students in the classroom (grade 5)	-0.019	-0.145
	(0.028)	*(0.097)*
Number of female students in the classroom (grade 5)	0.011	**0.154** **
	(0.022)	*(0.008)*
Change in the classroom characteristics between grade 5 and 6		
Delta classroom socio-economic background	**0.659*** **	**0.430***
	(0.135)	*(0.500)*
Constant	14.66	16.54
	(1.065)	*(2.88)*
Pseudo R^2	0.3	0.73

NOTES: *** is statistically significant at 1%; ** 5%, * 10%. Robust standard errors in parentheses.

Table 53b. Results of propensity score matching – class-level, Rome

Rome	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	320	405	Unmatched	0.230	0.119	0.111	0.029	3.89
				ATT	0.216	0.119	0.097	0.029	3.31
Mathematics	DUM_RESI	35	619	Unmatched	0.273	0.029	0.244	0.076	3.23
				ATT	0.273	0.000	0.273	0.018	15.23

Table 54a. Results from the probit regression to estimate propensity score – class-level, Turin

	Reading N=225	Mathematics N=236
Variable		
Student-level characteristics	coeff.	coeff.
ESCS (grade 5)	0.163	0.28
	(0.197)	*(0.223)*
Female student	-0.154	0.238
	(0.210)	*(0.208)*
First generation immigrant	0.046	0.257
	(0.291)	*(0.286)*
Second generation immigrant	0.228	0.081
	(0.268)	*(0.260)*
Characteristics of the classroom attended at grade 5		
Classroom average socio-economic background (grade 5)	**1.85*** **	**0.765*** **
	(0.358)	*(0.310)*
Classroom average score (grade 5)	**-0.088*** **	**-0.076*** **
	(0.013)	*(0.310)*
Number of immigrant students in the classroom (grade 5)	-0.003	**-0.098*** **
	(0.039)	*(0.036)*
Number of female students in the classroom (grade 5)	**0.090***	-0.034
	(0.048)	*(0.432)*
Change in the classroom characteristics between grade 5 and 6		
Delta classroom socio-economic background	**1.78*** **	**1.24*** **
	(0.284)	*(0.268)*
Constant	17.23	16.2
	(2.53)	*(2.18)*
Pseudo R^2	0.34	0.33

NOTES: *** is statistically significant at 1%; ** 5%, * 10%. Robust standard errors in parentheses.

Table 54b. Results of propensity score matching – class-level, Turin

Turin	**Output**	**#Untreated**	**#Treated**	**Sample**	**Treated**	**Controls**	**Difference**	**S.E**	**T-stat**
Reading	DUM_RESI	109	116	Unmatched	0.060	0.037	0.024	0.029	0.82
				ATT	0.064	0.037	0.028	0.030	0.93
Mathematics	DUM_RESI	97	139	Unmatched	0.036	0.010	0.026	0.021	1.23
				ATT	0.036	0.000	0.036	0.016	2.27

Table 55a. Results from the probit regression to estimate propensity score – class-level, Entire sample

	Reading N=1,735	Mathematics N=1,608
Variable		
Student-level characteristics	coeff.	coeff.
ESCS (grade 5)	0.071	**-0.174****
	(0.064)	*(0.095)*
Female student	0.022	0.068
	(0.073)	*(0.103)*
First generation immigrant	-0.089	-0.235
	(0.107)	*(0.145)*
Second generation immigrant	-0.067	0.05
	(0.098)	*(0.129)*
Characteristics of the classroom attended at grade 5		
Classroom average socio-economic background (grade 5)	**1.27*****	**0.870*****
	(0.103)	*(0.144)*
Classroom average score (grade 5)	**-0.071*****	**-0.060*****
	(0.003)	*(0.003)*
Number of immigrant students in the classroom (grade 5)	**-0.035*****	**-0.057*****
	(0.014)	*(0.017)*
Number of female students in the classroom (grade 5)	**0.024***	0.028
	(0.014)	*(0.020)*
Change in the classroom characteristics between grade 5 and 6		
Delta classroom socio-economic background	**1.06*****	**1.21*****
	(0.088)	*(0.130)*
Constant	14.38	12.23
	(0.705)	*(0.748)*
Pseudo R^2	0.31	0.48

NOTES: *** is statistically significant at 1%; ** 5%, * 10%. Robust standard errors in parentheses.

Table 55b. Results of propensity score matching – class-level, Entire sample

Entire sample	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	760	975	Unmatched	0.171	0.088	0.083	0.016	5.06
				ATT	0.170	0.088	0.082	0.017	4.78
Mathematics	DUM_RESI	310	1,298	Unmatched	0.225	0.052	0.173	0.025	7.07
				ATT	0.225	0.005	0.220	0.078	2.83

Table 56. Results of propensity score matching – school-level, Bologna

Bologna	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	61	88	Unmatched	0.148	0.082	0.066	0.054	1.21
				ATT	0.164	0.082	0.082	0.059	1.38
Mathematics	DUM_RESI	8	121	Unmatched	0.165	0.000	0.165	0.132	1.25
				ATT	0.165	0.000	0.165	.	.

Table 57. Results of propensity score matching – school-level, Milan

Milano	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	276	357	Unmatched	0.129	0.083	0.046	0.025	1.82
				ATT	0.130	0.083	0.047	0.026	1.79
Mathematics	DUM_RESI	142	426	Unmatched	0.094	0.063	0.031	0.027	1.12
				ATT	0.094	0.002	0.092	0.065	1.41

Table 58. Results of propensity score matching – school-level, Rome

Rome	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	327	398	Unmatched	0.221	0.131	0.090	0.029	3.14
				ATT	0.211	0.131	0.080	0.029	2.71
Mathematics	DUM_RESI	46	608	Unmatched	0.276	0.043	0.233	0.067	3.5
				ATT	0.276	0.003	0.273	0.305	0.89

Table 59. Results of propensity score matching – school-level, Turin

Turin	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	130	109	Unmatched	0.101	0.108	-0.007	0.040	-0.17
				ATT	0.101	0.092	0.009	0.040	0.23
Mathematics	DUM_RESI	111	138	Unmatched	0.10	0.05	0.06	0.03	1.67
				ATT	0.10	0.000	0.10	0.03	3.93

Table 60. Results of propensity score matching – school-level, Entire sample

Entire sample	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	797	952	Unmatched	0.168	0.110	0.058	0.017	3.45
				ATT	0.163	0.110	0.053	0.017	3.07
Mathematics	DUM_RESI	307	1,314	Unmatched	0.231	0.059	0.172	0.025	6.90
				ATT	0.231	0.007	0.224	0.112	1.99

3.5 Robustness check

As a robustness check, we firstly change our definition of treatment. More specifically, we restrict the definition assuming that the treatment is *"to be assigned to a class where peers have achievement gains (with respect to grade 5) higher than the average gain in each city"*. Secondly, we maintain the original definition of treatment estimating the first stage regression through a logit model instead of a probit regression.

Table 61-65 present results obtained when implementing the first robustness check. As we did for the original model, for the mathematics sample we employ a model with replacement, while PSM without replacement is implemented for the reading group. Increase in the probability to be resilient are in the same order of magnitude of the original model except for the mathematics group in the city of Rome (where the probability raises from 27.3 to 30.5%) that make arguably raise the probability of the entire sample from 22 to 25.6%. Considering the significance level of results, the city of Turin is still not significant for reading, while Milan loses statistical relevance in mathematics. All other cities maintain comparable level of significance.

Results obtained using a logit regression to derive propensity scores show that predictive pre-treatment characteristics are the same of the original model (tables 66-70a). Moreover, PSM results in tables 66-70b present similar findings, except for a decrease in the probability to become a resilient student in mathematics for Rome (from 27.3 to 11.3%).

Table 61. Restricting the definition of the treatment, Bologna

Bologna	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	72	78	Unmatched	0.205	0.069	0.136	0.056	2.42
				ATT	0.194	0.069	0.125	0.056	2.24
Mathematics	DUM_RESI	7	124	Unmatched	0.194	0.000	0.194	0.150	1.29
				ATT	0.194	0.000	0.194	0.036	5.43

Table 62. Restricting the definition of the treatment, Milan

Milan	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	336	297	Unmatched	0.141	0.080	0.061	0.025	2.47
				ATT	0.141	0.077	0.064	0.026	2.51
Mathematics	DUM_RESI	252	316	Unmatched	0.136	0.091	0.045	0.027	1.66
				ATT	0.136	0.066	0.070	0.097	0.72

Table 63. Restricting the definition of the treatment, Rome

Rome	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	347	378	Unmatched	0.238	0.118	0.120	0.028	4.24
				ATT	0.233	0.118	0.115	0.029	4.03
Mathematics	DUM_RESI	84	570	Unmatched	0.328	0.071	0.257	0.052	4.89
				ATT	0.328	0.023	0.305	0.096	3.18

Table 64. Restricting the definition of the treatment, Turin

Turin	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	111	114	Unmatched	0.061	0.036	0.025	0.029	0.88
				ATT	0.063	0.036	0.027	0.029	0.93
Mathematics	DUM_RESI	129	107	Unmatched	0.047	0.008	0.039	0.021	1.9
				ATT	0.047	0.000	0.047	0.020	2.28

Table 65. Restricting the definition of the treatment, Entire sample

Entire sample	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	862	873	Unmatched	0.178	0.092	0.086	0.016	5.28
				ATT	0.175	0.092	0.084	0.016	5.14
Mathematics	DUM_RESI	798	810	Unmatched	0.304	0.078	0.226	0.019	12.01
				ATT	0.304	0.048	0.256	0.093	2.76

Table 66a. Propensity scores defined through a logit model, Bologna

	Reading N=150	Mathematics N=131
Variable		
Student-level characteristics	coeff.	coeff.
ESCS (grade 5)	-0.286	-1.86
	(0.421)	*(1.54)*
Female student	0.166	-1.11
	(0.502)	*(1.97)*
First generation immigrant	0.138	-1.96
	(0.713)	*(2.39)*
Second generation immigrant	-0.629	-1.483
	(0.644)	*(2.176)*
Characteristics of the classroom attended at grade 5		
Classroom average socio-economic background (grade 5)	**2.662*** **	3.68
	(1.05)	*(2.781)*
Classroom average score (grade 5)	**-0.166*** **	**-0.118*** **
	(0.032)	*(0.057)*
Number of immigrant students in the classroom (grade 5)	**-0.393*** **	0.526
	(0.145)	*(0.548)*
Number of female students in the classroom (grade 5)	0.146	0.497
	(0.122)	*(0.475)*
Change in the classroom characteristics between grade 5 and 6		
Delta classroom socio-economic background	**3.355*** **	**6.30*** **
	(0.681)	*(2.719)*
Constant	32.89	18.48
	(6.59)	*(8.76)*
Pseudo R^2	0.48	0.64

Table 66b. PSM results, Bologna

Bologna	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	61	89	Unmatched	0.191	0.066	0.125	0.057	2.2
				ATT	0.197	0.066	0.131	0.060	2.17
Mathematics	DUM_RESI	6	125	Unmatched	0.192	0.000	0.192	0.162	1.18
				ATT	0.192	0.000	0.192	0.035	5.43

Table 67a. Propensity scores defined through a logit model, Milan

	Reading	Mathematics
	N=633	N=569
Variable		
Student-level characteristics	coeff.	coeff.
ESCS (grade 5)	0.193	**-0.467***
	(0.196)	*(0.221)*
Female student	0.205	0.163
	(0.216)	*(0.249)*
First generation immigrant	-0.169	-0.484
	(0.309)	*(0.359)*
Second generation immigrant	-0.072	0.073
	(0.263)	*(0.301)*
Characteristics of the classroom attended at grade 5		
Classroom average socio-economic background (grade 5)	**2.58***	**1.656***
	(0.323)	*(0.354)*
Classroom average score (grade 5)	**-0.133***	**-0.125***
	(0.011)	*(0.011)*
Number of immigrant students in the classroom (grade 5)	**-0.112***	**-0.167***
	(0.041)	*(0.043)*
Number of female students in the classroom (grade 5)	0.037	0.079
	(0.042)	*(0.049)*
Change in the classroom characteristics between grade 5 and 6		
Delta classroom socio-economic background	**2.32***	**2.592***
	(0.286)	*(0.358)*
Constant	27.19	25.7
	(2.40)	*(2.369)*
Pseudo R^2	0.35	0.36

Table 67b. PSM results, Milan

Milan	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	270	363	Unmatched	0.132	0.078	0.054	0.025	2.18
				ATT	0.148	0.078	0.070	0.027	2.59
Mathematics	DUM_RESI	172	396	Unmatched	0.131	0.081	0.050	0.029	1.71
				ATT	0.131	0.033	0.098	0.091	1.08

Table 68a. Propensity scores defined through a logit model, Rome

	Reading N=725	Mathematics N=660
Variable		
Student-level characteristics	coeff.	coeff.
ESCS (grade 5)	0.247	0.077
	(0.176)	*(0.190)*
Females	-0.078	-0.194
	(0.191)	*(0.203)*
First generation immigrant	-0.213	-0.512
	(0.294)	*(0.355)*
Second generation immigrant	-0.131	-0.249
	(0.303)	*(0.344)*
Characteristics of the classroom attended at grade 5		
Classroom average socio-economic background (grade 5)	**1.89*** **	2.39
	(0.277)	*(0.361)*
Classroom average score (grade 5)	**-0.125*** **	**-0.123*** **
	(0.009)	*(0.010)*
Number of immigrant students in the classroom (grade 5)	-0.032	-0.070
	(0.049)	*(0.059)*
Number of female students in the classroom (grade 5)	0.022	**0.082** **
	(0.037)	*(0.044)*
Change in the classroom characteristics between grade 5 and 6		
Delta classroom socio-economic background	**1.119*** **	**1.926*** **
	(0.246)	*(0.311)*
Constant	25.17	24.48
	(1.99)	*(2.05)*
Pseudo R^2	0.3	0.33

Table 68b. PSM results, Rome

Rome	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	320	405	Unmatched	0.230	0.119	0.111	0.029	3.89
				ATT	0.216	0.119	0.097	0.029	3.31
Mathematics	DUM_RESI	327	327	Unmatched	0.352	0.239	0.113	0.035	3.19
				ATT	0.352	0.239	0.113	0.035	3.19

Table 69a. Propensity scores defined through a logit model, Turin

	Reading N=225	Mathematics N=236
Variable		
Student-level characteristics	coeff.	coeff.
ESCS (grade 5)	0.321	0.491
	(0.342)	*(0.380)*
Female student	-0.249	0.37
	(0.368)	*(0.355)*
First generation immigrant	0.145	0.434
	(0.469)	*(0.492)*
Second generation immigrant	0.446	0.152
	(0.461)	*(0.443)*
Characteristics of the classroom attended at grade 5		
Classroom average socio-economic background (grade 5)	**3.057*** **	**1.297*** **
	(0.691)	*(0.533)*
Classroom average score (grade 5)	**-0.163*** **	**-0.128*** **
	(0.025)	*(0.019)*
Number of immigrant students in the classroom (grade 5)	-0.007	**-0.173*** **
	(0.067)	*(0.064)*
Number of female students in the classroom (grade 5)	**0.180** **	-0.056
	(0.092)	*(0.079)*
Change in the classroom characteristics between grade 5 and 6		
Delta classroom socio-economic background	**3.354*** **	**2.194*** **
	(0.568)	*(0.503)*
Constant	31.62	27.36
	(4.99)	*(3.965)*
Pseudo R^2	0.35	0.33

Table 69b. PSM results, Turin

Turin	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	109	116	Unmatched	0.060	0.037	0.024	0.029	0.82
				ATT	0.064	0.037	0.028	0.030	0.93
Mathematics	DUM_RESI	97	139	Unmatched	0.036	0.010	0.026	0.021	1.23
				ATT	0.036	0.000	0.036	0.016	2.27

Table 70a. Propensity scores defined through a logit model, Entire sample

	Reading	Mathematics
	N=1735	N=1,608
Variable		
Student-level characteristics	coeff.	coeff.
ESCS (grade 5)	0.118	**-0.285****
	(0.111)	*(0.166)*
Female student	0.054	0.118
	(0.125)	*(0.181)*
First generation immigrant	-0.152	-0.319
	(0.182)	*(0.259)*
Second generation immigrant	-0.105	0.13
	(0.168)	*(0.225)*
Characteristics of the classroom attended at grade 5		
Classroom average socio-economic background (grade 5)	**2.248*****	**1.502*****
	(0.185)	*(0.253)*
Classroom average score (grade 5)	**-0.124*****	**-0.111*****
	(0.006)	*(0.007)*
Number of immigrant students in the classroom (grade 5)	**-0.056*****	**-0.119*****
	(0.024)	*(0.030)*
Number of female students in the classroom (grade 5)	**0.042***	0.005
	(0.025)	*(0.036)*
Change in the classroom characteristics between grade 5 and 6		
Delta classroom socio-economic background	**1.907*****	**2.11*****
	(0.161)	*(0.236)*
Constant	24.9	22.6
	(1.325)	*(1.480)*
Pseudo R^2	0.32	0.48

Table 70b. PSM result, Entire sample

Entire sample	Output	#Untreated	#Treated	Sample	Treated	Controls	Difference	S.E	T-stat
Reading	DUM_RESI	760	975	Unmatched	0.171	0.088	0.083	0.016	5.06
				ATT	0.168	0.088	0.080	0.017	4.71
Mathematics	DUM_RESI	310	1,298	Unmatched	0.225	0.052	0.173	0.025	7.07
				ATT	0.225	0.005	0.220	0.091	2.42

Chapter 4

Conclusions and Policy Implications

In this study, we provide a definition of resilient students based on children ability to overcome their disadvantaged condition obtaining good academic performance in the switch from grade 5 to 6. In particular, we are interested in investigating the role that class and school factors have in this process, as this could entail important policy implications. In fact, the socioeconomic background has been identified as one the major factor influencing students' achievement, failure and drop-out. The ability of an educational system to provide equity of opportunity, overcoming a disadvantaged socioeconomic condition, is a major point for policymakers as it would result in lower social and economic costs.

In this research, we use INVALSI data from reading and mathematics standardized test taken in school year 2011/12 and 2012/13. The shift of students from primary to lower secondary school, allows us to identify a *variation* in class and school-level characteristics that can arguably affect students' achievement. Through different regression models, we implement educational production functions where inputs at individual, class and school level are combined to determine student's academic outcomes. As preliminary study on our sample of students living in five cities of Northern and Central Italy, we investigate which factors have a stronger correlation with students' achievement from both a static and a dynamic point of view. Firstly, we implement an education production function in order to investigate factors related to students' achievement in grade 6. Secondly, we employ value added models aiming at identifying which time variant characteristics between grade 5 and 6 mostly influence test scores variation. From the first model, we find that a higher socioeconomic background have a positive correlation with test score, highlighting the influence that the SES condition have on student's achievement. Similarly, results show the positive correlation with prior achievement (grade 5 score) consistently with our hypothesis about student's endowment and carried-over effect of past years of education expressed by the achievement in grade 5. Moreover, female students are more likely to perform better in reading than in mathematics, while being an immigrant student has a negative correlation with test score, especially in reading. Considering school's features, the most significant variables are related to the school environment, namely the school average socioeconomic background and the average test score. However, the school-average ESCS index is negatively related with the achievement in grade 6, which is a debatable result, but persistent in our research. On the contrary, the school-average score has a positive correlation with pupil's achievement in grade 6, entailing that the higher is peers' achievement, the better is the student's performance.

Value added models show similar results at individual level, but they provide additional information at class and school level. In particular, regression models highlight the importance that the characteristics of the class/school attended in grade 5 have on score's variation, especially when considering the average socioeconomic background and score. Furthermore, these models directly take into consideration the variation in class/school characteristics between grade 5 and 6, showing that attending a class/school in grade 6 where peers are wealthier have a negative effect on score's improvement, while attending a class/school with more proficient peers has a positive impact on this variation. This findings suggest that peers' performance and background can highly influence student's achievement not only when a precise time period is considered, but also when a variation in these factors happens. Finally these models underlines two interesting aspects: the fewer is the number of students per class in grade 6, the higher is the test score improvement, and the higher is the number of immigrant students in grade 6, the higher is the achievement variation. The first result can be interpreted in the stream of literature debating on class size effect, while the second is particularly interesting in the Italian context, where generally immigrant students are found to pull down natives' achievement.

Subsequently, we define resilient students among those whose ESCS index and score is in the bottom 33th percentile of the distribution of each city in grade 5, and perform above the average of the city in grade 6. We then compare this group of students with those who still perform below the average of the city in grade 6, who represent our control group. In this comparison, we find similar distribution of the socioeconomic index in grade 5, but an average higher performance of resilient students not only in grade 6 (when resilient students are defined) but also in grade 5, with a greater density of distribution close to the cut-off point. Thus, resilient students are not those marginally wealthier, but those who are relatively closer to the cut-off point. Consistently, when testing the sensitivity of our definition , we find that variations in the score threshold have a greater impact on the sample size than variations in the ESCS index threshold, especially in the mathematics test. Moreover, overall comparing the two groups, we find that immigrant students are generally under-represented in the resilient group and that resilient students attend classrooms where the average score variation between grade 5 and 6 is greater than the variation of the control group.

Focusing on class and school-level factors, we also identify resilient schools and classes as those where the number of resilient students is greater than the mean (of resilient students per school/class) plus a standard deviation. Descriptive statistics show how resilient students tend to be concentrated in classes characterized by a lower socioeconomic background, especially when considering the mathematics group. However, it comes to light that all disadvantaged students (both resilient and control group) are more likely to attend schools and classes where, on average, the socioeconomic condition is sensibly lower than the average of the city, arising the issue of the geographical segregation of disadvantaged students. Finally, we implement probit regressions in order to investigate which factors are more related to resiliency. Not surprisingly, among individual factors, prior achievement (in grade 5) and the socioeconomic background play a fundamental role in predicting resiliency. Focusing on school factors, having attended a class or a school (in grade 5) where peers perform averagely high has a positive correlation with the probability to become a resilient student. More importantly, attending in grade 6 a class or a school where peers achieve better performances than previous year has a positive relationship with resilience. Nevertheless, when considering this finding as a treatment through a propensity score matching, results are significant only at class level. This result could be related to the fact that the high significance level previously observed at school-level is actually a consequence of the overlap of classroom's effects, which are the only significant indeed. From our analysis, we find that the effect of a positive variation in peers' performance at class level between grade 5 and 6 is to increase the probability to become a resilient student between 3 and 27%, especially in mathematics. This finding is particularly interesting, since mathematics is an academic subject that tends to be influenced more by differences across schools than differences across families (Borman & Overman, 2004). The robustness check bolsters these results, showing good level of statistical significance of the model both when modifying the definition of the treatment and the calculation of propensity scores. To the extent that class assignment is random, this effect is causal. On the contrary, families, making decisions about the school attended by their children, affect the probability to be part of the treatment group at school-level, making the effect not recognizable. Therefore, the only effect we are able to observe is a within-school phenomenon that happens at class-level.

Our results can be interpreted in the stream of literature that affirm how peer ability affect students' achievement (Zimmer & Toma, 2000, Hanushek *et al.*, 2003). In addition, we find that peers' influence not only fosters children's improvement, but also affects the performance of a particular group of students, i.e. disadvantaged students, raising their probability to "beat the odds" and to become a resilient student.

In a policymaking perspective, our findings support the importance of promoting the possibility for disadvantaged students to attend classes characterized by a greater diversification. In particular, disadvantaged students benefit from attending a class where peers obtain good academic results, increasing the probability that resilience may occur.

Further researches could interestingly study resilient students' outcomes after grade 6, in order to verify if their performances keep constant over time. Moreover, our findings refer to five Italian cities selected in Northern and Central Italy and they cannot be immediately generalized to the Italian context.

Nevertheless, these results are worth to be taken into account when designing policies fostering greater equality in the educational system, as they highlight that class factors can play an important role in supporting resilience.

References

Agasisti, T., & Falzetti, P. (2013, September). Between-classes sorting within schools and test scores: an empirical analysis of the Italian junior secondary schools. In *XXVIII National Conference of Labour Economics, Rome* (Vol. 4, No. 1, pp. 33-53).

Agasisti, T., Ieva, F., & Paganoni, A. M. (2014). Heterogeneity, school-effects and achievement gaps across Italian regions: further evidence from statistical modeling. *Submitted [online] http://mox. polimi. it/it/progetti/pubblicazioni/quaderni/07-2014. pdf*.

Agasisti, T., & Longobardi, S. (2014a). Educational institutions, resources, and students' resiliency: an empirical study about OECD countries. Economics Bulletin, 34(2), 1055-1067.

Agasisti, T., & Longobardi, S. (2014b). Inequality in education: can Italian disadvantaged students close the gap? Journal of Behavioral and Experimental Economics, 52(1), 8-20.

Ammermueller, A., & Pischke, J. S. (2009). Peer effects in European primary schools: Evidence from the progress in international reading literacy study. *Journal of Labor Economics*, *27*(3), 315-348.

Angrist, J. D., & Lang, K. (2004). Does school integration generate peer effects? Evidence from Boston's Metco Program. *American Economic Review*, 1613-1634.

Azzolini, D., Schnell, P., & Palmer, J. R. (2012). Educational Achievement Gaps between Immigrant and Native Students in Two "New" Immigration Countries Italy and Spain in Comparison. *The ANNALS of the American Academy of Political and Social Science*, *643*(1), 46-77.

Ballatore, R., Fort, M., & Ichino, A. (2013). The Tower of Babel in the Classroom?, mimeo.

Becker, S. O., & Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *The stata journal*, *2*(4), 358-377.

Benard, B. (1991). Fostering resiliency in kids: Protective factors in the family, school, and community.

Bertoni, M., Brunello, G., & Rocco, L. (2013). When the cat is near, the mice won't play: The effect of external examiners in Italian schools. Journal of Public Economics, 104, 65-77.

Béteille, T., & Loeb, S. (2002). Teacher quality and teacher labor markets. *Age*, 3.

Borman, G. D., & Overman, L. T. (2004). Academic resilience in mathematics among poor and minority students. The Elementary School Journal, 177-195.

Borman, G. D., & Rachuba, L. T. (2001). Academic Success among Poor and Minority Students: An Analysis of Competing Models of School Effects.

Bradley, S., & Taylor, J. (1998). The effect of school size on exam performance in secondary schools. *Oxford bulletin of economics and statistics*, *60*(3), 291-324.

Bromberg, M., & Theokas, C. (2014). Falling out of the lead: Following high achievers through high school and beyond.

Browning, M., & Heinesen, E. (2007). Class Size, Teacher Hours and Educational Attainment*. *The Scandinavian Journal of Economics*, *109*(2), 415-438.

Brunello, G., & Checchi, D. (2005). School quality and family background in Italy. Economics of Education Review, 24(5), 563-577.

Chevalier, A., Harmon, C. P., O'Sullivan, V., & Walker, I. (2005). The impact of parental income and education on the schooling of their children.

Coleman, J. (1968). The concept of equality of educational opportunity. *Harvard Educational Review*, *38*(1), 7-22.

Contini, D. (2013). Immigrant background peer effects in Italian schools. *Social science research*, *42*(4), 1122-1142.

Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *The Journal of Economic Perspectives*, 79-102.

Corcoran, S. P. (2010). Can Teachers Be Evaluated by Their Students' Test Scores? Should They Be? The Use of Value-Added Measures of Teacher Effectiveness in

Policy and Practice. Education Policy for Action Series. *Annenberg Institute for School Reform at Brown University (NJ1)*.

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.

Doran, H. C., & Izumi, L. T. (2004). Putting education to the test: A value-added model for California. *San Francisco: Pacific Research Institute*.

Ermisch, J., & Francesconi, M. (2000). *The effect of parents' employment on children's educational attainment* (No. 215). IZA Discussion paper series.

Falzetti, P., & Ricci, R. Equity in school: a challenge for regional based educational systems.

Faubert, B. (2012). A literature review of school practices to overcome school failure.

Gamoran, A., & Long, D. A. (2006). Equality of Educational Opportunity: a 40 Year Retrospective (WCER Working Paper No. 2006-9). *Wisconsin Center for Education Research, University of Wisconsin–Madison, Madison. Retrieved April, 12*, 2009.

Hanushek, E. A. (2002). Evidence, politics, and the class size debate. *The class size debate*, 37-65.

Hanushek, E. A., & Woessmann, L. (2010). *The economics of international differences in educational achievement* (No. w15949). National Bureau of Economic Research.

Haveman, R., & Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of economic literature*, 1829-1878.

Heckman, J. J. (2008). The case for investing in disadvantaged young children. *Big ideas for children: Investing in our nation's future*, 49-58.5.

Hoxby, C. M., & Weingarth, G. (2005). *Taking race out of the equation: School reassignment and the structure of peer effects*. Working paper.

Izumi, Y. A. M. A. S. A. K. I. (2013). The Effect of Parental Involvement on Student Achievement in Junior Secondary School: Examining Data from the Botswana TIMSS 2007. *Journal of International Cooperation Studies*, *21*(1).

Jensen, P., & Rasmussen, A. W. (2011). The effect of immigrant concentration in schools on native and immigrant children's reading and math skills. *Economics of Education Review*, *30*(6), 1503-1515.

Johnson, G. M. (1997). Resilient at-risk students in the inner-city. *McGill Journal of Education/Revue des sciences de l'éducation de McGill*, *32*(001).

Krueger, A. B. (1997). *Experimental estimates of education production functions* (No. w6051). National Bureau of Economic Research.

Krueger, A. B. (2003). Economic considerations and class size*. *The Economic Journal*, *113*(485), F34-F63.

Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee Value Added Assessment System. *Educational evaluation and policy analysis*, *25*(3), 287-298.

Lauer, C. (2003). Family background, cohort and education: A French–German comparison based on a multivariate ordered probit model of educational attainment. *Labour Economics*, *10*(2), 231-251.

Lavy, V. (2002). Evaluating the effect of teachers' group performance incentives on pupil achievement. *Journal of political Economy*, *110*(6), 1286-1317.

Levin, H. M. (1988). *Accelerated schools for at-risk students*. CPRE.

Lubienski, S. T., Robinson, J. P., Crane, C. C., & Ganley, C. M. (2013). Girls' and Boys' Mathematics Achievement, Affect, and Experiences: Findings from ECLS-K. *Journal for Research in Mathematics Education*, *44*(4), 634-645.

Meunier, M. (2011). Immigration and student achievement: Evidence from Switzerland. *Economics of education review*, *30*(1), 16-38.

Muijs, D., Harris, A., Chapman, C., Stoll, L., & Russ, J. (2004). Improving schools in socioeconomically disadvantaged areas–A review of research evidence. *School effectiveness and school improvement*, *15*(2).

OECD, 2012, "Equity and Quality in Education: Supporting Disadvantaged Students and Schools", OECD, Paris.

OECD, 2012, "Untapped skills: realizing the potential of immigrant students", OECD.

Productivity Commission, (2012), "School Workforce", Research Report, Canberra.

Rasbash, J., Leckie, G., Pillinger, R., & Jenkins, J. (2010). Children's educational progress: partitioning family, school and area effects. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *173*(3), 657-682.

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. *Whither opportunity*, 91-116.

Reis, S. M., Colbert, R. D., & Hébert, T. P. (2004). Understanding resilience in diverse, talented students in an urban high school. *Roeper Review*, *27*(2), 110-120.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417-458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 247-252.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

Rothstein, J. (2008). *Teacher quality in educational production: Tracking, decay, and student achievement* (No. w14442). National Bureau of Economic Research.

Rothstein, R. (2004). Class and schools: Using social, economic, and educational reform to close the achievement gap. *Washington, DC: Economic Policy Institute*.

Schnepf, S. V. (2007). Immigrants' educational disadvantage: an examination across ten countries and three surveys. *Journal of Population Economics*, *20*(3), 527-545.

Schwartz, A. E., & Stiefel, L. (2006). Is there a nativity gap? New evidence on the academic performance of immigrant students. *Education*, *1*(1), 17-49.

Sewell, W. H., & Shah, V. P. (1968). Parents' education and children's educational aspirations and achievements. *American Sociological Review*, 191-209.

Simon, F., Malgorzata, K., & Beatriz, P. (2007). *Education and Training Policy No More Failures Ten Steps to Equity in Education: Ten Steps to Equity in Education*. OECD Publishing.

Shea, J. (2000). Does parents' money matter?. *Journal of public Economics*, *77*(2), 155-184.

Tajalli, H., & Opheim, C. (2005). Strategies for Closing the Gap: Predicting Student Performance in Economically Disadvantaged Schools. *Educational Research Quarterly*, *28*(4), 44-54.

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement*. *The Economic Journal*, *113*(485), F3-F33.

Todd, P. E., & Wolpin, K. I. (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human capital*, *1*(1), 91-136.

Van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, *5*(2), 134-150.

Vignoles, A., Levacic, R., Walker, J., Machin, S., & Reynolds, D. (2000). *The relationship between resource allocation and pupil attainment: a review*. Centre for the Economics of Education, London School of Economics and Political Science.

Walsh L., Black R., (2009), "Overcoming the barriers to engagement and equity for all students", Biennial Conference Curriculum: a national conversation, Canberra, 2-4 October 2009.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1997). Fostering Educational Resilience in Inner-City Schools. Publication Series No. 4.

Zimmer, R. W., & Toma, E. F. (2000). Peer effects in private and public schools across countries. *Journal of Policy Analysis and Management*, *19*(1), 75-92.