Doctoral School of Social Sciences
University of Trento
Doctoral Programme in
Sociology and Social Research
XXXII cycle

The classroom as a sorting machine: The influence of teachers, friends, and peers on students' outcomes

Ph.D. Candidate:
Emanuele Fedeli

Supervisor:
Prof. Moris Triventi

Committee members:
Prof. Dr. H.G. (Herman) van de Werfhorst
Prof. Fabrizio Bernardi
Prof. ssa Stefani Scherer

Trento, 11 September 2020

Acknowledgements

I am grateful to my PhD supervisor, Professor Moris Triventi, for his guidance, support, and proactivity during this experience. You taught me how to look at the overall picture and you have always been present when I could not see the wood for the trees. Thanks, Moris.

I am also thankful to the members of my PhD committee, Prof. Dr. H.G. (Herman) van de Werfhorst, Prof. Fabrizio Bernardi, and Prof. ssa. Stefani Scherer. Your comments, feedback, and suggestions have made this work a better manuscript. To Stefani, I am further grateful for your advice in the last years.

Voglio ringraziare la mia "tipica" famiglia italiana, senza il cui supporto sarebbe stato tutto più difficile. Ognuno di loro ha avuto un ruolo importante nella mia vita. Mia madre mi ha sempre insegnato ad essere resiliente nelle avversità e a crederci sempre. Mio padre mi ha trasmesso una forte curiosità, a pensare con la propria testa, ed andare per la propria strada. Mio zio mi ha insegnato ad essere empatico. Infine, mia nonna che mi ha cresciuto come un figlio.

Infine, gli amici. Prima di tutto, voglio ringraziare Chiara e Flavio, i miei più cari amici di infanzia. Spesso passano mesi senza sentirci, ma quando ci sentiamo è come se non fosse passato nemmeno un giorno. Voglio ringraziare anche tutta la banda degli amici dell'università: Fulvio, Ornella, Marco, Peppe, e Matteo. Amici incredibili che mi hanno sempre supportato in questo percorso e con cui ho passato momenti indimenticabili.

Il percorso di dottorato è stato intenso e duro ma ho avuto la fortuna di incontrare delle bellissime persone e amici sinceri. In particolar modo, voglio ringraziare Gabriele a cui mi lega una fraterna amicizia e di cui ho una stima sconfinata. Ringrazio anche Francesca che è sempre stata schietta e franca. Uno speciale ringraziamento va a Giulia, Mauro, Saverio, ed Ilaria che hanno offerto un mix di supporto emozionale e "abitativo". Infine, voglio ringraziare quelli che considero la "mia coorte" a Trento e con cui ho passato bei momenti o accese discussioni teoriche: Alex, Nico, Elisa, Filippo, Mattia, e Yue.

Summary

Page

Introduction

On a twitter thread, Charles Murray, one of the most influential U.S. social scientists, claims that life course inequality will be genes-oriented in the next decades, once removed social mobility obstacles. This dissertation does not debate on this idea of an upcoming GATTACA[1] world, but it argues that it will take some time before to eradicate social mobility constraints. After WWII, international organizations and mass media have supported an impressive campaign to reduce worldwide inequality through a monitoring system of macro indicators. The risk of this approach is to underestimate the root of inequality patterns, still present, and lead the policymakers to not see the wood for the trees.

In 2016, OECD outlined that many students are trapped in a vicious spiral of poor performance and motivation, preventing them from investing in more human capital. It is a critical issue because educational attainment is broadly recognized as an essential life-course predictor. To this extent, an impressive body of literature outlines that educational attainment is associated with numerous outcomes such as occupational positions, higher wages, and better health status. A better understanding of which characteristics foster the chance to reach the highest educational levels is sophisticated. Indeed, individual attainment depends not only on their attitudes, beliefs, personality traits, performance, and choices but also on the interplay between such dimensions and the interaction with other persons. This interaction happens every day as seeing individuals interact with several actors such as parents, mates, friends, teachers, colleagues, in distinct social environments such as family, school, classroom, and workplace.

A systematic understanding of how, and to what extent, schools and classrooms shape the cognitive and non-cognitive skills of students are entirely neglected. The educational system should be an environment where students enter to increase their competences, open their minds, and find a road to express themselves. Unfortunately, it is not like that. On the one hand, educational systems work as great equalizers to reduce inequalities arising from the social origin (Raudenbush & Eschmann 2015). On the other hand, this system behaves as a sorting machine or social machinery (Spring, 1945), creating new sources of inequality or

[1] It is a 1997 American science fiction film written and directed by Andrew Niccol.

resulting in a "sounding board" of socioeconomic origin. The idea of educational systems as social constructs where multiple hierarchies take place is not entirely new. Domina, Penner, and Penner (2017) argue that educational systems create internal categories such as grades, classrooms, course-taking patterns, and academic tracks as well as they impose labels to students with the risk to reinforce external categories such as race, class, and gender. Such modes of categorical inequality might affect students' beliefs, attitudes, performance, and aspirations.

The work emphasizes that the roots of inequality find fertile breeding grounds on the educational systems and focuses on classroom aiming to understand possible sources of inequality among mates because it is an environment where students interact, sharing much time together. In this dissertation, I investigate to what extent this "social machinery" affects several students' outcomes, how *hierarchies, network of friends, and classroom peers* influence students' motivations, aspirations, performance, and educational choices.

In chapter I of this work, I outline a theoretical framework arguing that classroom inequality is a result of varying characteristics of the interacting actors such as their gender, age, ethnic origin, socioeconomic background as well as academic competencies. The classroom sorting of students with specific characteristics broadly depends on formal and informal institutional rules. To shed light on these patterns of educational systems, I rely on three distinct concepts, such as *inequality, diversity,* and *sorting*, as theoretically and empirically debated by Roberto (2015). In the dissertation, I conceive *inequality* as the uneven distribution of outcomes across students with specific characteristics, *diversity* as a variety of student's "types" in the classroom, and *sorting* as the uneven distribution of students' and teachers' characteristics across distinct educational environments. The theoretical framework is a roadmap to plan a tailored empirical inquiry and identification strategies for my empirical chapters. For my analyses, I have chosen three topics where I can test the presence of this social machinery in a classroom environment exploiting distinct "sources" of inequality such as teachers, peers, and friends.

In chapter II of this work, I will test whether teachers' grading is an inequality-enhancing factor in Italy. Previous contributions suggest that teacher's grading is biased by preferences and stereotypes. My idea is that teachers' grading standard might produce a hierarchy among students, even among equally able students. This hierarchy, in turn, could have a pervasive influence on students' perception of their own competencies, thereby influencing their academic achievement, motivation, and self-stigma.

In chapter III, I investigate the extent to which extent smoking and drinking friends lead to emulate the same behavior in a critical age like the adolescence. Unhealthy habits dramatically affect life expectancy, above all, when rooted in the early stage of individual development. In addition, I analyze if non-reciprocal friendship matters more or not as a driver of the behavior emulation because adolescents desire to be accepted.

In final chapter IV, I test to what extent the presence of students with a migration background affects several outcomes in classrooms, including students' attitudes and anti-social behavior. Italy is dealing with a dramatic increase of immigrant students since the late '80s, but a series of data suggest that the school is not well equipped for this challenge.

Overall, the thesis aims to contribute to important theoretical debates in the sociology and economics of education, such as the role of relative positions in the social environment (chapter II), peer effects in critical developmental stages (chapter III), and the social integration in heterogeneous contexts (chapter IV). However, it aims also to inform policy makers on possible side effects of current widespread educational practices such as grading on a curve (chapter II), the actual role of peers in the spreading of unhealthy behaviors among adolescents (chapter III), and the need of imposing interventions devoted to optimizing classrooms compositions (chapter IV).

References

Domina T., Penner A., & Penner E. (2017). Categorical Inequality: Schools As Sorting Machines. *Annual Review of Sociology*, *43*, 311-330

OECD (2016). *Education at a Glance*, OECD Publishing, Paris.

Raudenbush, S. W., & Eschmann, R. D. (2015). Does schooling increase or reduce social inequality?, *Annual Review of Sociology*, *41*, 443-470.

Roberto, E. (2015). The Divergence Index: A Decomposable Measure of Segregation and Inequality. Working paper, https://arxiv.org/abs/1508.01167

Spring, J. (1945). The Sorting Machine: National Educational Policy Since 1945. *The Journal of Education, 159*(3), Review of Bruce C. Beezer.

Every time a friend succeeds, I die a little
Gore Vidal[2]

[2] Quote reported: 1975 January, Ms., "Can Friendship Survive Success?" by Thomas Powers, Start Page 16, Quote Page 16, Column 2, Published by Ms. Magazine Corp., New York, New York

Sorting, diversity, and inequality in educational systems

Chapter I

Abstract

Cognitive and non-cognitive skills are key for individual life chances, and much of skill development occurs within schools. Against this background, it is worth noting that education systems work as sorting machines influencing students' socio-emotional skills unequally, academic competences, and choices. This work posits that student outcomes depend critically on students' relative positions, their network of friends, and the characteristics of their peers. It happens because "the machine" exposes students to a classroom with varying characteristics of friends, peers, and teachers upon which then they make their friendship and draw on hierarchies. Classrooms are environments of pivotal importance, where students spend a lot of time with their friends, peers, and teachers. However, the work theoretically considers that the sorting of students and teachers among classrooms and the consequent degree of diversity within classrooms are the comprehensive result of national policies such as teacher training, tracking, and residential housing.

1. Introduction

Education is an “evergreen” topic both in academia and in public debates. Every day we are overwhelmed by information on topics such as the unequal attainment and achievement of students across countries, the gender divide, the burden of school investment, and attempts to reform the educational system (OECD, 2012). This “special interest” in education owes much to the role education plays in improving living standards. Indeed, individual educational attainment is the most important single predictor of later occupational attainments (Kempel & Wilner, 2008), wages (Mincher, 1958), higher social mobility (Bernardi & Ballarino, 2016; Bukodi et al., 2018), better health status, a lower probability of smoking and heavy drinking, and of becoming overweight or obese (Brunello & Schlotter, 2011; Conti, Heckmand, and Urzua, 2010; Cutler & Lleras-Muney, 2006; Kemptner, Jurges, & Reinhold, 2011).

Educational attainment has thus been identified as an essential life-course predictor, and a burgeoning number of studies analyzes which student characteristics foster educational attainment (Heckman, 2000). Economic theory (Becker, 1964) has first focused on the role of cognitive skills in the accumulation of human capital, while sociology has investigated the role of academic performance in driving educational attainment and inequalities in educational outcomes (Jackson, 2013). More recently, economics has also turned its attention to the role of so-called socio-emotional or non-cognitive skills in education and the labor market (Brunello & Schlotter, 2011). In this respect, following studies both in economics and psychology, there is a growing consensus that students’ beliefs and attitudes may be important factors influencing educational aspirations and expectations (Fuligni, 1997; Kao & Tienda, 1998). These are, in turn, crucial to explain the educational choices of students over distinct educational stages and trajectories.

Previous studies (e.g. Festinger, 1954) have always pointed out that beliefs, attitudes, and choices depend on social actors and social environments. In this sense, students are not an exception because the formation of their cognitive and non-cognitive skills is the result of interactions with diverse actors such as parents, friends, and teachers across several social environments such as the family, the school, and the workplace (Babad, 2009). However, this result is not random, but it depends - within the educational environment - on the *hierarchies, network of friends, and peers*. Even though several studies have enriched the debate on the extent to which the role of peers and educational environment affects motivation (Rao, 2019), achievement (Hoxby & Weinghart, 2005), and self-stigma (Borman & Pyne, 2016), this work makes it clearer how and to what extent educational environments shape unequally cognitive and non-cognitive skills of students.

Horace Mann (1957, p. 145), a US educator and politician, said in 1848 that the "*Education then, beyond all other devices of human origin, is the great equalizer of the conditions of men (and women), the balance-wheel of the social machinery*." In this perspective, education systems are social constructs, behaving as a parallel "social machinery" and aiming to reduce inequalities arising from ascriptive characteristics, especially social origins. However, this works debunks this view because this machinery stratifies the school population and condition the classroom diversity, resulting in social hierarchies among students, conditioning the friend network formation, and exposing students to varying peer characteristics. One of the guiding ideas of the present work is that classroom is an ideal environment to investigate these patterns since it is a "little society" where students have different characteristics such as wisdom, leadership, and popularity (Bar-Tal, 1979) and interact with friends, mates, and teachers. (Babad, 2009). Except for the seminal work of Spring (1945), the "social machinery" frame has been relatively neglected in the literature until the work of Domina, Penner, and Penner (2017). They expand the idea arguing the education systems sort students, creating internal categories such as grades, classrooms, course-taking patterns, and academic tracks, imposing related labels associated with students, and reinforcing external categorization processes based on salient individual traits such as race, class, and gender.

This work argues that student's inequality depends on "classroom interaction" because "the machine" exposes students to a classroom with different characteristics of friends, peers, and teachers upon which then they make their friendship and draw on hierarchies. In turn, the classroom diversity degree depends on a mix of formal and informal rules underlying the sorting process. To explain the extent to which the "classroom interaction" results in patterns of inequality, the work relies on three distinct concepts – *inequality, diversity,* and *sorting* (Roberto, 2015) – to describe the extent to which educational systems sort and might harm students. *Inequality* refers to the uneven distribution of outcomes across students, classroom *diversity* refers to the variety of "types" in the student population, and *sorting* refers to the uneven distribution of students' and teachers' characteristics across distinct learning environments.

Once accounted for the sorting of students and the diversity in classrooms, this work investigates the multilayer effect of "peers" and asks: *do and to what extent hierarchies, the network of friends,* and *classroom peers condition students' motivations, aspirations, behavior, habits, performance, and the choices of students?* In section 2, the work discusses the outcomes of interest to explain inequalities, considering the current debate on cognitive and non-cognitive skills. In sections 3 and 4, the works explain the main literature on peer comparison, the tentative theoretical mechanism behind, and the identification of hierarchies, networks of friends, and classroom peers. In sections 5, 6, and 7, the work discusses hierarchies, networks of friends, and classroom peers separately, indicating the contribution to the literature within the current literature. In section 8, the work discusses the role of the education system, bringing out institutional features of the classroom environment and providing the rationale for the empirical chapters. In the last section 9, the work provides a theoretical framework and concludes the discussion.

2. The Tower of Babel: personal traits, attitudes, beliefs, expectations and cognitive skills

In the social sciences, terms such as achievement, expectations, and choices find a common and well-known definition (Almlund, Duckworth, Heckman, & Kautz 2011). Hence, achievement is a proxy of how a task is accomplished, and the expectations report the desire that something can occur, and choices indicate an effective decision. It is not difficult to transpose these definitions to the education setting to measure, for instance, academic competences, education expectations, and track choice. Instead, it is more complex to find a common agreement on terms such as personal traits, attitudes, and beliefs. The literature is vast, but there is still no clear stance on how to distinguish such dimensions within social sciences. Economic research relies on a binary distinction between cognitive skills, frequently measured with standardized tests or marks, and non-cognitive skills such as beliefs, attitudes, aspirations, and expectations (Almlund et al. 2011). In contrast, social and developmental psychology is more skeptical about this classification and frames these dimensions as distinct stages of a continuum to explain individual behavior (Ajzen, 1991).

Roberts (2009, page 140) writes that "*Personality traits are the relatively enduring patterns of thoughts, feelings, and behaviors that reflect the tendency to respond in certain ways under certain circumstances*." However, personality traits do not fully explain attitudes, beliefs, and behavior. Indeed, for instance, to explain attitudes toward math, Grigutsch and Torner (1998) suggest that attitudes cover three components: an emotional response, a belief on the subject, and an intent to behave. This hint gives the idea that some psychological constructs depend on the context and finds more general support in the work of Almlund and colleagues (2011), Eccles (1993), and Dasgupta (2013). They stress the idea that psychological constructs depend on the relative context, but the extent to which personal traits change and context plays a role are open questions in the literature.

Indeed, social sciences debate the malleability of non-cognitive skills to context and dose stimuli. On the one hand, some support the *situational specific hypothesis* (Almlund et al., 2011) in economics for which every dimension can be molded, even the personality traits when they are responsive to situations, tasks, and incentives. In contrast, the situational approach does not gain support from the psychological literature as outlined by Lucas and Donnellan (2009, page 147), who write that "*not all behavior is simply a function of the situation*." Thus, the time stability of these characteristics is a salient and controversial issue (McGue, Bacon, Lykken 1993). It is widely accepted that personality traits, attitudes, and beliefs, as well as expectations, cannot be treated in the same way. Personality traits are more stable over time, and long-term intervention is necessary to eventually change them (UNESCO, 2016), whereas attitudes and beliefs are less stable and more susceptible to be molded by families, schools, or networks of friends. However, a plausible litmus test is the dose exposure to certain treatments. A markable example is a work of Cobb and Clark (2013), where they argue that a trait such as locus of control[3] may be modified in early childhood (3-8 years old) if exposed to the high-dose treatment.

This work recognizes the critical role played by cognitive skills, personal traits, attitudes, beliefs, and choices in life. Such a role is undisputed in the literature. Indeed, contributions find causal relations between personal traits such as openness to experience,

[3] The degree to which people believe that they have control over the outcome of events in their lives.

conscientiousness, extraversion, agreeableness, and neuroticism and school attainment (Anger, 2013; Lenton, 2014), labor market trajectories (Heckman & Stixrud, 2006), and performance in the labor market (Brunello & Schlotter 2011). Streams of research more focused on non-cognitive mechanisms single out a relation between locus of control and investment in human capital and rewards in the labor market (Lekfuangfu, Cornaglia, Powdthavee, & Warrinnier 2014). Others suggest a positive effect of students' self-perception and motivation on performance as well as educational attainment (Borman & Pyne, 2016). In turn, academic performance comes out as a substantial signal of "quality" transmitted by students to families, employers, and universities and is a crucial factor in further educational investment (Bobba & Frisancho, 2014; OECD, 2012; Gasperoni; 1998). Since the malleability of a psychological construct is controversial in the literature, the work is agnostic on this, and it investigates the link between the peer comparison (Eccles, 1993) and a broad array of student's outcomes in a school setting.

3. Theoretical lens: different (similar) frames on peer comparisons

A comprehensive assessment of student's attitudes, beliefs, and choices is of paramount importance even if troublesome. Although it is recognized that actors such as friends, mates, teachers, and parents influence student outcomes in related social environments, previous literature dispute on the complexity of peer comparisons. Eccles (1993) and Wilkinson Hattie, Parr, and Thrupp (2000) theorize that any student outcome is the result of multiple influences between actors and the social environment, and they stressed two main dynamics. The former points out mostly a plausible mechanism of how social actors influence the self-reinforcing association between self-esteem, motivation, and academic performance. The latter agnostically does not suggest a mechanism but highlights the role played by the characteristics of contextual levels - families, schools, and neighborhoods – in shaping student outcomes.

In the '60s, a line of research suggests that people tend to create hierarchies in a social environment and compare themselves adopting an ordinal criterion because it is quite informative and simple (Henson, 1964; Parducci, 1965). Even if several empirical works

have dealt, formally and informally, with hierarchies among students, the literature has quite neglected this line in the educational context.
Indeed, Gamoran (1989) pointed out how relative positions in an ability grouping affect the performance. More recently, Borman and Pyne (2016) outline how performance and self-stigma depend on the "internalized group position" of students within the school, while other contributions start to address the role of ability hierarchies on academic performance (Lavy, Weinhardt, & Silva, 2012), dropout (Murphy & Weinhardt, 2016), self-esteem (Marsh, 2008), and enrolment in more advanced courses (Elsner & Isphording, 2015). Nevertheless, there is not clear and explicitly conceptualization of hierarchies, nor are there connections to the extensive literature on peer comparison or peer effects.

A plausible connection to the previous literature is the Big Fish in Little Pond Effect (BFLPE, hereafter Marsh, 2008) that introduces the concept of relative position in the debate. The idea suggests that once controlled for their own ability, students have higher self-esteem or self-efficacy in lower-ability classes. Thus, it is better to be a "big fish" in a "little pond" that in a "large one." In the beginning, Marsh (2008) suggests that the linchpin of comparison is the intimate ability measured with standardized tests, but then, he recognizes the importance of other sources of comparison such as teacher's assigned marks, teacher's preferences, or parental preferences.

The previous and landmark contributions such as the Social Comparison Theory (SCT, hereafter) of Festinger (1954) as well as Reference Group Theory (RFG, hereafter) of Kelley (1952) and Merton (1968) have focused at most on the contextual/group characteristics of social environments, neglecting the role of hierarchies. Indeed, the SCT (Festinger, 1954) suggests that individuals are prone to compare themselves to others to balance their own opinions and abilities. While some studies do not specifically debate extensively on the reference of comparison (Festinger 1954), Kelley (1952) and Merton (1968) develop a framework in which people choose a group as a reference point for their comparison, conforming to or contrasting the group behavior. Besides, Merton (1968) emphasizes that individuals choose a group because they aspire to that group's position and may use different reference groups for conformity or contrasting patterns. Following this lens, Collins (1995) makes the case that the choice of individuals or group comparisons depends on personal strategies regarding interpersonal relations and self-identity.

However, this works is aware that some ascribed characteristics such as gender, ethnic background, and socio-economic status enhance the challenges of peer effect studies according to two profiles. First, the starting points of students are different, and the gap worsens during schooling. For instance, girls underperform in mathematics compared to boys in several countries. This performance gap frequently is due to low motivation, students' drive, and self-beliefs among the girls (OECS, 2013). Holding constant the socio-economic profile, immigrant underperform in science compared to non-immigrant students (OECD, 2016a). Finally, international reports confirm every year that socio-economically disadvantaged students lie behind (OECD, 2016b). Second, students might react asymmetrically to the peer comparison. In this view, girls are less competitive and over-confident than boys (Croson & Gneezy, 2009). Once accounted for the socio-economic level, this gap is still present even if literature outlines that socio-economically disadvantaged students are less competitive and risk-takers (Almås I., Cappelen, Salvanes, Sørensen, & Tungodden, 2015). Finally, immigrant students are not risk-takers and show less willingness to compete (Halek & Eisenhauer, 2001). Hence, the work will theoretically consider such issues in the development of empirical contributions.

4. Status construction theory, the role of networks, and classroom peers

It is quite a Gordian knot to disentangle the multilayer features of peer comparison in a social environment. This work proposes that students draw on and internalize the hierarchy, associating it to an indicator of success or competence. This hint is at the foundation of Social Construction Theory (SCT, hereafter) (Ridgeway, 2006; Correl & Ridgeway, 2006). Indeed, people from one category may have an advantage for some characteristics such as ability, socio-economic origin, ethnic background, BMI, popularity, and teacher's preferences. If other social actors internalize this advantage as a social norm, a hierarchy takes place, conditioning the student relations and shaping the formation of friend networks. To this view, categorical differences turn into inequality patterns, altering the perception of peers in an environment (Berger & Webster, 2006).

Henceforth, it is a salient issue to understand the spreading and persistence of a hierarchy and the underlying social norm. Previous contributions are quite broad on this topic, and these works consider the framework provided by Heffetz and Robert (2011). A social norm executes its function only if four features are present such as punctuated equilibrium, long-run stability, conformity warp, and local conformity-global diversity (LC-GD). The first argues that a norm shift is due to little incremental behavioral changes or sudden exogenous shocks. The second indicates that the long-lasting stability of a social norm depends on the extent to which a norm is rooted in people's mindsets. The third attempts to disentangle the extent to which a social norm is due to ex-ante social preferences or endogenous preferences in the environment. Finally, the last feature points out that the once set a social norm, the spreading depends on the density of population, cluster, or networks of people, and the degree of social acceptance.

The spreading feature of a norm is of cardinal importance because it opens the debate on the structure of a social environment and the interaction within it. So far, a hierarchy may arise in a social environment influencing students' perceptions about others, but it cannot alone fully explain network formation. Thus, the role of peers as contextual characteristics - what Merton (1968) and Kelley (1952) generally label reference group – has to be considered. This work proposes to distinguish between two kinds of the reference group in an educational context: one endogenous and one exogenous. Reference groups are endogenous when students choose own friends and create a cluster in a population, and they are exogenous when reference groups are assigned by construction such as peers in course-taking patterns, classrooms, and schools. Hence, inequality depends not only on hierarchies among students but also on a network of friends and peers.

Among the educational environments, the classroom is a natural candidate to detect the effect of hierarchies, networks of friends, and peers on students' outcomes. Indeed, as outlined by Babad (2009), the classroom is a sort of "two-sides" environment where students interact with their peers and teachers. In this way, hierarchies arise from this interaction, jointly with the effect of the network of friends and peers. Indeed, in such an environment, students select among the peers' own friends drawing on a network of strong and weak ties according to characteristics such as gender, socio-economic origin, ability, and health habits. Bar-Tal (1979) supports this view of the classroom as an environment in which students differ from each other according to several characteristics. Thus, the classroom is an ideal unit because it is an environment physically bounded, where students spend a lot of time, around six hours per day in OECD countries (2018). To conclude, this works outlines that peer comparison is always present in a social environment and that *inequality* in cognitive and non-cognitive skills among students depends on relative position in a hierarchy, networks of friends, and classroom peers.

5. Hierarchies

The Merriam Webster dictionary defines status as a "position or rank in relation to others." Relying on this definition, Weiss and Fershtman define social status as "*a ranking of individuals in a given society*" according to their traits, assets, and actions (1998, page 72) while Ball and colleagues define it as "*a ranking in a hierarchy that is socially recognized and typically carries with it the expectation of entitlement to certain resources*" (2001, page 15). In general, the social status indicates the position of an individual along a socially recognized hierarchy, based on some characteristics and entitled to certain resources. According to Heffetz and Robert (2011), social status is made of three dimensions: positionality, desirability, and non-tradability. First, social status is – by definition – a position with a negative externality: an increase in someone's relative status means a decrease in the relative status of the others in the group. Second, social status is desired by others because it entitles one in having tangible and un-tangible resources such us money or respect. Third, social status is not tradable, and this enhances the internalization of a hierarchy, stimulates a status-seeking behavior in the environment, or a strategic behavior such as moving to another school, classroom, or neighborhood to achieve a better position, once pondered the cost-opportunity. (Bault, Coricelli, & Rustichini 2007).

Analogous streams of research have pointed out other dimensions related to the hierarchy formation. For instance, development psychologists frame social status at most as peer's likability. Indeed, Moreno (1934) goes more in depth, breaking down social status into stages: attraction, peer acceptance, socio-metric popularity (positive or negative nominations), and perceived popularity. Furthermore, a small amount of literature has investigated social status bringing out more dimensions of social distance such as "nearness" and "distance." This affective distance might be the result (Park, 1924) of negative stereotypes, commonalities, and differences (Tarde, 1903). Despite these distinctions, this work adopts the definitions of social status provided by Heffetz and Robert (2011) because it is more general and applicable to distinct contexts.

As mentioned above, the classroom is an ideal environment where two sides interact daily. On the one hand, students influence each other, shaping hierarchies; on the other hand, teachers might create hierarchies among students as a result of teacher-student interaction. Still, today, despite the increasing interest in the study of the role of hierarchies, contributions are quite limited to the ability concerns of students and neglect the presence of other hierarchies. In a pioneering work, Tincani (2015) notes that the classroom peer effect is moderated by relative ability rank, using as natural experiment an earthquake in Chile. Students, being aware of their own ability, exhibit rank-concern behavior, internalizing their own ordinal position in a social environment. In experimental studies, Azmat & Iriberri (2010) show that information on relative group rank before a test can shape students' performance, effort, and decision. More recent studies examine how school ordinal rank affects achievement in the UK (Murphy & Weinhardt, 2016) and college completion in the US (Elsner & Isphording, 2015). On the same topic, Gamoran (1989) investigates how ordinal position in a students' reading group, based on a teacher's perceptions with a secondary check by a standardized test, influenced student achievement over time in six schools in Chicago. Finally, Ball & Newman (2013) propose a synthetic index based on network ties to proxy "*social status*."

So far, the role of teachers in shaping hierarchies among students is quite neglected, but several studies move in this direction. Several contributions highlight the presence of an asymmetric student-teacher relation in which students internalize teachers' beliefs, attitudes, and behavior in light of their experiences over time (Schizzerotto & Barone, 2006). Previous research finds undisputed that teachers' assigned mark reflects student's skills, but also teachers' subjective preferences and contextual factors. Indeed, teachers consider achievements, participation, motivation, effort, and behavior (Brookhart, 1993) but also classroom characteristics such as socio-economic or ethnic backgrounds (Westphal, Becker, Vock, Maaz, Neumann & McElvany, 2016). Finally, it is widely accepted that teachers discriminate students through grading students according to some characteristics such as ethnic status (Bowles & Gintis, 1976), gender (Di Liberto & Casula, 2016), and BMI (Braningan, 2017). Since mark could signal productivity (OECD, 2012), grading subjectivity might trigger a hierarchy with consequences on students' outcomes.

This work considers that hierarchies arise driven by endogenous interaction among student and social actors. As a contribution to the literature, the work will examine, in chapter II, the extent to which teacher's grading shapes a hierarchy among students with consequences on student's outcomes. It explicitly addresses an idea quite central in the BFLPE literature and applied research on human resources incentives (Gill, Zdenka, Jaesun, & Prowse, 2018):

asymmetric positions among students driven by endogenous interaction of external actors affects student' educational outcomes such as performance, expectations, and choice.

6. Networks

The formation of networks is embedded in human behavior. People aim at drawing on relations, and an extensive literature has analyzed reciprocity and exchange (Blau, 1964), triadic relations (Simmel, 1955), and strong and weak ties (Granovetter, 1973). In the last thirty years, researchers have investigated the network position of individuals in terms of popularity and the extent to which people set a group (homophily) according to characteristics such as ability, socio-economic status, ethnic background, shared outdoor or indoor activities, love stories, drugs, and alcohol (Lusher, Koskiner, and Robins 2013).

In a classroom, students tend to befriend some more than others, and it can happen according to several characteristics. The existing literature has developed several hypotheses about friendship formation. According to the propinquity effect, students become friends just from interacting with each other (Lusher et al., 2013; White, Boorman & Breiger, 1976). Then, students may get involved in friendships because they share the same values, stereotypes, ethnic identities, have similar incomes, or simply because they play indoor or outdoor activities in the same team or club (Lusher et al., 2013). Among others, the contact hypothesis is frequently used to single out self-stigma or ethnic pride in multi-ethnic schools, to investigate the impact and formation of gender stereotypes using indicators of friendship (van der Vleuten, Steinmetz, & van de Werfhorst, 2018) and real networks (Smith, 2015; Leszczensky, 2018). Networks of friends are important in explaining the aspirations of students (Burgess & Aponte, 2011), their mental behavior (Plenty & Mood, 2016), the consumption of tobacco and alcohol (Alm & Laftam, 2016), being unemployed (Daraganova & Pattison, 2012), the sharing of similar views and attitudes (Lusher et al., 2013), and the choice of enrollment in basic or advanced courses (Heck, Price & Thomas, 2004).

Network formation follows an endogenous pattern, which confronts scholars with the proper identification of a network effect on students' outcomes. In line with this, in recent years, there has been a growing literature focusing on possible solutions to tackle technical issues such as reflection, endogenous effect, and contextual effect (Paloyo, 2020). At the moment, there is a divide in social sciences. On the one hand, sociology is more oriented toward simulation studies and ABMs with a stronger emphasis on patterns underlying network formation. On the other, economics is more oriented toward a mix of network studies and spatial econometrics with an emphasis on modeling of exogenous variation on network formations (Scott & Carrington, 2011). Against this backdrop, the work of Bramoullé and colleagues (2009), and Patacchini and colleagues (2017) is notable because they reconcile these two swings of social sciences.

Once accounted for this rich literature and opposite perspectives, the work will revive, in chapter III, an old topic like network effect on adolescent smoking and drinking. By distinguishing between reciprocal and non-reciprocal ties, it adds to the current debate a neglected approach to causal identify the effect of a friend's network (Bramoullé and colleagues, 2009).

7. Classroom peers

Wilkinson and colleagues (2000) define compositional effects as aggregated characteristics of "groups" affecting learning outcomes such as the mean level of ability. Compositional effects are generally analyzed for measuring the heterogeneity of students' composition resources, climates, and practices. In contrast, peer effects are based on a strict concept of interaction between students and other individuals. Several mechanisms have been proposed in the literature to regulate peer effects. *Normative explanation* (Erbring & Young, 1979) states that students internalize the norms of their educational environment; *comparative explanation* focuses on the importance of a reference group (e.g., classrooms, friends, or individuals) for comparisons (Kelly, 2009; Borman & Pyne, 2016); *instructional explanation* (Beckerman & Good, 1981) suggests that teachers regulate their practices according to the features of the classroom and/or groups of students; and *language explanation*, emphasizes the role of language constraints in magnifying or reducing communication among students and teachers (Vygotsky, 1978; Rogoff, 1990).

Typically, contributions use these theoretical lenses to systematize how classroom characteristics such as the share of non-native or the share of low ability students influence cognitive and non-cognitive skills. A normative explanation can lead to a negative vicious circle in the presence of disadvantaged classrooms. Comparative explanations pave the way to multidirectional peer pressure among students in the function of several characteristics. Instructional explanation means that teachers calibrate their instructions according to the classroom level, giving to the disadvantaged classrooms a lower quality of instruction. Language explanation suggests that ethnic and socio-economic minorities perceive themselves out of context if they are segregated, with negative effects on their self-esteem (Borman & Pyne, 2016). In sum, not only peers' characteristics and the way their variation is associated with students' outcomes, but also the role of teachers and how they affect classroom outcomes emerge to be of crucial significance. Many studies focus on peer effects based on peer ability (Sacerdote, 2001; Lavy et al., 2012; Murphy & Weinhardt, 2016), socio-economic origin (Holtmann, 2016; van Ewijk & Sleegers., 2010; Hornstra., van der Veen, Peetsma, & Volman 2014) and ethnic background (Jensen & Rasmuessen, 2011; Contini, 2013). They tend to show that a high percentage of ethnic minorities or low SES students may be prevented from learning the national language to a high standard, reducing motivation and self-efficacy.

The literature on the role of teachers at the classroom level is relatively less extensive but heterogeneous, focusing on topics such as the consequences of the uneven distribution of teachers' characteristics across schools (Hanushek & Rivkin, 2012), teachers' expectations on students (Rosenthal, 2002), teaching strategies on performance and attitudes (Korbel & Paulus, 2017), and teachers' evaluations on expenditure for remedial tutorial lessons (Kiss, 2017). Results show that a teacher's characteristics such as job stability, in-field teaching, seniority, tertiary degree graduation mark, and possession of a tertiary degree are unevenly distributed among schools (Hanushek & Rivkin, 2012) or classrooms (Abbiati, Argentin, & Gerosa 2017), with an impact on students' performance. Besides, beliefs and attitudes of teachers have been reported to harm students' performance or motivation (Sansone, 2016), and teaching styles may affect students' participation (Korbel & Paulus, 2017).

Once accounted for the rich debate on causal identification of classroom peers effect (Angrist & Lang, 2004), the work will analyze, in the chapter IV, the role played by ethnic peers on student outcomes. It enriches the debate with a comprehensive array of socio-emotional skills and academic competences, and it brings out more the role played by linguistic diversity. Finally, it debates the sorting policy on student and teacher sides within schools.

8. Educational systems as "sorting machines."

The aim of the work is to investigate peer effects within the classroom in which "the machine" exposes students to varying characteristics of friends, peers, and teachers. Henceforth, it is critical to account for the degree of diversity across classrooms since such diversity has a role in shaping the formation and intensity of hierarchies, a network of friends, and classroom peers. Indeed, where does the *diversity* came from? Namely, it comes from the teacher's and student's sorting, but it actually depends on the interplay of macro policies regarding teacher training, tracking, and residential housing. Even if the work does not analyze this level, the debate of such policies – in particular the tracking one – theoretically enriches the analyses and support the identification approach of empirical chapters. It happens because "the machine" exposes students to a classroom with varying characteristics of friends, peers, and teachers upon which then they make their friendship and draw on hierarchies.

In the educational systems, students sort across schools and, in turn, across classrooms according to their characteristics (i.e., socio-economic origin, ethnic background, and academic performance) and the sorting of students - governed by *formal* and *informal rules* – conditions the related diversity degree. Explicit school enrolment procedures (e.g., selection based on previous academic performance) and residence criteria are an example of *formal rules*, whereas principals' decisions and parents' choices exemplify *informal rules* (Harris, 2011; Bohlomark, Holmlund & Lindahl, 2016). Since these rules might show high heterogeneity across municipalities, regions, and countries, the education system looks like a complex multilayer *Matryoshka* composed of students nested in classrooms, schools, and neighborhoods. Thus, national legislation is a key factor in giving more or less stance to informal regulations on school enrollment. The more binding the legislation of policymakers, the lower the stance of parents in deciding the children's school of enrolment. Likewise, principals or school boards may formally or informally manage classroom composition, producing homogeneous or heterogeneous groups in terms of ability, ethnic background, and socio-economic origin. All things considered, the cognitive and non-cognitive inequalities among students depend not only *directly* on classroom diversity but also *indirectly* on formal and informal sorting of students across educational systems and broad national policies.

In the educational system literature, it is undisputed that several cross-country differences emerge and tracking entry age is a good criterion to discriminate among systems (Hanushek & Woessman, 2006; Mons, 2007; Bol, Witschge, Van de Werfhorst & Dronkers, 2014). Based on the main model of secondary education, Blossfeld, Buchholz, Skopek, & Triventi (2016) have identified four dominant types of educational systems: Early tracking, Nordic inclusive, Individual choice, and Mixed tracking models. The *early tracking model* includes countries such as Germany, Hungary, the Netherlands, and Switzerland, which are characterized by a common early stratification even if there are some differences regarding the binding of standardized tests and teacher's recommendations to allocate in distinct school tracks. The *Nordic inclusive model* is used in Denmark, Finland, and Sweden. These countries are characterized by late tracking (at 16 years old) into a two-stream system distinguishing between academic and vocational paths, even though shadow grouping ability might be present like in Sweden. The *Individual Choice Model* present in Australia, England, Ireland, Scotland, and the US, is based on a mix of formal and informal differentiation, such as student's preferences for courses or access to educational stages conditional on prior achievement. Finally, the *mixed tracking model* found in countries such as Estonia, France, Israel, Italy, and Russia is based on academic, technical, vocational paths with strong school and regional divides.

However, the canonical distinction between late or early tracked systems is not sufficient to explain the sorting patterns because educational systems vary in other, not least important, facets such as *course-taking patterns*, the *composition of the classroom within the school*, the *stability of classroom composition* over the specific educational stage, the *chance of changing classroom*, *course*, *track*, and *school policy enrollment*. Such dimensions alter the general sorting toolkit, affecting, in turn, classroom diversity and student inequality. Hence, the personalization of subjects or levels, the option of tracking choice postponement, and the grouping ability option are an actual example - among many others - of how institutional features have pervasive "sorting" effects and affect the heterogeneity degree of learning environments in term of characteristics. Although the work is exclusively focused on what happens in a classroom, it recognizes that the classroom is only the last stage of a broader structure, and it accounts for more details for the Netherlands and Italy educational systems because the thesis exploits data from these countries.

In *the Netherlands,* the educational system does not allow regional differences in curriculum or quality, and it is highly centralized with a strong achievement orientation across educational stages. Compared to other educational systems, however, schools have a higher level of autonomy on the content of the curriculum, didactic methods, religious and philosophical backgrounds. The Dutch secondary education system provides an early selection into four main tracks at age 12 (Dronkers & Korthals, 2016). These are practical education (PO, secondary education on elementary school level), pre-vocational track (VMBO), and two general tracks. The first general track provides access to the university of applied sciences (pre-college track, HAVO), the second to university (pre-university track or grammar school, VWO). The prevocational track further includes four sub-tracks: VMBO-B is the most practical vocational track, VMBO-K is the foreman track, VMBO-G and VMBO-T have an increasingly theoretical focus. It is worth noting that secondary school principals decide the track placement of students, relying on an elementary school exit test and a track recommendation from the student's elementary school teacher.

The *Italian* education system is an example of late tracking, occurring around age 14 when students must choose between three main tracks: lyceum, technical, and vocational (Triventi & Contini, 2016). Before high school, students follow the same stages, such as elementary (5 years) and lower secondary schools (3 years). Despite each track providing distinct programs, the tracks remain distinct in terms of curricula, academic standards, composition, and prestige. In contrast with other systems, formally, students are free to choose any track. There is neither achievement orientation as in the Netherlands nor teacher's recommendation as in Germany. This policy leaves more room for family and neighborhood rules of enrollment, and the choice is driven by social origin (Argentin, Barbieri, & Barone, 2017). During high school, classroom composition is highly stable, and there are no options such as internal tracking or course differentiations: students are grouped for five years, learning the same subjects very frequently from the same teachers. The main source of variation is due to drop out or grade repetition, which mainly occurs in the vocational track. Track change is conditional upon the principal's decision and characterized by a "flight" to lower tracks. Also, Italian law omits to indicate classroom composition rules in detail, giving the principal a great autonomy on the assignment of students and teachers to the classroom. Therefore, even if any form of unequal sorting on the teaching and student side is formally forbidden, patterns of inequality occur along three so far tested dimensions: ethnic, in primary and lower secondary schools (Contini, 2013), socio-economic background, in primary schools (Agasisti

& Falzetti, 2017), and the teacher's characteristics, in secondary schools (Abbiati et al., 2017).

The ways of organizing the educational system may have important consequences for the characteristics of students. More recently, Domina and colleagues (2017) have summed up how organizational school practices such as internal tracking, group ability, grading, and course-taking patterns lead to categorical inequality, but previous studies have already discussed the phenomenon. Among several examples, Brookover and Schneider (1975) identify a sense of school futility to explain differences in levels of achievement among schools and tracks (Brookover & Schneider, 1975). In contrast, Van Houtte and Stevens (2015) suggest that when students enrolled in lower tracks are confronted with students enrolled in higher tracks in the same school; they tend to lose faith in the school system. Additional examples are the works of Trautwein, Ludtke, Koller, and Baumert (2006) and Mijs (2016) in which the homogeneity or heterogeneity of learning environment affects the self-esteem and the internalization of failure.

So far, this work has extensively discussed the role of *tracking policies,* but there are two complementary policies, such as *teaching staff* and *residential housing policies*. The role of teaching staff does not find a well-painted systematization in the literature on educational systems. Nonetheless, the characteristics of teachers play a pivotal role in student's outcomes (Hanushek & Rivkin, 2012). Particularly important for student educational performance is teacher quality and teacher training. To date, Eurydice (2008; 2013) stresses the presence of two initial teacher training models in Europe: a concurrent model, where the practical skills are developed during the general courses, and a consecutive model, where professional skills are acquired in-field directly. In contrast, residential housing policies are more investigated for the interplay with school segregation from the pioneering works of Rivkin (1994). He started a debate on to what extent racial segregation in the US was exacerbated by housing segregation, jointly with the more recent contributions of Scott (2005). Hence, residential housing policies might operate as increasing or decreasing unequal sorting factors, above all, when interacting with neighborhood enrolment rules (Calsamiglia, Martinez-Mora, & Miralles, 2020). In addition, a flourishing stream of research points out how housing policies may affect and be affected by educational systems at all. Among some examples, Figlio and Lucas (2004) show how high school grades are a good predictor of housing prices in Florida, and Battistin and Neri (2020) argue that this link is one of the possible mechanisms of grade inflation in the UK.

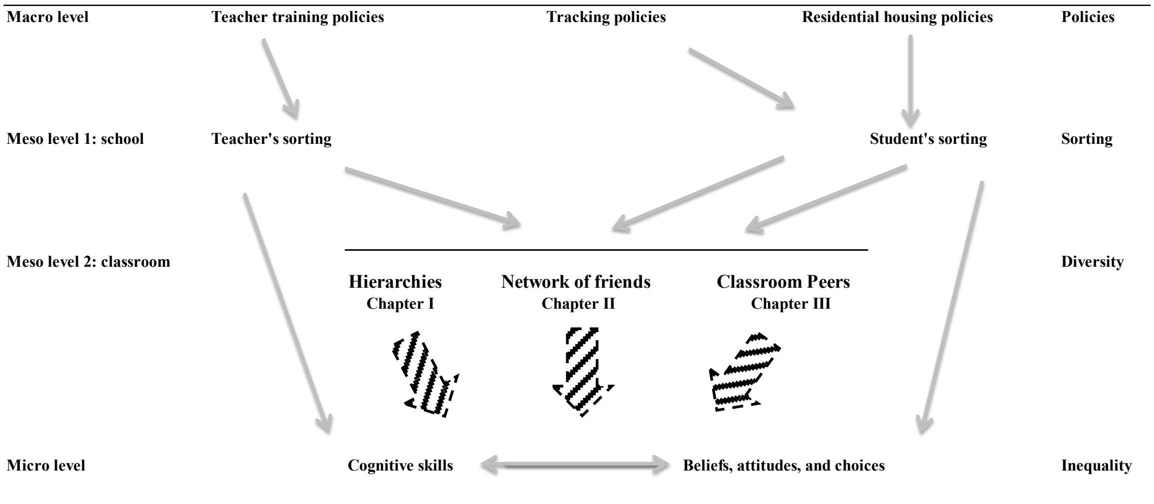
Countries, therefore, show distinct policies that shape students' paths: teacher training, tracking, and residential housing. Teacher training policies affect the distribution of teachers across schools and classrooms. Tracking policies influence students' educational stages. Residential housing policies shape the distribution of people across space, and in turn, the enrollment of students in a commuting area. Broadly, these distinct policies depend on "*the accepted mode of upward mobility that shapes the school system directly and indirectly through its effects*" (Turner, 1960, p. 855) across society.

9. Theoretical framework

This work has developed a theoretical framework upon which it analyzes the effect of hierarchies, network of friends, and classroom peers on congnitive skills, non congnitive skills, education decisions, and helthy behaviours. To make it possible, it exploits three related but mutually exclusive concepts: inequality, diversity, and sorting. Reformulating the contribution of Roberto (2015), *inequality* refers to the uneven distribution of resources, opportunities, or outcomes across students; *diversity* describes the variety of "types" in the student population and *sorting* refers to the uneven distribution of students and teachers across distinct schools. The allocation of students and teachers with specific characteristics depends on broad *policies* at the macro level. Thus, the inequalities of outcomes among students depens on the related diversity across classrooms, the underlying sorting, and the macro country policies.

In the current chapter, the work has proposed a conceptual framework (Figure I). This framework is based on three levels: macro, meso, and micro. At the *macro level,* teacher training, tracking policies, and residential housing policies affect the sorting of students and teachers. At the *meso level*, the teacher's and student's sorting conditions the classroom diversity. In the classroom, hierarchies, networks of friends, and classroom characteristics affects the *micro-level* composed by beliefs, attitudes, performances, and choices. The black dashed arrows indicate what the thesis investigate in the empirical chapters.

Figure I: Theoretical framework



Note: black dashed arrows indicate issues under investigation in the trhee empirical chapters

In chapter II of the thesis, the work will address the role played by teacher's grading in shaping hierarchies among students and it will test the pervasive effect on socio-emotial skills, academic performance, and educational choices. Along this setting, it will analyze also possible differences among boys and girls regarding the reaction to such hierarchy. In chapter III of the thesis, the work will debate the role played by friends in the adoption of unhealthy habits such as smoking and drinking. It contributes to the literature, analyzing reciprocal and non-reciprocal ties and accounting for the rich debate on causal identification. Finally, in the last IV chapter, the work will deal with the emerging increase of non-native students in the school. It enrich the current debate by focusing on a comprehensive array of socioemotional skills, behaviours, and academic performance. Besides, it considers more in depth the role played by principals in the sorting policy.

References

Abbiati G., Argentin G. & Gerosa T. (2017). Different teachers for different students? Evidence on teacher-student matching and its consequences in the Italian case. *Politica Economica*, *33*(1), 13-58.

Agasisti, T., & Falzetti, P. (2017). Between-classes sorting within schools and test scores: an empirical analysis of Italian junior secondary schools. *International Review of Economics*, *64*(1), 1-45.

Ajzen I. (1991). The Theory of Planned Behavior, *Organizational Behavior And Human Decision Processes*, *50*(2), 179-211

Alm, S. & Låftman S.B. (2016). Future orientation climate in the school class: Relations to adolescent delinquency, heavy alcohol use, and internalizing problems. *Children and Youth Services Review*, *70*, 324-331.

Almås I., Cappelen W. A., Salvanes G. K., Sørensen E. Ø., & and Tungodden B. (2015). Willingness to Compete: Family Matters*, Management Science, 62*(8), 2149-2162

Almlund, M., Duckworth A., Heckman J., & Kautz T. (2011). Personality psychology and economics. IZA Discussion Paper 5500.

Argentin G., Barbieri G., & Barone C. (2017), Origini sociali, consiglio orientativo e iscrizione al liceo: un'analisi basata sui dati dell'Anagrafe Studenti, *Politiche Sociali,* 1, 53-74

Anger S. (2013). Personality and educational attainment. Working Paper, DIW Berlin http://www.iwaee.org/papers%20sito%202013/Anger.pdf

Angrist, J.D., & Lang, K., (2004). Does school integration generate peer effects? Evidence from Boston's Metco Program. *American Economic Review 94* (5), 1613–1634.

Azmat, G. & Iriberri, N. (2010). The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics*, *94*(7-8), 435–452.

Babad E. (2009). *The social psychology of the classroom*. Routledge Research in Education.

Ball, S.B., Eckel, C.C., Grossman, P.J., & Zame, W. (2001). Status in Markets. *Quarterley. Journal of Economics*, *116* (1), 161–188.

Ball B. & Newman M.E.J. (2013). Friendship networks and social status. *Network Science 1*(1), 16-30.

Bar-Tal, D. (1979). Interactions of teachers and pupils. In I. Frieze, D. Bar-Tal & J.S. Carroll (Eds.) *New approaches to social problems: Applications of attribution theory* (pp. 337-358). San Francisco: Jossey-Bass.

Battistin E. & Neri L. (2020). School Performance, Score Inflation and Economic Geography ZA Discussion Paper No. 11161

Bault, N., Coricelli, G., & Rustichini, A., (2007). Interdependent utilities: How social ranking affects choice behavior, mimeo. University of Minnesota.

Beckerman, T. M., & Good, T. L. (1981). The classroom ratio of high- and low-aptitude students and its effect on achievement. *American Educational Research Journal*, *18*, 317–327.

Becker, G. S., (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship.

Berger, J & Webster M. (2006). "Expectations, Status, and Behavior." Pp. 268–300 in *Contemporary Social Psychological Theories*, edited by P. J. Burke. Stanford, CA: Stanford University Press.

Bernardi F. & Ballarino G. (2016), Education as the great equalizer: a theoretical framework, in Education, Occupation and Social Origin. In Bernardi F. & Ballarino G. *A Comparative Analysis of the Transmission of Socio-Economic Inequalities*, Editors: Fabrizio Bernardi and Gabrielle Ballarino, Elgar Publishing

Blau, P. M. (1964). *Exchange and power in social life*. New York: Wiley.

Blossfeld, H.- P., Buchholz, S. Skopek J. & Triventi M. (Eds.) (2016). *Models of Secondary Education and Social Inequality. An International Comparison*. Cheltenham, UK: Edward Elgar.

Bobba M., & Frisancho V., (2014). Learning About Oneself: The Effects of Signaling Academic Ability on School Choice, Inter-American Development Bank, mimeo.

Böhlmark, A., Holmlund, H., & Lindahl, M. (2016). Parental choice, neighbourhood segregation or cream skimming? An analysis of school segregation after a generalized choice reform. *Journal of Population Economics*, *29*, 1155–1190.

Bol, T., Witschge J., Van de Werfhorst H. G., & Dronkers J. (2014). Curricular Tracking and Central Examinations: Counterbalancing the Impact of Social Background on Student Achievement in 36 Countries. *Social Forces, 92*(4), 1545–72.

Borman G.D. & Pyne J. (2016). What If Coleman Had Known About Stereotype Threat? How Social-Psychological Theory Can Help Mitigate Educational Inequality. *The Russell Sage Foundation Journal of the Social Sciences* (RSF) 164-185

Bowles, S., & Gintis, H. (1976). *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life.* New York: Basic Books.

Bramoullé, Y., Djebbari, H. & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150(1), 41–55.

Braningan, A. R. (2017). (How) Does Obesity Harm Academic Performance? Stratification at the Intersection of Race, Sex, and Body Size in Elementary and High School. *Sociology of Education*, *90*(1), 25–46.

Brookhart M. S. (1993). Teachers' Grading Practices: Meaning and Values. *Journal of Educational Measurement*, *30*(2), 123-142.

Brookover, W. B. & Schneider, J. M. (1975). Academic environments and elementary school achievement. *Journal of Research and Development in Education*, 9, 82–91.

Brunello, G., & Schlotter M. (2011). Non-Cognitive Skills and Personality Traits: Labour Market Relevance and Their Development in Education & Training Systems. Discusssion Paper IZA N° 5743

Bol T, Witschge J., Van de Werfhorst H., & Dronkers J., (2014) Curricular Tracking and Central Examinations: Counterbalancing the Impact of Social Background on Student Achievement in 36 Countries, *Social Forces*, *92*(4), 1545–1572

Burgess S., Sanderson E., & Umana-Aponte M. (2011). School ties: An analysis of homophily in an adolescent friendship network, The Centre for Market and Public Organization, Department of Economics, University of Bristol, UK.

Bukodi E., Eibl F., Buchholz S., Marzadro S., Minello A., Wahler S., Blossfeld Hp., Erikson R., & Schizzerotto A. (2018). Linking the Macro to the Micro: A Multidimensional Approach to Educational Inequalities in Four European Countries. *European Societies*;. *20*(1), 26-64.

Calsamiglia C. & Loviglio A. (2020). Grading On A Curve: When Having Good Peers Is Not Good, Barcelona GSE Working Paper Series.

Carlana, M. (2019). Implicit Stereotypes: Evidence from Teachers' Gender Bias. *Quarterly Journal of Economics*, *134*(3), 1163–1224

Cobb-Clark, D. A., & Schurer S. (2013). Two economists 'musings' on the stability of locus of control. *Economic Journal*, *123* (570), F358-F400

Collins, R. L. (1995). For better or worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin*, *119*(1), 51–69.

Contini, D. (2013). Immigrant Background Peer Effects in Italian Schools. *Social Science Research, 42*(4), 1122–42.

Conti G, Heckman, J., & Urzua S., (2010). The Education Health Gradient. *American Economic Review: Papers & Proceedings*, *100*, 234–238

Correll S.J. & Ridgeway C.L. (2006). *Expectation States Theory*. In: Delamater J. (eds) Handbook of Social Psychology. Handbooks of Sociology and Social Research. Springer, Boston, MA

Croson, R. & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*. *47*(2), 1–27.

Cutler M. D. & Lleras-Muney A. (2006). Education and Health: Evaluating Theories and Evidence. NBER.

Daraganova, G., & Robins, G. (2012). Autologistic Actor Attribute Models. In D. Lusher, J. Koskinen, & G. Robins (Eds.), *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications* (Structural Analysis in the Social Sciences, pp. 102-114). Cambridge: Cambridge University Press.

Dasgupta N. (2013) "Chapter Five - Implicit Attitudes and Beliefs Adapt to Situations: A Decade of Research on the Malleability of Implicit Prejudice, Stereotypes, and the Self-Concept" *Advances in Experimental Social Psychology*, *47*, 233-279

Di Liberto, A. & Casula, L., (2016). Teacher Assessments versus Standardised Tests: Is Acting “Girly” an Advantage? IZA Discussion Papers 10458

Dronkers R. & Korthals R. (2016). Tracking in the Netherlands – Ability selection or social reproduction? in Blossfeld, H.-P., Buchholz, S. Skopek J. & Triventi M. (Eds.) (2016). *Models of Secondary Education and Social Inequality. An International Comparison.* Cheltenham, UK: Edward Elgar.

Domina T., Penner A., & Penner E. (2017). Categorical Inequality: Schools As Sorting Machines. *Annual Review of Sociology*, *43*, 311-330

Eccles, J., S. (1993). School and family effects on the ontogeny of children’s interests, self-perceptions, and activity choice. In J. Jacobs (Ed.), *Nebraska symposium on motivation, 1992: Developmental perspectives on motivation* (pp.145-208) Lincoln: University of Nebraska Press.

Erbring, L., & Young, A. A. (1979). Individuals and social structure: Contextual effects as endogenous feedback. *Sociological Methods & Research*, *7*(4), 396–430.

Elsner B., & Isphording I. (2015). “A big fish in a small pond: Ability rank and human capital investment”. *Journal of Labor Economics, 35*(3), 787-828

Eurydice (2008). *Levels of Autonomy and Responsibilities of Teachers in Europe.* European Commission

Eurydice (2013). *Key Data on Teachers and School Leaders in Europe.* European Commission

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140.

Figlio, D. & M. Lucas (2004). Do high grading standards affect student performance?, *Journal of Public Economics*, *88*(9-10)1815-1834

Fuligni J. A. (1997). The Academic Achievement of Adolescents from Immigrant Families: The Role of Family Background, Attitudes, and Behavior. *Child Development*, *68*(2), 351-363

Gamoran, A. (1989). Rank, Performance, and Mobility in Elementary School Grouping. *Sociological Quarterly*, *30*(1), 109-123.

Gasperoni, G. (1998). *Il rendimento scolastico.* Bologna: Il Mulino.

Gill, D., Zdenka, K., Jaesun, L. & Prowse, V. (2018). First-Place Loving and Last-Place Loathing: How Rank in the Distribution of Performance Affects Effort Provision, *Management Science*, *65*(2), 459-95

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380.

Grigutsch S., & Törner G. (1998) World views of mathematics held by university teachers of mathematics science, Schriftenreihe des Fachbereichs Matematik, Preprint 420. Duisburg : Gerhard Mercator University

Jackson M. (2013). *Determined to succeed?: performance versus choice in educational attainment*. Stanford University Press.

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.

Halek, M., & Eisenhauer, J. (2001). Demography of Risk Aversion. *The Journal of Risk and Insurance*, *68*(1), 1-24.

Hanushek, E. A., & Woessman, L. (2006). Does Educational Tracking Affect Performance And Inequality? Differences In Evidence Across Countries. *The Economic Journal, 116*(510), 63–76.

Hanushek, E. A. & Rivkin G. S. (2012). The Distribution of Teacher Quality and Implications for Policy. *Annual Review of Economics*, *4*, 131-157.

Harris, R. (2011). Measuring social segregation between London's secondary schools. Centre for Market and Public Organization Bristol Institute of Public Affairs University of Bristol.

Heckman, J. J., & Stixrud J. (2006). The effects of cognitive and non-cognitive abilities on labor market outcomes and social behavior. *National Bureau of Economic Research*, *53*, (9), 1689–1699.

Heckman J.J. (2000). Policies to foster human capital. *Research in economics, 54*(1), 3-56.

Heffetz O. & Robert H. F. (2011) Preferences for Status: Evidence and Economic Implications. In J. Benhabib, A. Bisin, M Jackson, *Handbook of Social Economics*, North Holland

Heck R. H. Price C. L. & Thomas S. L. (2004). Tracks as Emergent Structures: A Network Analysis of Student Differentiation in a High School. *American Journal of Education*, *110*(4) 321-353

Henson, H.. (1964). *Adaptation-Level Theory: An Experimental and Systematic Approach to Behavior*. Harper and Row: New York.

Holtmann C. A., (2016). Excellence through equality of opportunity – Increasing the social inclusiveness of education systems benefits disadvantaged students without harming advantaged students. In Blossfeld, H.- P., Buchholz, S. Skopek J. and Triventi M. (Eds.) (2016). *Models of Secondary Education and Social Inequality. An International Comparison*. Cheltenham, UK: Edward Elgar.

Hornstra L., van der Veen I., Peetsma T., & Volman M., (2014) Does classroom composition make a difference: effects on developments in motivation, sense of classroom belonging, and achievement in upper primary school. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice Publication, 26*(2), 125-152

Hoxby, C. & Weingarth, G. (2005). Taking race out of the equation: school reassignment and the structure of peer effects. Technical report, Department of Economics, Harvard University.

Kao G. & Tienda M. (1988). Educational Aspirations of Minority Youth. American Journal of Education, *106*(3), 349–384.

Kelly, S. (2009). Social identity theories and educational engagement. *British Journal of Sociology of Education*, *30*(4), 449–4.

Kelley, Harold H. (1952). The Two Functions of Reference Groups. pp 410-14. In G. E. Swanson, T. M. Newcomb and E. L. Hartley *Readings in Social Psychology*, editors. Hoit.

Kempel, J., & Willner, C. (2008). *Career Academies: Long-term impacts on labor market outcomes, educational attainment, and transitions to adulthood*. New York: MDRC.

Kemptner D., Jurges H., Reinhold S., (2011). Changes in compulsory schooling and the causal effect of education on health: Evidence from Germany, *Journal of Public Health*, 30(2), 340-54

Kiss D. (2018) How do ability peer effects operate? Evidence on one transmission channel, *Education Economics, 26*,3, 253-265

Korbel, V. & Paulus M. (2017). Do Teaching Practices Impact Socio-emotional Skills? Working Papers IES, https://ideas.repec.org/p/cer/papers/wp591.html

Lavy V. Weinhardt F., & Silva O. (2012). The Good, The Bad and The Average: Evidence on Ability Peer Effects in Schools. *Journal of Labor Economics*, *20*(2), 367-414.

Lekfuangfu N.W., Cornaglia F., Powdthavee N., & N. Warrinnier (2014). Locus of Control and Its Intergenerational Implications for Early Childhood Skill Formation. Center for Economic Performance, Discussion Paper No 1293.

Lenton P. (2014). Personality Characteristics, Educational Attainment and Wages: An Economic Analysis Using the British Cohort Study. Sheffield Economics Research Papers Series.

Leszczensky, L, Flache A., Stark T. & Munniksma A. (2018). The Relation between Ethnic Classroom Composition and Adolescents' Ethnic Pride. *Group Processes & Intergroup Relations*, *21*(7), 997-1013.

Lucas, R. E. & Donnellan, M. B. (2009). If the Person-Situation Debate Is Really over, Why Does It Still Generate So Much Negative Affect?, *Journal of Research in Personality, 43*(2), 146-149.

Lusher D., Koskinen J. & Robins G. (2013). *Exponential graphs random models*. Cambridge University Press.

Marsh W. E. (2008) "The Big-fish–little-pond-effect Stands Up to Critical Scrutiny: Implications for Theory, Methodology, and Future Research". *Education Psychology Review*, *20*, 319–350.

Mann, H. (1957). Report No. 12 of the Massachusetts School Board. In: Cremin, LA., editor. The republic and the school: Horace Mann on the education of free men. Teachers College Press; New York: 1957

Merton, Robert K. (1968). *Social Theory and Social Structure*. Free Press.

McEwan, P. J. (2003). Peer effects on student achievement: Evidence from Chile. *Economics of Education Review*, *22*(2), 131–141.

McGue, M., Bacon, S., & Lykken, D. T. (1993). Personality stability and change in early adulthood: A behavioral genetic analysis. *Developmental psychology*, *29*(1), 96.

Mijs B. J. J. (2016). Stratified Failure: Educational Stratification and Students' Attributions of Their Mathematics Performance in 24 Countries. *Sociology of Education*, *89*(2). 137–153.

Mincher, Jacob (1958). "Investment in Human Capital and Personal Income Distribution". *Journal of Political Economy, 66*(4): 281–302.

Mons N. (2007). *Les nouvelles politiques éducatives*, Paris: PUF.

Moreno, J. L. (1934). *Who shall survive? A new approach to the problem of human interrelations*. Washington, DC: Nervous and Mental Disease Publishing.

Murphy, R. & Weinhardt, F. (2016). Top of the Class: the Importance of Ordinal Rank. Working Paper, https://ideas.repec.org/p/ces/ceswps/_4815.html

OECD (2012). *Grade Expectations. How marks and education policies shape students' ambitions*, OECD Publishing, Paris.

OECD (2013), *Gender and Socio-Economic Disparities in Students' Engagement, Drive and Self-Beliefs*, in PISA 2012 Results: Ready to Learn (Volume III): Students' Engagement, Drive and Self-Beliefs, OECD Publishing, Paris.

OECD (2016a), PISA 2015 Results (Volume I): Excellence and Equity in Education, PISA, OECD Publishing, Paris.

OECD (2016b), *Socio-economic status, student performance and students' attitudes towards science*, in PISA 2015 Results (Volume I): Excellence and Equity in Education, OECD Publishing, Paris.

OECD (2018). *Education at a Glance*. OECD Publishing, Paris.

Patacchini E. Rainone E., & Zenou Y. (2017). Heterogeneous Peer Effects in Education *Journal of Economic Behavior and Organization*, *134*(C), 190-227

Parducci, A. (1965). Category Judgment: A Range-Frequency Model. *Psychological Review* 76(2):40

Wilkinson G.A.I., Hattie J. Parr M.J. & Thrupp M., (2000). Influence of Peer Effects on Learning Outcomes: A Review of the Literature. *International Journal of Educational Research, 37*, 395–401.

Park, Robert E. (1924). The Concept of Social Distance as Applied to the Study of Racial Attitudes and Racial Relations. *Journal of Applied Sociology*, *8*, 339–44.

Paloyo. A (2020) Peer effects in education: recent empirical evidence, in Braedley S. Grenn C. The economics of education. A comprehensive overview. AP. *Forthcoming*

Plenty S. & Mood C. (2016). Money, Peers and Parents: Social and Economic Aspects of Inequality in Youth Wellbeing. *Journal of Youth Adolescence*, *45*, 1294–1308.

Gautam R. (2019). Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools. *American Economic Review, 109*(3), 774-809.

Ridgeway, C. L. (2006). Status Construction Theory. In P. J. Burke (Ed.), *Contemporary social psychological theories* (p. 301–323). Stanford University Press.

Rivkin, S. (1994). Residential Segregation and School Integration. *Sociology of Education, 67*(4), 279-292.

Roberts, B. W. (2009). Back to the Future: Personality and Assessment and Personality Development. *Journal of Research in Personality 43*(2), 137-145.

Roberto, E. (2015). The Divergence Index: A Decomposable Measure of Segregation and Inequality. Working paper, https://arxiv.org/abs/1508.01167

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York, NY: Oxford University Press.

Rosenthal, R. (2002). The Pygmalion effect and its mediating mechanisms. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (p. 25–36). Emerald Group Publishing Limited

Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics*, *116*(2), 681–704.

Sansone D. (2016). Teacher characteristics, student beliefs and the gender gap in stem fields. CERP Working Paper 165/16.

Schizzerotto A. Barone C. (2006). *Sociologia dell'istruzione.* Il Mulino.

Scott, John, Carrington, Peter J. (2011) *The Sage Handbook of Social Network Analysis*. The Sage Publications, London, pp. 514–532.

Scott J., (2005). *School Choice and Diversity: What the Evidence Says. Front Cover*. Teachers College Press

Simmel, G. (1955). *Conflict and the web of group-affiliations*. New York: Free Press.

Smith S. (2015). *Ethnic Segregation in Friendship Networks. Studies of its Determinants in English, German, Dutch, and Swedish School Classes*. Ph.D. dissertation.

Spring J. (1945) The Sorting Machine: National Educational Policy Since 1945. *The Journal of Education*, *159*(3). Review of Bruce C. Beezer.

Tarde, G. (1903). *The Laws of Imitation*. New York: Henry Holt.

Trautwein, U., Ludtke O., Koller O., & Baumert J. (2006). Self-esteem, Academic Self-concept, and Achievement: How the Learning Environment Moderates the Dynamics of Self-concept. *Journal of Personality and Social Psychology, 90*(2), 334–49.

Tincani M. (2015). Heterogenous peer effect in classrooms, Working paper, https://ideas.repec.org/p/ces/ceswps/_6331.html

Contini D. & Triventi M. (2016). Between formal openness and stratification in secondary education: Implications for social inequalities in Italy in Blossfeld, H.-P., Buchholz, S. Skopek J. and Triventi M. (Eds.) (2016). *Models of Secondary Education and Social Inequality. An International Comparison*. Cheltenham, UK: Edward Elgar.

Turner, R.H. (1960). Sponsored and Contested Mobility and the School System. *American Sociological Review, 25*(6), 855-67.

Unesco (2016). *Non-cognitive skills: Definitions, measurement and Malleability*. Global Education Monitoring Report.

van Ewijk, R., & Sleegers, P. (2010). The effect of peer socio-economic status on student achievement: A meta –analysis. *Educational Research Revie*w, *5*, 134–150.

van der Vleuten M., Steinmetz S. & van de Werfhorst H. (2018) Gender norms and STEM: the importance of friends for stopping leakage from the STEM pipeline, Educational Research and Evaluation, *24*(6-7), 417-436

Van Houtte M. & Steven A. J. P. (2015). Tracking and sense of futility: the impact of between-school tracking versus within-school tracking in secondary education in Flanders (Belgium). *British Educational Research Journal*, *41*(5), 782–800.

Vygotsky, L. S. (1978). *Mind and society: The development of higher psychological processes. Cambridge*, MA: Harvard University Press

Weiss, Y., & Fershtman, C., (1998). Social status and economic performance: A survey. *European Economic Review, 42*(3-5), 801–820.

Westphal, A., Becker, M., Vock, M., Maaz, K., Neumann, M. & McElvany, N., (2016). The link between teacher-assigned grades and classroom socio-economic composition: The role of classroom behavior, motivation, and teacher characteristics. *Contemporary Educational Psychology* ,*46*, 218-227.

White, H. C., Boorman, S. A., & Breiger, R. L. (1976). Social Structure from Multiple Networks. Blockmodels of Roles and Positions. *American Journal of Sociology*, *81*(4), 730-780.

Wilkinson, I.A.G., Hattie, J.A., Parr, J.M., Townsend, M.A.R., Fung, I., Ussher, C., Thrupp, M, Lauder, H., & Robinson, T., (2000) *Influence of Peer Effects on Learning Outcomes: A Review of the Literature*. Ministry of Education, New Zeland.

"School is a society in miniature" preparing us for life in wider society

Émile Durkheim[4]

[4] Quote reported in Mitchell, M. (1931). Emile Durkheim and the Philosophy of Nationalism. *Political Science Quarterly*, 46(1), 87-106.

Teachers’ evaluations, student rank, and pupils’ educational outcomes

Chapter II

Abstract

Students in compulsory education spend a considerable amount of time in the classroom interacting with the peers and teachers, from whom they receive feedback and signals about their academic competence. I argue that teachers – when attributing marks to students – create a classroom hierarchy and the perception of one's position in this hierarchy might affect his/her academic outcomes and socio-emotional skills. To test this hypothesis, I use population data on two cohorts of students in Italy, who are followed from the lower to upper secondary education. To identify the causal effect of student rank, I exploit the idiosyncratic variation generated by differential teachers' grading standards and student composition across classrooms. Fixed effects regression models show that student's position in the classroom rank rises math confidence and self-esteem, increases the probability of enrolling in the academic track, and reduces the risk of school dropout.

1. Introduction

Across educational systems, student assessment is a core practice. However, it takes different forms, such as teachers' assigned marks, standardized tests organized by the school, general feedback to parents via written or oral communication, and standard evaluations of student cognitive and non-cognitive skills by external agencies. Although the choice of the predominant practices might depend on teachers' strategies, school regulations, stages of education, and the features of educational systems, teachers' assigned marks are the leading method for assessing students in most educational systems of Western countries (OECD, 2012).

On teacher's assigned marks, previous literature has stressed three main topics, such as the impact of marks on social actors, drivers of grading framework, and grading standards.

First, marks possess a "*signalling*" dimension, reflecting an evaluation of pupils' potential skills, which could affect parental investment in the human capital of their offspring, admission to the university, and employment prospects. Secondly, personal beliefs and stereotypes were found to bias teachers' judgments, resulting in gender discrimination (Di Liberto & Casula, 2016; Carlana, 2019) or ethnic discrimination (Bowles & Gintis, 1976; Triventi 2019). Also, individual and classroom characteristics such as socioeconomic or racial backgrounds modify the dimensions and related weights of evaluations (Westphal, Becker, Vock, Maaz, Neumann & McElvany, 2016). Finally, due to the signalling dimension and drivers of grading framework, grading may result in higher or lower standards affecting educational performance and attainment in the labor market (Betts & Groegger, 2003; Figlio & Lucas, 2004; Babcock, 2010).

Nevertheless, grading standards are only measured as a gap exposure between the average classroom or school mean GPA and standardized tests. It is an open and unexplored question if and how teacher's grading – via other channels – affects student's socio-emotional skills, performance, educational expectations, and choices. I propose a novel perspective based on a "status approach" combining the Big Fish in Little pond Effect (BFLPE, hereafter) (Marsh, Seaton, Trautwein, Ludtke, Hau, Craven, & O'Mara, 2008) and the Social Construct Theory (SCT, hereafter) (Ridgeway, 2006). I argue that teachers' grading produces a classroom hierarchy where each student covers a specific rank along with the classroom mark distribution, once controlled for individual ability, socio-demographic characteristics, and classroom characteristics. Across two classrooms, two comparable students might obtain the same mark in a given subject but might have different status depending on how the teacher evaluated the other classmates.

If students internalize the teacher's hierarchy, this relative position may stimulate – explicitly or implicitly – a peer comparison as outlined in the BFLPE (Marsh et al. 2008). In turn, this relative position might be seen as a competitive advantage associated with specific individual characteristics (Ridgeway, 2006), which results in a peer effect, increasing or decreasing socio-emotional skills, educational expectations, and affecting choices. In other terms, students perceive themselves ranked according to a *teacher's ability distribution,* at the net of their competences.

In this work, I ask: How does the position in the percentile distribution of teachers mark affect students' outcomes, once controlled for individual and classroom characteristics? Relying on INVALSI census, I adopt a longitudinal perspective and exploit institutional features of the Italian school system to provide a credible estimate of the causal effect of student rank position on a variety of outcomes, by tackling a set of issues typical of the peer effect literature such as reflection, contextual effects, correlated effects as well as selection bias due to drop out. My contribution adds to the literature the extent to which teachers construct hierarchies among students affecting students' outcomes and the presence of the gender divide. Finally, it shed lights on possible differences between hierarchies based on teachers' mark and standardized tests.

2.Theoretical background

2.1 Teacher's hierarchies

From kindergarten to high school, Italian students spend, on average, 6 hours every day in the classroom (OECD, 2018). After family and before afternoon activities, the classroom is a crucial social environment where students interact, shaping each other's expectations, aspirations, and decisions (Wilkinson, Hattie, Parr, Townsend, Fung, Ussher, Thrupp, Lauder, & Robinson, 2000). One of the guiding ideas of this work is that education systems are social constructs in which multiple hierarchies are present. Indeed, Babad (2009) argues that an educational environment such as a classroom is a "little society" in which people interact and compare themselves. Along with this frame, Bar-Tal & Bar-Tal (1979; 1990) argue that students have the same status, but they differ from each other regarding other characteristics in a group such as leadership, expertise, or popularity.

Domina, Penner, and Penner (2017) expand upon such ideas, framing educational systems as "sorting machines" following early insights by Spring (1945). According to them, education systems sort students into different kinds of "categories" defined by grades, classrooms, course-taking patterns, and academic tracks, imposing related labels with consequences on student's outcomes. I investigate the extent to which classroom behaves – via teacher's grading – as "social machinery" influencing socio-emotional skills, performance, educational expectations, and choices.

Teachers might play a leading role in shaping classroom hierarchies because they usually interact, asymmetrically, with students not as peers but as adults with an authoritarian approach (Schizzerotto & Barone 2006) whose beliefs, actions, and implicit and explicit behaviors (Sun, Pennings, Mainhard, & Wubbels, 2019) shape student's mindset resulting in a Pygmalion or Golem effect[5] (2002).

To this perspective, grading synthesizes this interaction, adding up the plausible dimensions of a teacher's grading policy, such as individual preferences, grading dimensions, classroom characteristics, and grading standards. It is widely accepted that teachers have their preferences toward students' characteristics. Among many examples, teachers assign a mark premium or penalty to girls (Di Liberto & Casula, 2016), to immigrants (Kiss, 2013), and obese individuals (Braningan, 2017). These preferences depend on personal beliefs and stereotypes held by teachers (Carlana, 2019). Beyond these stereotypes, teachers also assess dimensions such as effort, inter-disciplinarity, responsibility, capability, past scores, and starting point (Brookhart, 1993). On the verge of grading a student, teachers are free to give more weights to some dimensions than others, merely discriminating students according to their preferences. Besides, teachers take into account classroom composition, being more generous in grading students when classrooms show a high share of low ow socioeconomic background or high share of ethnic peers (Westphal and colleagues, 2016).

Since the seminal work of Bowles and Gintis (1976) on grading standards, the literature has leveraged the classroom or school mean difference between teacher's marks and blind standardized tests as a proxy of grading standards. Notable findings show that harsh standards improve educational performance (Betts & Grogger, 2003), educational attainment, right

[5] The Pygmalion effect is the phenomenon whereby others' expectations of a person affect the person's performance. High expectations lead to better performance (Pygamalion) and low expectations to a worst performance (Golem).

conduct (Figlio & Lucas, 2004), and, later, wages (Babcock, 2010). However, these contributions assume a dichotomous grading standard topology, masking the fact the grading policy might be quite varying. Against this scheme, previous theoretical hints have suggested the presence of distinct grading policies such as grading as a form of punishment, zero grading, valedictorians, and grading on a curve (Guskey, 2004). In recent work, Iacus and Porro (2011) count up to fifteen grading standards in a set of Italian schools, implicitly arguing that each teacher set up a classroom grading policy.

Once accounted for individual preferences, grading dimensions, classroom characteristics, and grading standards, teachers shape a classroom mark distribution. I argue that the classroom mark distribution is an enhancing-inequality factor because it sticks labels to students through marks drawing a hierarchy in the classroom, once controlled for student and classroom characteristics.

In general, the reliability of a hierarchy depends on the extent to which it is (1) socially recognized, (2) desirable, and (3) non-tradable (Ball, Eckel, Grossman, & Zame 2001; Heffetz & Robert, 2011). Henson (1964) and Parducci's (1965) suggest that individuals infer and internalize their relative position for every character in an environment. This recognition leads to a status desire because individuals compare themselves to friends, mates, siblings, in every social environment (Festinger, 1954; Kelley, 1952; Merton, 1968; Marsh et al., 2008). The non-tradable feature exacerbates the internalization resulting in a status-concern with consequences on beliefs, attitudes, achievements, and choices. A mark-based hierarchy met all these conditions because students deal daily with the teacher's grading policy.

The consequences of this internalization are twofold. First, high ranked students perceive themselves as the best in the classroom compared to the others via the BFLPE (Marsh et al., 2008), enhancing their socio-emotional skills. Secondly, as time wears on, students start to "construct" this mark-based hierarchy as an indicator of competence (Correll and Ridgeway, 2006). Via mark-based hierarchy, teachers sort students into mark categories, producing a systematic influence over people. In consequence, categorical differences turn into patterns of educational inequality, regardless of "true student ability." and relative contexts. Students struggle between an external assessment of their competences and their self-evaluation. This categorical inequality triggers a "causal ladder" of human behavior and choices. Indeed, modification of socio-emotional skills such as self-esteem leads to less trust in their capacities, and in turn, it affects attitudes, achievements, and decisions. (Eccles, 1993).

The internalization degree of a hierarchy also depends on the extent to which such marks are public knowledge, the "signalling" of a teacher's assigned marks, and classroom interaction. Indeed, without public transparency on teacher's assigned marks, it is troublesome to support the effectiveness of a mark-based hierarchy. Second, several findings point out the "signalling" effect of marks since teachers spend more time with high-ranked students (Rosenthal, 2002), parents use marks as a criterion to invest in the human capital of their child, universities condition enrolment on specific grade averages, and employers use marks to measure a jobseeker's productivity in the early stages of the school-to-work transition (Bobba & Frisancho, 2014). Finally, classroom interaction is the key driver or hierarchy of peer pressure. The stronger and more frequent are the within environment interactions, and higher is the probability students will internalize their relative position. Italian education system adheres to these features since teacher's marks are public, either through oral announcements and intermediate or/and final table, "signalling" effect works, and Italian students are sorted in the same classroom for the all year.

An extensive literature investigates the effect of peer characteristics such as gender, social class, ethnic status, or ability on student cognitive and non-cognitive skills. (see, for a review Paloyo, 2020; Lavy, Weinhardt, & Silva, 2012), but this stream has always conceptually neglected the role of a hierarchy. Investigation of student's hierarchy is not entirely new, but such contributions rely at most on experiments using grading schemes, the use of a score unknown to peers, and a lack of distinction between standardized tests and teachers' assigned marks. Finally, they give little attention to institutional constraints. Azmat and Iriberri (2010) show that information on relative group rank before a test can shape students' performance and decisions. Pekrun, Cusack, Murayama, Elliot, and Thomases (2014) report that positional anticipated feedback affects attitudes toward motivation and performance goals. Using administrative data and a natural occurrence such as an earthquake in Chile, Tincani (2015) shows that rank affects academic performance and moderates the impact of peer effect. Recent studies examine how ordinal rank in school affects performance, attitudes, and the choice of advanced STEM courses in the UK (Murphy & Weinhardt, 2016), and college completion in the US (Elsner & Isphording, 2015). A close contribution to my paper is a work of Gamoran (1989), who investigates how the ordinal position of a students' reading group – based on teachers' perceptions with a secondary check via standardized test – influenced the students' achievement over time in six Chicago schools. Gamoran shows how grouping by ability based on a teacher's decision is a crucial dimension rarely pointed out before now.

In the previous contributions, the critical point is to not distinguish between GPA and standardized tests. This work relied on the idea that they are different from each other and that teachers draw on hierarchy not strictly "ability" based. This discrepancy arises since they measure different things but also to other factors such a grading discrimination and classroom characteristics. In previous baseline settings, the source of their randomness is "school random sorting," fixing with classroom fixed effects when the randomness does not hold. Usually, their rank identification is conditional on the standardized test. In contrast here, the source of randomness is the teacher's grading policy (in detail, the mark distribution), and my Conditional Independence Assumption (CIA, hereafter) needs to control for marks and standardized tests. Although all previous works interact with a school subject to proxy a classroom environment, the unit of analysis is the school. In contrast, this work focuses on the real classroom, and I exploit that a teacher belongs to the classroom for the entire academic year. Finally, this works debates more in-depth the role of non-cognitive skills, except for one contribution that accounts for individual fixed effects singling out personal traits and genetic factors (Murphy & Weinhardt, 2016) and it distinguishes between Mark-based and Test-based Hierarchy. Finally, this paper enrich the debate with a broader array of socio-emotional skills, academic competences, and educational expectations and choices.

As outlined by Puerta, Valerio, and Bernal (2016), some define socio-emotional competencies as skills to accomplish tasks such as managing their emotions and interacting easily with mates, but there is not a strict consensus. Some define socio-emotional skills as non-cognitive or soft skills (Almlund, Duckworth, Heckman, & Kautz 2011); meanwhile, others use them as proxies of Big-5 personality traits (Brunello, Crema, & Rocco, 2018), pointing out the difference – not always stressed – between BIG-5 mean-level change within a trait and BIG-5 rank-order overall change (personality change) (McCrae, Costa Jr, Terracciano, Parker, Mills, Fruyt, Mervielde, 2002). Finally, the malleability of socio-emotional skills is a controversial topic since psychology and economics literature have opponent perspectives. The former suggests that some skills are not context-dependent (Lucas and Donnellan (2011), whereas the latter stand for situational specific hypothesis in which also personality traits may be modified (Almlund, Duckworth, Heckman, & Kautz 2011). The keystone of this debate might be the length of treatment exposure and timing of analysis. The longer the treatment, the stronger the impact on personality traits and socio-emotional skills. Also, the timing of analysis is crucial because childhood and adolescence are key stages for individual development, and existing studies find an average change in BIG-5 traits, mainly neuroticism and openness (Borghuis 2017; Branje et al. 2007) in the 12–16 age span.

As outlined by Eccles (1993), socio-emotional skills represent the first block of a "causal ladder" to explain human behaviour and choices. A broad array of students' outcomes makes possible to investigate students' socio-emotional skills, performance, educational expectations, and choices comprehensively. I expect that high ranked students show higher self-esteem, less self-stigma, more motivation. This results in better performance and higher educational expectations with less chance of dropout and higher chance of enrolling in more prestigious educational pathways.

> H1– The higher the rank in 8^{th} grade, the higher the positive changes of socio-emotional skills in 10^{th} grade.
>
> H2 - The higher the rank in 8^{th} grade, the higher the performance in 10^{th} grade, investment in further education, choice of academic track.
>
> H3 – The higher the rank in 8^{th} grade, the lower the dropout probability in 10^{th} grade.

2.2 Gender

An extensive body of literature outlines the differences between boys and girls in terms of attitudes and overall expectations toward school. On attitudes toward mathematics, existing empirical contributions show that boys have a higher maths self-conception and more positive attitudes towards the subject (Skaalvik & Skaalvik, 2004). On dropout and investment in tertiary education, a historical gender divide is present, and it is due to the intersection of family characteristics, non-cognitive skills, and behaviors (Almas, Cappelen, Salvanes, Sorensen & Tungodden, 2016; Goldin, Katz & Kuziemko, 2006). A large part of the gender divide is due to boys' disruptive behavior and their low levels of self-application regarding homework. Only a greater sense of confidence reduces boys' dropout risk, whereas more informed beliefs about the labor market mitigate girls' one (Almas and colleagues, 2016). Concerning academic performance, boys tend to perform better than girls on standardized tests in maths, but the opposite is true when the teachers' assigned marks are used to proxy ability (Enzi, 2015). Finally, boys and girls display some differences in personality traits as well. Finally, females tend to score significantly higher on measures of Agreeableness, Conscientiousness, and Openness (Vecchione Alessandri, Roccas,, & Caprara 2019).

Relying on these baseline differences, it is interesting to understand whether there is a heterogeneous effect of mark-based hierarchy by gender. The impact of the hierarchy can vary among boys and girls due to gender-oriented socialization processes. These can lead to differences between boys and girls in (1) attitudes toward competition, (2) network depth, and (3) familiarity dimensions. However, as we will see, insights from the literature covering these aspects lead to somewhat contrasting expectations.

In the first aspect, laboratory experiments show gender differences in commitment: girls on average are more motivated to finish a task and more persistent in their choices, even if their willingness to compete and to take risks is lower than that of boys (Croson & Gneezy, 2009). In the educational field, for instance, some contributions suggest that boys are more confident in their abilities (Cho, 2017) and relatively more obsessed with "social status" than girls (Niederle & Vesterlund, 2007; Buser et al. 2014). They report that boys emerge as more competitive than girls when choosing between more or less prestigious tracks in the Netherlands. Following these arguments and research findings, one can expect that boys are more sensitive to rank when there is a prestigious track in play.

> H4a – The effect of student rank in the classroom on the probability of enrolling in the more prestigious academic track (versus lower prestigious technical/vocational schools) is larger among boys

> H4b - The effect of student rank in the classroom on socio-emotional skills is larger among boys

Nevertheless, network studies seem to draw an alternative picture. While girls give more importance to local social hierarchies and friendship ties, boys are more sensitive to larger social hierarchies (Crosnoe, 2008). According to the review of Crosnoe (2008), girls tend to create less than boys extensive networks of friends. Hence, it appears they are more concerned about their social status in a more delimited reference population than boys are. Drawing on these insights, I can formulate a contrasting hypothesis:

H4c – The effect of student rank in the classroom on the probability of enrolling in the more prestigious academic track (versus lower prestigious technical/vocational schools) is larger among girls

H4d - The effect of student rank in the classroom on socio-emotional skills is larger among girls

In a recent contribution, Joensen and colleagues (2015) add the dimension of familiarity to the broad literature of gender differences. Exploiting Swedish course-taking patterns, they report that peer pressure is stronger among girls when females interact with their peers for a longer time in the classroom. Unfortunately, this kind of analysis relies on formal and informal rules shaping classroom composition in the educational systems. Assuming a continuum where the left pole is the total flexibility of classroom organization and the right pole is a fully rigid system, Italy relies on the latter. Students attend every subject with the same peers at all educational stages, and variation in classroom composition is due to student transfer and dropouts alone.

2.3 Italian educational features

Italy shows a mixed tracking model (Triventi, Kulic, Skopek & Blossfeld, 2016) with a degree of high centralization and formal external tracking in upper secondary schools. The level of within or between school transfer is lower than in other countries such as Germany, and the classroom composition is, to a large extent, stable during the academic year. On the teacher's side, at the beginning of the academic year, principals optimize the teacher to classroom assignment[6], reducing the within year teacher turnover to personal reasons such as a family leaving or sick leaving, not to systemic one.

Until 8th grade, where a decentralized national examination takes place[7], a standard curriculum is in place. Then, students choose between lyceum, technical, and vocational tracks. In high school, such student sorting is, to a large extent, based on parent and student preferences, and each track varies in terms of socioeconomic, ethnic, and ability composition (Contini & Triventi 2016). Lyceum has a higher share of high SES-students and high ability students compared to technical and vocational tracks. Until some years ago, teachers made

[6] "Circolare ministeriale" Number 4, comma 10.2, of January 15, 2009

oral and informal suggestions; in recent years, teachers provide formal joint track advice, though still non-binding. However, as outlined by Argentin, Barbieri, and Barone (2017), teacher's advice is not a stronger predictor of track choice compared to other individual characteristics such as socioeconomic and ethnic origin.

In Italy transparency on marks is notable, as (1) evaluation of exams and homework is public for every subject; (2) a mid-term evaluation report for every subject is made public usually in February every academic year; (3) for each subject a final evaluation report is made public in late June. Despite several reforms changing the features of national examinations, the 8th-grade national examination has preserved the same structure. In such an examination, students are tested by written (in mathematics, Italian, and foreign language subjects) and oral examinations (all subjects starting from the discussion of a short conceptual essay) with a committee made up of internal professors and an external president. As usual, the final mark is public, with a score from 0 to 10 with a 6/10 pass threshold. In constrast INVALSI standardized tests are not public. his feature allows us to inquire to what extent public mark-based hierarchies matter more than test-based non-public hierarchies based only on an intimate awareness of their own competence. Theoretically, the interaction between assessment (marks and tests) and information (public and private) draws an ideal matrix summed up in Table I. The present work focuses only on the available yellow cases (1 & 3). Indeed, case 4 is not present in the Italian education system. Besides, even if in the last years, INVALSI agency makes the test public on a coarse scale (case 2), but there are some issues to exploit this reform.

Table I: Scheme of Mark and Test based hierarchies

		ASSESSMENT TOOL	
		Teacher's assigned mark	*Standardized Test*
INFORMATION	*Public*	CASE 1	CASE 2
	Not Public	CASE 4	CASE 3

On the one hand, a mark-based hierarchy might have a stronger effect than a test-based hierarchy because (1) teachers play an advisory role during the school years; (2) teachers' hierarchies affect student mindsets on an everyday basis; and (3) not having an objective, reliable measure, parents might rely on a teacher's judgment above all in critical turning points such as track choice.

> H5a – Mark-based hierarchies have a stronger effect than test-based hierarchies on educational outcomes

On the other hand, parents and students do not rely a lot on the advisory role of teachers, as debated in the mentioned work of Argentin and colleagues (2017).

> H5b – Test-based hierarchies have a stronger effect than mark-based hierarchies on educational outcomes

2.4 Non-linear effects

So far, this work embraces an implicit assumption that students react in the same ways along this mark-based hierarchy. Thanks to the literature on income inequality, redistribution, and social status (Kuziemko et al., 2014) and the systematic generalization of the median player literature in the work of Hotelling in his Stability in Competition (Hotelling, 1929), it is well-known that this assumption does not hold since the reaction to treatment might depend on the relative position in the distribution, drawing an inverted U-shaped. Findings show that the attitudes to redistributive policies depend on the "distance" between personal income and country median income (Kuziemko et al., 2014). Under the median, individuals show positive but diminishing redistributive attitudes as quickly as we are close to the country median income. Above the median, negative but increasing attitudes take place as we move away. The human resources stream transposes such modeling hints to the analysis of hierarchies among workers in firms. Empirical evidence of non-linear hierarchy effects finds support on the work of Gill, Zdenka, Jaesun, and Prowse (2018). They investigate the workplace hierarchy on effort and performance, also looking to the heterogeneous effect of feedback policies, workplace structures, and incentive schemes. It comes out that workers are status seekers, but they exhibit 'first-place loving' and 'last-place loathing' behavior with a neutral stance on intermediate places. Similarly, I argue that students behave in the same way with high ranked students more sensitive due to top competition in the upper distribution and low ranked students more sensitive due to bottom competition in the lower distribution. This asymmetric effect is owing to higher marginal revenue to change position in both tails of the distribution.

> H6 – Students on either tail of distributions are more sensitive to rank.

3. Identification strategy

3.1 Main rationale

This research aims to estimate the effect of student hierarchy generated by the teachers' assigned marks on students' outcomes. As debated above, daily exposure to classroom interaction spreads information on the relative position of the students, once account for peers and teachers' roles. The paper relies on three assumptions:

A. *Students internalize the teacher's hierarchy.*

   It is a realistic assumption because students interact every day in the classroom. They compare themselves internalizing what teachers say in the classroom (Festinger, 1954).

B. *Students are aware that the teacher's hierarchy is not a true litmus of their ability.*

   How could they know? Students might understand their value (Elsner & Isphording, 2015) and that classroom grading policy depends on teacher's proclivities. Suppose a teacher evaluates all students with 6 and very few with 9. Several examples of grading schemes come out. Iacus and Porro (2011) find at least 15 grading schemes. I think that this internalization is easier in mathematics than the Italian language because math is a more objective language. Hence, students can ponder better discrepancies between teacher's hierarchy and how they effectively perform. In addition, it is a widely accepted that medium and high perfomers students show higher self-evaluation score and we control for previous ability (Kruger & Dunning, 1999; Hacker, Bol, Horgan, & Rakow, 2000).

C. *Mark is a synthetic measure of cognitive and non-cognitive skills.*

   It is a widely accepted fact that teachers also evaluate non-cognitive skills (Brookhart, 1993)

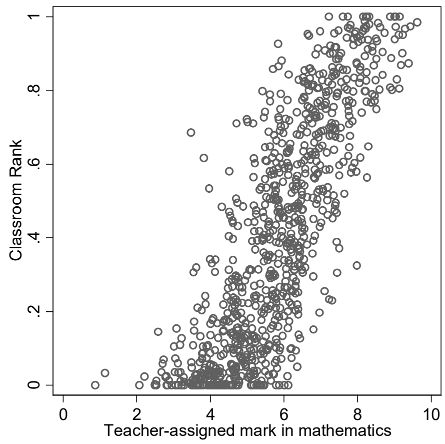
As debated in the background literature, a teacher's grading policy relies on individual and collective dimensions. In the following equation, you find a reduced form of such function:

$$1)\ R_{ics} = Mark_{ics} + Size_{cs}$$

$$2)\ Mark_{ics} = \alpha_1 E_{ics} + \alpha_2 Mo_{ics} + \alpha_3 Pa_{ics} + \alpha_4 A_{ics} + \alpha_5 C_{cs}$$

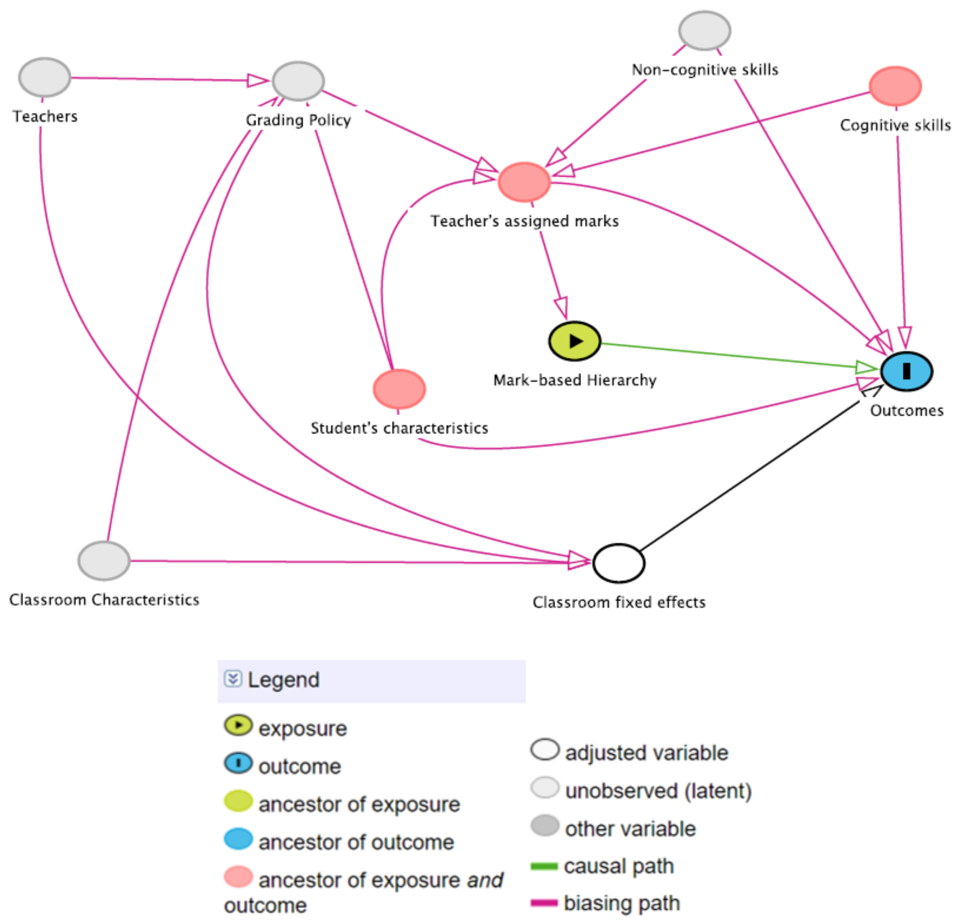
Equation (1) outlines the teacher's grading policy where hierarchy position (*R*) depends on the underlying ranking dimension (*Mark*, in this case) and contextual size of the environment (*Size*). In turn, grading functions reflect the individual effort (*E*), motivation (*Mo*), participation (*Pa*), ability (*A*), and contextual characteristics such as teaching styles or grading schemes (E). These characteristics may vary at the individual (*i*), classroom (*c*), and school level (*s*).

Figure I: Scatter dot (based on a random sampling) between classroom rank and teacher's mark in mathematics (mean of the oral and written mark)



Empirically, we leverage the fact that in each classroom, a teacher sets up their grading policy. This policy depends on teacher's proclivities, student's socio-demographic characteristics, and classroom characteristics. Once accounted for these factors, I exploit this classroom grading policy, and I build upon them a hierarchy that depends on the underlying marks and class size. This hierarchy measures a relative position because despite students show an identical teacher's assigned mark, they have distinct relative positions across classrooms. In figure I, two empirical facts come out: (1) a positive association between mark and classroom rank, and (2) a variation within mark across classroom rank. This maximum variation is mainly present in the center of the distribution (4^{th} and 8^{th} deciles) and less in the tails. Intuitively, I want to exploit this "quasi random" grading variation, once blocked plausible biasing paths.

Figure II: Directed acyclic graph



In this kind of work, selection bias due to panel attrition, reflection problem, sorting effects, and correlated effects are severe and notorious issues for a causal identification of the treatment. I rely on a DAG (Figure II) to make more transparent the set of minimal conditions for causal identification. Briefly, the DAG shows that teacher's mark depend on several factor but it is necessary and sufficient condition to control for teacher's assigned marks and classroom fixed effects. In the next paragraphs, I discuss in detail the main issues.

A typical issue of longitudinal data is to lose cases over time, resulting in a selection bias due to the fact that those who dropout are different compared to those who are always observed. In administrative data, attrition is usually due to the dropout and grade repetition of students over time. Theoretically, such kind of bias crowds out low performing students, so the sample would be positively selected and will underestimate the effect of mark-based hierarchy. Despite, Italian law alleviates the selection bias resulting from drop out because school is mandatory until the age of 16, there is a dramatic grade repetition in the technical track and especially in the vocational one. I approach these issues relying on inverse probability weighting (IPW, hereafter) (Seaman & White, 2013). IPW is ideal in such cases where the design relies on two time-points, and it is essentially a prediction of missing cases as a function of a number of relevant antecedent characteristics before the treatment, then the inverse of this probability is used to weight the observations.[8]

The reflection problem is always present when the behavior of agents introduces perfect collinearity between the expected mean outcome of the group and its mean characteristics. Hence, it is troublesome to identify a "true" peer effect (Patacchini, Rainone, & Zenou 2017). However, reflection problem is less present. Indeed, outcomes and treatment of interests are not the same, framing the reflection issue, if at all, as an omitted variable issue. Second, the hierarchy is measured in the 8th grade while the outcomes are collected two years later in a new environment where the likelihood of being enrolled for a student on the same track, in the same school, in the same classroom, is extremely low. This time lag breaks any residual reflection issue.

Another plausible issue is that non-fully observable characteristics drive the sorting of students and teachers within the classroom leading to biasing paths. Although the Italian education system stands for an equal sorting of students and teachers, parents may strategically choose schools, or principals might sort students within schools, creating "ghetto classrooms." Indeed, several findings support the presence of segregating patterns between and within schools based on socioeconomic (Agasisti & Falzetti, 2017) and ethnic characteristics (Contini, 2013). In addition, the teacher's allocation is random only at the beginning of a career, turning into a "flight to school quality" over time (Barbieri, Rossetti, & Sestito 2010) and resulting in an unequal sorting is mainly present in the lower and upper secondary school (Abbiati, Argenti, & Gerosa, 2017). The non-random teachers might open a biasing path a cautious approach calls for classroom fixed effects with the assumption that

[8] An alternative strategy is to recover the missing students from datasets of the following cohorts, in which students with delays should be ideally present. However, the follow-up investigation is possible but it shows more limitations. Indeed, given the institutional setup in Italy and the data available, I can recover only students who were retained until age 16 because after it school is not anymore mandatory. After this age, students may be out of the system and these follow-up solutions need an IPW approach too. Another issue related arises regarding the assignment of the repeating students and the proper academic year. Indeed, looking at the data, students may reappear – in the extreme case – in the INVALSI test and survey for four years because they repeat the 10th grade four times. Finally, this follow-up procedure is possible only for some outcomes such as performance, dropout, investment in further education, and track choice.

teacher characteristics, and classroom characteristics are time-invariant variables. A realistic assumption in the Italian education system where the chance of changing a professor[9] is almost negligible during an academic year and the classroom composition does not change by law.

In any estimation of peer effects, results might be flawed due to the presence of peer-group unobservable characteristics. Among many examples, students exposed to common factors such as good teachers, external school activities, and school clubs. Despite the identification strategy can exploit a quasi-random variation of the rank across classrooms, a basic set of socio-demographic characteristics is adopted.

My variable of interest is causally identified because the rank is individually assigned within sections. Relying on the classroom fixed effect, I provide a classroom within-transformation. It is crucial to point out that this transformation does not change the shape of a distribution. Hence, the impact of rank is identified from differences across classrooms in higher moments of the ability distribution (variance, skewness, kurtosis, etc.). Thus, I compare students with the same mark across classrooms but different ranks, resulting in a *within marks across classroom strategy.*

3.2 Minority reports

The paper aims to single out a causal effect of Mark-based Hierarchy at the net of underlying marks and classroom characteristics. Teachers' marks measure not only – in some way – "ability" but also dimensions such as effort, participation, resilience, misconduct behavior. In addition, the teacher's marks convey a teacher's proclivities and preferences. The work does not disentangle what - among these dimensions – drive more or not this hierarchy, and *agnostically* adjust for mark fixed effect. In this way, any biasing (confounding) path due to the teacher's mark is blocked. However, teacher’s marks depends on classroom peers too, and classroom fixed effects block this path, making it possible to adjust for selection bias into the classrooms. However, one can argue that the strategy does not control for non-cognitive skills at all. Hence, running a robustness check and relaxing a bit of our baseline strategy, I include

9 Indeed, at the beginning of each academic year, the principal manages teachers’ schedules to optimize classroom coverage. See decret N. 37381 of Ministry of Education, Universities and Research

mark as simple control, and I draw on a within students across subjects strategy (WSAS hereafter, Lavy, 2015). This strategy frame a pseudo-pane where students are counted twice, exploiting the different subjects. This approach draws on within students across subjects variation (individual fixed effects), singling out personal traits, and non-cognitive skills. In addition, I relax this approach, including attitudes toward subjects as control, assuming that some non-cognitive skills might be subject-oriented. Both robustness checks show a complete adherence to the baseline approach, reinforcing the idea that rank is assigned as good as random, or at least, personal traits, a proxy of non-cognitive skills, and genetic traits do not bias the estimation. Nevertheless, as outlined in the DAG, non-cognitive skills affect the teacher's assigned mark and the outcomes but not the rank. Once blocked the biasing path using mark fixed effects or marks in the WSAS strategy, the introduction of non-cognitive skills is redundant to identify an effect causally.

4. Research designs

To address my questions, I rely on two main research designs exploiting the possibility to link specific editions of the INVALSI data and following the same students, along with their educational career. In the I design, I follow the pupils enrolled in the 8th grade in 2012/13 through to the 10th grade (2014/15). To deal with selection bias due to dropout between the 8th and 10th grades, I use IPW (Seaman and White, 2013). This design makes it possible to investigate an extensive array of students' outcomes from socio-emotional skills to educational choices.

Table II: Design I – One cohort

	8th GRADE (2012/13)	**10th GRADE (2014/15)**
Concepts	Mark-based hierarchy Ability control	Educational outcomes
Variables	Mid-term teachers' marks Standardized test scores	Socio-emotional skills, expected university enrolment, expected and effective dropout, track choice, and performance.
Source	Administrative data	Administrative data and student questionnaires

The II design follows a single cohort over three grades of the Italian educational system: 5th grade of primary school (2012-13), 8th grade of lower secondary education (2015-16), and 10th grade of upper secondary (2017-18). To account for selection bias due to dropout and retention, I opt for IPW as well. The II design works as a backup design to address possible concerns on identification strategy, such as the ability to measure or to test possible differences between the mark and test-based hierarchies. Due to changes in the questionnaire survey, this strategy relies only on three outcomes, such as academic performance, track choice, and effective dropout.

Table III: Design II – One cohort

	5th GRADE **2012/13**	**8th GRADE** **2015/16**	**10th GRADE** **2017/18**
Concepts	Ability control	Mark and test-based hierarchy	Educational outcomes
Variables	Mid-term teachers' marks Standardised test scores	Mid-term teachers' marks	Actual dropout, track choice, and performance.
Source	Administrative data and student questionnaires	Administrative data	Administrative data and student questionnaires

5. Analytical strategy

5.1 Data

The paper relies on INVALSI data, an official census of Italian students, a mix of administrative information, and surveys longitudinally collected each year from 2012/13. I use two cohorts of students, namely 143,420 students for the design I and 264,172 students for design II, around 22,708 classrooms, and approximately 6,000 schools, following students with information from administrative archives and survey questionnaires from the 5th grade, through the 8th grade, to the 10th grade enrolled in the same schools.

5.2 Variables

I address my research questions using different designs aimed at controlling for various possible sources of bias. The outcomes of interest include 1) socio-emotional skills; 2) academic performance; 3) educational choices and expectations.

Table IV: Scheme of dependent variables with alpha

SOCIO-EMOTIONAL SKILLS		
Indexes	**Original variables**	**Alpha**
Agreeableness	1. How many mates interact with you? 2. With how many of your mates, do you feel good? 3. How many of your mates, do you help if in trouble? 4. How many of your mates, do you consider friends?	0.74
Extrinsic Motivation	1. For me, it is important to show others that I am a good student 2. For me, it is important to show others that I go well in the assessments 3. For me, it is important to show others that I go well at schools 4. For me, it is important to show others that I look more intelligent than my mates	0.84
Intrinsic Motivation	1. My goal is to learn as much I can 2. I want to learn new things for me 3. I want to understand what I study 4. I want to increase my skills	0.77
Confidence in Mathematics competence	1. I like mathematics 2. I learn interesting topics in math lessons 3. Mathematics is interesting	0.85
Neuroticism	1. I am nervous for examinations 2. I was anxious for the assessment 3. I have the feeling that I was not so good during the assessment 4. I was calm in the assessment	0.81
Self-stigma	1. Sincerely, I do know what I am doing at school. it is not for me 2. I am unfit for these school things 3. I do not understand my role at school, and I do not care 4. I do not know what I am doing at school	0.83
Dropout Intention	1. I am not sure to go to school next year 2. I am thinking about leaving school 3. I have the intention to leave school	0.80
ACADEMIC PERFORMANCE		
Ability	1. Standardized test scores in mathematics 10^{th} grade	
EDUCATIONAL CHOICES AND EXPECTATIONS (DUMMY)		
Track choice (Academic Vs. Others)	1. Administrative information on track attended in 10^{th} grade: lyceum vs technical/vocational	
Actual dropout until 10° grade	1. ID student 8^{th} grade does not find a match in the 10^{th} grade. Hence, the sample size is larger, see note and appendix	
Expected university enrolment	4. I do not want to finish high school 5. Three years professional diploma 6. High school 7. Bachelor 8. Master and PhD	

Socio-emotional skills are measured, creating seven quantitative indexes; each one summarizes the information provided by a set of questions using Likert scales for students' answers. These variables have been standardized with an average of zero, and a standard deviation equals to one. The internal reliability of these indexes has been checked, computing the Cronbach's alpha, with a double-check using factor analysis. Brunello and colleagues (2018) argue that INVALSI items are useful to measure Five-Factor Models (Big-5). I follow them, but I am more skeptical about speaking of personality changes, as I have debated in the literature review, and because INVALSI is school-oriented.

Academic performance is measured using students' scores in the standardized mathematics test administered, usually in May. The scores were provided already by INVALSI as the outcome of IRT analysis on the students' patterns of answers to the test.

Educational choices and expectations are measured using three dummy variables indicating 1) whether the student is enrolled in an academic track (versus technical/vocational schools) in upper secondary education; 2) whether the student dropped out (it is no longer enrolled in the educational system in the 10th grade);[10] 3) whether the student declares to aim to attend university after high school.

I rely on basic control variables such as gender, ethnic origin, socioeconomic background (education and occupation of parents), ability, and repetition of grade. I use gender to identify boys and girls, whereas I exploit administrative information for ethnic status. To control for the enrollment, I use the year of birth, taking into account what the age of students should be given Italian school rules. I create a dummy with three codes: early, regular, or late students. I control for education and jobs of parents by adopting a dominance criterion. The former is built using the three classic response variables: primary, secondary, and tertiary education. The latter is measured by a recodification of a scale following Campodifiori, Figura, Papini, and Ricci (2010). Finally, I rely on multiple measures of academic competences such as the INVALSI standardized test in mathematics at 8° and 5° grade

[10]The variable actual dropout shows three limitations. It does not account for students who repeat a grade or go abroad for the high schools, and INVALSI collection data procedure. However, in the Design II, the mismatch is less present.

5.3. Methods

My identification strategy relies on a *two-way fixed effects approach* along with two research designs. As debated above, I include teachers' marks dummies to account for commonalities among students who receive the same teacher's mark and classroom fixed effect to hold constant unobserved heterogeneity at the classroom level. The following equations sum up our two main empirical strategies. One, based on design I, follows around 143,420 students, whereas the other is based on design II and follows 264,172 students. The only difference is when the proxy of ability is collected. In the first equation, it is collected in the same academic year with the other individual characteristics. In the second equation is collected 3 years before in primary school.

$$1)\ \text{(Design I)}\ Y_{ict_{+2}} = \alpha_{ict} + \mu_{ct} + \gamma_{mt} + \beta Rank_{ict} + \beta Ability_{ict} + \beta Z_{ict} + \varepsilon_{ic}$$

$$2)\ \text{(Design II)}\ Y_{ict_{+2}} = \alpha_{ict} + \mu_{ct} + \gamma_{mt} + \beta Rank_{ict} + \beta Ability_{ict_{-3}} + \beta Z_{ict} + \varepsilon_{ic}$$

Where $Rank_{ic}$ represents the variable of interests (rank of students of i in classroom c at time *t*), $Ability_{ict}$ represents the objective performance of students, Z_{ict} represents a vector of control variables, μ_{ct} indicates the classroom fixed effects, γ_{mt}indicates the marks fixed effects, and ε_{ic} is the error term. Following Elsner and Isphording (2015), the rank is built on the average of oral and written marks in mathematics, which is transformed into a percentile rank, using the following formula:

$$(2)\ \ R_{ic} = \frac{(r_{ic}-1)}{(N_c-1)}, R_{ic} \in \{0,1\}$$

Where r_{ic} is an ordinal rank[11] of student in the classroom and R_{ic} is an adjusted ordinal rank controlled for the size of the classroom ($\mathrm{N_c}$). Using this formula, I obtain a normalized local rank which is not biased by the different size of the classroom. The highest take value 1, and the lowest takes value 0.

[11] There are several specifications to rank, and the main differences among them concern how to treat cases with the same value in a group. Take the extreme approaches. One, quite arbitrary, assigns the rank randomly, and we do not adopt it because it leads to a measurement error in the independent variable. Another one simply removes the position on the "podium place." Let's do an example. If there are three top students in a classroom, we give each of them an ordinal rank of 2, and the next in line has a rank of 4. Previous literature follows this approach (Elsner & Isphording, 2015).

The estimation strategy is based on OLS linear regression with clustered errors at the classroom level. I use reghdfe (Correia, 2017) in STATA to run the two-way fixed effect approach. For the analysis of the II design, I use only a linear specification for estimating the relationship between rank and the outcome variables, whereas for the I design, I rely on four specifications:

- ✓ Linear Specification
- ✓ Rank-gender interaction
- ✓ Squared non-linear specifications with a secondary check for cubic and quadratic relations
- ✓ Squared non-linear specifications

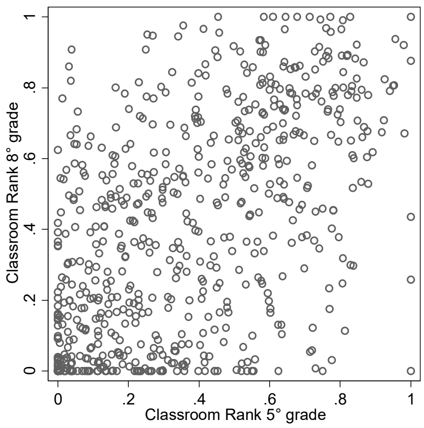
It is worth noting that when I run a comparison between a hierarchy based on marks and one based on a standardized test, the approach is similar. Indeed, I adopt classroom fixed effect and test fixed effects. Note that test fixed effects amounts at 86 categories.

6. Empirical findings

6.1 Descriptive statistics

In figure III, I run a sort of stability test, namely to what extent rank changes between the 5thgrade and 8th grades of the educational system. It emerges that the variation is present in the overall distribution even if bottom-ranked students in 5th grade are more likely to persist in a lower position in the following academic years.

Figure III: Scatter dot (based on a random sampling) between classroom rank in the 8th grade and 5th grade (mean of the oral and written mark)



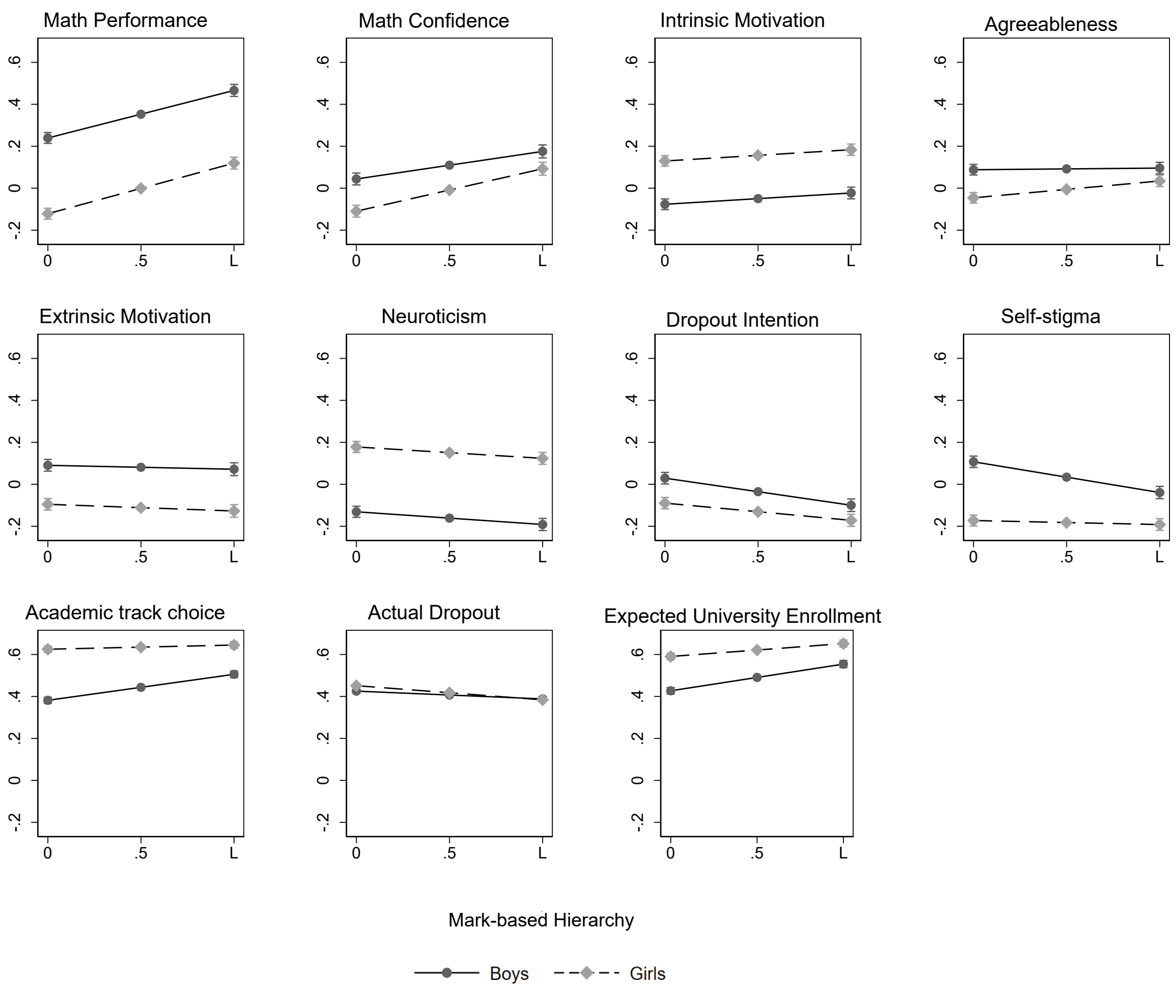
6.2. Main findings

Overall, my main analysis suggests that teachers' grading is an inequality-enhancing factor because it affects socio-emotional skills, educational expectations, and academic choices. Regarding H1-H3 (Figure IV), I find that student rank (being higher in the classroom mark hierarchy) positively affects academic performance and confidence mathematics in 10th grade, and it reduces the intention to leave school and self-stigma. In contrast, students' positions in the mark-based hierarchy affect only to a limited extent their socio-emotional skills – such as intrinsic motivation and neuroticism –, while it does not influence extrinsic motivation and agreeableness at all. Regarding expectations and education choices, it comes out that one standard deviation of the rank position increases expected university enrollment by 9 percentage points (p.p., hereafter), and an academic track choice by 7 p.p., whereas it reduces effective dropout by 5 p.p.

Figure IV: Average Marginal Effect, Classroom And Mark Fixed Effect Model - Main Effect Of Rank On Socioemotional Skills, Performance, And Educational Choices (95% Conf. Int. Reported). Ordered by coefficient size.

Looking at the heterogeneous effect over gender (H4 a, b, c & d), we find partial empirical evidence of a gender divide (figure V). A significant divide is present for math confidence, agreeableness, self-stigma, academic track choice, and expected university enrollment. To my perspective, H4b finds support because boys appear to be more reactive to their position in the mark-based rank than girls when important issues are in play, such as academic track choice, expected university enrollment, and or dropout. In contrast, girls are more reactive to rank on an array of socio-emotional skills such as math confidence, and agreeableness. The only exception to this pattern is represented by self-stigma, where boys are more affected by their rank position than girls. Since from the literature we know that boys are more prone to leave school, my findings suggest that obtaining a higher position in the classroom rank is able to boost their educational expectations.

Figure V: Rank by Gender (Predicted Probabilities) On Socioemotional Skills, Performance, And Educational Choices; Classroom And Mark Fixed Effect Model (95% Conf. Int. Reported)



Looking at the comparison between the mark and standardized test-oriented hierarchy (H5-a & H5-b), I find that the mark-based hierarchy matters more than test-based hierarchy only for performance (figure VI & VII). No significant differences are present for academic track choice and effective dropout. This pattern does not change if I control for ability at 8^{th} or 5^{th} grade. When I check at possible gender heterogeneity, no great differences are present. It comes out that boys look more reactive than girls to both hierarchies, mainly for the mark one. For the main effect on math performance, a possible explanation is that mark-based hierarchy plays an important signaling role, increasing the investment of the parents on children's cognitive and non-cognitive skills. Besides, a sort of self-fulfilling prophecy can work as well. High mark ranked students' trust more in their competencies, and this is connected to the effects I find on self-stigma.

Figure VI: Average Marginal Effect, Classroom And Mark Fixed Effect Model - Main Effect Of Test and Mark based Hierarchy On Socioemotional Skills, Performance, And Educational Choices (95% Conf. Int. Reported)

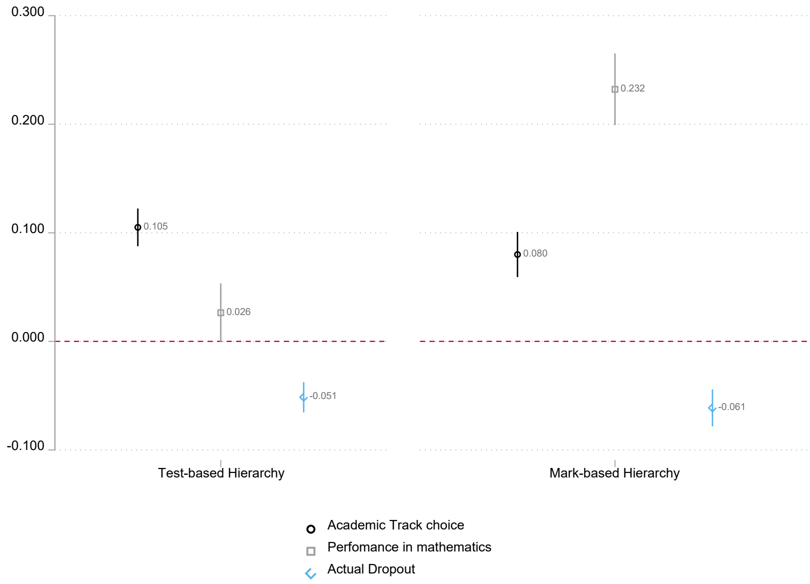
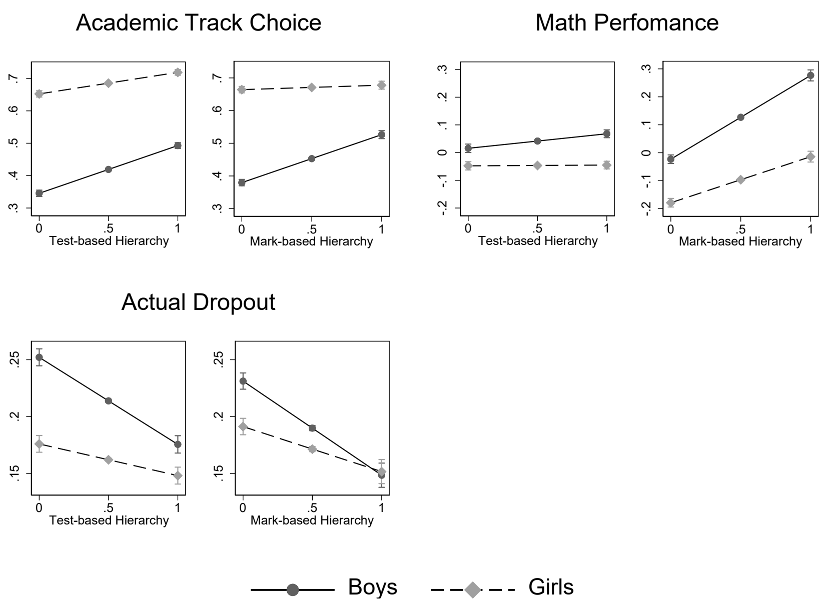
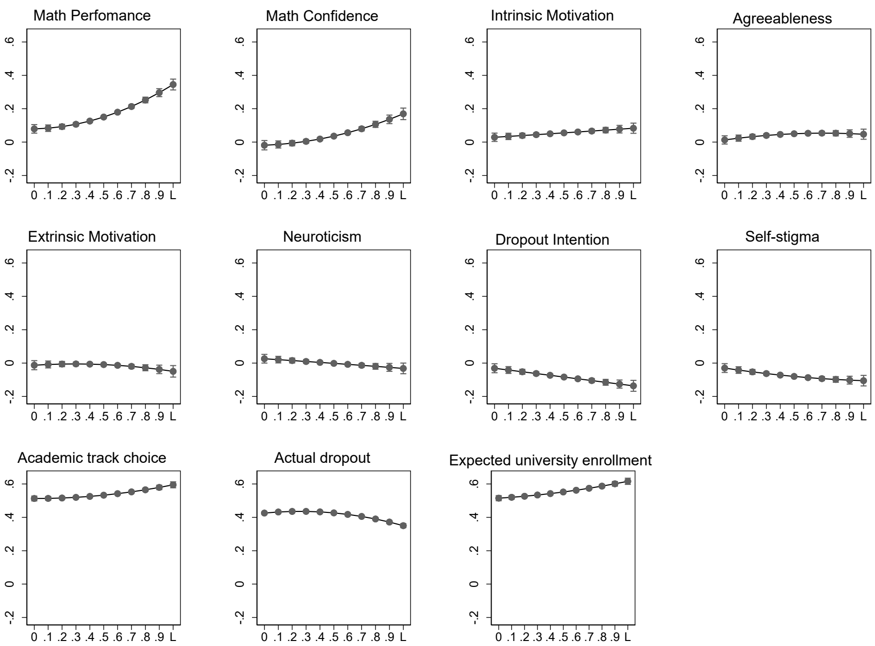


Figure VII: Mark and Test based hierarchies by Gender (Predicted Probabilities) On Socioemotional Skills, Performance, And Educational Choices; Classroom And Mark-Test Fixed Effect Model (95% Conf. Int. Reported)



Looking at the non-linear effects (Figure VIII), I report only squared specifications because cubic and quartic models reduce the goodness of fit, measured in terms of Akaike and Bayesian Information Criteria. Regarding H6, the non-linear effect is present only on the upper tail of the distribution, mainly on math performance and confidence, self-stigma, academic track choice, effective dropout, and expected university enrollment. It is quite visible a common turning point around the center of the rank distribution, where the slope starts to be higher.

Figure VIII: Non-linear Effect (Predicted Probabilities) On Socioemotional Skills, Performance, And Educational Choices; Classroom and Mark Fixed Effect Model (95% Conf. Int. Reported)



7. Further analyses

This work tackles the main perils of peer effect literature and takes into account several designs to deal with biases such as selection, reverse causality, and confounding variables. Nevertheless, some further limitations might arise. I see five main concerns: (1) possible reverse causality between marks-based hierarchy and standardized tests; (2) multiple hierarchies; (3) individual confounding characteristics; and (4) use of mathematics marks.

As outlined above, a critical point of the teacher's grading function is controlling for individual ability. In designs I, a standardized test is collected almost simultaneously with the teacher's assigned marks (March-May). As baseline modeling, I opt for such contemporaneous proxy of ability because I aim to capture the main effect of the teacher's hierarchy holding constant current ability and individual characteristics. Nevertheless, estimations might be flawed by partial reverse causality insofar as hierarchy affects further performance in the standardized tests. To solve this issue, I rely on design II to exploit three points over time: the 5th, 8th, and 10th grades of the Italian education system. Thus, I document that hierarchy matters even if controlling for 5th-grade ability on academic performance in the 10th grade, dropout decision, and track choice.

This work theoretically accepts the presence of multiple hierarchies in a classroom. Indeed, the interplay between the theory of categorical inequality and social construct might result in multiple hierarchies based on other individual characteristics. In this work, I focus on the teacher's hierarchy, but I run a comparison with another hierarchy based on standardized tests.

In extensive works on classroom peer effect, a critical pitfall is the lack of a proper strategy to address individual characteristics such as genetic factors, beliefs, and attitudes. It is not easy to adopt individual fixed effects within peer effect studies. A possible solution is to exploit a within-student across-subjects strategy following Lavy (2015). The idea is to build a pseudo panel wherein each student is observed twice, namely in mathematics and Italian. Hence, subject-oriented characteristics vary across individuals but not the fixed ones. In this way, it is possible to control for individual fixed effects. Unfortunately, this strategy in my setting does not allow a proper inclusion of the teachers' marks fixed effect. Indeed, it would imply collapsing the two subject-related marks fixed effects together. At the moment, while I would be able to rely on differences in the rank between two subjects, I would not be able to control for differences in unobserved teachers' characteristics. Indeed, the estimates of the rank effects obtained from this alternative identification strategy are, in general, greater than my baseline results.

I am aware that the exclusive use of mathematics marks might be controversial. Hence, I am re-running some preliminary analyses – with the design I – with Italian language marks. I do not find divergent patterns on the effect of the rank. There is a common direction on all outcomes of interest, with the exception of confidence on related subjects. It could be that hierarchies based on marks in Italian and mathematics affect differently the propensity to enroll in specific curricula within the academic track (e.g., scientific or classic lyceums), but unfortunately, INVASI does not provide information on the detailed type of school attended in upper secondary education; therefore, it is not possible to test this expectation.

8. Discussion and conclusion

Due to daily classroom interaction, teachers play an essential role in a student's development. This work focuses its attention on daily grading and the extent to which such a practice sorts students into categories and shapes a hierarchy among them. As theoretically expected, mark-based hierarchy triggers a reaction in the student mindsets with consequences on the socio-emotional skills in a critical span of adolescence (Cunha & Heckman, 2009). Indeed, top-ranked students show higher confidence in math and less attitude to stigmatize themselves compared to bottom-ranked students. Besides, they slightly reveal less neuroticism and motivation more intrinsic-oriented.

The first block of findings confirms the malleability of some socio-emotional skills to the relative position in the classrooms. This confirms the expectation of the Eccles framework (1993), but from a policy perspective, the critical question is whether this malleability of socio-emotional turns into a "self-fulfilling prophecy" harming academic performances and actual conditioning decisions. Unfortunately, findings confirm this scenario. Compared to the bottom-ranked students, top-ranked students perform better and have lower intentions do leave school. Also, they have higher chances to attend an academic track, to enroll in a tertiary degree, and not to drop out. In view of the results, mark-based hierarchy is an increasing inequality factor because it shapes a hierarchy among students with essential consequences on students' life and educational paths.

This picture gest even more complicated when looking at gender differences with mixing findings. In fact, boys and girls have different reactions to hierarchies. Boys are more sensitive to their position in the mark-based hierarchy when choosing track in upper secondary education or when shaping their own expectations in further human capital. In contrast, girls appear to be more sensitive to rank position, especially in terms of socio-psychological attitudes, such as neuroticism and agreeableness. What emerges is an interesting but moderate picture of a female penalty. Boys are more responsive to the hierarchy for crucial choices such as academic track and expected enrollment.

When looking at the comparison between mark-based hierarchy and test one, it comes out that Mark-based hierarchy impacts more than the test one on academic performance, whereas it is the opposite for actual dropout. Further, boys look more reactive than a girl to test-based hierarchy. Finally, findings partially confirm non-linear patterns on the outcomes of interest. Indeed, top-ranked students react more than bottom ones for academic performance, academic track, actual dropout, and university enrolment.

To the best of my knowledge, this is the first work aiming to measure teachers' hierarchies due to grading among students. Previous contributions focus more generally on ability hierarchies, without singling out the difference between marks and standardized tests. In contrast with other contributions, this work uses the classroom as a unit of analysis, exploiting Italian institutional features. Moreover, this study follows one cohort of students in three critical stages of the education pipeline: primary, lower, and upper. It is important because it examines the exposure of students to teacher hierarchies over time. Finally, looking at past contributions, this work sheds more light on the role played by personal traits, self-stigma, and classroom pressure.

The effect size is not negligible in absolute terms but also when compared to the previous close contributions. As investigated in the literature section, previous contributions exploit more standardized tests or do not take into account a difference between GPA and standardized tests. However, lines of comparison can be drawn. On the gender divide, this work is close to the contribution of Murphy and Weinhardt (2016), who identify an impact in favor of boys. Nevertheless, the size of that impact is greater by far compared to that found in my work. I suggest two explanations. First, the stronger effect might be due to the stronger element of competition in the English education system from primary education onwards. Second, my work adopts a more conservative strategy by introducing a mark fixed effect. Indeed, I find more similarities with the contribution of Elsner and Isphording (2015), who do not find a gender divide. In addition, it is stronger than works on the effect of popularity (in-degree only) on academic performance (Mihaly, 2009) and it is not trivial compared to the all literature on peer effect (see the review of Paloyo 2020 and the work of Tincani 2015)

A minority report might argue that teacher's assigned marks evaluate other dimensions such as patience, behavior, or the resilience of students. My empirical design and strategy deal with these issues and strengthen my findings. Indeed, rank is identified in a way that it is as good as random, exploiting grading standards. In addition, I exploit two-way fixed-effect strategies at mark and classroom level. As a sensitivity check, the use of a within-students across-subjects strategy sets aside these critics by including individual fixed effects. A possible limitation is that lower secondary school unobserved (subject-specific) characteristic impact later, but I think that classroom fixed effect in 8th grade partially accounts for this issue.

I am far from calling for the abolition of the current grading practice in educational systems, but a thoughtful reflection is needed. Even if this work does not aim to abolish grading, it outlines how it harms educational trajectories and advocates the consideration of alternative approaches toward grading, such as those employed in Sweden (Facchinello, 2016).

However, this work outlines three crucial implications. First of all, grading increases categorical inequality among students, shaping their trajectories over time. Second, teachers have to be aware that their decisions may affect not only a choice of the track but also self-esteem, self-stigma, and the self-identity of students. Third, I believe that this work takes the same furrowed path as Facchinello (2016). He argues that a lack of grading and less harsh grading schemes might enhance the educational outcomes of students at risk, such as low performing students, low SES students, and ethnic minorities.

References

Abbiati G., Argentin G. & Gerosa T. (2017). Different teachers for different students? Evidence on teacher-student matching and its consequences in the Italian case. *Politica Economica*, *33*(1), 13-58.

Agasisti, T. & Falzetti P. (2017). Between-classes sorting within schools and test scores: an empirical analysis of Italian junior secondary schools. *International Review of Economics*, *64*(1), 1-45.

Almås, I., W. Cappelen, A., Salvanes, K., G., Sørensen E. & Tungodden B. (2016). What Explains the Gender Gap in College Track Dropout? Experimental and Administrative Evidence, *American Economic Review*: Papers & Proceedings, *106*(5), 296–302.

Almlund, M., Duckworth A., Heckman J., & Kautz T. (2011). Personality psychology and economics. IZA Discussion Paper 5500.

Argentin G., Barbieri G., and C. Barone (2017), Origini sociali, consiglio orientativo e iscrizione al liceo: un'analisi basata sui dati dell'Anagrafe Studenti, *Politiche Sociali, 1*, 53-74

Azmat, G., & Iriberri N. (2010). The importance of relative performance feedback information: Evidence from a natural experiment using high school students, *Journal of Public Economics*, *94*(7), 435–452.

Babad E. (2009). *The social psychology of the classroom.* Routledge Research in Education.

Babcock, P. (2010). Real costs of nominal grade inflation? New evidence from student course evaluations. *Economic Inquiry*, *48*(4), 983-996.

Ball, S.B., Eckel, C.C., Grossman, P.J., & Zame W. (2001). Status in Markets. *Quarterley. Journal of Economics*. *116*(1), 161–188.

Bar-Tal, D. (1979). Interactions of teachers and pupils. In I.Frieze, D. Bar.Tal & J.S.Carroll (Eds.) New approaches to social problems: Applications of attribution theory (pp. 337-358). San Fransisco: Jossey-Bass

Bar-Tal Y. & Bar-Tal D. (1990) Social psychology analysis of classroom interaction. In Feldman R. S. (1990). The Social Psychology of Education, Cambridge University Press

Barbieri, G., Rossetti, C., & Sestito P. (2010). The Determinants of Teacher Mobility. Evidence from a Panel of Italian Teachers. Banca d'Italia Working Paper 761.

Betts, J. R. & Grogger J. (2003). The impact of grading standards on student achievement, educational attainment, and entry-level earnings. *Economics of Education Review*, *22*(4), 343–352.

Bobba M. & Frisancho V. (2014). Learning About Oneself: The Effects of Signaling Academic Ability on School Choice, Inter-American Development Bank, mimeo, https://econpapers.repec.org/paper/idbbrikps/7968.htm

Borghuis, J., Denissen, J.J.A., Oberski, D., Sijtsma, K., Meeus, W.H.J., Branje, S., Koot, H.M. & Bleidorn W. (2017). Big Five personality stability, change, and codevelopment across adolescence and early adulthood. *Journal of Personality and Social Psychology*, *113*(4), 641–657.

Bowles, S. & H. Gintis (1976). *Schooling in capitalist America*. New York: Basic Books 1976.

Braningan, A. R. (2017). (How) Does Obesity Harm Academic Performance? Stratification at the Intersection of Race, Sex, and Body Size in Elementary and High School. Sociology of Education, *90*(1), 25–46.

Branje, Susan & Lieshout, C.F.M. and J. Gerris (2007). Big Five personality development in Adolescence and Adulthood. *European Journal of Personality*, *21*(1), 45–62.

Brookhart, M. S. (1993). Teachers' Grading Practices: Meaning and Values. *Journal of Educational Measurement*, *30*(2), 123–142.

Brunello G., Crema, A. and L. Rocco (2018). Testing at Length If It Is Cognitive or Non-Cognitive, IZA DP No. 11603.

Buser, T., Niederle, M. & Oosterbeek H. (2014). Gender, Competitiveness, and Career Choices. *Quarterly Journal of Economics, 129*(3), 1409–47.

Carlana, M. (2019). Implicit Stereotypes: Evidence from Teachers' Gender Bias. *Quarterly Journal of Economics*, *134*(3), 1163–1224

Campodifiori, E., Figura E., Papini M. & Ricci R. (2010). Un indicatore di status socio-economico-culturale degli allievi della quinta primaria in Italia. INVALSI Working Paper.

Cho, S. (2017). Explaining Gender Differences in Confidence and Overconfidence in Math. Joint Discussion Paper Series in Economics MAGKS.

Croson, R. & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*(2), 1–27.

Crosnoe, R., Riegle-Crumb R., Frank K, Field S., & Muller C., (2008). Peer Group Contexts of Girls' and Boys' Academic Experiences, *Children Development, 79(1)*, 139–155.

Contini, D. (2013). Immigrant Background Peer Effects in Italian Schools. *Social Science Research*, *42*(4), 1122–42.

Contini D. & Triventi M. (2016) "Between formal openness and stratification in secondary education: Implications for social inequalities in Italy" In Blossfeld, H.-P., Buchholz, S., Skopek, J. & Triventi, M. (Eds.). *Models of Secondary Education and Social Inequality. An International Comparison*. Cheltenham, UK: Edward Elgar

Correia, S. (2017). Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator, Working Paper, http://scorreia.com/research/hdfe.pdf

Correll S.J, & Ridgeway C.L. (2006). Expectation States Theory. In: Delamater J. (eds) Handbook of Social Psychology. H*andbooks of Sociology and Social Research*. Springer, Boston, MA

Cunha F., & Heckman J.J. (2009). The Economics and Psychology of Inequality and Human Development, *Journal of the European Economic Association*, 7(2-3), 320–364

Di Liberto, A. & Casula, L., (2016). Teacher Assessments versus Standardised Tests: Is Acting "Girly" an Advantage?, IZA Discussion Papers 10458, Institute for the Study of Labor (IZA).

Domina T., Penner A., & Penner E. (2017). Categorical Inequality: Schools As Sorting Machines. *Annual Review of Sociology*, *43*, 311-330

Eccles, J., S. (1993). School and family effects on the ontogeny of children's interests, self-perceptions, and activity choice. In J. Jacobs (Ed.), Nebraska symposium on motivation, 1992: Developmental perspectives on motivation (pp.145-208) Lincoln: University of Nebraska Press.

Elsner B., & Isphording I. (2015). A big fish in a small pond: Ability rank and human capital investment. *Journal of Labor Economics, 35*(3), 787-828

Enzi, B. (2015). Gender Differentials in Test Scores and Teacher Assessments: Evidence from Germany. Working Paper, http://www.edge-page.net/jamb2014/papers/Enzi%20-%20Draft.pdf

Figlio, D. & M. Lucas (2004). Do high grading standards affect student performance?, *Journal of Public Economics*, *88*(9-10)1815-1834

Facchinello, L. (2016). The Impact of Early Grading on Academic Choices: Mechanisms and Social Implications. Job Market Paper.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, Relations, *7*, 117–140.

Gamoran, A. (1989). Rank, Performance, and Mobility in Elementary School Grouping. *Sociological Quarterly*, *30*(1), 109–123.

Goldin, C., Katz, K.L. & Kuziemko, I. (2006). The Homecoming of American College Women: The Reversal of the College Gender Gap. *Journal of Economic Perspectives*, *20* (4), 133–56.

Gill, D., Zdenka, K., Jaesun, L. & Prowse, V. (2018). First-Place Loving and Last-Place Loathing: How Rank in the Distribution of Performance Affects Effort Provision, *Management Science*, *65*(2), 459-95

Guskey, T. (2004). *Zero Alternatives*. Principal Leadership: High School Edition, *5*(2), 49–53.

Joensen J.S., Jalava N, & Pellas M.E. (2015). Grades and Rank: Impacts of Non-Financial Incentives on Test Performance, *Journal of Economic Behavior & Organization*, *115*(C), 161-196

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology, 92*(1), 160-170.

Heffetz O. & Robert H. F. (2011). Preferences for Status: Evidence and Economic Implications. In J. Benhabib, A. Bisin, M Jackson, *Handbook of Social Economics*, North Holland

Henson, H.. (1964). *Adaptation-Level Theory: An Experimental and Systematic Approach to Behavior*. Harper and Row: New York.

Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, *39*(153), 41–57.

Horvath H. (2017). Classroom Assignment Policies and Implications for Teacher Value-Added Estimation, Working paper, https://sites.google.com/site/hedvigh/home

Iacus M. S. & Porro G. (2011). Teachers' evaluations and students' achievement: a 'deviation from the reference' analysis, *Education Economics*, *19*(2), 139-159

Kelley, H. H. (1952). The Two Functions of Reference Groups. pp 410-14. *Readings in Social Psychology*. Guy E. Swanson, Theodore M. Newcomb and Eugene L. Hartley, editors. Hoit.

Kiss, David (2013) Are immigrants and girls graded worse? Results of a matching approach, *Education Economics, 21*(5), 447–463.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121-1134.

Kuziemko, I., Buell, R. W., Reich T., & Norton M. I. (2014). Last-Place Aversion: Evidence and Redistributive Implications. *Quarterly Journal of Economics*, 129.1.

Lavy V. Weinhardt F., & Silva O. (2012). The Good, The Bad and The Average: Evidence on Ability Peer Effects in Schools. *Journal of Labor Economics*, *20*(2), 367-414.

Lavy V. (2015). Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries, *The Economic Journal*, *125*(588), F397–F424,

Lucas, R & Donnellan, M. (2011). Personality Development Across the Life Span: Longitudinal Analyses With a National Sample From Germany. *Journal of personality and social psychology*, *101*(4), 847-61

Marsh, H., Seaton, M., Trautwein, U., Ludtke, O., Hau, K.T, Craven, R., & O'Mara, A. (2008) The Big-fish-little-pond-effect Stands Up to Critical Scrutiny: Implications for Theory, Methodology, and Future Research. *Educational Psychology Review*, *20*, 319-350.

McCrae, R R., Costa Jr P.T., Terracciano A., Parker, W. D., Mills C. J., De Fruyt F., & Mervielde I. (2002). Personality Trait Development From Age 12 to 18: Longitudinal, Cross-Sectional, and Cross-Cultural Analyses. *Journal of Personality and Social* Psychology, *83*(6), 1456–1468.

Merton, R.K. (1968). *Social Theory and Social Structure*. Free Press.

Mihaly, K., (2009) Do More Friends Mean Better Grades?: Student Popularity and Academic Achievement, RAND Working Paper Series WR- 678.

Murphy, R. & Weinhardt, F. (2016). Top of the Class: the Importance of Ordinal Rank. Working Paper, https://ideas.repec.org/p/ces/ceswps/_4815.html

Niederle, M. & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*. *122*(3), 1067–101.

OECD (2012). G*rade Expectations: How Marks and Education Policies Shape Students' Ambition*, OECD Publishing, Paris.

OECD (2018), *How much time do students spend in the classroom?*, in Education at a Glance 2018: OECD Indicators, OECD Publishing, Paris.

Paloyo. A (2020). Peer effects in education: recent empirical evidence, in Braedley S. Grenn C. The economics of education. A comprehensive overview. AP. *Forthcoming*

Parducci, A. (1965). Category Judgment: A Range-Frequency Model. *Psychological Review* 76(2), 407–418

Patacchini E. Rainone E., & Zenou Y. (2017). Heterogeneous Peer Effects in Education *Journal of Economic Behavior and Organization*, *134*(C), 190-227

Pekrun R., Cusack A. Murayama K., Elliot A.J., & Thomase K. (2014) The power of anticipated feedback: Effects on students' achievement goals and achievement emotions" *Learning and Instruction*, *29*, 115-124

Puerta MLS, Valerio A., & Bernal M.G. (2016). *Taking Stock of Programs to Develop Socioemotional Skills: A Systematic Review of Program Evidence*, Directions in development. Washington, D.C., World Bank Group.

Ridgeway, C. L. (2006). Status Construction Theory. In P. J. Burke (Ed.), *Contemporary social psychological theories* (p. 301–323). Stanford University Press.

Rosenthal, R. (2002). The Pygmalion effect and its mediating mechanisms. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (p. 25–36). Academic Press.

Skaalvik, S., & Skaalvik, E.M. (2004). Gender Differences in Math and Verbal Self-Concept, Performance Expectations, and Motivation. *Sex Roles, 50*, 241–252.

Seaman SR, & White IR. (2013). Review of inverse probability weighting for dealing with missing data. *Statistical Methods Medical Research*, *22*(3), 278-95

Schizzerotto, A. & Barone, C. (2006). *Sociologia dell'istruzione*. Il Mulino.

Spring, J. (1945). The Sorting Machine: National Educational Policy Since 1945. *The Journal of Education*, *159*(3). Review of Bruce C. Beezer.

Tincani M. (2015). Heterogenous peer effect in classrooms, Working paper, https://ideas.repec.org/p/ces/ceswps/_6331.html

Triventi, M., Kulic, N., Skopek, J. & Blossfeld, H.P. (2016) "Secondary school systems and inequality of educational opportunity in contemporary societies." In Blossfeld, H.-P., Buchholz, S., Skopek, J. & Triventi, M. (Eds.). *Models of Secondary Education and Social Inequality. An International Comparison*. Cheltenham, UK: Edward Elgar

Triventi, M. (2019). Are Children of Immigrants Graded Less Generously by their Teachers than Natives, and Why? Evidence from Student Population Data in Italy. *International Migration Review.*

Vecchione M., Alessandri, G., Roccas, S. & Caprara V.G. (2019). A look into the relationship between personality traits and basic values: A longitudinal investigation. Journal *of Personality, 87*(2), 413–427.

Sun X.., Pennings J.M.H., Mainhard T., & Wubbels T. (2019) Teacher interpersonal behavior in the context of positive teacher-student interpersonal relationships in East Asian classrooms: Examining the applicability of western findings. *Teaching and Teacher Education, 86*, 102-898.

Weiss C., & Fershtman Y. (1998). Social status and economic performance: A survey. *European Economic Review*, *42*(3-5), 801-820.

Westphal, A. Becker, M., Vock, M., Maaz, K., Neumann, M. & McElvany, N. (2016). "The link between teacher-assigned grades and classroom socioeconomic composition: The role of classroom behavior, motivation, and teacher characteristics." *Contemporary Educational Psychology, 46*, 218–227.

Wilkinson, I.A.G., Hattie, J.A., Parr, J.M., Townsend, M.A.R., Fung, I., Ussher, C., Thrupp, M, Lauder, H., & Robinson, T., (2000). *Influence of Peer Effects on Learning Outcomes: A Review of the Literature*. Ministry of Education, New Zeland.

“I'd like to know who's behind all my blunders”

Francesco Tullio Altan[12]

[12] Quote reported in Gino e Michele, Matteo Molinari, (1997) Anche le formiche nel loro piccolo s'incazzano, *Arnoldo Mondadori Editore*

The effect of best and non-reciprocal friends on smoking and drinking behaviour

Chapter III

Abstract

Legal drugs like tobacco and alcohol create dependence and physical harm in the short and long term, dramatically reducing life expectancy. These unhealthy habits root in the early stages of adolescence when the transition to the adult world takes place. Disentangling the early-adoption mechanisms of these habits has ever been of policy interest, and a stream of research has focused on the extent to which social actors such as relatives, parents, friends, and mates exert pressure on the adoption of non-healthy behaviors like smoking and drinking. Surprisingly, findings are quite contrasting in sociological and economic literature so far. Several contributions find no significant effect, whereas others indicate a weak effect. Tracing back and combining sociological and economic literature, I focus on best friends as a reference group to identify a peer effect in a sample of Dutch high school students. In addition, I also investigate the presence of a peer effect in case of non-reciprocal friendship. Once accounted for plausible sources of bias such as reflection, selection bias, contextual effects, correlated effects, and measurement error, I show that the smoking and drinking peer effect is present, mainly in the context of best friendship.

1. Introduction

Since the pioneering studies on the diffusion of innovations (Coleman, Katz, & Menzel, 1957) and aspirations (Duncan, Haller, & Portes, 1968), social science has been interested in analyzing peer effects. In any case, the famous Coleman report (1960) gave a boost to the flourishing peer effect literature on several topics such as job attainment (Granovetter, 1973), spreading of unhealthy behaviors (Gaviria & Raphael, 2001), academic performance (Hoxby, 2000), or agreement on political preference (Yamaguchi, 2013). Regarding the attitudes toward health, studies have investigated extensively the degree to which peers influence each other in adopting risky health behaviors like smoking and drinking.

On smoking and drinking, two paths of research stand out. The first is more focused on the analysis of social actors such as relatives, parents, peers, and mates who play a role in enhancing or mitigating the risk of smoking and drinking. The second investigates the peer effect relying on a friend as a unit of analysis and analyzing individuals' network ties. The focus on friends mitigates ecological fallacy and measurement errors present in the first stream, but it makes more troublesome a causal identification of peer effects. Indeed, such literature always struggles with reflection, selection bias, contextual effects, correlated effects, and measurement error issues. The use of friends increases the selection bias dramatically because individuals tend to choose people as friends with similar characteristics such as socioeconomic origin, ethnic group, and ability (Manski, 1993), generating homophily patterns (Lusher, Koskinen, & Robins, 2013).

In this work, I want to investigate the effect of having a best friend who smokes and drinks on the chance to adopt similar behaviors in high school and I address two research questions: Do smokers and drinkers' best friends enhance the adoption of similar behavior? Do non-reciprocal ties matter in the adoption of similar behavior?

The research questions are not new, but I contribute to the literature in two ways. First, I distinguish between two dimensions of friendship formation: reciprocal and not reciprocal ties. I do focus on the latter because most of the existing studies focus on reciprocal ties, and it is not uncommon that some friendships are not fully reciprocal. From a sociological perspective, it is important to test if a peer effect works still in this case, as theoretically argued by Merton (1968) and Kelley (1952) and empirically shown by Aloise, Graham, and Hansen (1994). Second, I aim to improve the identification strategy compared to most existing works by tackling the possible biases above mentioned and therefore provide a credible estimate of peer effect.

To address my research question, I rely on data from CILS4EU (2016a; 2016b), by focusing on the high school survey conducted in Dutch school classrooms that collects information on multiple dimensions of friendship such as best friends, family friends, and friends to do homework or a walk. To deal with peer effect issues, I follow Bramoullè and colleagues (2009), exploiting indirect friends in a network as instrumental variables. The use of indirect friends is ideal because indirect friends are not exposed to distinct networks and related characteristics, and they are not affected by selection bias due to homophily. Usually, this research stream exploits indirect friends in the same environment provided by the sampling survey, such as school. In contrast, I attempt to collect indirect friends outside the classroom and the schools. With an ego perspective, the aim is to exploit quasi-random variation in the assignment of an indirect friend to the direct one to single out a credible causal effect of having smokers and drinkers as friends. Besides, I exploit the sampling structure of the dataset, where networks and classrooms broadly overlap. Hence, I check for classroom fixed effect to account for unobserved contextual characteristics at the network level. The use of the IV approach shows how the peer effect is underestimated compared to previous contributions in the overall analysis. Non-reciprocal friendships seem not to matter much in contrast with previous literature, but the estimates, in this case, might suffer from a problem of the number of cases available, which enhances the sampling uncertainty. Finally, I debate possible explanations in light of the work of An (2015) on measurement error and the work of Angrist and Pischke (2008) on the IV issues.

2. Theoretical background

2.1 Network effects

Historically, several fields of studies embrace peer effect literature to explain social phenomena, such as the adoption of technological innovations, behaviors, or performance (Mouw, 2006). The idea behind a peer effect is quite simple. Social actors in a social environment influence each other, resulting in similar or dissimilar behavior. Festinger (1954) adopts the term social pressure to indicate a possible mechanism about the internalization of behavior due to a social norm. So, people tend to conform to the social norm in the larger part of cases because they do not want to be isolated from a group. However, social conformity involves a certain degree of conflict among individuals; some reject and others accept a specific behavior. Other authors minimize the presence of conflict and argue that behaviors can spread among people without claims and conflicts (Thorton and Arrowood, 1966), resulting in social facilitation or imitation in a context of no ties among individuals.

What makes arduous an analysis of peer effect is that it is difficult to decompose the peer effect between endogenous and exogenous ones. The former is due to "true" peer pressure, whereas the latter is owing to characteristics of the peers such as age, gender, ability (Moffit, 2001). In addition, peer effect may be heterogeneous regarding the sign, the direction, the categories involved, or how it works. An (2011) proposes a brief classification of possible directions of peer effect. Given a reference group, a peer effect is negative when the effect is socially recognized as negative, like smoking or drinking, while it is positive in a case as ability peer effect. An effect is active when there is a connection or ties among individuals while is passive when there is not a connection. In addition, a tie may not be reciprocal. For instance, A is a friend of B, but B does not have the same feeling toward A resulting in a non-reciprocal friendship. Sometimes, the effect of non-reciprocal friendship is used to explain when individuals start smoking and drinking just because they aim to be a member of a group. (Aloise et al., 1994; Mora & Oreopoulos, 2011). Finally, an effect is asymmetric over some groups. For instance, the effect of having smoker friends is stronger for boys than for a girl, for non-native than for native

Finally, it is worth noting that the identification of what is a reference group is far from being easy. Merton (1968) suggests that people tend to have different reference groups depending on the social actors in town and social environments. Indeed, the literature indicates that reference groups in a network rely on a mix of individual characteristics such as gender, ability, socioeconomic origin, and beauty (Lusher and colleagues, 2013) due to homophily, the selection of a reference group is exclusively a researcher choice. Here, it is essential to discern between the endogenous formation of references like friends and exogenous formation of references like classmates and schoolmates. Indeed, the cluster of friends is at high risk of selection bias, whereas the other clusters are less at risk because they depend on the formal and informal assignment rules of education systems within and across schools (Scott & Carrigton, 2011). The selection of a good reference group and the selection bias are critical issues of network studies together with historical issues of peer effects such as reflection, contextual effects, and correlated effects (Manski, 1993).

Reflection is always present when it is not possible to distinguish between endogenous and exogenous effects of the peers. The endogenous effect of peers is confounded by characteristics of peers such as age, ability, and gender (Paloyo, Mendolia, & Walker 2018). It is hard to solve it because individuals are affected by all individuals belonging to their group and by nobody outside the group. Hence, it is not possible to exploit a variation within the group, and this has previously lead literature to rely on a within-school across-cohort or IVs approach. In contrast, network data are extremely useful for reducing the severity of this issue because the reference group is always individual-specific, allowing the overlapping of cliques across individuals. However, the clique of friends is not randomly formed, enhancing the risk of selection bias. In classrooms and schools, selection bias is present because there is not a random sorting across and within schools resulting in schools with a higher share of students from high-income families, for instance. In the cliques of friends, the selection arises because students befriend peers with similar characteristics. Not accounting for this selection leads to an overestimation of the peer effect since, within cliques, variation is lower than across ones. A standard solution is to adopt network fixed effects, as shown by Mouw (2006).

Even if a proper strategy deals with reflection and selection, unobservable factors may still bias the peer effect estimation. On this issue, one solution is to control student's characteristics as much as possible. Another solution is to exploit a credible exogenous variation, once accounted for basic individual characteristics of students. Network data give the chance of using indirect friends as instrumental variables, but this is not enough. Indeed, indirect friends in the same social environment are not ideal due to transitivity property (Flashman, 2014), according to which it is likely that a guy knows or is a friend of indirect friends.

Thus, I argue it is better to exploit indirect friends across distinct social environments. The use of network data is a value-added because it avoids ecological fallacies of classrooms and schools, but it can lead to a measurement error. Indeed, students may adopt strategic behavior, not reporting each tie of a network. Usually, the bias on the left side of the equation reduces the accuracy of the estimation, but OLS is still unbiased and consistent. However, the simultaneity of the peer effect, it not only attenuates the estimated peer effect but also inflating their standard errors (Wooldridge, 2012). This suggests that previous critiques that peer effects have been overestimated are less grounded when measurement error is present (An, 2015).

2.2 Perspective on network analysis

Previous contributions follow two patterns of research. The first stream, largely based on network studies, focuses on friendships and related cliques as the primary reference group. The use of such a reference group is notable, but it leads to selection bias in the way that students selects friends not randomly but according to some preferences. Hence, it is challenging to identify a credible effect within a friendship network, singling out the selection bias and the other perils of the identification of peer effects. Indeed, such studies are more committed to distinguishing between peer selection and peer influence and are less prone to issues related to causal identification (Lusher and colleagues, 2013).

The second stream of research investigates peer effect exploiting reference groups "by default," such as course-taking patterns, classrooms, and schools. In this way, they alleviate the selection bias due to friendship formation. Usually, they exploit fixed effect, cohort variation, and instrumental variables to deal with the main threats to the identification of peer effects such as reflection, contextual, and correlated effects. They pave the way to a causal identification at the cost of rigid assumptions on network formation.

However, some contributions, such as Bramoulle, Djebbari, and Fortin (2009) and Patacchini, Rainone, and Zenou (2017), have established a new protocol to find a compromise between these two streams: friendship and causal identification. The hint is to exploit the logic of spatial econometrics and to relax some rigid assumptions of network formations. In the beginning, the emerging approach was to exploit the lagged time value of friendship, but this works just if assuming independence of individual preferences over time. Indeed, this is not credible, as outlined by Wang and Bellamare (2020). In contrast, the ongoing approach is to exploit a spatial "lagged" value, namely indirect friends, possibly in different dominions compared to the primary dominion under scrutiny. The idea is to exploit indirect friends to model an exogenous variation to single out possible correlated effects. Some contributions exploit indirect friends in the same social environment (Patacchini and colleagues, 2017; Flashman, 2014), while only a few exploit different dominions (Büyükkeçeci, Leopold, van Gaalen, & Engelhard, 2020).

Using a sub-sample from the Add Health data and network analysis, Fuijmoto and Valente (2010) find a strong peer influence of close friends on their attitudes toward smoking and drinking but find a stronger influence from non-close friends. Gaughan (2006) explores a gender divide in the attitudes toward drinking. Using Add Health and network analysis, it comes out that mixed-sex best friendships influence girls, whereas girls do not show an impact on male friend behavior. Alia and Dwyerb (2010) show that an increase of 10% in the proportion of classmates who drink will increase the alcohol use of four percentage points. In contrast, they find a lower effect on close friends. This finding is linked to the pioneering work of Aloise and colleagues (1994). They suggest that non-reciprocal friendship may exert stronger peer pressure because individuals want to avoid rejection and tend to behave like their "friends." It comes out an emulation effect to be a member of a group. Other observational studies find no pattern of smoking peer effect overall adolescence but only at the beginning (Engels, Knibbe, De Vries, Drop, & Van Breukelen 2006).

Many contributions use family characteristics as instrumental variables, but we report here the more important for the development of the literature. Gaviria and Raphael (2001), using the National Educational Longitudinal Survey in the U.S as a cross-section, estimate the two-stage least square model where peer smoking is instrumented by peer family factors such as parental education and a single-parent household indicator. They report that a one percentage point increase in peer smoking prevalence is associated with a one-sixth percentage point increase in the probability of individual smoking. With Audits and Surveys and similar IV strategy, Powell and colleagues (2005) find a larger effect. Using Add-Health data and Fletcher (2010) find evidence for peer effects in smoking, with an impact between what was found by Gaviria and Raphael (2001) and Powell and colleagues. (2005). On drinking and illicit drug use, Lundborg (2006) exploits school and grade fixed effects in Sweden and founds positive peer effects and alleviates endogeneity by relying on IVs based on family characteristics.

However, these contributions do not convince at all regarding the exclusionary restrictions of instrumental variables. Hence, other contributions move toward other approaches. With the same dataset, Clark and Loheac (2007) exploit the longitudinal design to replace contemporaneous peer smoking with lagged peer smoking, and they adopt school fixed effects. They find no significant peer effects in smoking at all, and they do not discuss the drawbacks of this lagged approach. Using Add Health and NELS, Eisenberg (2004) adopts two different strategies. In the first one, he exploits a longitudinal perspective and cliques of friends as peers unit. Then, he exploits the fact that friends change census resulting in quasi-random variation. In the second one, he exploits the age variation within course-taking patterns. Hence, students of lower age are exposed to students of higher age, in turn, more exposed to risky behavior. However, these strategies show no significant findings. The work of Eisenberg (2004) has on him the development of different perspectives modeling a quasi-random variation, but concerns arise about this random variation. Indeed, the sorting of students is not random, as well as a change of census.

In 2009, Bramoulle and colleagues proposed a novel IV based on indirect ties. According to them, it is possible to exploit indirect ties as an instrument of direct ones. This contribution has lead to growing research in several fields, such as the chance to vote a bill in the congress (Battaglini, Sciabolazza, & Patacchini, 2020), social norms (Ushchev & Zenou, 2020), juvenile delinquency (Lee, Patacchini, & Zenou, 2020). However, the critical point is the specification of the IV equation. Patacchini and colleagues (2017) adopt a massive IVs strategy instrumenting every characteristic of the baseline equation to investigate the peer effect of investment in education in the US using Add Health. With the same dataset, Flashman (2014) choose some variables as instruments to investigate the peer effect of friends on academic performance in the US. Differently, Paloyo and colleagues (2018) use only one instrument to investigate the peer effect of friends on the chance to pass final year certification in England with administrative information.

Aim of this work is to combine as much as possible the two close but distinct streams of research on smoking and drinking peer effect, shedding light on possible gaps I have reviewed previously. Tracing back the contribution on the role of inside and outside group of Aloise and colleagues (1994), I focus on the role played by best friendship and non-reciprocal friendship in shaping individual behavior. The expectation is that smoking and drinking friends enhance the chance of adopting similar unhealthy behavior. Besides, I exploit that students nominate some friends, but they do not have a reciprocal nomination. The expectation is that these non-reciprocal friends might matter more on behavior.

- ✓ H1: Higher share of smoking and drinking best friends leads to similar behavior
- ✓ H2: Non-reciprocal ties lead to the adoption of similar behavior

3. Identification strategy

3.1 Main rationale

A simple model of individual behavior is described in equation 1, where the chance of smoking depends on individual characteristics and clique characteristics.

$$(1) \qquad Y_{ic} = \alpha_{ic} + \mu_c + \beta Clique_{\ ic} + \beta G_i + \beta W_{ic} + \varepsilon_{ic}$$

Where $Clique_{ic}$ represents the variable of interests, the share of students who smoke and drink (clique of friends for students i in classroom c), G_{ic} represents a vector of individual characteristics, W_{ic} represents a vector of clique characteristics, μ_c indicates the classroom fixed effects, and ε_{ic} is the error term. As debated earlier, some issues came out, such as reflection, selection bias, contextual effects, correlated effects, and measurement error. In-network data, cliques overlap each other reducing the reflection issues. However, selection bias is always present because individuals select their friends with similar characteristics to themselves (McPherson and colleagues, 2001), and the risk is to overestimate the effect of friends on achievement (Mouw, 2006). In addition, students may be exposed to unobserved characteristics such as smoking and drinking role models (older friends, brothers, and sisters). Finally, simultaneity between Y and clique of interest increases the risk to overestimate or to underestimate the effect due to the measurement error. As mentioned before, the adoption of a longitudinal design does not solve any issues because the time change between the two waves is yet biased by selection bias. Hence, I rely on an instrumental variable to model an exogenous variation in the smoking clique of friends, such as an indirect friend. I instrument the beta of equation (1) with equation (2) that is the first stage of 2SLS:

$$(2) \qquad Clique_{ic} = \alpha_{ic} + \mu_c + Z\ Indirect\ Clique_{\ ic} + \beta G_i + \beta W_{ic} + \varepsilon_{ic}$$

Where $ZIndirect\ Clique_{ic}$ represents the variable of interests (indirect clique of friends for students i in classroom c), βG_{ic} represents a vector of individual characteristics, βW_{ic} represents a vector of clique characteristics, μ_c indicates the classroom fixed effects, and ε_{ic} is the error term. Equation (1) and (2) show the two-step approach to model an IV approach. The only concern for the reliability of this strategy is that equations 1 and 2 must have the same kind of controls. Then, two assumptions must hold in an IV strategy, as mentioned in Angrsit, Imbens, and Rubin (1996):

1) the indirect friends' characteristics must be independent of the error term

2) one's friends' characteristics must be correlated with her indirect friends' characteristics.

Assumption 1) is not testable, but indirect friends are good candidates because they do not share the same cliques, and they are not exposed to the same unobserved factors. Previous contributions such as Flashman (2014) and Patacchini and colleagues (2017) exploit indirect friends in the same schools. This paves the way to possible issues because they are exposed to the same context. In contrast, we walk a bit away from these contributions, relying on indirect friends in a different social environment compared to the one under scrutiny. Indeed, indirect friends may be in the same classroom, schools, and the workplace. To avoid any issue (Figure I and II), I use only indirect friends from other schools or other contexts (green and yellow). Besides, it is worth noting that A is a friend of B, but it is a friend of F, thanks to the transitivity. I explore more in-depth this issue, analyzing non-reciprocal ties. It is a way to hold constant transitivity between two friends, assuming that transitivity is less likely in the context of non-reciprocal ties. Finally, I adopt a basic set of controls such as gender, ethnic origin, and socioeconomic origin to account for possible confounders.

Figure I: Student-centered network scheme across social environments

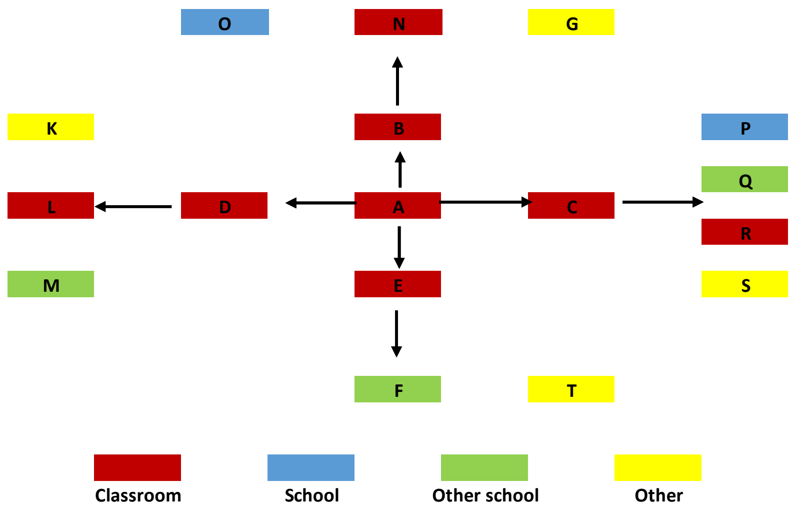
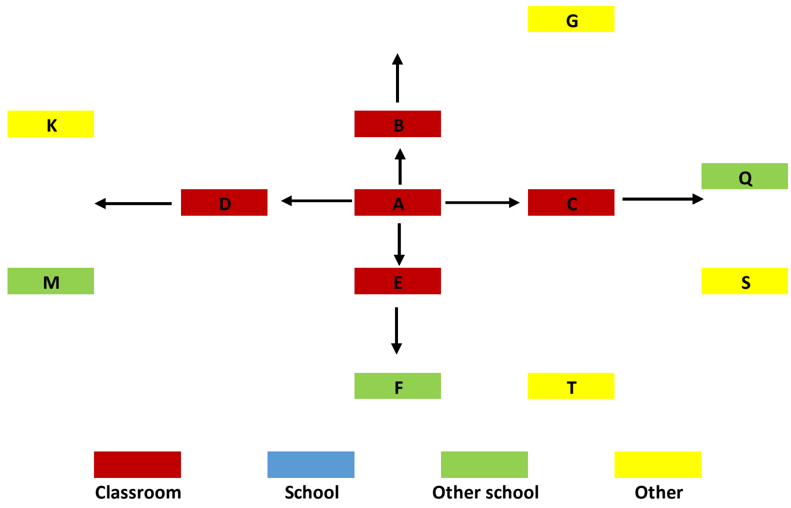


Figure II: Student-centered network scheme across social environments after cleaning for IVs strategy



Assumption II is testable in the first stage of the IVs approach. The idea is that indirect friends are similar to direct ones except for the cliques of origin and – in this work – the social environment of origin. In the first stage, characteristic of indirect friends is tested on the outcomes of interest. Usually, the rule of thumb is to accept every instrument with an F statistic above 10 (Stock, Wright, and Yogo Test) and with a weakness test (Kleibergen Test) close to 0. Things in practice go differently, and several papers deal with weak instruments and F-statistic under ten, as discussed by Andrews, Stock, and Sung (2018). Then, the same authors discuss the importance of distinguishing between the presence of one instrument for one endogenous variable or more instruments for one or more endogenous variables. It can happen to have one or more instruments but weakly correlated to the variables of interest. At the moment, it is possible to account only for one instrument in the context of one endogenous variable. This theoretical contribution is linked to the previous contribution to indirect friends, where there are three main empirical strategies on IVs (Table I). First, Paloyo and colleagues (2018) instrument only the endogenous clique of interest (for instance, smoking clique). A second approach is followed by Flashman (2014), who aggregates direct friends according to a characteristic of interest such as ability and then instruments the endogenous clique of interest and other clique's characteristics introduced as controls (based on the literature such as ethnic, gender, and socioeconomic background). The third approach is outlined by Patacchini and colleagues (2017). Adopting spatial econometric techniques, they instrument every friend with multiple indirect friends for every endogenous characteristic.

Currently, my work follows the first and the second strategy, respectively, in the designs A and case B-I, selecting some specific characteristics such as gender, ethnic, and socioeconomic origin.

Table I. Resume of IV approaches in the literature

	Design A	**Design B-I**	**Design B-II**
Patacchini, Rainone, and Zenou (2017)			**All** endogenous cliques have own instrument
Paloyo Mendolia and Walker (2018)	**One** endogenous clique and one instrument		
Flashman (2014)		**Some** endogenous clique and related instruments	

Empirical strategy

	Model A: introduction of a clique of interest and related IV	Model B-I: Inclusion of control cliques and related IVs	Model B-II: Inclusion of control cliques and some related IVs

3.2 Main threats to the current identification

To clarify the identification strategy, a directed acyclic graphs (DAG, hereafter) may help in singling out the main threats to a causal identification. In figure III, a synthetic DAG reports the main issues already mentioned and discussed in the previous sections.

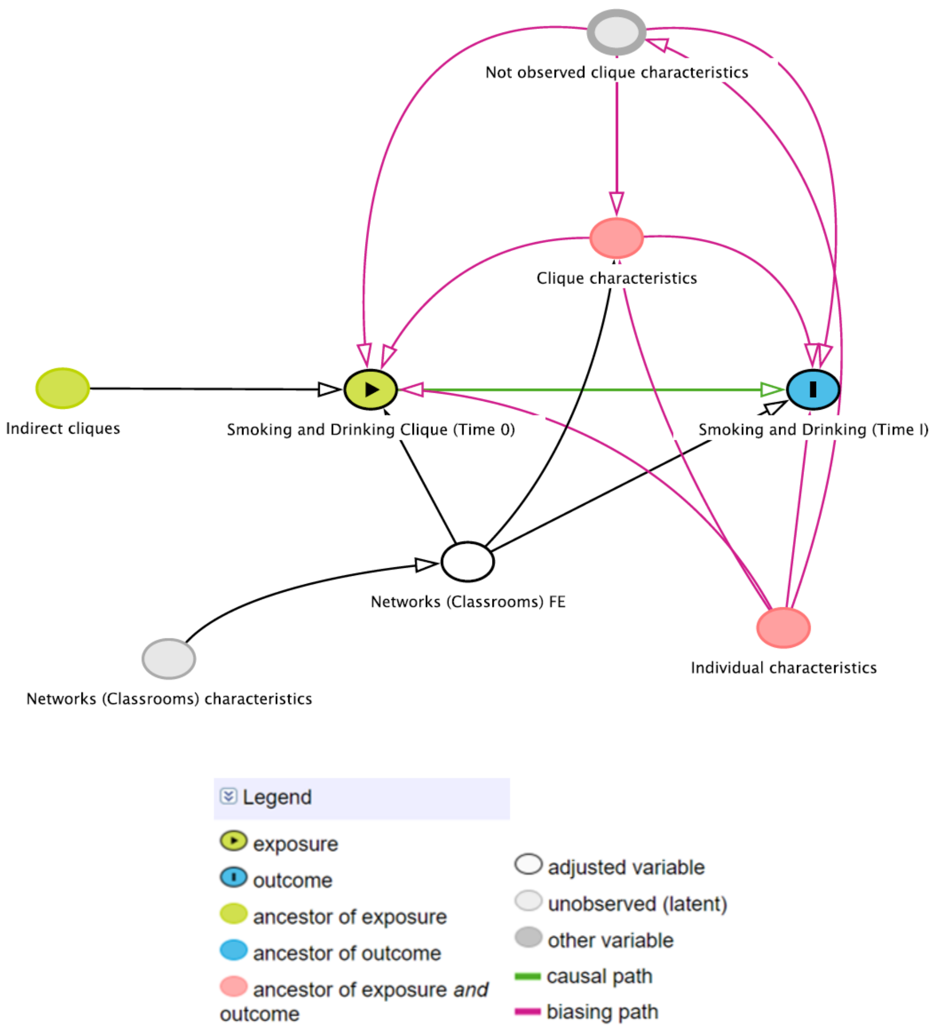
Hence, unobserved characteristics open a biasing path since they affect both treatment and outcome. In this perspective, network fixed effects may alleviate selection bias but are not a panacea for correlated effects. The unobserved variables might be exposed, for instance, to external schooling activity. A solution is given by an instrumental variable that correlates with the outcomes only through the treatment and does not correlate with the error term. Once accounted for these constraints, a credible causal identification is possible.

1a. An indirect friend of another school does not correlate with the error term but only with the treatment.

Nevertheless, someone could claim that "the friend's friend is my friend." In this perspective, the use of a sub-analysis with non-reciprocal ties may help the identification.

> 1b. An indirect friend - of another school and friend of not reciprocal ties - does not correlate with the error term but only with the treatment.

Figure III: Main identification with Directed acyclic graph



Hence, unobserved characteristics open a biasing path since they affect both treatment and outcome. In this perspective, network fixed effects may alleviate selection bias but are not a panacea for correlated effects. The unobserved variables might be exposed, for instance, to external schooling activity. A solution is given by an instrumental variable that correlates with the outcomes only through the treatment and does not correlate with the error term. Once accounted for these constraints, a credible causal identification is possible.

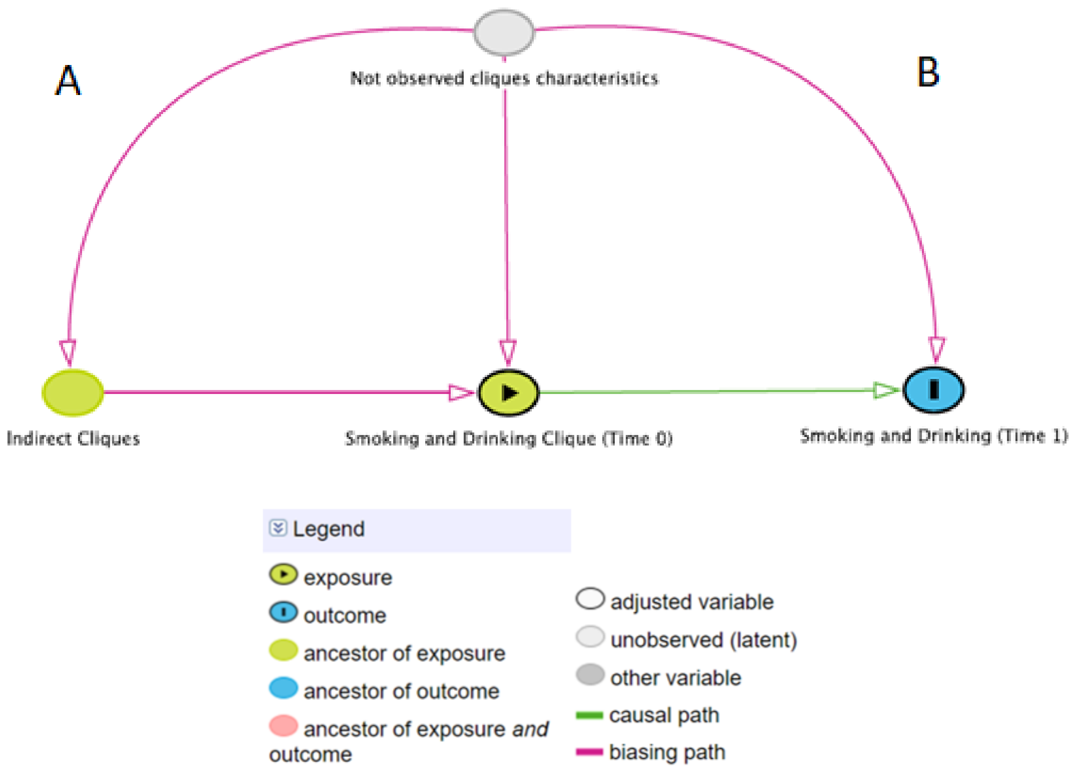
1a. An indirect friend of another school does not correlate with the error term but only with the treatment.

Nevertheless, someone could claim that "the friend's friend is my friend." In this perspective, the use of a sub-analysis with non-reciprocal ties may help the identification.

1b. An indirect friend - of another school and friend of not reciprocal ties - does not correlate with the error term but only with the treatment.

This work models three behaviors: the behavior of students, of friends, and indirect friends. The dag (figure +) sums up the issues. If these behaviors are both exposed to not observed characteristics, a biasing path arises, and Indirect Cliques are not anymore good instruments. Reasonably, actual threats are mainly two: neighborhood or family effects. Both affect the estimation in section A & B. Indeed, students, friends, and indirect friends might have some relatives in a common, parents might be friends, or they live in the same neighborhoods. This link exposes all to something out of the researcher's control.

Figure IV: Lack of identification due to common exposure



At this current research stage, the indirect friends are not - by construction - of the same schools. However, this approach does not solve all of the mentioned threats. Currently, the work relies on two implicit assumptions regarding indirect cliques: (1) family connections are rare and (2) the sorting of high schools makes low the chance of having a classmate of the same neighborhoods. The former might be reasonable, but the latter is a strong institutional assumption. Hence, in the next stages, the *work is planning to exploit* more in-depth family connections and neighborhood exposure.

Thanks to socio-metric data, it is possible to know where direct and indirect cliques interact, places such as school, neighborhoods, clubs, at home, online, at work, elsewhere. The answer is not mutually exclusive, and the idea is to drop who tick neighborhoods.

> 2a. A ticked environment of interaction is a proxy of living in the same neighborhoods. If direct and indirect live in the same neighborhoods, the use of this proxy breaks the biasing path in section A.

In addition to this information, it is possible to exploit another source, such as if friends live within a 5-minute walk from the home of the students.

> 2b. If we assume that it is a reliable proxy of the neighborhood, the use of this proxy breaks the biasing path in section B.
>
> The joint adoption of 2a & 2b is a way to identify neighborhoods and control for them. The risk is that 5-minutes is "arbitrary," and it can inflate or deflate the estimation.

Thanks to socio-metric data within the classroom, it is possible to know if students and direct clique friends' parents communicate with each other. This information ideally works as control variables or fixed effect.

> 3. Parental communication is a proxy of parental friendship. If we assume it, it is possible to alleviate the biasing path of section B but not of section A. It is reasonable to think that once accounted for the neighborhood effects, the risk that students and indirect friend parents have connections is negligible.

3.3 Debate on the interpretation of IV effect

Finally, it is crucial to traceback of the current debate on instrumental variables. The interpretation of an IVs strategy partially depends on the treatment intensity, namely the distribution of the treatment variable and its interaction with the kind of variables such as continuous and categorical (Imbens & Angrist, 1994; Angrist & Pischke, 2008). Theoretically, when both variables are categorical, there is a loss of information, and what is estimated is a local average treatment effect. Having at least one continuous variable allows us to have more information, and the resulting estimate is a weighted average of the causal effect. An ideal situation is when both treatment and instrument are continuous because it is possible to exploit the whole distribution of the two variables resulting in the weighted average of the marginal causal effect. Hence, the features of the two variables (endogenous in play are important to understand the kind of effect estimated. (Table II). Although our work relies on two continuous variables, smoking and drinking *direct and indirect cliques* are unevenly distributed. In my analytical samples, on average, students are exposed to direct and indirect cliques where only the 25% smoke and the 55% drink. Despite the fact that our estimation tends to the identification of a weighted average of the marginal causal effect, it conveys the underlying distributions, already mentioned. In this way, students are exposed to a different "dose" of smoking and drinking cliques. Indeed, students may have the chance of having all smoker friends, only two, or only one. Hence, I argue that I identify a local weighted average of the causal effect (Angrsit & Imbens, 1992).

Table II: Scheme of interpretation of IVs strategy

		Treatment (endogenous beta)	
		Categorical	Continuous
Instrument	Categorical	LATE	Local Weighted average of the causal effect
	Continuos	Local Weighted average of the causal effect	The weighted average of the marginal causal effect

4. Analytical strategy

4.1 Data

For this work, data come from the Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU, 2016a & 2016b), which covers four European educational systems: England, Germany, the Netherlands, and Sweden. It comprises four countries, 400 schools, 800 classrooms and networks, and about 18.000 students. Indeed, the design is based on two randomly selected classrooms for each randomly sampled school. The main targets are 14-year-old students with an oversample of children with ethnic origin, information from co-ethnic, and interethnic peers.

CILS4EU (2016a, 2016b) follows students for three waves, complemented by interviews with teachers and parents in one wave. In detail, students fill up one questionnaire about socio-demographic characteristics, beliefs, attitudes, and behavior, one test about language and logic competencies, one questionnaire about ties in the classroom according to several dimensions (for instance, who are the best friends in the classroom), and another one, a "friends" questionnaire about ties across social environments (not only classrooms). Exploiting these questionnaires, it is possible to draw on a network of direct and indirect friends, inside and outside the classroom. However, it is worth noting that the two network questionnaires overlap regarding classroom ties in the UK, Germany, and Sweden. This is due to the lack of specific questions regarding the context of the tie if present in the classroom, school, or another context. In contrast, the Netherlands survey allows investigating in detail the ties with information on friendships and ties outside school, which is critical for my identification strategy. Also, the Dutch questionnaire in the "friends" questionnaire collects individual measures of friends' smoking and drinking in contrast to other country questionnaires, where the aggregate percentage is collected.

4.2 Variables

To capture our outcome of interest, I use self-reported behavior of smoking and drinking in wave I and wave II (Table III). To make easier the analysis and the interpretation of the effect, I recode these Likert scales in dummy variables at the time I and time II. In the smoking dummy, at times I and II, 0 indicates that a student does not smoke and 1 the opposite. I do the same recodification for drinking at times I and II.

My reference group is the best friend of each student in the classroom. I exploit the fact that each student nominates in wave I up to five friends, and I know these best friends smoke or drink. This is my reference on which I build my direct clique of smoking and drinking friends. Hence, I build a percentage of who smoke (*smoking direct clique*) and the percentage of those who drink (*drinking direct clique*) out of all friends in wave I. I followed the same procedure to build my main instrumental variables, with only a difference. Exploiting the friend questionnaire, I computed the percentage of smokers (*smoking indirect clique*) and drinkers (*drinking indirect clique*) among indirect friends outside the school, which implies excluding indirect friends related to classroom and school environments. All the share of smoking and drinking cliques (direct and indirect) have ranged from 0 to 1.

In the design I, I rely on a set of basic control variables, including gender, ethnic origin, socioeconomic background, body mass index (BMI), a proxy of student ability, and a proxy of grade retention. I use gender to identify boys and girls. To proxy ethnic origin, I created a dummy variable, distinguishing between native and non-native children (self-reported answer). For socioeconomic background, I use the SIOP scale present in the dataset. To proxy ability, I rely on measures of academic competencies such as standardized tests in logic competencies measured in wave I. To capture late or early students, I use it as a proxy the year of birth, taking into account the age students should have, given the Dutch school rules. However, I note the presence of a not negligible age variation within classrooms. Hence, I create a variable with three codes: regular, late, or and early students. BMI is built, relying on self-reported measures of height and weight. Finally, I use aggregate computing the mean (continuous variables) or proportion (for dummy variables) of these individual controls at the clique level using a reference group the best friends nominated by each student in wave I. In detail I use the share of girls, the share of non-natives, the share of late students[13], the average SIOP score, the average ability, and the average BMI. All the "control" computed as shares vary from 0 to 1.

[13] All born in 1993 and 1994.

Table III: Description of the questions from which the main variables are derived

Variables	Original questions
Dependent variables	
Smoking	How often do you smoke cigarettes? (every day, once or several times at week, once or several times at month, less often, never). **After the recoding:** Best friend: Smokers time I: 22,33%; Smokers time II: 25,76% Non reciprocal: Smokers time I: 22,33%; Smokers time II: 19,80%
Drinking	How often do you drink alcohol? (every day, once or several times at week, once or several times at month, less often, never). **After the recoding:** Best friend: Drinkers time I: 43,84%; Drinkers time II: 70,73% Non reciprocal: Drinkers time I: 56,68% Drinkers time II: 54,5%
Item to build direct clique of best friends (Classroom ties questionnaire)	
Best friends	Who are your best friends?
Item to build cliques of indirect friends (Friends questionnaire)	
Gender	Is this friend a boy or a girl?
Ethnic	What is his/her background?
Socio-economic origin	1. What type of education does he/she do (If he/she is no longer in school: What type of education did he/she do?): vmbo-basis, vmbo-kader, vmbo-gt, vmbo-t, havo, vwo, mbo, hbo, university
Smoking	Does he/she smoke cigarette?
Drinking	Does he/she drink alcohol?
IV – selection criteria	2. Does he/she go to your school? 3. Same classroom 4. Same school 5. Another school 6. No schooling

To carry on my second design, I need to build other indirect cliques to use as instruments. Following the contribution of Flashman (2014), I built “indirect friends instrumental variables” for other characteristics than the share of smokers and drinkers, such as share of female, the share of non-native students, and share of low socioeconomic origin among out-of-school indirect friends. More precisely, I use the share of out-of-school indirect friends that do not attend the same schools (see table III). Note that in the questionnaire, there is not a direct question on the socioeconomic origin of friends, but I proxy it using the track attended in school[14] (see table III). One should note that this variable is both a proxy of both the socioeconomic origin and academic ability of these peers (Blossefeld, Buchholz, Skopek, & Triventi 2016). One should note that a re-mixing of students due to field choice takes place between the two waves in the Netherlands school system. This leads to a loss of cases because students are not anyone tracked by the questionnaire. To mitigate the loss of information due to panel attrition in wave II, I rely on inverse probability weights, which is a relatively simple technique that allows for correcting missing values as a function of observed covariates (Seaman & White 2013). To predict the missing values, I use basic demographic characteristics such as gender, ethnicity, socioeconomic index, and ability.

I run network data management with R because it provides suitable packages, but I run all main analyses with STATA. First, I compute the cliques of direct and indirect friends separately, then I marge all information in one dataset. In doing so, each student is exposed to the share or average of each clique characteristics, and the overall share or average of indirect clique characteristics.

4.3 Methods

My first empirical strategy relies on OLS estimates with and without classroom fixed effects and clustered standard errors at the school level. Regarding the IV estimation, I developed two specifications following case A, B - II mentioned in the identification strategy section. I report herewith equations the main idea with smoking as an example.

- **IV A:** I instrument smoking and drinking clique with only an indirect clique

$$Smoking\ Clique_{ic} = \alpha_{ic} + \mu_c + Z\ Smoking\ Indirect\ Clique\ _{ic} + \beta G_i + \beta W_{ic} + \varepsilon_{ic}$$

- **IV B-I**: I instrument smoking and drinking clique and all control cliques with related indirect clique used as control.

$$Smoking\ Clique_{ic} = \alpha_{ic} + \mu_c + Z\ Smoking\ Indirect\ Clique\ _{ic} + \beta G_i + \beta W_{ic} + \varepsilon_{ic};$$

[14] All students who report an enrollment in the VMBO + MBO path. However, in my minority reports and preliminary analyses, I create and use different combinations such as share of students enrolled in HAVO or VMBO + MBO and HAVO + HBO. In addition, I planned an "informal prestige score" assigning a score to each possible path to have a continuos variable. For instance, VMBO + MBO (0), HAVO + HBO (1), VWO + WO (2), University (3). However, I can anticipate that the use of different socioeconomic indicator does not change the results of the analyses.

$$Share\ of\ female\ Clique = \alpha_{ic} + \mu_c + Z\ Share\ of\ female\ Indirect\ Clique_{\ ic} + \beta G_i + \beta W_{ic} + \varepsilon_{ic};$$

$$Share\ of\ non-natives\ Clique = \alpha_{ic} + \mu_c + Z\ Share\ of\ non-natives\ Indirect\ Clique_{\ ic} + \beta G_i + \beta W_{ic} + \varepsilon_{ic};$$

$$Socioeconomic\ Index\ Clique = \alpha_{ic} + \mu_c + Z\ Socioeconomic\ Index\ Indirect\ Clique_{\ ic} + \beta G_i + \beta W_{ic} + \varepsilon_{ic};$$

Currently, my work relies on four main working samples (see table IV). For the best analysis, the sample amounts to 3,317 students at the time I, while at time II to 2.655 cases. It is worth noting that the original sample is 4,618 cases. In contrast, the sample of the non-reciprocal friends comprises 1752 in time I and 1390 in time II. The low case number of non-reciprocal ties makes the joint estimation of fixed effects with IV difficult. For this reason, I decided to adopt just the IV strategy to investigate the peer effects of non-reciprocal friends.

Table IV: Description of analytical samples

	Time I	Time II
Best	3310	2655
Non-reciprocal ties	1752	1390

As mentioned above, the collection of information about best friends happens in the classrooms. Hence, it is possible to exploit classroom fixed effects that broadly overlap with networks. My approach estimates a Linear Probability Model because I OLS with dummies as dependent variables. Given that my core independent variables vary from 0 to 1, a unit variation means that student is exposed to a clique of *all smokers or drinkers*.

5. Empirical findings

5.1 Descriptive statistics

On the working samples, around 22,3% of students answer that they have smoked and 43,84% that they have drunk in the last weeks or months. Compared to the other countries of the survey, the smoking percentage is in line with England and Germany (around 25-30%) but quite higher than in Sweden (15%). In contrast, the drinking pattern is quite different because Sweden shows a share of drinking students, around 30%, whereas England and Germany display a dramatic share of around 60-70%. Adopting a network perspective, each student nominates on average, three close friends out of 5, and 2 nominations on average are non-reciprocal. In turn, each friend might nominate up to 5 friends met not only in classrooms but also in another social environment outside the classroom. On average, two nominations out of 5 came from the workplace or other contexts.

Figure V: Lowess plot between smoking direct clique and smoking indirect clique (bandwidth =0.6; r=0,54)

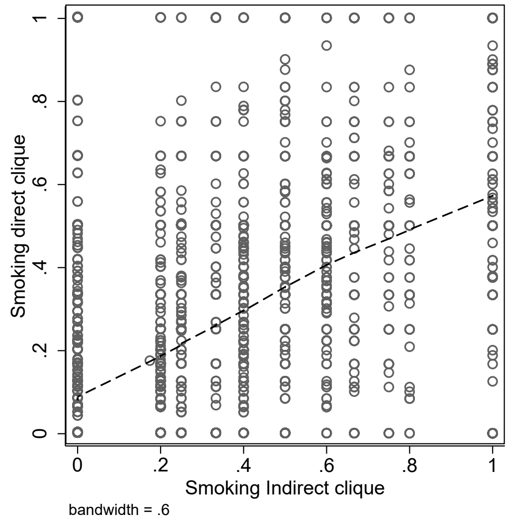
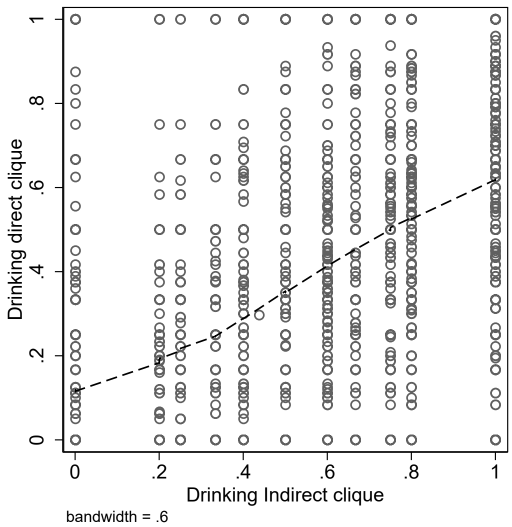


Figure VI: Lowess plot between drinking direct and indirect clique (bandwith =0.6; r=0,51)



In addition, a preliminary descriptive check supports the hint behind the identification strategy. As outlined in figure V and VI, it comes out enough variation to be exploited between an indirect clique and direct clique both for smoking and drinking. Variation is present along with all distribution, even if between 0 and 0.2, and 0.8 and 1, there are no observations in the indirect clique.

5.2 Main findings

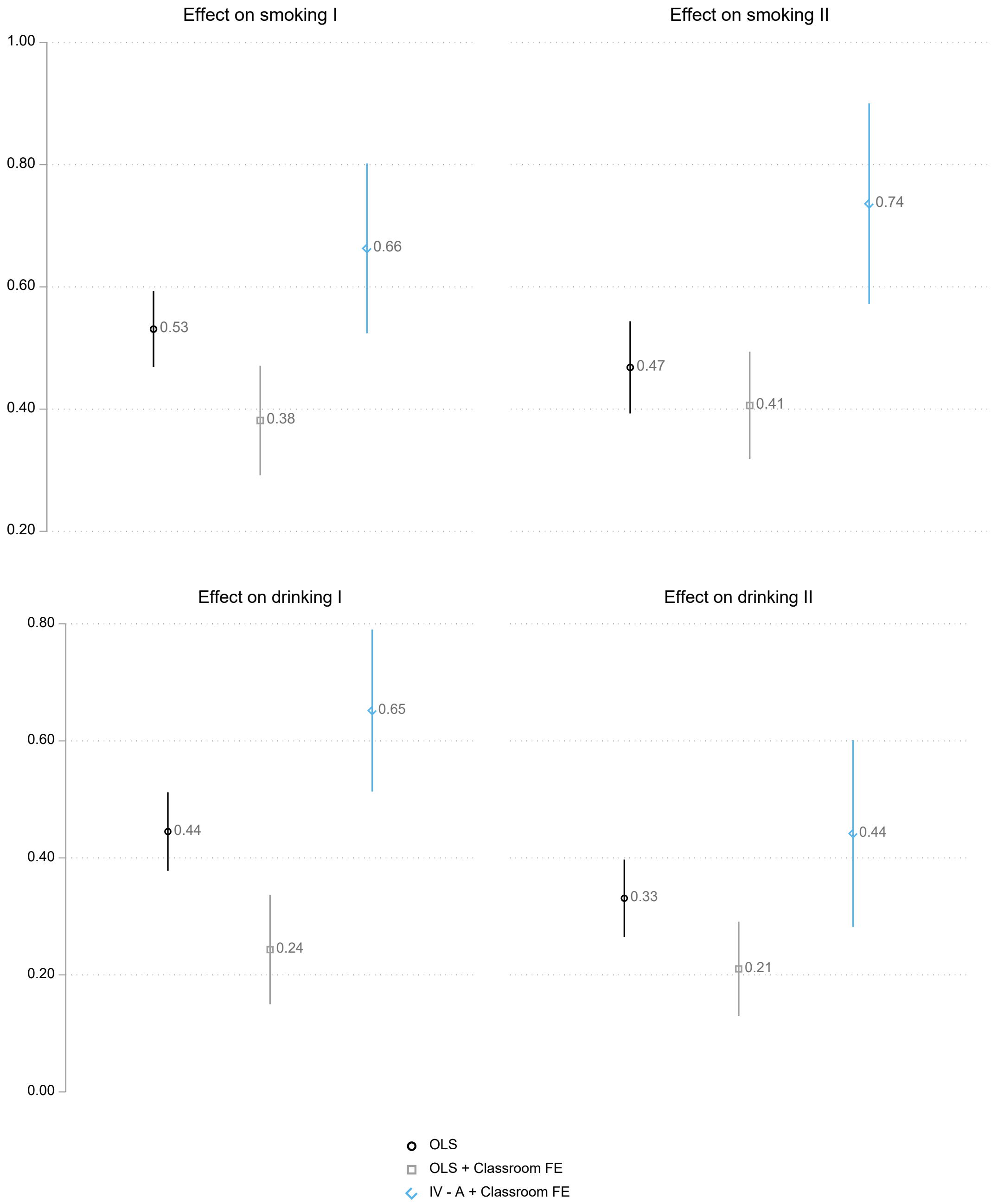
To display the main findings, I report OLS estimation graphically with and without classroom fixed effect, and the results from the IV-A. I do not report the other two here since they do not differ from each other significantly both for the analysis of best friends and non-reciprocal ties. As visible in figure V, OLS estimation of peer effects with and without classroom fixed effects is statistically significant, but the classroom FE reduces the magnitude of the effect since it shrinks the overestimation bias due to self-selection processes across schools and classrooms. The introduction of IV with strategy A increases the main coefficient but also the confidence interval.

This pattern is quite present in all the analyses. On the one hand, it follows the pattern unmasked by An (2015), on the other hand, the larger confidence interval might suggest that there is huge heterogeneity in the working sample, which could stem from some characteristics not yet addressed. The increase of main effect and confidence intervals is common in the IV strategy, and it happens in close contributions such as An (2015) or Paloyo and colleagues (2018). In the latter case, also the magnitude of the effect is similar, and this is owing to a similar approach for building direct and indirect cliques. Finally, it comes out also that in drinking at times I and II, the IV effect is significantly different from the OLS estimation or OLS FE classroom estimation.

Regarding the diagnostic of IV-A, tests are negative in the way that indirect clique works well in terms of correlations (look at the appendix for F-statistic of first stages, page 170). In addition, IV B-I displays confirm the results with a good Kleibergen test only in some cases. However, as visible by the first stages, there are some issues about the first stage of some indirect cliques. This is due to the fact that other cliques' characteristics, such as the share of low-socioeconomic origin,[15] may have a low F-statistics in the first stage, in some cases up to 3. As discussed by Andrews and colleagues (2018), weak instruments depress the overall diagnostics, and at the moment, there is no solution in cases of one weak instrument out of *n* instruments.

[15] The use of different socioeconomic indicators does not change the main results and the related F-statistics. Actually, the share of enrolled in "low tracks" is the more conservative (in term of First stage).

Figure VII: Average Marginal Effect of Best friends; main effect of smoking and drinking direct clique on smoking and drinking at time I and II, Classroom FE, (95% Conf. Int. Reported)

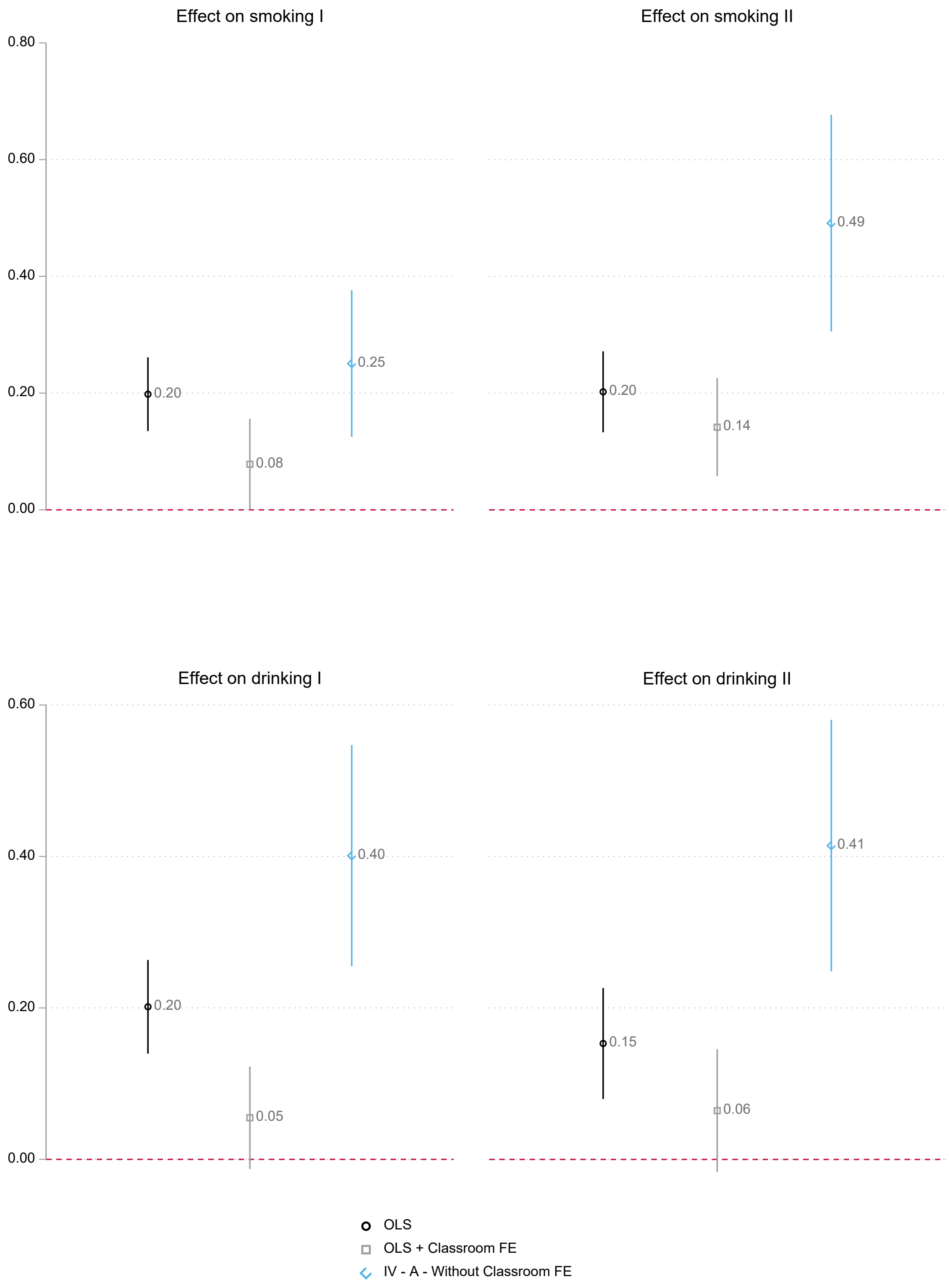


The effect size of having a clique with all smokers or drinkers is not negligible. A unit variation in the clique enhances the chance of smoking of 0.66 percentage points (hereafter, p.p.) in time I, 0.74 p.p. in time II. Regarding drinking, the effect is lower but still sizeable, 0.65 p.p. in time, I and 0.44 p.p. in time II. Is this effect credible? I point out four lenses to read these effects. First, I am analyzing best friends nominated in a network questionnaire. It is reasonable to argue that the best friends have a pervasive influence on student's behavior. Second, I find here a range effect because the above-mentioned changes of smoking and drinking are a cap where all clique members smoke or drink. In detail, students are exposed to the highest dose of treatment. Third, I work with dummy variables where there is a sort of aggregation of smoking and drinking preferences, not distinguishing between one or more cigarettes and one or more drinks. Finally, smoking and drinking are habits embedded in the society and framed as a rite of passage in adulthood.

Compared to the previous contributions discussed in the theoretical background, the effect found here is larger than the effect found with IVs (Eisenberg, 2004) or observational studies (Engels, 2006). The effect is still present when I account for classroom fixed effects, and when I adopt the IV approach. Nevertheless, this kind of comparison is arduous because reference group definition is crucial and frequently changes across contributions. My reference is best friends of students and so it is reasonable to think that they have a stronger pervasive influence than others on behavior patterns.

To hold constant the transitivity, I analyze non-reciprocal ties, namely students who experience a non-reciprocal nomination in the survey. The effect of non-reciprocal friends is quite smaller compared to the reciprocal one with OLS and classroom FE. In addition, the estimation looks significantly different from the best ones. It suggests that there is a low emulation pattern in this way, in contrast with what was found by Fuijmoto and Valente (2010). When I adopt the IV approach without classroom fixed effects, the estimates remain lower but not anymore statistically different from the analysis of best friends. Indeed, confidence intervals are larger and overlap with the estimates of best friends. As discussed in the method section, I drop classroom fixed effects because the estimation shows not informative confidence interval, and this is due (possibly) to the low cases of the analytical samples. These findings seem to go against the previous literature that theorized a stronger emulation effect due to non-reciprocal ties (Aloise et al., 1994). However, this analysis suffers for the presence of a few cases resulting in a lack of statistical power.

Figure VIII: Average Marginal Effect of non-reciprocal ties; main effect of smoking and drinking clique on smoking and drinking at time I and II, Classroom FE, (95% Conf. Int. Reported)



6. Discussion and conclusion

In this work, I show the extent to which best friends and non-reciprocal friends influence the adoption of risky behaviors such as smoking and drinking. It comes out that having smoking and drinking best friends enhances a lot the chance to smoke cigarettes and drink alcohol. In contrast, compared to the previous literature, there is not the same pattern for non-reciprocal ties when I run OLS with and without classroom FE. Nevertheless, the estimates differences are not anymore present when I run IV strategy.

I rely on OLS and IV estimation to account for correlated effect bias in the peer effect. The use of instrumental variables in this stream of research is not new, but to the best of knowledge, no one has exploited indirect friends to estimate smoking and drinking peer effect. Following An (2015), I report that the fixed effect alleviates selection bias due to homophily patterns, but IV strategy increases the estimated coefficients. This is due to a mix of explanations. Firstly, IV strategy eradicates measurement error, increasing beta by default, and secondly, as outlined by Angrist and colleagues (1996), I deal with a dose treatment, enhancing the risk to find large coefficients.

Nevertheless, this work deals with a specific set of limitations, such as lack of adequate family characteristics, exclusionary restrictions of indirect friends, a residual correlated effect due to neighborhood. Indeed, the empirical approach does not control for smoking and drinking habits of parents. Anyway, it would be a hot issue only if someone argues that parental habits and lifestyle influence both the outcome of children and student' friends. While this could be reasonable for the former (parents' behavior affecting children's behavior), it is much less plausible for the latter (parents' behavior affecting the behavior of their child's friends). However, accounting for the literature on role models, and social capital (Geven, S., & van de Werfhorst, 2020) as an additional check I run a preliminary analysis excluding students who know parents' friends[16], and I note a lower effect but still significant, above all when I account for the IV approach.

The approach of IV assumes that indirect friends do not have ties with the ego. I argue that indirect friends collected in environments distinct from the school make more supportive of

[16] In the classroom network questionnaire, several questions are present. At the moment, I use two questions: Who do your parents know? & Whose parents do your parents get together with once in a while or call each other on the phone?

this assumption. However, a possible bias is due to a neighborhood effect when one ego and indirect friends share the same neighborhood. I argue that this channel is not so much present. Indeed, I investigate students in upper secondary schools with a pervasive tracking system that alleviates the chance to find relevant patterns of neighborhood effect (when accounted for a plausible exogenous IV.

An *advocatus diaboli* could claim that the lower and significant effect of not-reciprocal ties is a litmus test that transitivity affects my indirect friends. On this side, I argue that this dynamic is still present at the wave II when best friend clique might be changing owing to selection and course-taking pattern choice. However, I am aware that this residual bias could be present and leads to another research stage. In the next step, I aim at adopting spatial econometrics techniques to relax some assumptions on the network ties using different measures of centrality and neighborhood. This stance allows us to investigate other indirect friends of a different order and to reduce the role played by transitivity but also to investigate more in-depth non-reciprocal ties. Another development will be to deal with the missing information present in the CILS4EU dataset by way of multiple imputations for both waves I and II in place of using IPW just for wave II.

The previous contributions report mixing findings regarding the effect size of IV compared to OLS one. Indeed, Flashman (2014) and Patacchini and colleagues (2017) report a lower IV effect compared to the OLS one. In contrast, Paloyo and colleagues (2018) have a larger effect. Once accounted for the fact that this work has not yet blocked the neighborhood effects, large coefficients might be due to (1) IV correlates with the error term, (2) measurement error ad debated by An (2015), (3) "late effect" (Mogstad, Torgovitsky, and Walters 2020). Preliminary findings suggest that the IV corrected by neighborhoods reduces the coefficient size but is still larger than what was found with OLS. In perspective, I think that the "late effect" plays a role.

In this work, I introduce an IV approach to tackle with reflection, correlated effects, and measurement errors. IV baseline strategy shows a higher effect than OLS estimation. Explanations of this larger effect are different. First, the IV approach eliminates correlated effects and possible measurement error resulting in higher estimation. Second, IV impact depends critically on its own distribution and the extent to which it interacts with the treatment. It can happen that distribution is not uniform, and it produces higher coefficients (Angrist, Imbens and Rubin 1996). Third, previous literature takes for granted that homophily patterns are always present for smoking and drinking too. That is partially true because the choice of best friends rests on a broad array of individual characteristics, beyond the only smoking and drinking. This means that selection bias is not the main issue, and it can explain the large coefficient when adjusted for the correlated bias by IV strategy.

References

Alia M.M. & D. S. Dwyerb (2010). Social network effects in alcohol consumption among adolescents, *Addictive Behaviors, 35*(4), 337-342

Aloise-Young, P., Graham, J., & Hansen W. (1994). Peer influence on group members and outsiders. *Journal of Applied Psychology*, *79*(2), 281–287

An, W., (2011). Models and methods to identify peer effects. In: Scott, John, Carrington, Peter J. (Eds.), *The Sage Handbook of Social Network Analysis*. The Sage Publications, London, pp. 514–532.

An W. (2015). Instrumental variables estimates of peer effects in social networks, *Social Science Research, 50*, 382–394

Andrews I, Stock J., & Sung L. (2018). Weak Instruments in IV Regression: Theory and Practice, Working paper, https://scholar.harvard.edu/stock/publications/weak-instruments-iv-regression-theory-and-practice

Angrist, J. D. & G. W. Imbens (1992). Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity, *Journal of the American Statistical Association, 90*(430), 431-442

Angrist, J.D. & Pischke J. S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ.

Angrist, J. D., Imbens, G W., & Rubin D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*(434), 444–455.

Battaglini M. Leone Sciabolazza V. & Patacchini E. (2020). Effectiveness of Connected Legislators" *American Journal of Political Science,* forthcoming

Bramoullé, Y., Djebbari, H. & Fortin B. (2009). Identification of peer effects through social networks". *Journal of Economet*rics, 150(1), 41–55.

Blossfeld, H.-P., Buchholz, S. Skopek J. & Triventi M. (Eds.) (2016). *Models of Secondary Education and Social Inequality. An International Comparison*. Cheltenham, UK: Edward Elgar.

Büyükkeçeci, Z., Leopold, T., van Gaalen, R., & Engelhardt H. (2020). Family, Firms, and Fertility A Study of Social Interaction Effects, *Demography*, *57*(1), 243-266

Clark, A. E., & Loheac Y. (2007). It wasn't me, it was them! Social influence in risky behaviour of adolescents. *Journal of Health Economics*, *26*(4), 763-784

Coleman, J. S., (1960). The adolescent subculture and academic achievement. *American Journal of Sociology*, *65*(4), 337–347.

Coleman, J, Katz, E., & Menzel H. (1957). The diffusion of an innovation among physicians. *Sociometry*, *20*(4), 253–270.

Duncan, O. D., Haller, O. A., & Portes A. (1968). Peer influences on aspirations: a reinterpretation. *American Journal of Sociology*, *74* (2), 119–137.

Eisenberg, D. (2004). Peer effects for adolescent substance use: Do they really exist? Working Paper UC Berkeley School of Public Health.

Engels C.M.E.R., Knibbe A. R., De Vries H. Drop M. J., & Van Breukelen P.J.G. (2006). Influences of Parental and Best Friends' Smoking and Drinking on Adolescent Use: A Longitudinal Study *Journal of Applied Social Psychology*, *29*(2), 337-361

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140.

Flashman J (2014). Friend Effects and Racial Disparities in Academic Achievement. *Sociological Science*, *1*(17), 260-276.

Fletcher, J. M. (2010). Social interactions and smoking: evidence using multiple student cohorts, instrumental variables and school fixed effects. *Health Economics*, *19*(4), 466-484.

Fujimoto T. & Valente W. (2010) Decomposing the Components of Friendship and Friends' Influence on Adolescent Drinking and Smoking, *Journal of Adolescent Health*, *51*(2), 136-143

Gaviria, A., & Raphael, S. (2001). School-based peer effects and juvenile behaviour. *Review of Economics & Statistics*, *83*, 257-268.

Gaughan, M. (2006). The Gender Structure of Adolescent Peer Influence on Drinking. *Journal of Health and Social Behavior*, *47*(1), 47–61

Granovetter, M. S., (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360–1380.

Geven, S., & van de Werfhorst H. G. (2020). The Role of Intergenerational Networks in Students' School Performance in Two Differentiated Educational Systems: A Comparison of Between- and Within-Individual Estimates. *Sociology of Education*, *93*(1), 40–64

Hoxby, C., (2000). Peer Effects in the Classroom: Learning from Race and Gender Variation. National Bureau of Economic Research Working Paper 7867.

Imbens, G., & Angrist J. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*(2), 467-475.

Kelley, H. H. (1952). *The Two Functions of Reference Groups*. Guy E. Swanson, Theodore M. Newcomb and Eugene L. Hartley, Readings in Social Psychology, 2° edition, pp 410-14. New York: Holt, Rinehart, & Winston

Lee, L-F., Liu, X., Patacchini, E. & Zenou Y. (2020), Who is the Key Player? A Network Analysis of Juvenile Delinquency". *Journal of Business and Economic Statistics*, forthcoming.

Lundborg P. (2006) Having the wrong friends? Peer effects in adolescent substance use, *Journal of Health Economics, 25*(2), 214-233

Lusher D., Koskinen J. & Robins G. (2013). *Exponential graphs random models*. Cambridge University Press.

Manski, C. (1993). Identification of endogenous social effects: the reflection problem. *Review of Economic Studies*, *60*(3), 531-542.

McPherson M., Smith-Lovin L., & Cook M. J. (2001). Birds of a feather: Homophily in social networks, *Annual Review of Sociology, 27*, 415-444.

Merton, Robert K. (1968). *Social Theory and Social Structure*. Free Press.

Moffitt, R., (2001). Policy interventions, low-level equilibria, and social interactions. In: Durlauf, S., Young, P. (Eds.), *Social Dynamics*. Massachusetts Institute of Technology Press, Cambridge, MA.

Mogstad M., Torgovitsky A., & Walters C. R. (2020). The Causal Interpretation of Two-Stage Least Squares with Multiple Instrumental Variables, NBER Working Paper No. 25691

Mora T. & Oreopoulosbc P. (2011). Peer effects on high school aspirations: Evidence from a sample of close and not-so-close friends, *Economics of Education Review, 30*(4), 575-581

Mouw, T. (2006). Estimating the Causal Effect of Social Capital: A Review of Recent Research. *Annual Review of Sociology, 32*, 79-102.

Paloyo A. Mendolia S. & Walker I. (2018). Heterogeneous effects of high school peers on educational outcomes, *Oxford Economic Papers*, *70*(3), 613–634.

Patacchini E. Rainone E., & Zenou Y. (2017). Heterogeneous Peer Effects in Education, *Journal of Economic Behavior and Organization*, *134*(C), 190-227

Powell, L. M., Tauras J.A.,& Ross H. (2005). The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behaviour. *Journal of Health Economics*, *24*(5), 950-968.

Ushchev, P. & Y. Zenou, (2020). Social Norms in Networks, *Journal of Economic Theory, 185*, 104-969.

Scott, J., & Carrington P. J. (2011). *The Sage Handbook of Social Network Analysis*. The Sage Publications, London

Seaman S.R., & White. I.R. (2013). Review of inverse probability weighting for dealing with missing data, *Statistical Methods in Medical Research*, *22*(3), 278-95.

Thorton, D. & Arrowood A. J. (1966). "Self-evaluation, self-enhancement, and the locus of social comparison". Journal of Experimental Social Psychology, *1*(1),40-48

Yamaguchi, K. (2013). New regression models with egocentric social network data: an analysis of political party preference in Japan. *Sociological Methodology*, *43*, 312–345.

Wang Y. & Bellemare M. (2020). Lagged Variables as Instruments, Working paper, http://marcfbellemare.com/wordpress/research

Wooldridge J. M. (2012). *Introductory Econometrics: A Modern Approach*, South-Western Pub

Dataset CILS4EU

Kalter, Frank, Anthony F. Heath, Miles Hewstone, Jan O. Jonsson, Matthijs Kalmijn, Irena Kogan, and Frank van Tubergen. (2016a). Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU) – Reduced version. Reduced data file for download and off-site use. GESIS Data Archive, Cologne, ZA5656 Data file Version 1.2.0

Kalter, Frank, Anthony F. Heath, Miles Hewstone, Jan O. Jonsson, Matthijs Kalmijn, Irena Kogan, and Frank van Tubergen. (2016b). Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU) – Reduced version. Reduced data file for download and off-site use. GESIS Data Archive, Cologne, ZA5656 Data file Version 2.3.0

"He who is different from me does not impoverish me – he enriches me."

Antoine de Saint-Exupéry[17]

[17] Quote reported in de Saint-Exupéry.A. (1942). Flight to Arras, Reynal & Hitchcock Edition

Ethnic concentration and Diversity in Italian primary schools: consequences for pupils' attitudes, social integration, and academic performances

Chapter IV

Abstract

Over the last few decades, a growing literature has investigated ethnic peer effect on student outcomes such as performance and inter-ethnic conflict. Surprisingly, less is known about the extent to which non-native origin may also affect socio-emotional skills and behavior. Additionally, the existing literature frames non-native stock as a homogenous group relying on a quantitative criterion of ethnic presence overshadowing the socio-linguistic diversity of ethnic groups. Armed with Italian administrative data from primary schools, the paper proposes a novel perspective, unmasking ethnic share patterns and spotlighting the role of diversity in shaping an array of students' outcomes. We show an adverse effect of ethnic share and diversity index on being bullied, doing bullyism, and extrinsic motivation, but only diversity negatively affects language competences. However, the effect sizes are very tiny. In addition, we document three critical facts. First, classroom diversity draws on a non-linear negative effect on language competences, mainly at the expense of CMO. Second, no significant thresholds are present to set a classroom maximum of ethnic students. Third, the impact of ethnic share and diversity is asymmetric, increasing extrinsic motivation for II generations of students with migration background but exposing students with Italian origins to a higher perception of conflict in the classroom and a negative learning effect.

1. Introduction

The issue of integration of ethnic minorities and immigrants has a long history in a number of Western societies such as the United States, France, or Germany, while it emerged more recently in other contexts, such as Southern European countries (Zincone, 2006). This is the result of new immigration fluxes involving not only of economic migrants and family reunifications, but also refugees and asylum seekers (Bratti, Deaiana, Havari, Mazzarella, & Meroni 2017). While labor markets and welfare state systems are two critical institutions for the short-term integration in the host society, the educational system is a crucial institutional arena for the successful integration of the children of immigrants in the destination countries.

Indeed, social scientists consider educational attainment to be a key driver of occupational prospects for immigrants and a way to boost their chances of upward social mobility (Borjas, 2006; Rossi & De Phillippis, 2020). Beyond the role of education in fostering labor market opportunities, the educational system is of pivotal importance because natives and immigrants interact daily in school, with the chance of influencing each other's beliefs, attitudes, expectations, behaviors, performance, and educational choices.

The increased presence of immigrants and their uneven geographical distribution due to residential segregation makes the presence of children of immigrants in school an issue with high visibility. Furthermore, in a historical period characterized by growing economic inequalities and an enlarged consensus of populist and far right-wing parties (Bratti and colleagues, 2017), the coexistence of children with heterogeneous ethnolinguistic origins became a salient issue and hot topic in public debates. Some consider the mix of natives and children with immigrant-origin to be an occasion to improve the reciprocal knowledge and mutual trust between ethnic groups, which is expected to foster social cohesion in the long run (Cheong, Edwards, Goulbourne, & Solomon 2007). Others, instead, believe that an excessive presence of immigrants in specific schools and heterogeneity in the student body poses severe difficulties to teaching and learning processes, and this could undermine native students' academic achievement (Angrist & Lang, 2004; Contini, 2013). If parents are aware of this problem, the more informed and socio-economically advantaged families will choose to send their children to expensive private schools or public schools in areas with a low incidence of migrants. This so-called "white-fly" phenomenon, in turn, could exacerbate school segregation and social inequalities in educational outcomes.

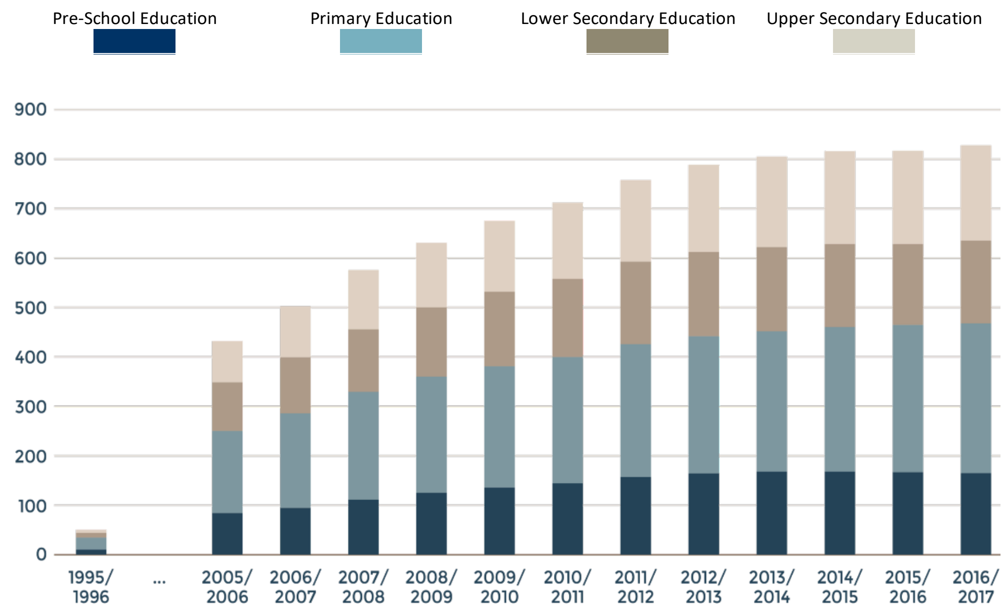
In this paper, we are interested in studying the consequences of the presence of students with a migration background on a variety of student outcomes in Italy. To date, an extensive literature has mainly investigated the link between the share of children with migration origin (CMO, hereafter) and academic performance, overlooking other critical dimensions for children's successful development, such as socio-emotional skills, behavior, and attitudes toward school. These dimensions play a critical role in promoting lifetime success, like better working conditions, good health status, and low criminal behavior (Heckman, Stixrud, & Urzua, 2006; Kautz, Heckman, Diris, Weel, & Borghans, 2014) and, as debated by Almlund, Duckworth, Heckman, and Kautz (2011), their development conveys the age and the social environment.

We address the following research questions: do students benefit, or are they negatively affected by the presence of non-native students in their classroom, in terms of behavior, social integration, socio-emotional skills, and academic competences? Is it possible account for the role of linguistic diversity in classroom? Is the effect of CMO non-linear? Are heterogeneous effects present among natives, second, and first-generation of CMO? Is it possible to identify optimal ethnic thresholds for each of these outcomes?

Past contributions have focused mostly on ethnic concentration – the proportion of immigrant students in the school – as the primary variable of interest, ignoring the socio-cultural Diversity of CMO. Notwithstanding, students with a migration background represent a heterogeneous group in terms of linguistic background and ethnic origins. This basic fact is recognized in the parallel literature on migrants' integration in the labor market and society (Schaeffer, 2013), but it has been surprisingly overlooked when analyzing immigrants in school. We address such limitation of previous works by first building a novel index of ethnic diversity and then testing whether (1) it has an independent effect on students' outcomes, net of ethnic concentration index; (2) whether the effect of ethnic concentration changes when the ethnic diversity in the students' classroom is taken into account; and (3) whether the effect of ethnic concentration depends on the level of ethnic diversity. Finally, we advance the current literature by constructing two indicators of ethnic concentration and diversity at the classroom level instead of the school level, as done by most previous studies. Since the classroom is the educational environment in which students spend most of the time in European school systems, its composition is likely to have paramount importance for pupils' socio-emotional and academic development.

We focus on Italy as a case study for a variety of reasons. First, since the late 1980s, Italy has experienced a growing influx of immigrants from many countries. This has rapidly accelerated in the last two decades, with the share non-native student population growing by about 67% from 2006 to 2017 (MIUR, 2017). Such rapid growth has produced an intense pressure on school principals and teachers to successfully integrate CMO in the educational system, avoiding native performance decline and inter-ethnic conflict. Nevertheless, the education system is still trying to adjust to new challenges. Indeed, there is evidence of biased teachers' evaluations against CMO (Triventi, 2019) and of lower academic performance of children of immigrants compared to natives (Azzolini & Barone 2013). Italy has the highest percentage of foreign early school leavers in Europe (around 35%; Bertozzi 2018), and a large share of teachers complain they did not receive specific training to tackle the challenges of teaching in a multicultural educational environment (OECD, 2012).

Table I: Number students (in thousands) with non-Italian citizenship across the year and by educational level. Italy, 2005-2017 Source: MIUR 2017



Second, CMO is not equally distributed across the country: they are more present in the northern regions and urban areas, in which there are better occupational opportunities and well-structured networks of early immigrants (Jajet, Ukrayinchuk, & De Arcangelis, 2010). Albeit in the Italian education system, an equal sorting of students and teachers across schools and classrooms is formally envisaged by law, parents have room to choose schools for their children, and principals might sort students within their schools in a non-random way. The existence of "hidden" segregation policies between and within schools linked to children's socioeconomic characteristics (Agasisti & Falzetti, 2017) and ethnic backgrounds (Contini, 2013) has been demonstrated in previous works. Moreover, in 2010/2011, the Minister of Education attempted to establish a maximum threshold of 30% of students without Italian citizenship in each classroom. However, this regulation was not systematically enforced, leading to a non-negligible number of schools and classrooms with a share of CMO exceeding the official limit.

In section 2, we will discuss the general theoretical background and then focusing separately on distinct facets of the literature, such as the link between ethnic share and academic performance, the effort for causal identification, non- linear patterns and asymmetric effects, and the role of diversity. In sections 3 and 4, we present our research questions and identification strategy. After presenting our analytical approach in section 5, we will discuss the main findings, and we open a discussion in the last section.

2. Theoretical background

Many empirical contributions investigate the effect of ethnic concentration on students' outcomes, but an autonomous and comprehensive theoretical model is currently lacking. Quantitative studies in this area usually rely on the peer effect literature and especially on the framework provided by Wilkinson, Hattie, Parr, & Thrupp (2000), which theoretically distinguishes compositional effects from what they call "true" peer effects. Compositional effects derive from "*measurement artifacts, differential school or classroom resources, differential school or classroom climates, and differential teacher practices*" (Wilkinson, et al. page 397). The "true" peer effects, instead, arise from the inter-individual processes of comparison, imitation, and influence, stemming from the identification of specific reference points, such as individual classmates or groups of students (Festinger 1954; Merton 1968). Daily interactions and comparison with peers might shapes children's worldviews, thereby influencing their socio-emotional skills, competencies, behaviors, and choices resulting in a virtuous or depressing spiral (Babad, 2009). We argue that the critical driver of social interactions is language, and, for this reason, we focus specifically on this dimension. The importance of social interactions regarding the link between language and cognitive development was already highlighted by Vygotsky (1978) and Rogoff (1990). Language is the primary tool for social interactions and communication (Halle et al. 2014). Indeed, students with proficient national language are more likely to engage in close peer relationships and to establish a relationship with teachers, whereas students not proficient might be labeled as shy and may experience peer rejection and self-stigma (Borman & Pyne, 2016).

Leveraging "Role theory" (Beckerman & Good, 1981) and "Social contagion theory" (Erbring & Young, 1979; Kelly, 2009), we outline arguments that can help us in understanding those processes by which ethnic composition of the classroom could affect children's outcomes. Role theory suggests that the minority group set as a reference group the majority, internalizing their behaviors and characteristics through peer interactions and comparison (Beckerman & Good, 1981). According to the Social contagion theory, students will become more alike through social interactions due to imitation processes (Erbring & Young, 1979; Kelly, 2009). While these theories furnish some elements for expecting that heterogeneity in the ethnic background of the student body might affect children's outcomes via peer effects, the direction of this relationship is not clear a priori.

On the one hand, interaction with pupils of different backgrounds might enhance the socio-emotional skills and performance of all. Exploiting a reform of admission access in the Indian elite schools, Rao (2019) identifies a positive effect of diversity on integration when students from different social classes are sorted together. Exposed to this diversity, rich students adapt their behavior developing higher prosocial skills, are more generous and egalitarian, and learn not to discriminate against the poor but to socialize with them.

Besides, Mijs (2016) outlines how heterogeneous contexts may reduce self-stigma across mates because students start to learn that academic performance outside of their control, such as socioeconomic origin, parental support, or luck. Similarly, we argue that diversity alleviates the self-stigma of students, increasing classroom participation, mainly driven by external motivation. Our reasoning relies on the idea that students internalize that other external factors play a role in achieving and their attitudes toward external motivation change in turn.

Some authors (Hoxby, 2000; Van Ewijk and Sleegers, 2010; Hermansen and Birkelund, 2015) argue in their final remarks that positive higher interaction between natives and CMO students might enhance a "network" of learning support and cohesion with a positive effect on the performance. Indeed, classrooms with an adequate dose of integration make the environment more profitable and enjoyable for learning. Nevertheless, dynamics might go in the opposite direction. High ethnic share may disrupt classroom interactions and communication because students do not integrate with each other, resulting in a sort of Tower of Babel (Ballatore, Ichino, & Fort 2019). The lower interaction among mates might hamper classroom communication resulting in misconduct behavior as a proxy of inter-ethnic conflict (Geven, Kalmijn, & Van Tubergen, 2016).

Previous contributions disseminate different perspectives but not conclusive either theoretically either empirically. Fortunately, the ethnic density hypothesis enhances the understanding of this phenomenon. It comes out that the presence of co-ethnics students improve social support and a buffer against non-inclusive behavior with byproducts such as self-esteem, alienations, and depressive symptoms. Indeed, Halpern (1993) points out that the perception and active presence of emotional support critically depend on co-ethnics mates. Depending on the ethnic share, native or CMO feel integrated or isolated. Hence, it is crucial to investigate the asymmetric effect among native and CMO to single out distinct patterns among cleavages.

In light of the above-mentioned distinction between true peer effect and compositional effects (Wilkinson, Hattie, Parr, & Thrupp, 2000), this theoretical frame might be shaped positively or negatively by teacher's characteristics and school resources. Indeed, teachers usually adapt their instruction quality relying on exploiting learning opportunities, classroom management, and individual support (Klieme, 2006) and the school resources such as rooms available, scientific laboratories, and reasonable financial funds. (Rumberger and Palardy, 2005).

For a long time, ethnic peer effect literature always has been privileged to investigate the link between CMO and students' performance or educational choices. Then, a growing interest was born about a causal identification of an ethnic peer effect. Finally, some stream of research was interested in some extensions, analyzing the presence of non-linear patterns or in a quite limited extent, the role played by the linguistic diversity.

2.1 The effect of ethnic concentration on students' academic achievement and socio-emotional skills

Peer effect literature on the role of ethnic density on students' academic achievement and educational choices is extensive both in the U.S. and Europe, but findings are not conclusive.

While some studies found no evidence of adverse effects of ethnic concentration or even small positive effects, many others report negative consequences of a higher share of CMO in school, albeit the magnitude of these effects is usually pretty modest.

Among the first group of studies, Angrist and Lang (2004) investigate the impact of the Metco de-segregation program on students in Boston and find little evidence that incoming black students negatively affect white students in high schools. In detail, they find a null effect of black students on white performance but a positive effect of white peers on black performance. It is important to bear in mind that they focused on a context with a low native to non-native ratio. These findings were echoed by Cebolla-Boado (2007), who finds non-significant effects of the share of foreigners in school on various children's outcomes in French lower secondary school.

In contrast, the second group of studies found adverse effects of ethnic concentration in school. For instance, in the US, Hoxby (2000) and Hanushek et al. (2009) show that the proportion of black students in a school negatively affects the achievement of blacks in the high schools. Similar negative results are found in Norway, considering other ethnolinguistic groups (Black, 2013). Dumay & Dupriez (2008) find adverse effects on language proficiency in Belgium. Ammenmuller and Pischke (2009), analyzing TIMMS data from Germany, France, Iceland, the Netherlands, Norway, and Sweden, report a negative effect of ethnic concentration on the academic performance of both native and non-native children in the fourth grade. Using the PISA survey, Brunello and Rocco (2011) find evidence of small but negative effects on the academic performance of natives across 27 countries. Other contributions focus on other outcomes, such as dropout and track choices. Gould and colleagues (2009) in the US and Hermansen and Birkelund (2015) in Norway find, respectively, adverse effects on dropout rate, track choice, and the chances of passing the high school exam. Even if such findings are not always consistent, peer effects related to the immigrant background are adverse, small, and sometimes not statistically significant.

As argued in the introduction, most research focused on a limited set of outcomes related to academic performance and later educational transitions. Only a few studies investigated the influence of ethnic concentration on children's behavior and socio-emotional skills. Finn and Voelkl (1993) find that students exhibit worse behavior when they attend a school with a higher share of ethnic minorities. In contrast, Gieling, Vollebergh, & van Dorsselaer (2010) report that ethnic minority students are less aggressive if they attend a school with a higher share of ethnic minorities. Demanet and Van Houtte (2014) point out that ethnic minority and majority students engage in fewer misconduct behaviors in schools with a higher share of ethnic minority students. Hornstra and colleagues (2014) show that non-native share is associated with higher motivation and attitudes toward school belonging. However, a critical point in most of the designs is the extent to which they account for school characteristics, such as the average socio-economic status of students at the school.

Relying on the literature mentioned above, we inquire as to the extent to which CMO share affects student outcomes such as being bullied, doing bullyism, social integration, and performance in mathematics and language. Driven by the several contributions that outline the overall negative association between CMO share and educational outcomes, we hypothesize that higher CMO share leads to a negative effect on integration, socio-emotional skills, and academic competences.

- ✓ Hypothesis 1: Higher CMO share might negatively affect "relational dynamics," "integration," and socio-emotional skills.

2.2 The role of diversity

The role of diversity attracted much attention in business and human resources studies, which examined the role of heterogeneity in employees' origin in fostering or hampering individuals' productivity in the workplace. In this perspective, diversity has been conceived mostly as a value-added for the business (Cox, 1993), satisfying customers' needs, improving the quality of products and services offered (Richard, 2000), or broadening employee perspectives with a positive return on problem-solving (Cox, 2001). Usually, in such studies, diversity means heterogeneity of the workforce in terms of gender, socio-economic background, or ethnic origin.

While the concept of diversity has been widely used in research on US higher education (Allen & Wolniak, 2019) and qualitative educational studies (Deering, 1996), surprisingly, the concept of diversity has been widely overlooked by quantitative studies, except for some work on the migrant's integration in the labor market (Schaeffer, 2013). As stated before, existing contributions only tangentially touched this issue, and they mostly equated diversity with the presence of non-natives in the school. Nonetheless, this approach neglects the fact that, in the school and the classroom, distinct ethnic minorities are present. This has two important implications. First, some ethnic minorities could be more easily integrated than others because of their socio-linguistic and cultural background. Second, two classrooms with the same share of non-natives students can differ in the degree of ethnic origin variety, which might lead to different sets of opportunities for interaction across single individuals and groups. To account for these aspects, we argue it is important first to conceptually distinguish between ethnic concentration (share of CMO) and ethnic diversity (relative heterogeneity within a group of students). Second, we develop an attempt to measure ethnic diversity at the empirical level, by building an "entropy index" that fully accounts for the richness of classroom-level ethnic composition. This index – which takes into account both quantitative and qualitative aspects of the student body – is particularly useful for countries such as Italy, with a large number of different ethnic groups and where ethnic minorities are heavily unequally distributed.

To best of our knowledge, few papers address the topic of diversity; they come mainly from the economics of education, development psychology, and sociology of education. Most of them focus on a kindergarten level before children enter school, whereas others on primary education. Cho (2012) investigates the link between classroom diversity in pre-school and children's academic-related competencies. However, given that this study used only two language categories to define groups of children, it measures a concentration index related to the share of pupils who have a primary language other than the dominant one. In contrast, Gottfried (2016) and Meng (2018) address diversity with a proper indicator, exploiting linguistic membership with more detailed information. Working on a sub-sample of Head Start Program, a survey for kindergarten in the U.S, they find a positive effect of ethnolinguistic diversity on cognitive skills, socio-emotional skills, and behavior in classrooms. Finally, using country of origin as a proxy of diversity, a line of research investigates the effect of ethnic share and diversity on academic performance (Veerman, 2015; Veerman, van de Werfhorst, & Dronkers, 2013; Veerman & Dronkers, 2016). Nevertheless, all these strategies do not account for unobserved fixed school characteristics and unequal sorting within schools

From our review of the empirical contributions, it comes out a research gap, since there is not study that analyzed student socio-emotional skills and behavior as an outcome along with academic competencies, by distinguishing between ethnic concentration and ethnic diversity. Hence, we decide to analyze outcomes such as being bullied, doing bullyism, social integration, extrinsic motivation, and academic competences in mathematics and reading. We argue that classroom ethnic share affects not only performance but also the extent to which students integrate, influencing the learning environment in different ways. A higher share of non-native students might lead to reduced communication and hostility, resulting in misconduct or increased social distance among students with detrimental byproducts on academic competences. However, we consider also that both ethnic share and ethnic diversity may have a different effect, depending on their size in the classroom. Indeed, the higher ethnic share can reduce inter-ethnic conflict because CMO students have drawn on a large "network" and compensate for the presence of natives students. The same reasoning refers to ethnic diversity because a higher diversity dose may make more granular the ethnic-composition, reducing barriers, and promoting mates communication. To synthetize such perspective we think that ethnic diversity works as a moderating variable increasing the effect of ethnic share because a higher dose of diversity might lead to a sort of Babel classroom, exacerbating the negative aspects of a learning environment ethnically concentrated.

- ✓ Hypothesis 2: If we control for the classroom diversity, the negative association between CMO and outcome increases.

2.3 Non-linear and asymmetric effects

To investigate more in-depth the ethnic density hypothesis Halpern (1993), we test some facets of a peer effect such as non-linearities, thresholds, and asymmetric effects. Our expectation is that higher ethnic share and diversity harm the educational outcomes but asymmetrically over the ethnic background. This asymmetry depends on the classroom ethnic and diversity share.

In Austria, Schneeweis and Winter-Ebmer (2015) account for a quadratic relationship between the share of immigrants and the choice of an academic track after primary school and grade repetition in primary school. Findings support this relationship only when the share of immigrants of those belonging to the same country of origin is adopted as the leading indicator. Also, as the CMO group is large, the negative effect is weaker because of the higher concentration of ethnic groups makes the classroom management and communication easier. Gould and colleagues (2009) find a significant quadratic relation between school ethnic share and the chance that natives pass the matriculation exam for the college in the US. It comes out with a non-linear and convex pattern. According to their calculations, an increase of the immigrant concentration from 0 to 10% reduces the mentioned chance by 4.2 percentage points. Tonello (2016) finds a negative and quadratic specification on the native performance in Italy, a finding similar to what was found by Szulkin and Jonsson (2007) in Sweden. They adopt ex-ante bins of ethnic share on academic performance, and they identify a sort of cut-off point about the 40% of school ethnic share. Above this threshold, the effect is negative and sizeable.

In contrast, Virdia (2018) identifies a threshold of around 20% in Italian schools with multilevel models, and he finds a negative effect between CMO and academic performance. Besides, this effect is stronger above the cut-off point. Finally, Andersen and Thomson (2011) systematically investigate the presence of a threshold and the presence of heterogeneous or asymmetric effects between CMO and Danish students. Relying on ex-ante bins, they find the threshold around 50% and that the negative effect is stronger among CMO. Also, they argue that the presence of non-linearity is a litmus test about the presence of asymmetric effects among students according to some characteristics. In Norwegian upper secondary education, Fekjaer and Birkelund (2007) highlight the presence of heterogeneous effects across groups of students: indeed, they find limited positive effects of ethnic concentration on achievement and the probability of university enrolment for native students and second-generation immigrants, but negative effects on achievement for first-generation immigrants.

Overall, these findings suggest that the negative effect on native are higher at lower levels of CMO concentration. This is due to the school's ability to absorb immigrants or that the integration of CMO is easier in context with a higher share of similar peers. This is in line with the work of Geven, Kalmijn, & Van Tubergen (2016) in Sweden and the Netherlands, where a higher share of co-ethnic concentration is associated with lowering misconduct behavior.

Looking at these contributions, it comes out that the presence of non-linear patterns and the asymmetric effects advocate for policy intervention. However, such linearities and asymmetric effect might be a mere ecological fallacy when the empirical strategy does not account for unobserved school characteristics (Paloyo, 2020) or a strategic sorting of students, as discussed in the previous section. In addition, the search for the threshold is biased when the detection cut-off relies on handwork because the researcher builds ex-ante model specifications. Now, the frontier of research approaches toward the LASSO procedure to detect regularities and discontinues between two or more variables (Backus & Peng, 2019) at n-polynomial specification controlling for individual and contextual characteristics. Relying on the mentioned perspectives about the non-linear patterns, asymmetric effects, and cut-off points, we develop the following hypotheses:

- ✓ Hypothesis 3: The effect of classroom ethnic is non-linear drawing on a concave curve
- ✓ Hypothesis 4: The effect of classroom ethnic share might be asymmetric over students' ethnic background.
- ✓ Hypothesis 5: There is a cut-off point in the relation between ethnic share and outcomes of interest.

2.4 Challenges to the identification of the effects of ethnic concentration

A credible identification of the causal effect of ethnic concentration on students' outcomes based on observational data has to tackle many challenges. The first issue refers to the fact that parents and principals might adopt strategic behavior, which leads to non-random sorting of students across schools and classrooms. Indeed, parents may strategically enroll their children in schools with specific characteristics, for instance, a low share of foreign children, better socio-economic composition, higher resources. This would make the identification of the ethnic concentration based on differences across schools problematic due to unobserved features of the schools (such as resources, geographical position, and teacher-student ratio) and the students attending different schools (such as gender, ethnic origin, and socioeconomic background). Additionally, school principals might rely on three main tools, such as manipulation of classroom composition, class size, and teacher-student matching, to tackle possible unbalances in student body composition across classrooms. Such manipulations might work simultaneously, making arduous a proper investigation of compositional effects also when one relies on within-school estimates. Some of the research already mentioned, especially studies conducted by economists, try to take into account some but not all of these threats to identification at the same time.

Angrist and Lang (2004) address the first two issues by exploiting the abovementioned METCO de-segregation program in the Boston area and non-native to native assignment rule, namely 23 natives for one non-native. In an attempt to account for non-random sorting, Ammenmuller and Piscke (2009) developed a statistical test to identify schools with a random assignment of students across the classrooms in primary education in four European countries.

Table II: Papers addressing endogeneity in Italy

Author/s	Year	Effect Size	Sign	Unit of analysis	Strategy
Tonello	2012	tiny	-	school	Within school across cohorts à la Hoxby
Contini	2013	tiny	-	classroom	School fixed effect and sorting test
Ballatore, Ichino, and Fort	2019	moderate	-	school	IV strategy à la Angrist

In Italy, three main contributions attempt to tackle issues related to the endogeneity of ethnic share about classroom composition (Contini 2013; Tonello 2016) and class size (Ballatore et al. 2019). Using the administrative INVALSI population data, Tonello (2016) exploits cohort variation overtime at the school level, and, in this way, he can take into account time-constant school unobserved characteristics. A negative causal link between ethnic share and academic performance was found. Differently, and more in line with Ammermueller and Pischkes (2009) 's approach, Contini (2013) investigates the effect of ethnic concentration on students' performance in the standardized test at the classroom level, by using school fixed-effects model and retaining only the schools that passed a statistical test reporting evidence of random sorting of children across classrooms. Similarly to Tonello (2016), she identifies a small negative causal effect of the share of CMO on students' competencies in math and Italian. Finally, Ballatore and colleagues (2019) used restricted-use data and exploited an instrumental variable approach to single out a "pure compositional effect" of ethnic share across classrooms. All of these studies find a negative effect, although of varying size.

2.5 Features of the Italian educational system

The Italian educational system has a high level of formal standardization concerning exams, curriculum, and financial budgets. It encompasses a pre-primary education not mandatory (3 to 6), a primary education (6-11), a lower secondary education (11-14), and an upper secondary education (14 to 19). Then, students may apply for tertiary education (19-24).

Historically, the Italian education system relies heavily on a neighborhood criterion for school choice (until secondary education), with available slots for students whose parents work in the school areas. Albeit this regulation partially attenuates the possibilities of strategic parents' choice, it mainly reflects the socioeconomic composition of the catchment areas exacerbating socio-economic and ethnic differences across schools. The enrolment follows an operative year window, with a strict neighborhood preference for students who apply before the end of January. Italian law states that schools have to submit their lists of students for each classroom by June, but in fact, school managers have more time to create their classrooms, and most make them public around the beginning of the school year in September.

Along with this institutional setting, it is worth noting the autonomy of the principals regarding students' assignments to the classroom. The formation of the classroom should be random, but school principals do not always respect such criterion. In general, the school principal has access to the relevant basic information of students, such as past marks of previous educational stages, geographical address, and whether the student is native or non-native to Italy. In some cases, regarding non-native students, the principal will know their citizenship status, and their country of origin and their age upon arrival if available. It is reasonable to think that these characteristics might be used for the assignment of students to the classroom. Indeed, findings support the moderate presence of segregating patterns between and within schools based on both socioeconomic in lower secondary schools (Agasisti & Falzetti, 2017) and ethnic characteristics in primary (Contini, 2013). Besides, principals may manipulate the teacher-students assignment, combining better teachers to the worst students or vice versa, as outlined by Abbiati, Argentin, and Gerosa (2017). In sum, although the Italian education system calls for an equal sorting of students and teachers, principals might sort students and teachers within their schools, creating "ghetto classrooms" according to essential socio-demographic characteristics such as socioeconomic background, ability, and ethnic origin.

In this work, we focus on primary education because it is crucial to a better understanding of the phenomenon in the early stages, as debated above. This choice allows us to block some biases due to the unequal sorting of students and teachers in the Italian education system.

Unequal sorting of the student according to ability, in primary education, is not present because principals do not have previous information on child career as seeing as pre-primary education is not mandatory and does not provide any internal school to the school report. In addition, empirical evidence suggests that, in primary education, unequal sorting of students based on socioeconomic origin is negligible (Agasisti & Falzetti, 2017). That teacher-student assignment is as good as random according to several characteristics of teachers and students (Abbiati, Argentin, and Gerosa, 2017). Hence, we need to deal with the unequal sorting due to the ethnic characteristics, as shown by Contini (2013).

3. Identification strategy

Working on peer effects studies implies a cautious design to tackle theoretical and empirical perils such as the reflection problem (reverse causality), correlated effects, and sorting effects (Payolo, 2020). Our level of interest is the classroom, and we aim at modeling an ideal random formation of the classroom, singling out strategic manipulation of (1) student's characteristics, (2) teacher's characteristics, and (3) class size (Angrist & Lang, 2004). Ideally, we can reach a credible causal effect if the following conditions are satisfied:

A. **Assumption on the reflection issue:**
   1. *It is possible to distinguish endogenous and contextual effects*

B. **Assumption on the correlated effects:**
   2. *No unobserved characteristics flaw the identification*

C. **Assumption on the sorting effects:**
   1. *Sorting of CMO to schools and classroom is random*
   2. *CMO assignment to the classroom is random*
   3. *Teacher assignment is random*
   4. *Classroom size is random*

The reflection problem **(A)** is always present when the behavior of agents introduces perfect collinearity between the expected mean outcome of the group and its mean characteristics (Patacchini, Rainone & Zenou, 2017). By definition, it is almost impossible to differentiate between the effect of peers' choice of effort (endogenous effects) and peers' characteristics (contextual effects) that have an impact on their choice of effort. Intuitively, there is a simultaneity between the mean outcome of the group and its mean characteristics. To date, three solutions come out (Paloyo, 2020): (1) inclusion of the endogenous peer characteristics at the net of student *j*, (2) omission of the endogenous peer characteristics, (3) exploiting idiosyncratic variation of network data.

In our setting, we want to estimate a fixed characteristic such as ethnic share on a large array of student's outcomes, such as extrinsic motivation, and option 3 is not available. A key question arises: Should we include classroom motivation in our analysis? The literature is quite divergent. On the one hand, Moffit (2001) suggest estimating the impact of class composition effects without attempting to separate the results due to peer achievement from other related impacts due to peer characteristics. According to him, their joint action is still of interest in public policy.
On the other hand, Tatsi (2020) suggests not to omit it as a rule of thumb, including endogeneity and relying on a random formation of class size. We adopt the standard practice of option 2 (Ballatore and colleagues 2019; Contini, 2013; Moffit, 2001) because the random formation of the classroom size is rare, and the omission does not lead to confounding in our case. Indeed, Tatsi (2020) debates the endogeneity between non-fixed characteristics such as classroom performance and motivation. In contrast, our treatments are share of CMO and language index, namely fixed ones.

Correlated effects **(B)** suggest that the presence of peer-group unobservable characteristics might flaw peer effects estimation; for example, students exposed to common factors such as club membership or a pleasant learning environment. To deal with this issue, we control for several variables such as gender, socioeconomic background, and ability at the individual and classroom level (Anelli, Shih, & Wlliams 2017)

In our view, the critical issue of this work is the selection bias – due to students' sorting – composed of two patterns, one between schools and the other within schools **(C 1 & 2)**. Regarding the between school sorting, parents might strategically choose one school or change neighborhoods because the quality is higher or CMO stock is lower than in other schools. Also, teachers might opt for schools with low CMO students because they are less problematic. Following this reasoning, we rely on school-fixed effects models to account for teaching differences at school level derived from strategic parents' choices such as a flight to the quality phenomenon (Barbieri, Rossetti, & Sestito 2010). Our identification strategy is immune to a neighborhood bias because it relies on a within-school variation. The fundamental intuition is that schools reflect neighborhood composition, and in turn, distinct ethnic stocks. The use of school fixed effects allows us to weight differently observed and unobserved school characteristics.

Nonetheless, it is possible to adopt a school fixed effect to provide estimates of the causal effects of class composition (ethnic peers) only if children are randomly assigned to the classroom within schools (Ammermueller and Pischke 2009). Such an assumption does not find strong evidence in the Italian setting because school principals might optimize the assignment of students to the classroom according to socio-demographic characteristics. To address this issue, we apply a statistical test to identify schools that use a random sorting of students. In the education field, there are two principal ways of testing random sorting within schools: (1) "empty" regression models with classroom fixed effect for each school (Horvath, 2017); or (2) adoption of a sorting test using Fisher's exact or Chi-squared test.

Finally, the last concern regards the random assignment of students according to the ethnolinguistic groups. Indeed, principals might discriminate according to ethnic characteristics but also to the language spoken at home using the country of origin as a proxy. To have a picture of this phenomenon, we also run a sorting test on language information.

We decided to opt for the Fisher's exact test, for several reasons. First, in our context, it is more efficient than the regression model with classroom fixed effect. It is also more conservative than the standard chi-square test since it calculates the exact p-value. Third, it works well with the small expected frequencies due to low non-native stock in some schools. Unfortunately, the adoption of the Fisher test is too demanding when we use language because it shows too many answers. In this case, we run a Chi-squared test with a more robust p-value threshold.

The assumption of random assignment of CMO features is tested at the school level. This test implies independence between the non-native background or language and the classroom the student is assigned to:

$$(1) \quad H_0: P_{CMO,classroom \mid school} = P_{CMO \mid school} * P_{classroom \mid school}$$

$$(2) \quad H_0: P_{Language,classroom \mid school} = P_{Language \mid school} * P_{classroom \mid school}$$

Where $P_{CMO \,\&\, Language\, ,classroom \mid school}$ is the joint probability that a randomly chosen child from a given school is non-native or speaks one language and is assigned to a classroom c_i, $P_{CMO \,\&\, Language \mid school}$ is the overall proportion of migrants in the school, and $P_{classroom \mid school}$ is the proportion of children in classroom m c_i. Following Contini (2013), we considered a prudential significance level of 0.10 for the Fisher and of 0.15 for the chi-squared. It comes out that 70% out of 6620 Italian primary schools do random assignments.

Table III: Results of sorting tests

OVERALL SORTING TEST		
Random	4,609	70 %
Non-random	2,011	30 %
Total	6,620	100

		SORTING ON LANGUAGE		
		Random	*Non-random*	*Total*
SORTING ON ETHNIC BACKGROUND	*Random*	4,609	480	5,089
	Non-random	1,132	399	1,531
	Total	5,741	879	6,620

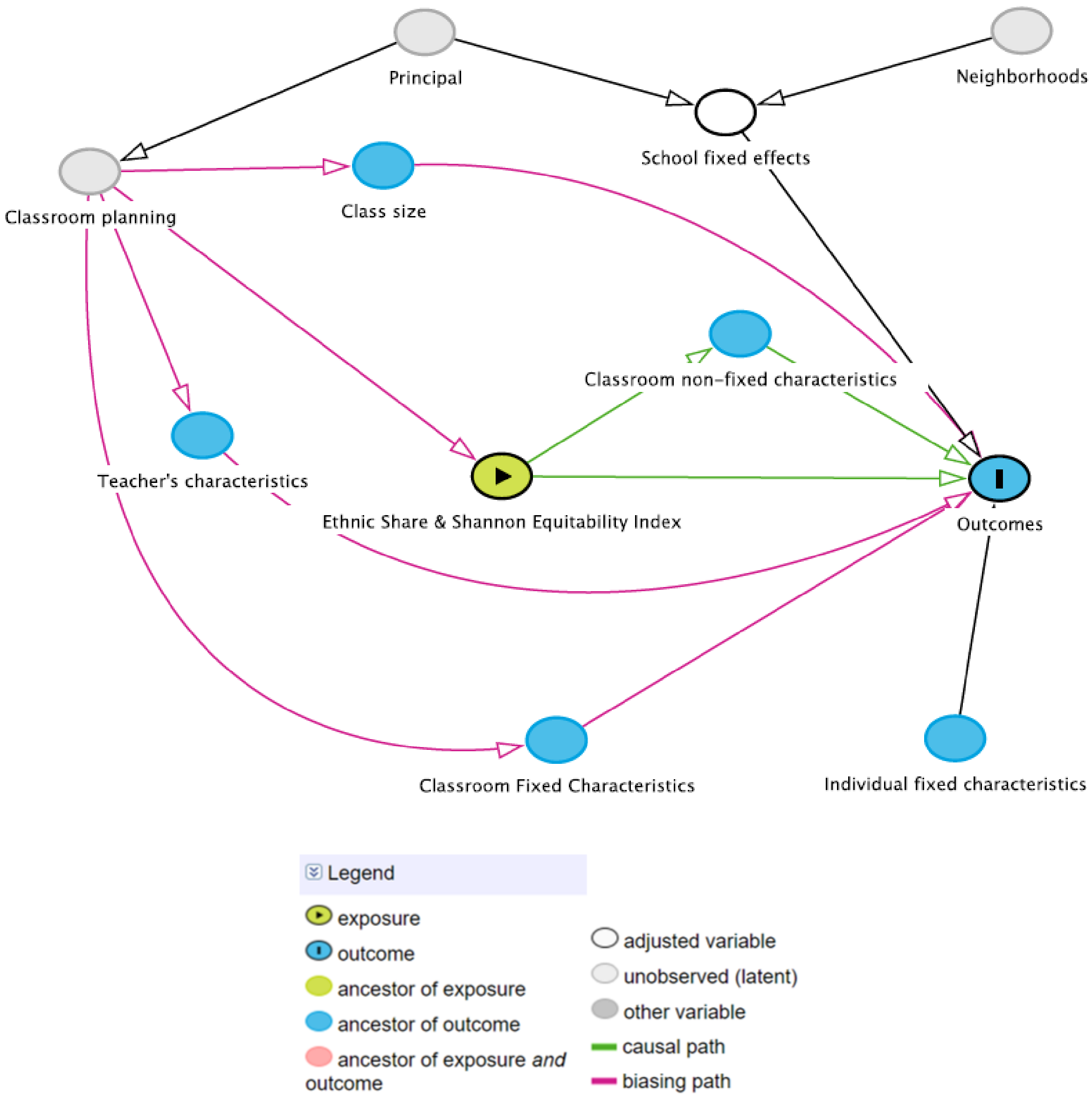
At this point, previous works restricted the estimation of ethnic concentration effects to the subsample of schools with random assignment of children across classrooms (Ammermueller and Pischkes, 2009; Contini 2013). While this strategy allows achieving unbiased estimates if the assumptions of the model are met, it suffers from issues of external validity, since the selected schools might differ in many respects from those excluded from the estimation.

To mitigate this problem and provide estimates with higher external validity, we applied a weighting procedure that allowed us to re-proportion the results obtained on our analytical sample to the whole population of schools. More specifically, we estimate weights at the school level using an entropy balancing procedure (Hainmueller, 2012). Entropy balancing is an algorithm that weighs the observations in the treatment group to balance mean, variance, and skewness of a set of covariates concerning the control group. In our application, the treatment group is constituted by schools that passed the Fischer's test, the control group is all the other schools, and the covariates used to create the weights are share of female, the share of low educated, share of low-SES, the share of late students, CMB share, diversity index, average performance in the mathematic and national language. The weights obtained from this procedure, then, can be used with standard estimators to identify the treatment effect of interest.

The last issues are teacher assignment and class size **(C 3 & 4)** assumption. Indeed, while this procedure can credibly tackle the non-random sorting of students, non-random sorting might happen on the teacher side as well. Assuming that ethnic concentration hurts students' outcomes, if there is compensatory allocation by school principals and high-quality teachers are allocated more often to classrooms with high ethnic share, this will lead to underestimating peer effects. The negative consequences of peer effects will be instead overestimated if lower-quality teachers were assigned to the more problematic classrooms. Nevertheless, even if plausible on a theoretical level, existing empirical findings reported by Abbiati, Argentin & Gerosa (2017) indicate that there is a random matching of teachers' and students' characteristics in Italian primary schools. This random matching allows us to exclude that teachers' characteristics bias identification of peer effect in our context, and – together with the other pieces of our identification strategy – it represents an improvement compared to the existing literature.

Finally, principals might manipulate class size, accommodating the classroom composition. Intuitively, if CMO is too high, principals might increase the share of natives or simply transfer some CMO to another classroom. Exploiting an idiosyncratic variation on CMO late enrolment, Ballatore and colleagues (2019) identify this phenomenon. However, they recognize in their discussion that principals behave as compensatory room and that their strategy singles out a specific sub-effect in contrast to our work where we identify an overall effect at gross of the principal behavior. Indeed, the result found from Ballatore and colleagues (2019) is a local effect for those school principals that cannot manipulate the class-size.

Figure I: Directed Acyclic Graph of main identification



In figure I, we briefly resume our identification strategy drawing on a within schools design. The DAG shows how classroom planning and principal characteristics are unobserved, leading to biasing paths. Once established that the classroom Ethnic share and Shannon index are as good as random, to estimate the total effect of ethnic share or Shannon equitability index, it is sufficient to adjust for class-size, teacher's characteristics, and classroom fixed characteristics. We block the biasing paths, controlling for class-size and classroom fixed characteristics such as share of females. Regarding the teachers, we do not have information, but we know that there is a random classroom assignment (Abbiati et al. 2017). In addition, unobserved principals and neighborhood characteristics might bias our estimation, but we block them with school fixed effects. Finally, it is worth noting that classroom non-fixed characteristics work as a mediator in our setting. Since it does not open biasing paths and mediation analysis is not a research goal, we do not investigate it in depth.

4. Analytical strategy

4.1 Data

Our empirical analysis uses data collected by the National Institute for the Evaluation of the Italian School System on the *whole population* of students enrolled in the 5th grade (primary education) in 2014–15 (INVALSI 2015). Our final analytical sample comprises 206.443 students across 4609 schools. We merge information from different sources, namely: administrative information from schools, performance on standardized tests in Italian and mathematics; and information from student questionnaires. We rely on this data for three reasons. Firstly, it is the only available information on outcomes of interest. Second, it covers a critical phase of a student's development, in particular just before the transition to lower secondary education. Finally, there is a higher share of non-native students in primary education than in lower and upper education, due to the increased presence of immigrants with children in the country in recent years and to the absence of student dispersion at this school stage.

4.2 Variables

Our dependent variables measuring socio-emotional skills are four quantitative indexes that were built by relying on Cronbach's alpha to assess their internal reliability. The average alpha score across the four indexes is 0.85 (ranging from a minimum of 71 to a maximum of 84), which is considered an acceptable and good internal reliability in the psychometric literature. Using factor analysis as a double-check confirms the good reliability of the indexes. The first index is "Being Bullied," which indicates the extent to which other students bully the respondent, whereas the second one ("Bullying"), measures the degree to which the student practices bullyism. The index of "Social Integration" measures how much the students perceive themselves as integrated (versus isolated) in the scholastic environment. Finally, the "Extrinsic Motivation" index measures the extent to which the student behavior is driven by external rewards such as money, fame, grades, and praise. This is usually considered a non-desirable trait in the educational psychology literature since it is related to less successful outcomes compared to intrinsic motivation (Walker, Green, & Mansell, 2006).

Table IV: Scheme of dependent variables with alpha

Index	Original variables	Alpha
Being bullied	1. How many times do your mates make fun of you? 2. How many times do your mates insult you? 3. How many times do your mates isolate you? 4. How many times do your mates beat you?	0.77
Bullying	1. How many times do you make fun of your mates? 2. How many times do you insult your mates? 3. How many times do you isolate your mates? 4. How many times do you beat your mates?	0.71
Integration	1. How many mates interact with you? 2. With how many of your mates, do you feel good? 3. How many of your mates, do you help if in trouble? 4. How many of your mates, do you consider friends?	0.74
Extrinsic Motivation	1. For me, it is important to show others that I am a good student 2. For me, it is important to show others that I go well in the assessments 3. For me, it is important to show others that I go well at schools 4. For me, it is important to show others that I look more intelligent than my mates	0.84
Performance in mathematics	1. Standardized Test corrected with IRT (provided by INVALSI	
Performance in language	Standardized Test corrected with IRT (provided by INVALSI	

We have two independent variables of interest. The first one is "Ethnic concentration," and the second is "Ethnic diversity." The first one is constructed as the proportion of students with a migration background in the classroom, including in the numerator, both first and second-generation students.[18] As denominator, we used the total number of students in the classroom provided in the INVALSI data.

[18] Indeed, previous studies show they perform very similarly to Italians with both parents born in Italy.

To construct the second independent variable of interest, namely "ethnic diversity," we take advantage of the information provided by the INVALSI's questionnaire in which students report the language they regularly speak at home (sixteen possible categories, see Table V). To this extent, we follow the work of Chiswick and Miller (2005), who outlines how linguistic dissonance might be helpful to investigate non-native effects in greater depth.

Table V: Share of the language spoken at home by the ethnolinguistic group.

	Freq	**%**
Italian	189205	91.65
Arabic	2646	1.28
Albanese	2500	1.21
Romanian	2444	1.18
English	1472	0.71
Spanish	1365	0.66
Chinese	950	0.46
French	475	0.23
Hindi	324	0.16
German	253	0.12
Portuguese	211	0.10
Croatian	127	0.06
Ladin	78	0.04
Greek	53	0.03
Slovenian	67	0.03
Other	4273	2.07
Total	206443	100

It is important to notice that our approach allows us to take into consideration both the variety of ethnolinguistic groups and the specific incidence of each group, which could vary widely across classrooms. In the existing literature, several diversity indices are present with different peculiarities. We opted for the Shannon-Equitability Diversity Index, which can reveal complexities due to small shares of contextual characteristics. The index is computed as follows:

$$(1) \qquad \frac{\left(-\sum_{i=1}^{i} s_i * log(s_i)\right)}{MAX_{S_I}}$$

Where s is the share of i ethnolinguistic group in a classroom and MAX_{s_i} is the maximum number of categories in a classroom. Each share multiplies itself by its logarithm transformation. This formula reduces the weight of higher shares and increases the weights of lower shares. The MAX_{s_i} correction makes the index independent from the categories avoiding a mechanical correlation between linguistic categories and diversity index.

In our statistical models, we used some individual and classroom characteristics as control variables. We use gender to identify boys and girls. To proxy ethnic origin, we use the language spoken at home, creating a dummy variable: native and non-native (Italian language spoken at home or not). Another control variable is whether students are late or early students. To capture this, we use the year of birth as a proxy, taking into account the age students should have, given Italian school rules. We created a variable distinguishing early, regular, or late students. We control at the beginning education and jobs of parents adopting a dominance criterion. The former distinguishing between primary, secondary, and tertiary education. Parents' occupation is measured six categories by recoding a scale provided in Campodifiori, Figura, Papini, and Ricci (2010). Then, we adopt the ESCS indicator provided by INVALSI. Besides, we rely on a measure of academic competence, namely the average of teachers' marks in mathematics and Italian obtained in the mid-term report in February,[19] two-three months before students sit to take the INVALSI standardized test. Control variables at the classroom level are the same used at the individual level. For the sake of clarity, jobs and education of parents are recoded as a percentage of low education and low jobs. Finally, we include class-size.

4.3 Methods

Our empirical strategy relies on linear regression school fixed-effects models, which are estimated on the re-weighted sample of the selected schools that passed the statistical test of random sorting of children across classrooms. All the models adjust for several individual and classroom characteristics presented in the previous section.

A critical point is the introduction of an ability proxy to control for academic performance. The fact that the students' performance in the standardized tests refers to the same days as the outcomes of interest might pave the way to reverse causality. Therefore, our strategy is to use the teachers' assigned marks attributed before the INVALSI data collection and to run

[19] The results are similar if we control for the two separate variables.

two analyses with and without such proxy of ability. The models in their more complete form are specified as follows:

$$(2) \quad Y_{ic} = \alpha_{ic} + \mu_s + \beta Non - Native\ Indicator_{ic} + \beta Ability_{ic} + \beta Z_i + \beta W_{ic} + \varepsilon_{ic}$$

Where μ_s is school fixed effect, βZ_i is a vector of individual characteristics, and βW_{ic} a vector of classroom characteristics. The primary independent variables are non-native indicators, such as ethnic share and diversity index. We run six model specifications:

- ✓ Linear specification for Ethnic Share and Diversity (included separately)
- ✓ Liner specification for Ethnic share and diversity (included together)
- ✓ Non-linear squared specifications as a check
- ✓ Linear specification with the interaction between Ethnic share and diversity. Here, we convert the diversity index into three groups (low, medium, and high)
- ✓ Interaction between Ethnic share and Diversity with Ethnic background.
- ✓ Threshold regressions & Piecewise analysis for Ethnic share

To accomplish the identification of possible discontinuities, we focus on ethnic share because the indicator is more informative and directly interpretable than ethnic diversity. In line with a few previous works (Ong & De Witte, 2013), our objective is to see whether the ethnic concentration begins or ends to exert negative effects on students' outcomes once the share of CMO exceeds a given threshold. We applied the threshold model, which relies on a procedure similar to LASSO to allow coefficients to differ across regions that are identified by a given variable being above or below one or more threshold values (Hansen 2000; 2011). These models – which are usually employed in macroeconomic time series – are good alternatives to linear models for capturing abrupt breaks or asymmetries in a quantitative variable in which the specific threshold is unknown a priori.[20] Compared to most previous works that found the threshold based on bivariate analysis (Backus & Peng, 2019), we used a threshold regression that makes it possible to find a break adjusting for individual and classroom characteristics.

[20] This contrasts with segmented or piecewise regression models in which the threshold is pre-determined by the researcher.

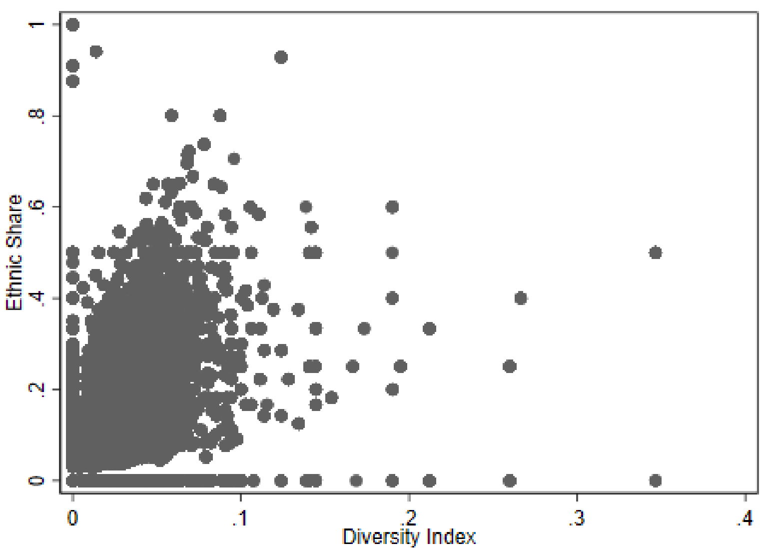
Given that this procedure is extremely time-intensive, we run it on a random sample of schools belonging to our analytical sample.[21] Finally, we run a piecewise analysis to test an intercept or slope break at one specific threshold. We choose 30% of CMO across classrooms since Ministery of Education has identified it as an ideal threshold.

5. Empirical findings

5.1 Descriptive statistics

In Figure II, we show that ethnic concentration and classroom diversity are two close but distinct concepts: they are positively correlated, but their linear correlation is far from perfect (0.52). This result finds support even with a lowess function.

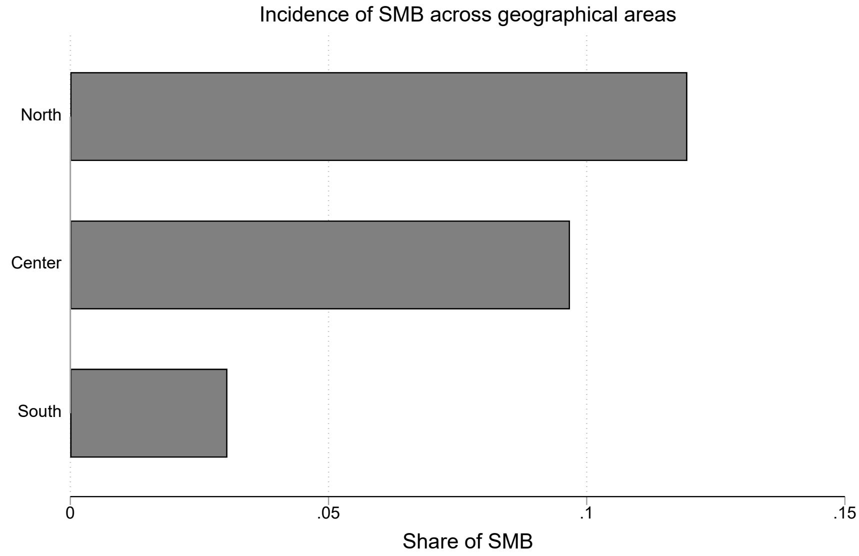
Figure II: Scatterplot between Ethnic Share and Diversity Index (r = 0,52)



In Figure III, we give empirical evidence that CMOs are uneven distributed across Italy. Indeed, North of Italy shows a higher share of CMOs compared to the center, and the south.

[21] We randomly selected 10% of schools from our analytical sample according to a vector of school characteristics such as socioeconomic background, share of girls, share of non-natives, share of dropout, mean of ability, and mean of ethnic diversity.

Figure III: Incidence of Student with migration background across geographical areas



In figure IV and V, we observe the distribution of ethnic share and diversity across the ethnic background. It is worth noting that it is a characteristic of a student matched with a classroom one. It comes out the presence of a skewed distribution that reflects the uneven distribution across classrooms.

Figure IV: Distribution of Ethnic share across the ethnic background

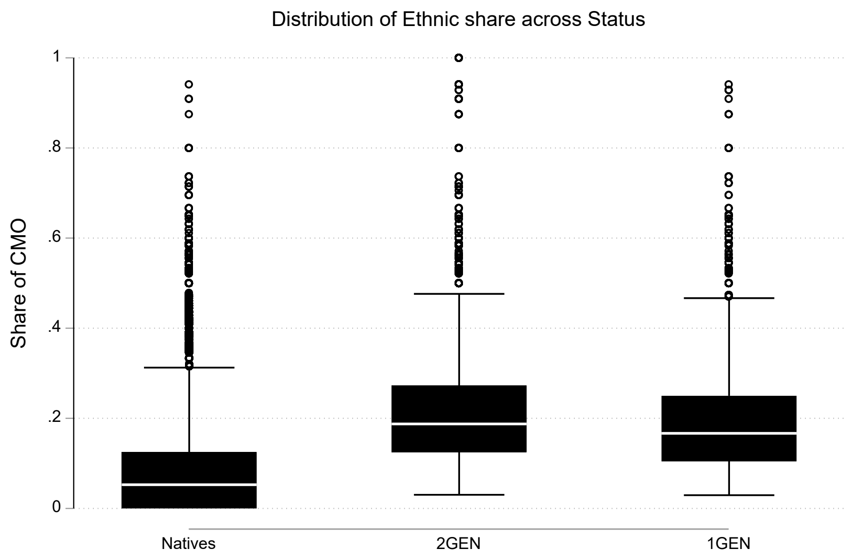
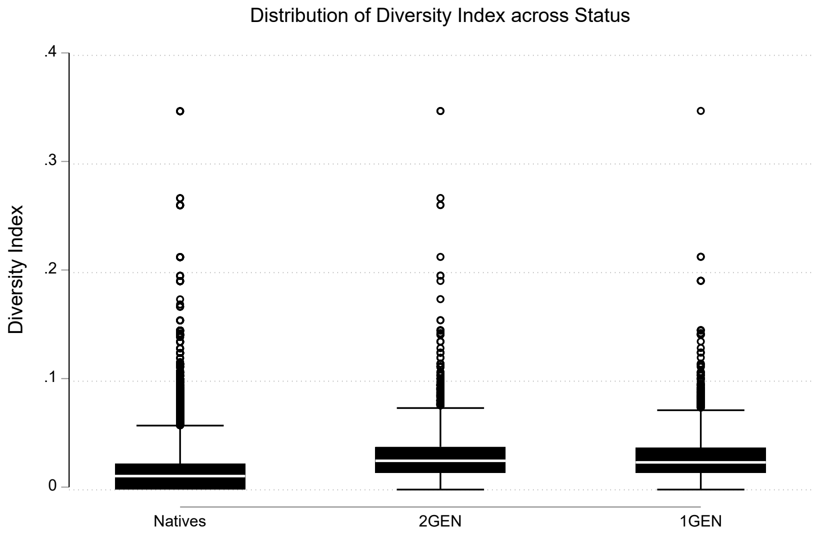


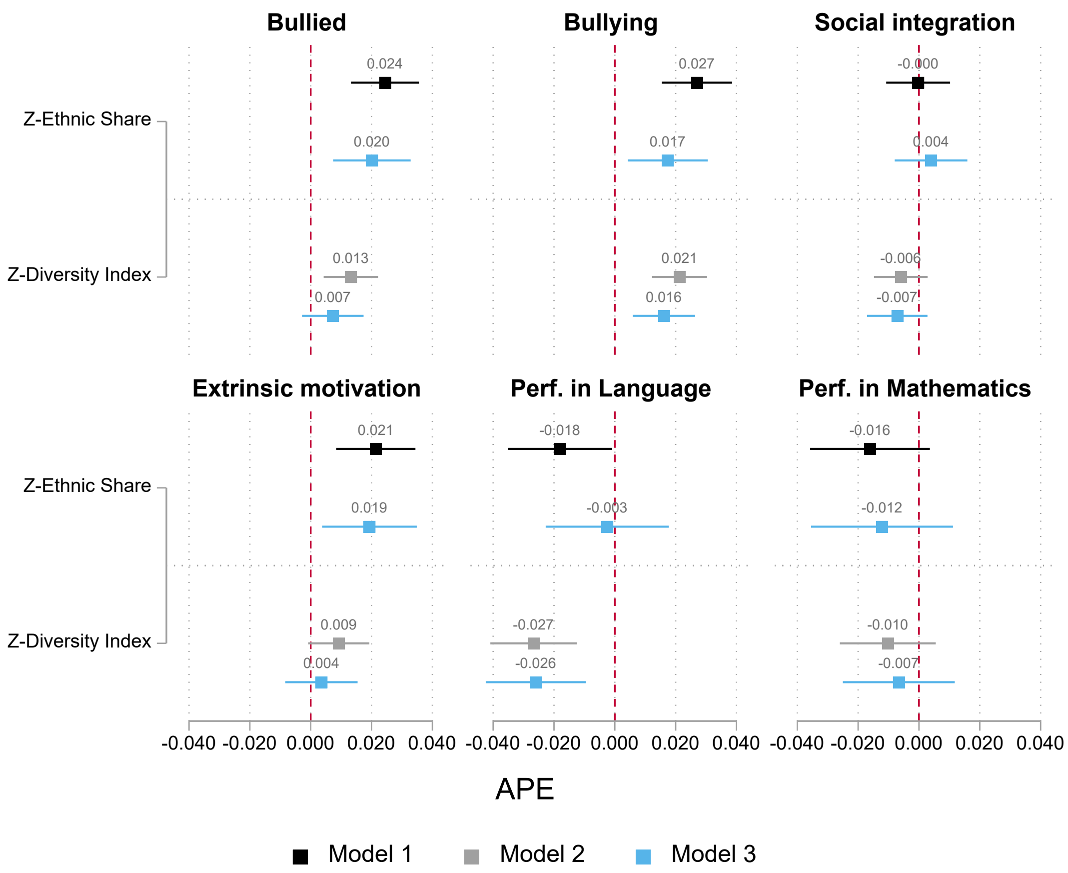
Figure V: Distribution of Diversity index across the ethnic background



5.2 Main findings

The overall findings show a tiny effect of ethnic share and diversity on cognitive and non-cognitive skills. Focusing on the first model specification (Figure VI), we find that H1 is not rejected. Ethnic share affects the exposure and perpetration of antisocial behavior since it is related to higher chances of being bullied and of doing bullyism within the classroom. We interpret these findings as proof that ethnic presence creates conflict within the classroom. Besides, we do not see evidence on social integration, but we document a positive effect of ethnic share on extrinsic motivation. Similarly, ethnic diversity matters as ethnic share, positively affecting being bullied, doing bullyism, but non-extrinsic motivation. However, estimates change when we jointly include ethnic share and diversity in the model. Ethnic share is still significant for bullied, bullying, and extrinsic motivation, whereas diversity matter just for bullying. In contrast, when looked at the effect on academic language performance, only diversity matters both in the alone and joint specification. It comes out that fragmentation of classrooms in distinct ethnolinguistic groups leads to a penalty on language skills.

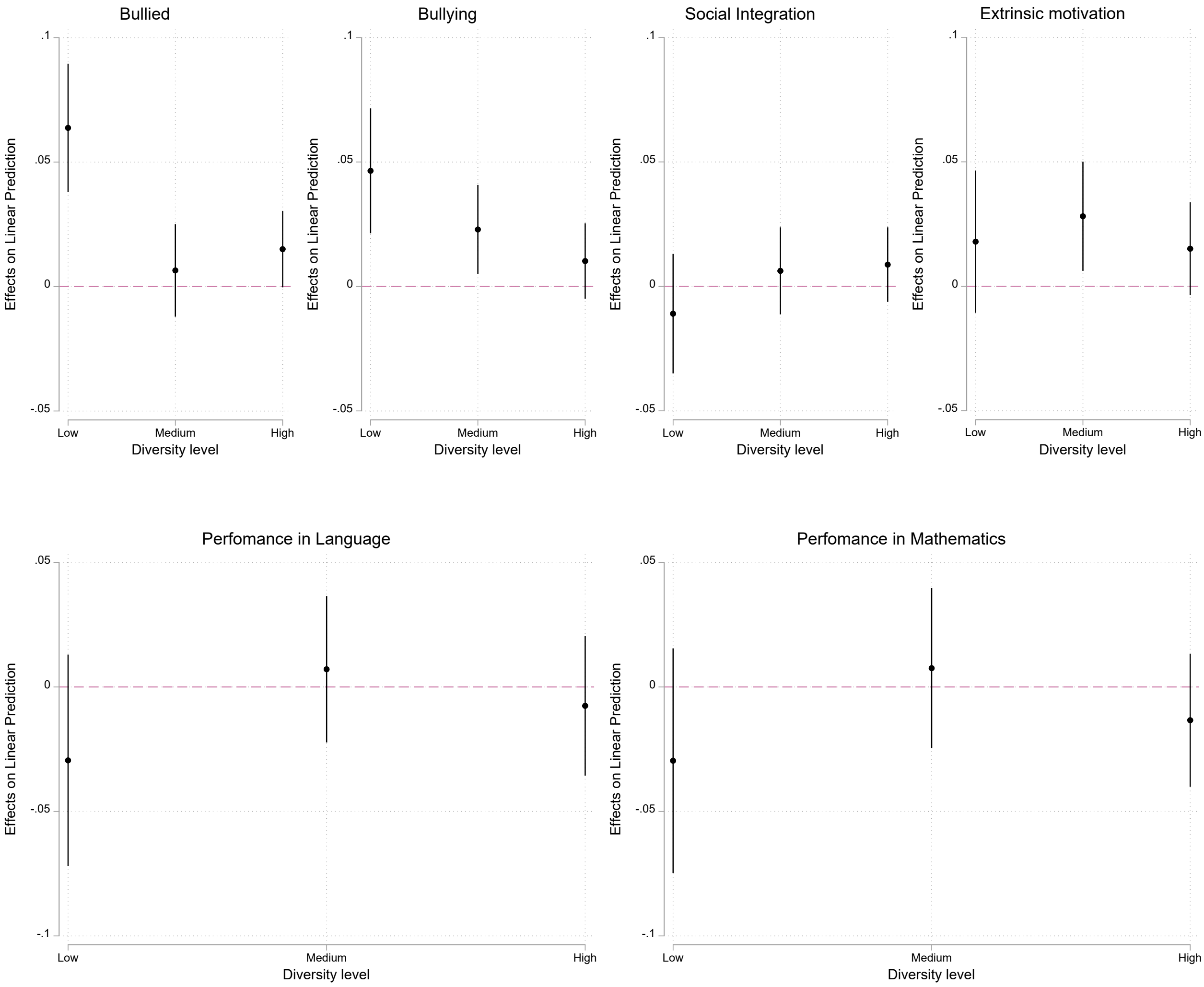
Figure VI: Average marginal effects (and 95% confidence intervals) of Ethnic share and Diversity Index across three model specifications.



Note: Model 1 includes only Ethnic share, Model 2 only Diversity index, Model 3, both independent variables. All the models adjust for the basic control variables and school fixed effects.

Now, we analyze more in-depth a possible interaction between ethnic share and diversity, testing H2 (Figure VII). We find partial confirmation of our hypothesis because the moderating effect is significantly present only for a low dose of diversity. We think that diversity is a granular increasing factor. One classroom with high ethnic share but low diversity might lead to a sort of ethnolinguistic polarization - Italians Vs. Others - and possible conflict in the classroom. This interpretation finds support on the clear trend of bullying and social integration, even if the last one shows no significant effects at all. A key message is that a higher dose of diversity reduces possible conflict, favoring integrations. In contrast, we do not find a moderating role of diversity regarding academic competences.

Figure VII: Interaction between Ethnic share (standardized values) and Diversity Index (three categories: low, medium, and high), School FE, (95% Conf. Int. Reported)



So far, we have imposed that the two indicators of ethnic concentration and diversity have a linear effect on the outcomes; this means that we assumed that one-unit increase on these variables would lead to the same variation in the outcomes irrespective of the part of the distribution in which this variation is considered. To relax this possibly strong assumption, we test (H3) if the effects of interest are substantially modified when using non-linear specifications (Figures VIII & IX). Indeed, we do not see strong evidence supporting non-linear patterns adopting quadratic models except for diversity on bullying and performance in language. After the cubic and quartic specification check, we can confirm the presence of a quadratic trend both for bullying and language, confirming the previously baseline findings where diversity penalizes language competences.

Figure VIII: Quadratic non-linear effects of Ethnic share (standardized values), School FE, (95% Conf. Int. Reported)

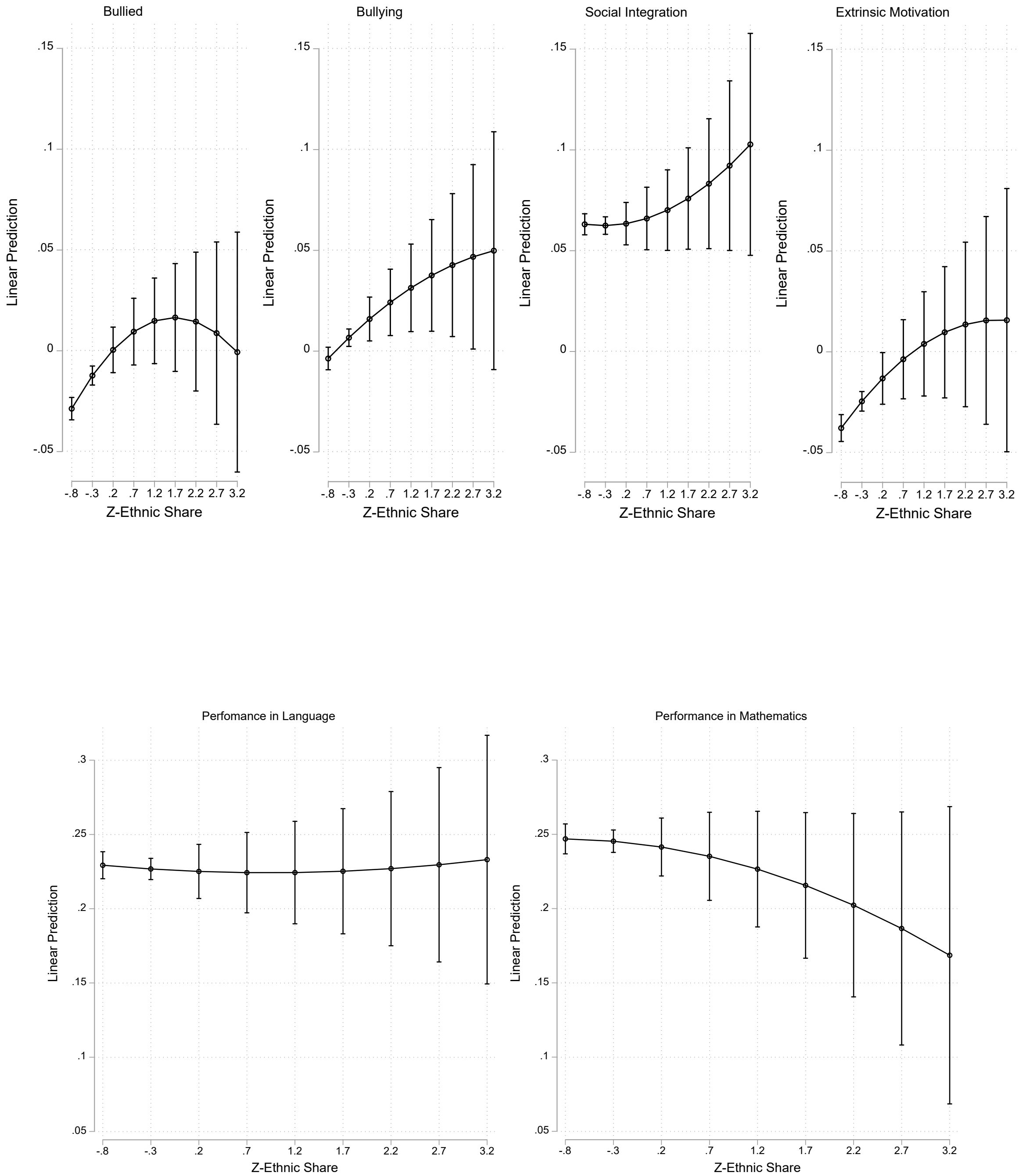
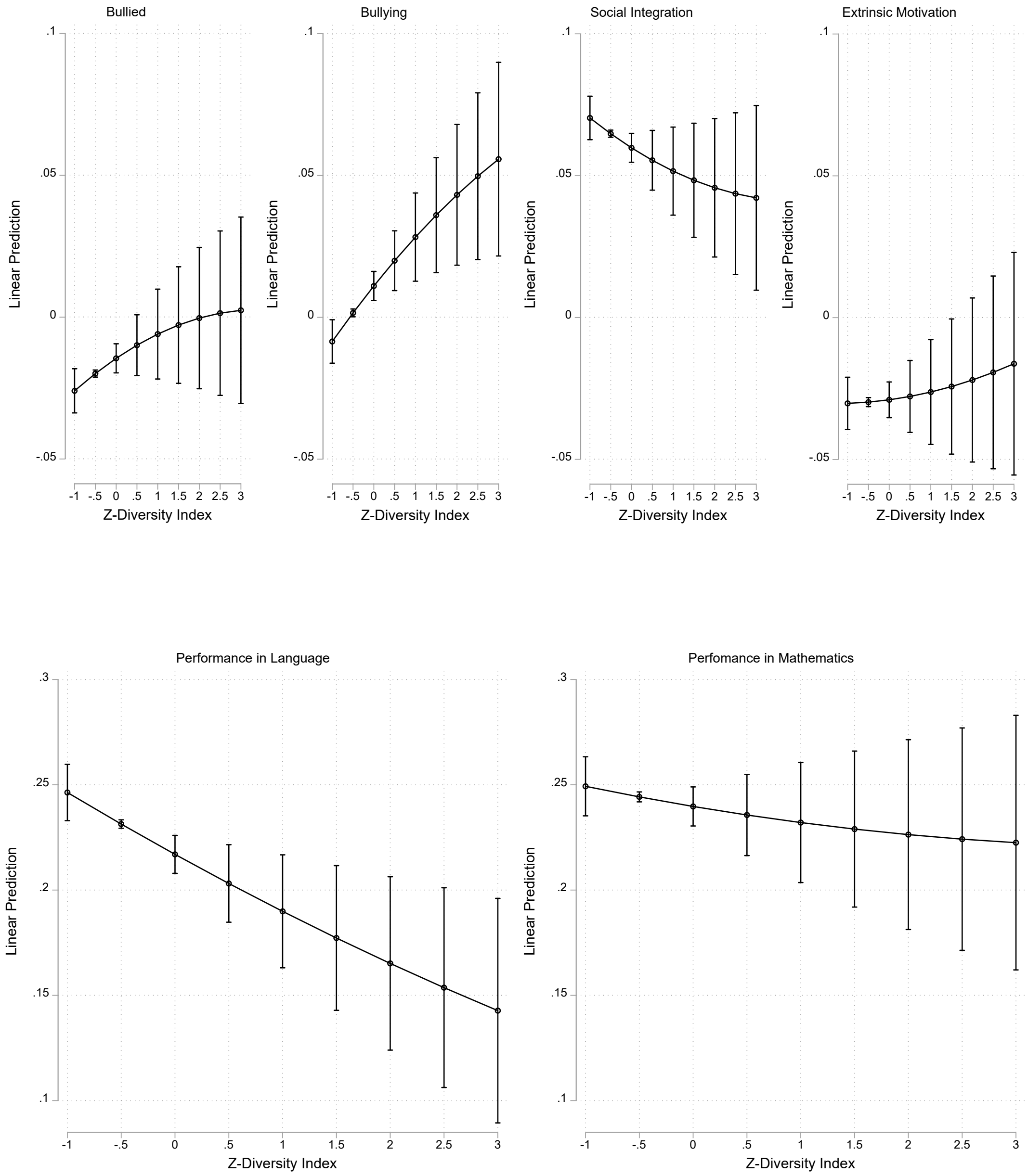


Figure IX: Quadratic non-linear effects of Diversity Index (standardized values), School FE, (95% Conf. Int. Reported)



Finally, we check (H4) our estimates are heterogeneous or asymmetric (Figure X & XI) in the words of Andersen and Thomesn (2011). We observe a penalty but asymmetric for each outcome. Looking at ethnic share, natives are more likely to being bullied and doing bullyism. Also, they show a positive effect on extrinsic motivation but lower compared to the I generation CMO. The effect on bullied and bullying is still present when we look at the impact of diversity. When we investigate the effect on academic performances, there is a

CMO penalty on language due to diversity. Both I & II generation deal with a loss in terms of competences.

Figure X: Average marginal effects, the heterogeneous effect of Ethnic share over ethnic background (Natives, First Generation, and Second Generation, School FE, (95% Conf. Int. Reported)

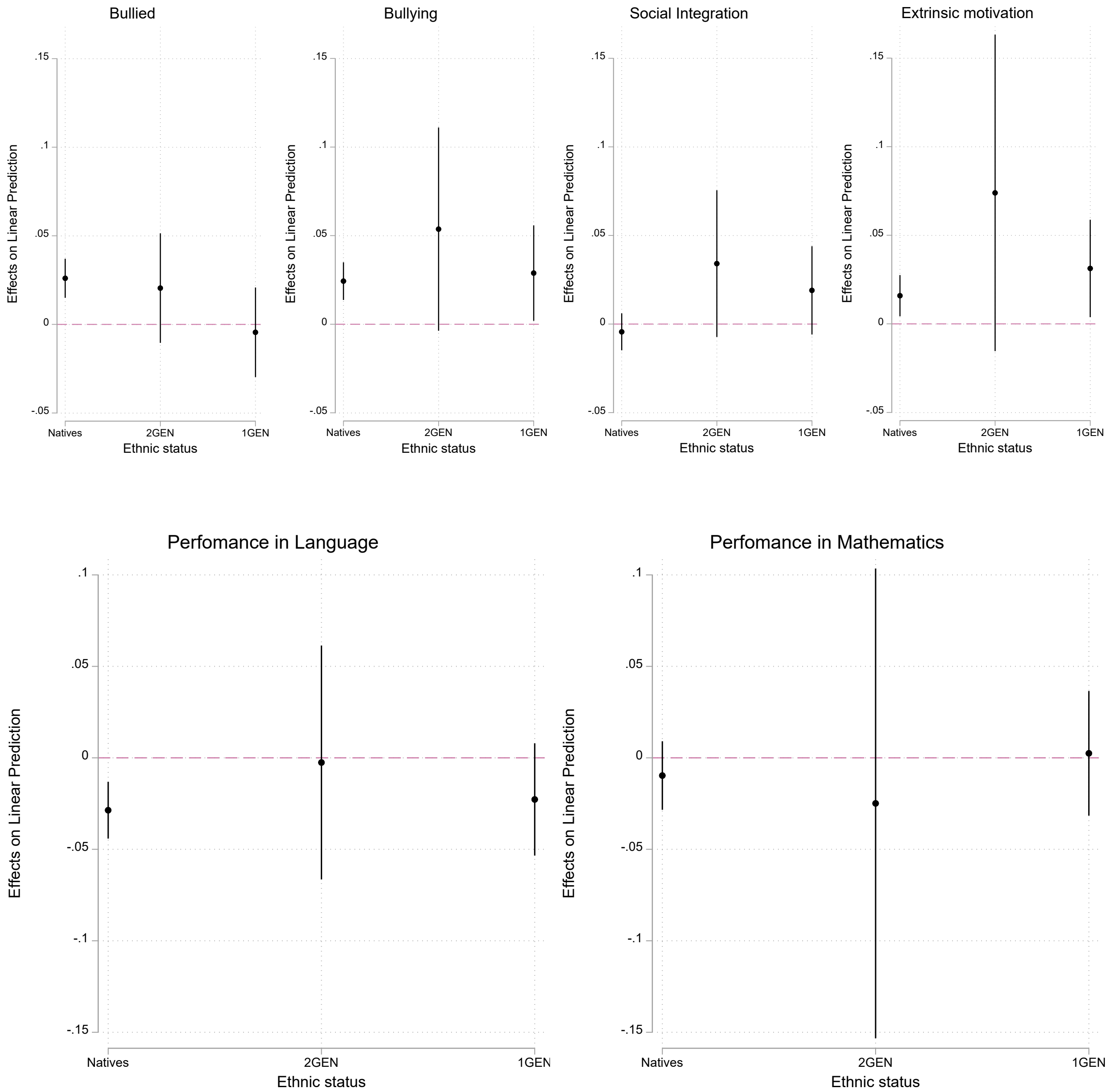
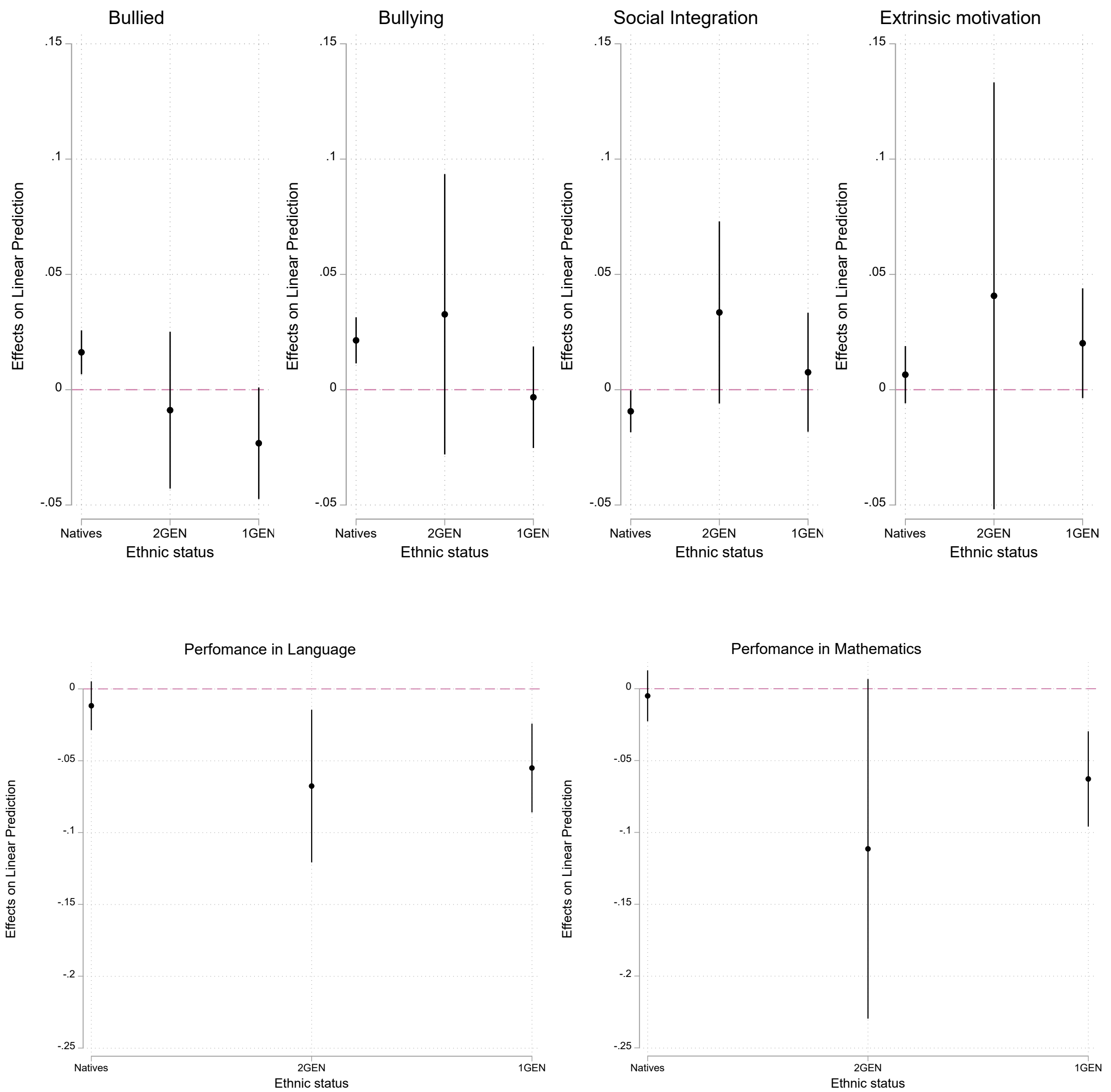


Figure XI: Average marginal effects, Asymmetric effect of Diversity Index over ethnic background (Natives, First Generation, and Second Generation, School FE, (95% Conf. Int. Reported)



Finally, accounting for possible non-linearities in our analyses, we investigate in a pervasive way the presence of possible breaks between outcomes and the main independent variable using the same controls of the main analyses. We adopt a LASSO procedure to evaluate possible discontinuities and test (H5) them between outcomes and ethnic share (see appendix, page 160). We do not observe significant thresholds for all our outcomes. It means that there is not a significant break in ethnic share. To gain a better picture, we run a piecewise analysis at a 30% threshold for ethnic share, but we do not find a significant intercept or slope break for all outcomes.

6. Discussion and conclusion

As mentioned above, previous contributions rely heavily on ethnic share computed as a percentage of non-native students in the total of students to identify the effect of non-native stock. Such a criterion focuses only on the evenness of native and non-native stock and underestimates the socio-linguistic richness of non-natives in a social environment.

We tried to revive an evergreen topic such as ethnic peer effect proposing a novel perspective, namely, the adoption of ethnolinguistic diversity in the student population of a country that has been exposed to increasing fluxes of immigrants in the recent decades. Since the '90s, analysis of diversity across social environments finds significant contributions in human resources studies, but only scarce application in an educational context. We fill up this gap drawing a theoretical perspective where ethnic concentration and ethnolinguistic diversity work together, affecting students' outcomes such as socio-emotional skills, behaviors, and academic performances.

On the empirical side, the contribution of this chapter can be summarized as follows. First, we developed a research design to tackle the main issues in the identification of contextual effects, by focusing on schools in which students' allocation to the classroom is as good as random. Second, we jointly analyze ethnic concentration and ethnic diversity, expanding the previous contributions, only focusing on the former measure. Indeed, we can rely on more than only two language categories as in Cho (2012) to proxy classroom diversity, and we control for unobserved fixed school characteristics, in contrast with the works of Gottfried (2016), Meng (2018), Veerman, (2015) Veerman, van de Werfhorst, and Dronkers, (2013), Veerman & Dronkers, (2016), and we debate the importance of teaching attributes. Regarding the ethnic indicators, we are aware that the diversity index is based on self-reported categories, up to 15 languages, plus a general "other" category. Of course, it would be better to have more details about such an "other" category," which could have allowed us a furthermore fine-grained differentiation of socio-linguistic background. Nevertheless, this paper is the first in using jointly ethnic concentration and diversity indicators as well as in exploiting ethnolinguistic groups in a country like Italy, where a collection of ethnolinguistic membership is rare.

Our work does not reach the same conclusions of Meng (2018) and Gottfriend (2016) regarding the effect of ethnic share and diversity on socio-emotional skills and behavior of students. Indeed, we find mixed-findings, although, on average, the effect is very tiny.

We delineate three important facts about CMO in the Italian education system. First, higher ethnic share leads to a higher chance of being bullied and doing the bullying, jointly with higher extrinsic motivation, but no effects are present on language and mathematic performance. When accounted for ethnic background, natives show more chance of being bullied and doing bullyism in the classroom, whereas CMO shows a higher penalty on language performance. Second, diversity index negatively affects the classroom environment, increasing the chance of doing bullyism. Besides, higher diversity leads to a competence language loss, and this loss is asymmetric at the expense of CMO. Besides, just diversity draws on a negative quadratic effect for bulling and language performance. Finally, we identify a moderating effect of diversity. Indeed, a low diversity classroom dose enhances the role of ethnic share in being bullied and doing bullyism. We think it is due to a sort of ethnolinguistic polarization frame: less diversity brings more out the ethnolinguistic groups leading to less integration.

Looking at the previous literature on peer effects, we should ask whether our strategy measures a credible peer effect and possible threats to our identification strategy. As outlined by Paloyo (2020), several strategies deal with the identification of peer effects. However, they frequently do not disentangle true peer effect from underlying characteristics of peers – compositional effects – in the words of Manski (1993) or true peer effects from class composition and size in terms of Ballatore and colleagues (2019). As discussed in our identification strategy, we do not disentangle these patterns following Moffit (2001) and Contini (2013). Nevertheless, the detected parameter remains interesting as well because it gives an overall effect of interest for the policymakers. At the moment, research is strictly oriented to decompose such an overall effect, but it neglects the role of teachers and related characteristics a lot.

The work of Ballatore and colleagues (2019) suggests that – once purged by class size effects and strategic behavior of the principal in assigning students to the classrooms – the true effect of ethnic concentration is stronger than the one usually found. However, their results apply only when the principals do not have room to accommodate and manipulate classroom composition. As stated by Ballatore and colleagues (2019), to our perspective, this is another policy indicator but distinct to what we found here.

Indeed, we find that the effect of ethnic share and diversity is not overall enough to advocate a new sorting policy. In our empirical approach, we estimate an overall tiny effect size.

Compared to the previous works, our contribution is quite in line regarding the effect size between an ethnic indicator and a student's outcome. We argue that it is owing to the principals that behave as clearing house to mitigate incoming issues for the academic years. Thanks to the threshold and piecewise analysis, we outline how the rationale of a classroom CMO threshold – frequently debated from mass-media – does not find empirical support.

This work strictly relies on the random assignment of students to the classroom within schools. Possible additional but close drawbacks of this strategy are omitted variables and measurement error (Hanushek et al., 2003; Ammermueller & Pischke, 2009; Contini, 2013). Regarding the first, we might argue that our specification model includes enough controls to avoid major issues of unobserved confounders. The inclusion of controls maybe not efficient. Indeed, and we come to the second drawback, the inclusion of antecedent family and school characteristics might overestimate the peer effect (Hanushek et al., 2003). A typical example is the collection of socioeconomic status, for example, using the number of books at home. In the empirical strategy, we avoid using a status classification based on books at home, reducing this specific measurement error entirely.

References

Abbiati G., Argentin G. & Gerosa T. (2017). Different teachers for different students? Evidence on teacher-student matching and its consequences in the Italian case. *Rivista di Politica Economica, 33*(1), 13-58.

Almlund, M., Duckworth A., Heckman J., & Kautz T. (2011). Personality psychology and economics. IZA Discussion Paper 5500.

Andersen, S & Thomsen, M., (2011). Policy Implications of Limiting Immigrant Concentration in Danish Public Schools, *Scandinavian Policy Studies, 34*(1), 27-52.

Angrist, J.D., & Lang, K., (2004). Does school integration generate peer effects? Evidence from Boston's Metco Program. *American Economic Review, 94*(5), 1613–1634.

Agasisti, T., & Falzetti, P. (2017). Between-classes sorting within schools and test scores: an empirical analysis of Italian junior secondary schools. *International Review of Economics, 64*(1), 1-45.

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.

Allen, D., & Wolniak, G.C. (2019). Exploring the Effects of Tuition Increases on Racial/Ethnic Diversity at Public Colleges and Universities. *Research Higher Educatio*n, 60(1), 18–43

Anelli M. K. Y. Shih, & Kevin Williams (2017). Foreign Peer Effects and STEM Major Choice, *Working Paper*

Azzolini, D., & Barone, C. (2013). Do they progress or do they lag behind? Educational attainment of immigrants' children in Italy: The role played by generational status, country of origin and social class. *Research in Social Stratification and Mobility, 31*(1), 82–96

Babad E. (2009). *The social psychology of the classroom*. Routledge Research in Education.

Barbieri, G., Rossetti, C., & Sestito, P., (2010). The Determinants of Teacher Mobility. Evidence from a Panel of Italian Teachers. Banca d'Italia Working Paper 761.

Ballatore, R. and Fort, M. & Ichino, A., (2019). Tower of Babel in the Classroom: Immigrants and Natives in Italian Schools, *Journal of Labor Economic*s, 36(4), 885-921.

Backus M, & Peng S. (2019). On Testing Continuity and the Detection of Failures, NBER Working Paper No. 26016

Beckerman, T. M., & Good, T. L. (1981). The classroom ratio of high- and low-aptitude students and its effect on achievement. *American Educational Research Journal, 18*(3), 317–327.

Bertozzi R., (2018). Studenti di origine immigrata che accedono all'università. Il ruolo del background migratorio e delle scelte scolastiche, in "Scuola democratica, Learning for Democracy, ISSN 1129-731X. - I(2018), pp. 5-22.

Black, S.E., Devereux, P.J., & Salvanes, K.G., (2013). "Under Pressure? The Effect of Peers on Outcomes of Young Adults". *Journal of Labor Economics, 31*(1), 119-153

Borman G.D. & Pyne J. (2016). What If Coleman Had Known About Stereotype Threat? How Social-Psychological Theory Can Help Mitigate Educational Inequality. *The Russell Sage Foundation Journal of the Social Sciences* (RSF) 164–185.

Borjas, George J. (2006). Making It in America: Social Mobility in the Immigrant Population. *The Future of Children, 16*(2), 55-71.

Bratti M, Deiana C., Havari E., Mazzarella G., & Merone E.C.. (2017). What Are You Voting For? Proximity to Refugee Reception Centres and Voting in the 2016 Italian Constitutional Referendum", IZA Discussion Papers 11060

Brunello, G., & Rocco, L., (2011). The Effect of Immigration on the School Performance of Natives: Cross Country Evidence Using PISA Test Scores. IZA Discussion Paper 5479.

Campodifiori, E., Figura, E., Papini, M., & Ricci, R., (2010). Un indicatore di status-economico-culturale degli allievi della quinta primaria in Italia. INVALSI Working Paper 02/2010.

Cebolla-Boado, H., (2007). Immigrant concentration in schools: peer pressures in place? *European Sociological Review, 23*(3), 341–356.

Cheong, P. H., Edwards, R., Goulbourne, H., & Solomos, J. (2007). Immigration, social cohesion and social capital: A critical review. *Critical Social Policy, 27*(1), 24–49.

Chiswick, Barry R. & Miller W. P. (2005). Linguistic Distance: A quantitative measure of the distance between English and other languages. *Journal of Multilingual and Multicultural Development, 26*(1), 1–11.

Cho M. R. (2012). Are there peer effects associated with having English Language Learner (ELL) classmates? Evidence from the Early Childhood Longitudinal Study Kindergarten Cohort (ECLS-K), *Economics of Education Review, 31*(5), 629– 643

Contini, D. (2013). Immigrant Background Peer Effects in Italian Schools, *Social Science Research, 42*(4), 1122–42.

Cox, T. (1993). *Cultural Diversity in Organizations: Theory, Research, and Practice*. San Francisco, CA: Berrett-Koehler

Cox, T. (2001). *Creating the Multicultural Organization: A Strategy for Capturing the Power of Diversity*. San Francisco, CA: Jossey-Bass.

Deering, P.D. (1996). An ethnographic study of norms of inclusion and cooperation in a multiethnic middle school. *Urban Review*, *28*, 21–39

Demanet, Jannick, & Mieke Van Houtte. (2014). Social–ethnic School Composition and Disengagement: An Inquiry into the Perceived Control Explanation. *The Social Science Journal*, *51*(4), 659–675.

Dumay, X., & Dupriez, V., (2008). Does the school composition matter? Evidence from Belgian data. *British Journal of Educational Studies, 56*(4), 440–477.

Erbring, L., & Young, A. A. (1979). Individuals and social structure contextual effects as endogenous feedback. *Sociological Methods & Research, 7*, 396–430.

Fekjaer, S.N., & Birkelund, G.E., (2007). Does the ethnic composition of upper secondary schools influence educational achievement and attainment? A multilevel analysis of the Norwegian case. *European Sociological Review, 23*(3) 309–323.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations,* 7, 117–140.

Finn J. D. & Voelkl K. E. (1993). School Characteristics Related to Student Engagement." *Journal of Negro Education*, *62* (3), 249–268.

Geven, S., M. Kalmijn & F. & Van Tubergen (2016). The ethnic composition of schools and students' problem behaviour in four European countries: the role of friends. *Journal of Ethnic and Migration Studies, 42*(9), 1473-1495

Gieling, M., Vollebergh W., & van Dorsselaer S. (2010). Ethnic Density in School Classes and Adolescent Mental Health. *Social Psychiatry and Psychiatric Epidemiology*, *45*(6), 639–646.

Gottfried, M. A. (2016). Peer sameness and peer diversity: The influence of breadth and depth of classmates with high academic needs in kindergarten. *Early Education and Development*, *27*(5), 655–675.

Gould, E.D., Lavy, V., & Paserman, M.D., (2009). Does immigration affect the long-term educational outcomes of natives? Quasi-experimental evidence. *The Economic Journal* *119*(540), 1243–1269.

Halle, T. G., Whittaker, J. V., Zepeda, M., Rothenberg, L., Anderson, R., Daneri, P., Wessel, J., & Buysse, V. (2014). The social–emotional development of dual language learners: Looking back at existing research and moving forward with purpose. *Early Childhood Research Quarterly*, *29*(4), 734–749

Halpern, D. (1993). "Minorities and Mental Health." *Social Science & Medicine, 36*(5), 597-607.

Hainmueller, J., (2012). Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies. *Political Analysis, 20*(1), 25-46.

Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica, 68*(3), 575–603.

Hansen, B. E. (2011). Threshold autoregression in economics. *Statistics and Its Interface*, *4*(3), 123–127.

Hanushek, E.A., Kain, J.F., Markman, J.M., & Rivkin, S.G., (2003). Does peer ability affect achievement? *Journal of Applied Econometrics, 18*(5), 527–544.

Hanushek, E.A., Kain, J.F., & Rivkin, S.G. (2009). New evidence about Brown v. Board of Education: The complex effects of school racial composition on achievement. *Journal of Labor Economics, 27*(3), 349–383.

Heckman J.j. Stixrud J. Urzua S. (2006) The Effects of Cognitive and Non-cognitive Abilities on Labor Market Outcomes and Social Behavior, *Journal of Labor Economics*, *24*(3), 411-482

Hermansen A. S., & Birkelund G. E. (2015) The Impact of Immigrant Classmates on Educational Outcomes, *Social Forces*, *94*(2), 615–646, Hornstra L., van der Veen I., Peetsma T., & Volman M., (2014) Does classroom composition make a difference: effects on developments in motivation, sense of classroom belonging, and achievement in upper primary school. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice Publication, 26*(2), 125-152

Horvath H. (2017). Classroom Assignment Policies and Implications for Teacher Value-Added Estimation, Working paper, https://sites.google.com/site/hedvigh/home

Hoxby, C., (2000). Peer Effects in the Classroom: Learning from Race and Gender Variation. National Bureau of Economic Research, Working Paper 7867.

Jayet, H., Ukrayinchuk, N., & De Arcangelis, G. (2010). The Location of Immigrants in Italy: Disentangling Networks and Local Effects. *Annals of Economics and Statistics*, (97-98), 329-350.

Kautz T., Heckman J.J., Diris R., ter Weel B., & Borghans L., (2014). Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success, NBER Working Paper No. 20749

Kelly, S. (2009). Social identity theories and educational engagement. *British Journal of Sociology of Education*, *30*(4), 449–462.

Klieme, E. (2006). Empirische unterrichtsforschung: aktuelleentwicklungen, theoretische Grundlagen und fachspezifische Befunde. *Zeitschrift fur Padagogik*, *52*(6), S. 765–773.

Manski, C., (1993). Identification of endogenous social effects: the reflection problem. *The Review of Economic Studies, 60*(3), 531–542.

Merton, Robert K. 1968. *Social Theory and Social Structure*. Free Press.

Meng C. (2018). Effect of classroom language diversity on head start ELL and non-ELL children's social-emotional development, *Applied Developmental Science*, *24*(3), 230-241

Mijs, J. J. B. (2016). Stratified Failure: Educational Stratification and Students' Attributions of Their Mathematics Performance in 24 Countries. *Sociology of Education*, *89*(2), 137–153.

MIUR (Ministero Istruzione Università e Ricerca) 2017. *La Scuola in Cifre 2009–10*. Sistan.

Moffitt, R., (2001). Policy interventions, low-level equilibria, and social interactions. In: Durlauf, S., & Young, P. (Eds.), *Social Dynamics*. Massachusetts Institute of Technology Press, Cambridge, MA.

OECD, 2012. *Untapped Skills. Realizing the Potential of Immigrant Students*. OECD Publishing, Paris.

Ong, C., & De Witte, K. (2013). The influence of ethnic segregation and school mobility in primary education on high school dropout – Evidence from regression discontinuity at a contextual tipping point. UNU-Merit Working Paper Series 2013-064, 1-35, Maastricht (Nederland): University of Maastricht. Graduate school of governance

Paloyo. A (2020). Peer effects in education: recent empirical evidence, in Braedley S. Grenn C. The economics of education. A comprehensive overview. American Press., *Forthcoming*

Patacchini E. Rainone E., & Zenou Y. (2017). Heterogeneous Peer Effects in Education *Journal of Economic Behavior and Organization*, *134*(C), 190-227

Rao, G. (2019). Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools. *American Economic Review*, *109*(3), 774-809.

Richard, O.C. (2000). Racial Diversity, Business Strategy, and Firm Performance: A Resource Based View. *The Academy of Management Journal, 43*(2),164–77.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York, NY: Oxford University Press.

Rossi F., & De Phillips M (2020). Parents, Schools and Human Capital Differences across Countries, *Journal of the European Economic Association*, Forthcoming

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

Schaeffer, M. (2013). Can Competing Diversity Indices Inform us About Why Ethnic Diversity Erodes Social Cohesion? A Test of Five Diversity Indices in Germany. *Social Science Research*, *42*(3), 755–774

Schneeweis, N., & Winter-Ebmer, R., (2015). Peer effects in Austrian schools. *Empirical Economics, 32*(2–3), 387–409.

Szulkin, R., & Jonnson, J.O., (2007). Ethnic Segregation and Educational Outcomes in Swedish Comprehensive Schools. Working Paper 2007:2 Stockholm University Linnaeus Center for Integration Studies.

Tatsi E. (2020). Common pitfalls in peer-effects models, Working paper, https://sites.google.com/site/etatsi/working-papers

Triventi, M. (2019). Are Children of Immigrants Graded Less Generously by their Teachers than Natives, and Why? Evidence from Student Population Data in Italy. *International Migration Review.*

Tonello M. (2016). Peer effects of non-native students on natives' educational outcomes: mechanisms and evidence, *Empirical Economics*, *51*(1), 383–414.

Van Ewijk, R. and Sleegers, P. (2010). Peer ethnicity and achievement: a meta-analysis into the compositional effect. *School Effectiveness and School Improvement*, *21*(3), 237–265.

Veerman, G. M. (2015). The Relationship between Ethnic Diversity and Classroom Disruption in the Context of Migration Policies. *Educational Studies, 41*(1–2), 209–25.

Veerman, G. M., van de Werfhorst H. G., & Dronkers J. (2013). Ethnic Composition of the Class and Educational Performance in Primary Education in The Netherlands. *Educational Research and Evaluation, 19*(5), 370–401.

Veerman, G. M. & Dronkers J. (2016). Ethnic Composition and School Performance in the Secondary Education of Turkish Migrant Students in Seven Countries and 19 European Educational Systems. *International Migration Review, 50*(3), 537–67.

Vygotsky, L. S. (1978). *Mind and society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press

Virdia, S. (2018). Ethnic Peer Pressure or School Inequalities? Ethnic Concentration and Performance in Upper-Secondary Schools. *Italian Journal of Sociology of Education*, *10*(2), 155-180.

Walker C. O., Greene B. A., & Mansell R. A., (2006). Identification with academics, intrinsic/extrinsic motivation, and self-efficacy as predictors of cognitive engagement *Learning and Individual Differences*, *16*(1), 1-12

Wilkinson G.A.I., Hattie J. Parr M.J. & Thrupp M., (2000). Influence of Peer Effects on Learning Outcomes: A Review of the Literature. *International Journal of Educational Research, 37*, 395–401.

Zincone G. (2006). The Making of Policies: Immigration and Immigrants in Italy, *Journal of Ethnic and Migration Studies*, *32*(3), 347-375

Conclusions

Educational systems are at the very heart of the public debate as the primary way to enhance living standards. Better education gives an opportunity to increase personal skills and has, in turn, positive returns on a virtually limited amount of individual and societal outcomes, first and foremost wages and occupations (Becker, 1964). Although this perspective is based upon undisputed findings, it masks and neglects an extensive literature according to which the consequences of the environmental features related to education do not always seem to be positive for the student.

The possibility that education can be harmful impels to reflect more broadly on the mechanisms that bind the educational environment to the student. This thought process was behind this thesis, which acknowledged such dynamics and pitched the idea that educational environment acts as parallel social machinery which entails harmful consequences on students' outcomes (Domina, Penner, & Penner, 2017).

Drawing upon this new strand of research, this work highlighted the critical role of a specific environment of this social machinery within which students daily compare themselves and interact with their mates and teachers: the classroom. Especially, classroom interaction and comparison a turning point were considered to explain possible detrimental consequences in term of student's opportunities and outcomes.

Existing literature is almost unanimous in suggesting that the comparison group operates at different levels. This was the key theoretical and empirical challenge faced by this work, which considered three main domains as a reference for the investigation: *hierarchies*, *networks of friends*, and *classroom peers*. Indeed, students can create a "social ladder" among themselves (Babad, 2009) in the classroom, they convey the opinion of friends (Lusher, Koskiner, & Robins 2013), and they interact with classroom mates of different characteristics such as age, ability, ethnic background, and socio-economic origin (Paloyo, 2020). All these significant relationships among students originate a multilayer structure of peer comparison, which could influence students' outcomes resulting in *inequality* patterns.

How do these forms of peer comparison affect the students? The works relied on the idea of classroom *diversity*. It is the exposure to varying characteristics of the mates such as gender, ethnic background, ability, socio-economic origin, habits as well as of teachers that result in hierarchies, condition the formation of friends, and define the classroom peers. However, such classroom diversity is not random, but it depends on the sorting process of students and teachers across schools. Such complexity was accounted for by relying on and reformulating the contribution of Roberto (2015). Indeed, *inequality* refers to the uneven distribution of resources, opportunities, or outcomes across students; *diversity* describes the variety of "types" in the student population and *sorting* refers to the uneven distribution of students and teachers across distinct schools. Finally, the allocation of students and teachers with specific characteristics depends on broad *policies* at the macro level.

In line with this, the work considered that the multilayer structure of peer comparison depends on classroom diversity, which, in turn, is affected by the sorting process and overall policies. Although the focus of the thesis was the within classroom peer effect, the work did not neglect the role played by institutional features of the classroom and plausible consequences. Indeed, this acknowledgement supports the rationale for the investigation. Henceforth, throughout the text, the education system was likened to a *Matryoshka,* to depict its complex multilayer structure where students are nested in classrooms, schools, and neighbourhoods.

In the last decades, social sciences have been dealing with growing attention on causal identification of social phenomena in distinct branches like economics, sociology, psychology, and politics. In addition, the collapse of research borders among such branches makes of paramount importance a "common language". The work took as much as possible these concerns into account with maximum transparency with regard to the aim of the thesis, the identification strategy, the underlying assumptions, and plausible threats to the reliability of findings.

The thesis was structured into four chapters. It started with a theoretical contribution followed by three empirical chapters which exploited a mix of administrative and survey data to disentangle the extent to which hierarchies, the network of friends, and classroom peers affect students' outcomes.

Congruent with the idea that educational systems behave as "social machinery" (Spring, 1945), Chapter 1 drawn the attention on how schools could create social inequalities. Indeed, Schools sort students, produce internal categories by grouping students on the basis of grades, classrooms, course-taking patterns, and academic tracks; they impose labels associated with students and reinforce external categorization processes based on salient individual traits such as race, class, and gender (Domina, Penner, and Penner 2017). Then, the chapter concentrates its effort on three important features. First, it discussed the importance of cognitive and, non-cognitive skills, behaviours, and choices in light of the current debate across sociology, economics, and psychology. Second, it proposed to investigate whether social forms of comparison within the classroom affects a large array of student's skills and behaviours by distinguishing between the formation of hierarchies, the network of friends, and classroom peers, and accounting for the relevant literature at play. Third, it briefly discussed the main feature of the educational systems with a final provision of a theoretical framework.

In chapter II, the work zoomed in on the role of *hierarchies*. In contrast to the previous literature in which the source of hierarchy has always been a student characteristic such as ability, I proposed a novel perspective by focusing on teachers' characteristics. Through a two-way fixed effect model, I provided empirical evidence that teacher's assessment via grading is an increasing inequality factor in Italy. Teachers' grading standard produces a hierarchy among students, irrespective of their ability and socio-demographic characteristics. Further, I show that grading hierarchy has a pervasive influence on a broad array of outcomes: not only students exposed to such hierarchy show less confidence and higher self-stigma, but this mindset also affects their academic performance, track choice, and investment in further human capital.

It also emerged that this mark-based hierarchy has a subtle relation to gender. Boys look more reactive when critical issues are in play, such as the choice of high school track or the expectation to invest in further education. Girls, in contrast, are more affected in their non-cognitive skills. In addition, the chapter reports two interesting facts. First, non-linear effects are present mainly in the upper tail of the distribution and not in the bottom. Second, mark-based hierarchy seems to influence academic performance more than the test-based one, meaning that an external judgement worths more that ability awareness. One of the more significant findings to emerge from this study is that a source of inequality happens every day in the classroom from the teacher's side, having harmful effects on students' skills, expectations, and choices. This result opens the possibility that such teacher's hierarchies might explain the rise of gender inequalities in a specific educational stage – i.e., the transition from lower to higher educational school – and calls for a deeper understanding of the consequences of grading practices.

In addition to the role of teachers, this thesis has taken into consideration a second important mechanism of social comparison, which takes place daily: best friends. Chapter III was dedicated to the specific role of friends as a linchpin of comparison, and it specifically examined the smoking and drinking peer effect of best friends. By means of an instrumental-variables approach that allow a proper identification strategy, and by distinguishing between reciprocal and non-reciprocal ties, I presented the effect of best friend behaviour on students' behaviour using data from the Netherlands.

Once considered the endogenous formation of the network and its perils, results provided evidence that the exposure to smoking and drinking of best friends strongly increases the adoption of similar behaviours, mainly in the context of reciprocal ties. While this study did not consider the heterogeneous effects in terms of socio-economic origins, ethnic background or gender (Huisman & Bruggeman, 2012), it offered valuable insights into possible policy implications. Indeed, someone can think that it is so unreasonable to think in term of clique policy intervention. This chapter informs the policymaker to adopt a network perspective. Indeed, a map of the classroom network can inform regarding the key actors in a classroom and makes possible a tailored intervention.

With chapter IV, I directly connected to important policy implications and the current political debate. The work addressed the presence and effect of children with a migration background in the classrooms. This is a central topic to be analysed because the exposure of the students to ethnic mates may have positive or negative consequences. The key difference with the previous forms of social comparison is those classroom peers are assigned by the "sorting machine" (Domina, Penner and Penner, 2017).

Once accounted for the sorting between schools, the empirical evidence provided by the study enriches the current debate on several aspects with a thorough investigation of factors affecting cognitive and non-cognitive skills. It provided empirical evidence of an effect of ethnic peers on a specific set of outcomes such as bullyism, integration, and academic competences. It also took up the torch of a research stream that gives importance not only to ethnic share but also to ethnolinguistic diversity. Results not only showed that ethnic share might affect student behaviour and motivation, but also that ethnolinguistic diversity plays a role in harming competences of Italian students. In addition, the study debunked the idea that children with migration background harm only themselves or only Italian, by showing a significant asymmetric effect. Finally, the overall effect remained modest, not advocating for a redistribution of children with migration background across classrooms.

The results of this chapter provide important points of discussion. First, invites the policymaker to think in term of linguistic differences in the classrooms and not only ethnic share. Second, the overall effect of children with a migration background is very tiny. Finally, the worrying finding emerged about the sorting within the school. Indeed, a not negligible body of schools operates an uneven sorting of children with migration background within schools across classrooms.

Taken together, the results of my dissertation suggest that that in the classroom, the multilayer structure of peer comparison operates with not negligible effect on a large array of student's outcomes. *Hierarchies, the network of friends, and classroom peers* are powerful components of the educational, social machinery. They affect a student's confidence, self-stigma, academic competences, educational expectations, investment in further educational stages, behaviour, and healthy habits. Hence, education systems may create inequality patterns every day in a social environment in which students should find a way to reduce cognitive and non-cognitive gaps and draw on a good way for their life. Policymakers should analyse more in-depth such patterns in order to make more effective educational policies.

References

Babad E. (2009). *The social psychology of the classroom*. Routledge Research in Education.

Becker, G. S., (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship.

Domina T., Penner A., & Penner E. (2017). Categorical Inequality: Schools As Sorting Machines. *Annual Review of Sociology*, *43*, 311-330

Huisman, C., & Bruggeman J. (2012). The Social Network, Socioeconomic Background, and School Type of Adolescent Smokers. *International Journal of Behavioral Development, 36*(5), 329–37.

Lusher D., Koskinen J. & Robins G. (2013). *Exponential graphs random models*. Cambridge University Press.

Paloyo. A (2020) Peer effects in education: recent empirical evidence, in Braedley S. Grenn C. The economics of education. A comprehensive overview. AP. *Forthcoming*

Roberto, E. (2015). The Divergence Index: A Decomposable Measure of Segregation and Inequality. Working paper, https://arxiv.org/abs/1508.01167

Spring J. (1945). The Sorting Machine: National Educational Policy Since 1945. *The Journal of Education*, *159*(3). Review of Bruce C. Beezer.

APPENDIX

Appendix Chapter II

Note that for the main analysis (Model 1-Model5), I report a limited stepwise approach to show how the effect depends on inclusion of mark fixed and classroom fixed effects. For the hierarchy comparison and non-linear effects, we report directly the analyses without a stepwise approach including always mark and classroom fixed effect. Note that in all models references are native, low parental education, with parents unemployed, and with late enrolment. We suggest to look at the following specifications:

Model 1: Rank with control variables

Model 2: Inclusion of mark fixed effect

Model 3: Inclusion of classroom fixed effect

Model 4: Rank-gender interaction

Model 5: Test Based Hierarchy (Test and Classroom FE)

Model 6: Mark Based Hierarchy (Mark and Classroom FE)

Model 7: Rank Squared (Mark and Classroom FE)

Model 8: Rank Cubic(Mark and Classroom FE)

Model 9: Rank Quartic (Mark and Classroom FE)

Scheme of Parental Job Classification	
Parental Job I:	Parents are unemployed
Parental Job II:	House-parents
Parental Job III:	At least one is Low level Employee
Parental Job IV:	At least one is Teacher, medium level employee, self-employed
Parental Job V:	At least one is Manager, professor, business man, land owner

Section 1: Stepwise approach for all outcomes

Agreeableness				
	Model 1	Model 2	Model 3	Model 4
Girls	-0.101***	-0.102***	-0.099***	-0.134***
	(0.00)	(0.00)	(0.00)	(-0.01)
Rank	0.141***	0.053***	0.044	0.008
	(-0.01)	(-0.02)	(-0.03)	(-0.03)
Rank * Girls				0.072***
				-0.01)
Non-natives	-0.057***	-0.055***	-0.025**	-0.025*
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parents with secondary education	-0.024	-0.026	-0.046	-0.046
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parents with tertiary education	-0.005	-0.01	-0.027	-0.028
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parental Job II	0.057***	0.056***	0.046***	0.046***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	0.015	0.014	0.032*	0.032*
	(-0.01)	-0.01)	(-0.01)	(-0.01)
Parental Job IV	0.051***	0.048***	0.063***	0.063***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	0.066***	0.064***	0.075***	0.075***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Regular enrollment	0.102***	0.100***	0.102***	0.104***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	0.148***	0.146***	0.116***	0.118***
	(-0.01)	(-0.01)	(-0.02)	(-0.02)
Math Test 8° grade	0.019***	0.018***	0.062**	0.063**
	(0.00)	(0.00)	(-0.02)	(-0.02)
constant	-0.094**	-0.047	-0.031	-0.016
	(-0.03)	(-0.03)	(-0.04)	(-0.04)
Mark Fixed Effects	No	Yes	Yes	Yes
Classroom Fixed Effects	No	No	Yes	Yes
R-sqr	0.013	0.014	0.171	0.171
F	0.000	0.000	0.000	0.000
BIC	317904.4	317819.1	288402.1	288381.3
AIC	317776.1	317690.7	288283.8	288253.1

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Intrinsic Motivation				
	Model 1	Model 2	Model 3	Model 4
Girls	0.207***	0.205***	0.206***	0.206***
	(0.00)	(0.00)	(0.00)	(-0.01)
Rank	0.389***	0.145***	0.054*	0.054*
	(-0.01)	(-0.01)	(-0.03)	(-0.03)
Rank * Girls				0.00
				(-0.01)
Non-natives	0.041***	0.044***	0.062***	0.062***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parents with secondary education	0.031	0.025	0.017	0.017
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parents with tertiary education	0.091***	0.075**	0.052	0.052
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parental Job II	0.016	0.014	0.017	0.017
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	-0.038**	-0.038**	-0.021	-0.021
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job IV	0.008	0.001	0.003	0.003
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	0.015	0.01	0.012	0.012
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Regular enrollment	0.067***	0.063***	0.056***	0.056***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	0.069***	0.066***	0.033*	0.033*
	(-0.01)	(-0.01)	(-0.02)	(-0.02)
Math Test 8° grade	0.013***	0.010**	-0.044*	-0.044*
	(0.00)	(0.00)	(-0.02)	(-0.02)
constant	-0.350***	-0.219***	-0.164***	-0.164***
	(-0.03)	(-0.03)	(-0.04)	(-0.04)
Mark Fixed Effects	No	Yes	Yes	Yes
Classroom Fixed Effects	No	No	Yes	Yes
R-sqr	0.053	0.057	0.198	0.198
F	0.000	0.000	0.000	0.000
BIC	313539.2	312887.2	285695.3	285707.1
AIC	313410.9	312758.8	285577	285579

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Extrinsic motivation				
	Model 1	Model 2	Model 3	Model 4
Girls	-0.191***	-0.190***	-0.192***	-0.186***
	(0.00)	(0.00)	(0.00)	(-0.01)
Rank	-0.128***	-0.040*	-0.025	-0.019
	(-0.01)	(-0.02)	(-0.03)	(-0.03)
Rank * Girls				-0.013
				(-0.02)
Non-natives	0.121***	0.120***	0.125***	0.125***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parents with secondary education	-0.062	-0.059	-0.025	-0.025
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parents with tertiary education	-0.041	-0.035	0.003	0.003
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parental Job II	0.005	0.005	0.007	0.007
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	-0.022	-0.022	-0.014	-0.014
	(-0.01)	(-0.01)	(-0.02)	(-0.02)
Parental Job IV	-0.005	-0.003	0.009	0.009
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	0.002	0.004	0.012	0.012
	(-0.01)	(-0.01)	(-0.02)	(-0.02)
Regular enrollment	-0.008	-0.004	0.003	0.003
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	0.033*	0.036*	0.014	0.014
	(-0.02)	(-0.02)	(-0.02)	(-0.02)
Math Test 8° grade	0.003	0.004	-0.055*	-0.055*
	(0.00)	(0.00)	(-0.02)	(-0.02)
constant	0.197***	0.147***	0.092*	0.089*
	(-0.04)	(-0.04)	(-0.04)	(-0.04)
Mark Fixed Effects	No	Yes	Yes	Yes
Classroom Fixed Effects	No	No	Yes	Yes
R-sqr	0.019	0.02	0.162	0.162
F	0.000	0.000	0.000	0.000
BIC	344886	344808.9	317515.2	317526.2
AIC	344757.6	344680.6	317396.9	317398

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Math confidence				
	Model 1	Model 2	Model 3	Model 4
Girls	-0.122***	-0.125***	-0.120***	-0.154***
	(0.00)	(0.00)	(-0.01)	(-0.01)
Rank	0.863***	0.407***	0.167***	0.131***
	(-0.01)	(-0.02)	(-0.03)	(-0.03)
Rank * Girls				0.071***
				(-0.02)
Non-natives	0.108***	0.115***	0.146***	0.146***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parents with secondary education	-0.041	-0.051	-0.080*	-0.080*
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parents with tertiary education	0.03	0.000	-0.065*	-0.066*
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parental Job II	-0.018	-0.022	-0.030*	-0.030*
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	-0.079***	-0.080***	-0.057***	-0.057***
	(-0.01)	(-0.01)	(-0.02)	(-0.02)
Parental Job IV	-0.042**	-0.055***	-0.051***	-0.051***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	-0.030*	-0.042**	-0.043**	-0.043**
	(-0.02)	(-0.01)	(-0.02)	(-0.02)
Regular enrollment	-0.006	-0.015	-0.029*	-0.027*
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	0.105***	0.097***	0.029	0.031
	(-0.02)	(-0.02)	(-0.02)	(-0.02)
Math Test 8° grade	0.040***	0.034***	0.095***	0.096***
	(0.00)	(0.00)	(-0.03)	(-0.03)
constant	-0.249***	-0.001	0.159***	0.174***
	(-0.03)	(-0.04)	(-0.04)	(-0.04)
Mark Fixed Effects	No	Yes	Yes	Yes
Classroom Fixed Effects	No	No	Yes	Yes
R-sqr	0.103	0.114	0.26	0.261
F	0.000	0.000	0.000	0.000
BIC	352535.1	350785.7	320162.7	320149.4
AIC	352406.7	350657.4	320044.4	320021.2

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Expected university enrollment				
	Model 1	Model 2	Model 3	Model 4
Girls	0.134***	0.132***	0.133***	0.164***
	(0.00)	(0.00)	(0.00)	(-0.01)
Rank	0.455***	0.173***	0.094***	0.127***
	(0.00)	(-0.01)	(-0.02)	(-0.02)
Rank * Girls				-0.066***
				(-0.01)
Non-natives	-0.003	0.001	0.012*	0.012*
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parents with secondary education	0.115***	0.107***	0.073***	0.074***
	(-0.02)	(-0.02)	(-0.02)	(-0.02)
Parents with tertiary education	0.297***	0.278***	0.200***	0.201***
	(-0.02)	(-0.02)	(-0.02)	(-0.02)
Parental Job II	0.000	-0.002	-0.002	-0.003
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	-0.030***	-0.030***	-0.015	-0.016
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job IV	0.078***	0.071***	0.060***	0.060***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	0.091***	0.085***	0.064***	0.064***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Regular enrollment	0.130***	0.122***	0.113***	0.111***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	0.170***	0.163***	0.123***	0.121***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Math Test 8° grade	0.021***	0.018***	0.057***	0.056***
	(0.00)	(0.00)	(-0.01)	(-0.01)
constant	-0.038*	0.119***	0.213***	0.199***
	(-0.02)	(-0.02)	(-0.02)	(-0.02)
Mark Fixed Effects	No	Yes	Yes	Yes
Classroom Fixed Effects	No	No	Yes	Yes
R-sqr	0.171	0.18	0.317	0.317
F	0.000	0.000	0.000	0.000
BIC	179790.4	178142.1	149768.3	149708.6
AIC	179662.1	178013.7	149650	149580.4

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Actual dropout				
	Model 1	Model 2	Model 3	Model 4
Girls	0.017***	0.014***	0.014***	0.026***
	(0.00)	(0.00)	(0.00)	(0.00)
Rank	0.311***	0.060***	-0.052***	-0.037***
	(0.00)	(-0.01)	(-0.01)	(-0.01)
Rank * Girls				-0.030***
				(0.00)
Non-natives	0.052***	0.054***	-0.074***	-0.075***
	(0.00)	(0.00)	(0.00)	(0.00)
Parents with secondary education	0.104***	0.093***	0.018***	0.018***
	(0.00)	(-0.01)	(0.00)	(0.00)
Parents with tertiary education	0.163***	0.142***	0.040***	0.041***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job II	-0.045***	-0.046***	0.015***	0.015***
	(0.00)	(0.00)	(0.00)	(0.00)
Parental Job III	0.078***	0.076***	0.002	0.002
	(0.00)	(0.00)	(0.00)	(0.00)
Parental Job IV	0.108***	0.098***	0.040***	0.040***
	(0.00)	(0.00)	(0.00)	(0.00)
Parental Job V	0.111***	0.102***	0.040***	0.041***
	(0.00)	(-0.01)	(0.00)	(0.00)
Regular enrollment	0.174***	0.156***	0.153***	0.152***
	(0.00)	(0.00)	(0.00)	(0.00)
Early enrollment	0.043***	0.026***	0.145***	0.144***
	(0.00)	(0.00)	(0.00)	(0.00)
Math Test 8° grade	-0.051***	-0.054***	-0.041***	-0.042***
	(0.00)	(0.00)	(0.00)	(0.00)
constant	0.023***	0.150***	0.271***	0.267***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Mark Fixed Effects	No	Yes	Yes	Yes
Classroom Fixed Effects	No	No	Yes	Yes
R-sqr	0.13	0.14	0.411	0.411
F	0.000	0.000	0.000	0.000
BIC	479680.1	475305.7	333828.2	333785.8
AIC	479539.3	475164.9	333698.2	333645

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Dropout Intention				
	Model 1	Model 2	Model 3	Model 4
Girls	-0.093***	-0.091***	-0.097***	-0.119***
	(0.00)	(0.00)	(0.00)	(-0.01)
Rank	-0.427***	-0.176***	-0.105***	-0.129***
	(-0.01)	(-0.02)	(-0.03)	(-0.03)
Rank * Girls				0.047**
				(-0.02)
Non-natives	-0.089***	-0.093***	-0.099***	-0.098***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parents with secondary education	-0.117***	-0.110***	-0.095**	-0.096**
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parents with tertiary education	-0.214***	-0.198***	-0.149***	-0.149***
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parental Job II	0.004	0.006	-0.008	-0.008
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	0.031*	0.032*	0.020	0.020
	(-0.01)	(-0.01)	(-0.02)	(-0.02)
Parental Job IV	-0.037**	-0.030*	-0.02	-0.02
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	-0.045**	-0.039**	-0.02	-0.02
	(-0.01)	(-0.01)	(-0.02)	(-0.02)
Regular enrollment	-0.221***	-0.214***	-0.209***	-0.208***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	-0.250***	-0.244***	-0.222***	-0.221***
	(-0.02)	(-0.02)	(-0.02)	(-0.02)
Math Test 8° grade	-0.008*	-0.005	-0.067**	-0.066**
	(0.00)	(0.00)	(-0.02)	(-0.02)
constant	0.536***	0.397***	0.339***	0.349***
	(-0.04)	(-0.04)	(-0.04)	(-0.04)
Mark Fixed Effects	No	Yes	Yes	Yes
Classroom Fixed Effects	No	No	Yes	Yes
R-sqr	0.046	0.049	0.189	0.189
F	0.000	0.000	0.000	0.000
BIC	335671.2	335152	307794.2	307794
AIC	335542.9	335023.7	307675.9	307665.8

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Math performance				
	Model 1	Model 2	Model 3	Model 4
Girls	-0.358***	-0.365***	-0.354***	-0.361***
	(0.00)	(0.00)	(0.00)	(-0.01)
Rank	1.491***	0.599***	0.234***	0.227***
	(-0.01)	(-0.02)	(-0.03)	(-0.03)
Rank * Girls				0.014
				(-0.01)
Non-natives	0.012	0.025**	-0.028**	-0.028**
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parents with secondary education	0.176***	0.151***	0.048	0.048
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parents with tertiary education	0.367***	0.307***	0.132***	0.132***
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parental Job II	-0.040**	-0.047***	-0.016	-0.016
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	0.035*	0.033*	-0.02	-0.02
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job IV	0.149***	0.125***	0.041**	0.041**
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	0.152***	0.130***	0.053***	0.053***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Regular enrollment	0.277***	0.255***	0.235***	0.235***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	0.200***	0.181***	0.293***	0.294***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Math Test 8° grade	-0.018***	-0.028***	0.550***	0.550***
	(0.00)	(-0.01)	(-0.03)	(-0.03)
constant	-0.897***	-0.405***	-0.049	-0.046
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Mark Fixed Effects	No	Yes	Yes	Yes
Classroom Fixed Effects	No	No	Yes	Yes
R-sqr	0.306	0.334	0.518	0.518
F	0.000	0.000	0.000	0.000
BIC	342035.8	336114.1	285090.6	285101.2
AIC	341907.4	335985.8	284972.3	284973

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

	Neuroticism			
	Model 1	Model 2	Model 3	Model 4
Girls	0.312***	0.313***	0.312***	0.309***
	(0.00)	(0.00)	(0.00)	(-0.01)
Rank	-0.274***	-0.142***	-0.057*	-0.061*
	(-0.01)	(-0.02)	(-0.03)	(-0.03)
Rank * Girls				0.006
				(-0.01)
Non-natives	0.114***	0.112***	0.110***	0.111***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parents with secondary education	-0.136***	-0.133***	-0.085**	-0.085**
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parents with tertiary education	-0.233***	-0.225***	-0.136***	-0.137***
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parental Job II	0.030*	0.030*	0.022	0.022
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	0.030*	0.030*	0.021	0.021
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job IV	-0.027*	-0.024	-0.011	-0.011
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	-0.045**	-0.043**	-0.024	-0.024
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Regular enrollment	-0.113***	-0.113***	-0.109***	-0.109***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	-0.141***	-0.142***	-0.127***	-0.127***
	(-0.01)	(-0.02)	(-0.02)	(-0.02)
Math Test 8° grade	-0.025***	-0.024***	-0.238***	-0.238***
	(0.00)	(0.00)	(-0.02)	(-0.02)
constant	0.222***	0.154***	0.046	0.048
	(-0.03)	(-0.03)	(-0.04)	(-0.04)
Mark Fixed Effects	No	Yes	Yes	Yes
Classroom Fixed Effects	No	No	Yes	Yes
R-sqr	0.068	0.069	0.218	0.218
F	0.000	0.000	0.000	0.000
BIC	323877.9	323715	294420.9	294432.5
AIC	323749.6	323586.6	294302.6	294304.3

* $p<0.05$, ** $p<0.01$, *** $p<0.001$
(Standard errors in parentheses)
Source: INVALSI

Self-stigma				
	Model 1	Model 2	Model 3	Model 4
Girls	-0.222***	-0.221***	-0.220***	-0.280***
	(0.00)	(0.00)	(0.00)	(-0.01)
Rank	-0.355***	-0.138***	-0.082**	-0.146***
	(-0.01)	(-0.02)	(-0.03)	(-0.03)
Rank * Girls				0.127***
				(-0.02)
Non-natives	-0.035***	-0.038***	-0.038***	-0.037***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parents with secondary education	-0.041	-0.033	-0.025	-0.026
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parents with tertiary education	-0.093**	-0.077**	-0.052	-0.053
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parental Job II	-0.005	-0.003	-0.014	-0.013
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	0.006	0.007	0.001	0.001
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job IV	-0.025	-0.019	-0.009	-0.009
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	-0.035*	-0.029*	-0.015	-0.014
	(-0.01)	(-0.01)	(-0.02)	(-0.02)
Regular enrollment	-0.147***	-0.136***	-0.119***	-0.116***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	-0.136***	-0.126***	-0.108***	-0.105***
	(-0.02)	(-0.02)	(-0.02)	(-0.02)
Math Test 8° grade	-0.002	0.000	-0.014	-0.012
	(0.00	(0.00)	(-0.02)	(-0.02)
constant	0.415***	0.289***	0.233***	0.260***
	(-0.03)	(-0.03)	(-0.04)	(-0.04)
Mark Fixed Effects	No	Yes	Yes	Yes
Classroom Fixed Effects	No	No	Yes	Yes
R-sqr	0.049	0.051	0.19	0.191
F	0.000	0.000	0.000	0.000
BIC	325987.3	325598.3	298265.7	298183.3
AIC	325858.9	325469.9	298147.4	298055.1

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Academic track choice				
	Model 1	Model 2	Model 3	Model 4
Girls	0.200***	0.198***	0.194***	0.243***
	(0.00)	(0.00)	(0.00)	(0.00)
Rank	0.450***	0.125***	0.071***	0.124***
	(0.00)	(-0.01)	(-0.01)	(-0.02)
Rank * Girls				-0.104***
				(-0.01)
Non-natives	-0.060***	-0.055***	-0.037***	-0.038***
	(0.00)	(0.00)	(-0.01)	(-0.01)
Parents with secondary education	0.139***	0.130***	0.099***	0.100***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parents with tertiary education	0.379***	0.357***	0.254***	0.255***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job II	0.017*	0.014	0.012	0.011
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	-0.042***	-0.043***	-0.018*	-0.018*
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job IV	0.098***	0.090***	0.076***	0.076***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	0.128***	0.121***	0.092***	0.092***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Regular enrollment	0.123***	0.115***	0.107***	0.105***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	0.191***	0.184***	0.117***	0.114***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Math Test 8° grade	0.029***	0.025***	0.048***	0.046***
	(0.00)	(0.00)	(-0.01)	(-0.01)
constant	-0.125***	0.055**	0.142***	0.120***
	(-0.02)	(-0.02)	(-0.02)	(-0.02)
Mark Fixed Effects	No	Yes	Yes	Yes
Classroom Fixed Effects	No	No	Yes	Yes
R-sqr	0.232	0.245	0.423	0.423
F	0.00	0.00	0.00	0.00
BIC	169668.4	167275.2	126857.8	126658.7
AIC	169540.1	167146.9	126739.5	126530.5

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Section 2: Mark and Test comparison, Mark/Test and Classroom fixed effects

	Actual Dropout		Math Perfomance		Academic track choice	
	Model 5	Model 6	Model 5	Model 6	Model 5	Model 6
Girls	-0.053***	-0.023***	-0.089***	-0.215***	0.265***	0.226***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Test-based Hierarchy	-0.051***		0.026		0.105***	
	(-0.01)		(-0.01)		(-0.01)	
Mark-based Hierarchy		-0.061***		0.232***		0.080***
		(-0.01)		(-0.02)		(-0.01)
Non-natives	0.084***	0.070***	0.019**	0.046***	-0.046***	-0.034***
	(0.00)	(0.00)	(-0.01)	(-0.01)	(0.00)	(0.00)
Parents with secondary education	-0.148***	-0.128***	0.064***	0.034*	0.116***	0.095***
	(-0.01)	(-0.01)	(-0.02)	(-0.02)	(-0.01)	(-0.01)
Parents with tertiary education	-0.178***	-0.142***	0.177***	0.132***	0.275***	0.237***
	(-0.01)	(-0.01)	(-0.02)	(-0.02)	(-0.01)	(-0.01)
Parental Job II	-0.031***	-0.023***	0.002	-0.019**	0.010*	0.004
	(0.00)	(0.00)	(-0.01)	(-0.01)	(0.00)	(0.00)
Parental Job III	-0.012**	-0.009*	-0.024**	-0.027***	-0.018***	-0.018***
	(0.00)	(0.00)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job IV	-0.068***	-0.053***	0.039***	0.011	0.084***	0.070***
	(0.00)	(0.00)	(-0.01)	(-0.01)	(0.00)	(0.00)
Parental Job V	-0.052***	-0.042***	0.029***	0.011	0.084***	0.075***
	(0.00)	(0.00)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Regular enrollment	-0.092***	-0.095***	0.110***	0.137***	0.043***	0.041***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	-0.098***	-0.099***	0.169***	0.206***	0.053***	0.051***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Math Test 5° grade	-0.024***	-0.020***	0.220***	0.304***	0.044***	0.037***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
constant	0.515***	0.470***	-0.207***	-0.213***	0.132***	0.206***
	(-0.01)	(-0.01)	(-0.02)	(-0.02)	(-0.01)	(-0.01)
R-sqr	0.237	0.269	0.621	0.613	0.361	0.39
F	0.000	0.000	0.000	0.000	0.000	0.000
BIC	233817.31	198214.736	509640.696	469177.267	271484.288	235432.941
AIC	233688.553	198087.127	509514.44	469052.145	271358.033	235307.82

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Section 3: Non-linear effeects, Mark and Classroom fixed effects

	Agreeableness			Intrinsic Motivation			Extrinsic Motivation		
	Model 7	Model 8	Model 9	Model 7	Model 8	Model 9	Model 7	Model 8	Model 9
Girls	-0.100***	-0.100***	-0.100***	0.206***	0.206***	0.206***	-0.192***	-0.192***	-0.192***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Rank	0.114**	0.044	-0.027	0.052	0.164*	0.106	0.052	-0.04	-0.174
	(-0.04)	(-0.07)	(-0.12)	(-0.04)	(-0.07)	(-0.13)	(-0.05)	(-0.08)	(-0.14)
Rank^2	-0.079*	0.119	0.494	0.003	-0.317	-0.013	-0.089*	0.172	0.88
	(-0.04)	(-0.17)	(-0.54)	(-0.04)	(-0.17)	(-0.56)	(-0.04)	(-0.19)	(-0.60)
Rank^3		-0.138	-0.753		0.223*	-0.275		-0.182	-1.342
		(-0.11)	(-0.84)		(-0.11)	(-0.87)		(-0.13)	(-0.94)
Rank^4			0.314			0.254			0.592
			(-0.42)			(-0.43)			-(0.47)
Non-natives	-0.025**	-0.025**	-0.025**	0.062***	0.062***	0.062***	0.125***	0.125***	0.125***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parents with secondary education	-0.046	-0.046	-0.046	0.017	0.017	0.017	-0.025	-0.025	-0.025
	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parents with tertiary education	-0.027	-0.028	-0.028	0.052	0.052	0.052	0.003	0.003	0.003
	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parental Job II	0.046***	0.046***	0.046***	0.017	0.017	0.017	0.007	0.007	0.007
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	0.031*	0.031*	0.032*	-0.021	-0.021	-0.021	-0.014	-0.014	-0.014
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.02)	(-0.02)	(-0.02)
Parental Job IV	0.062***	0.062***	0.062***	0.003	0.003	0.003	0.009	0.009	0.009
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	0.074***	0.074***	0.074***	0.012	0.012	0.012	0.012	0.012	0.012
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.02)	(-0.02)	(-0.02)
Regular enrollment	0.101***	0.101***	0.101***	0.056***	0.056***	0.056***	0.003	0.003	0.003
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	0.115***	0.115***	0.115***	0.033*	0.033*	0.033*	0.014	0.014	0.014
	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)
Math Test 8° grade	0.061**	0.061**	0.061**	-0.044*	-0.044*	-0.044*	-0.056*	-0.056*	-0.056*
	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)
constant	-0.038	-0.035	-0.034	-0.164***	-0.168***	-0.167***	0.085*	0.088*	0.090*
	(-0.04)	(-0.04)	(-0.04)	(-0.04)	(-0.04)	(-0.04)	(-0.04)	(-0.04)	(-0.04)
R-sqr	0.171	0.171	0.171	0.198	0.198	0.198	0.162	0.162	0.162
F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BIC	288408.1	288418.2	288429.3	285707.1	285714.1	285725.5	317521.1	317530.4	317540.3
AIC	288280	288280.1	288281.5	285579	285576.1	285577.6	317393	317392.4	317392.4

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

	Math Confidence			Neuroticism			Self-stigma		
	Model 7	Model 8	Model 9	Model 7	Model 8	Model 9	Model 7	Model 8	Model 9
Girls	-0.120***	-0.120***	-0.120***	0.312***	0.312***	0.312***	-0.220***	-0.220***	-0.220***
	(-0.01)	(-0.01)	(-0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Rank	0.029	0.199*	0.185	-0.053	-0.038	0.137	-0.125**	-0.160*	-0.244
	(-0.05)	(-0.08)	(-0.14)	(-0.04)	(-0.08)	(-0.13)	(-0.05)	(-0.08)	(-0.13)
Rank^2	0.158***	-0.323	-0.248	-0.005	-0.049	-0.968	0.049	0.15	0.591
	(-0.04)	(-0.19)	(-0.60)	(-0.04)	(-0.17)	(-0.56)	(-0.04)	(-0.18)	(-0.57)
Rank^3		0.336**	0.213		0.031	1.535		-0.07	-0.792
		(-0.13)	(-0.95)		(-0.11)	(-0.87)		(-0.12)	(-0.88)
Rank^4			0.063			-0.768			0.369
			(-0.48)			(-0.43)			(-0.44)
Non-natives	0.146***	0.146***	0.146***	0.110***	0.110***	0.110***	-0.038***	-0.038***	-0.038***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parents with secondary education	-0.080*	-0.079*	-0.079*	-0.085**	-0.085**	-0.085**	-0.025	-0.025	-0.026
	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parents with tertiary education	-0.065*	-0.065*	-0.065*	-0.136***	-0.136***	-0.136***	-0.052	-0.052	-0.052
	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parental Job II	-0.030*	-0.030*	-0.030*	0.022	0.022	0.021	-0.013	-0.013	-0.013
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	-0.057***	-0.057***	-0.057***	0.021	0.021	0.021	0.001	0.001	0.001
	(-0.02)	(-0.02)	(-0.02)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job IV	-0.051***	-0.051***	-0.051***	-0.011	-0.011	-0.012	-0.009	-0.009	-0.009
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	-0.042**	-0.042**	-0.042**	-0.024	-0.024	-0.024	-0.015	-0.015	-0.015
	(-0.02)	(-0.02)	(-0.02)	(-0.01)	(-0.01)	(-0.01)	(-0.02)	(-0.02)	(-0.02)
Regular enrollment	-0.028*	-0.029*	-0.029*	-0.109***	-0.109***	-0.109***	-0.119***	-0.119***	-0.118***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	0.03	0.029	0.029	-0.127***	-0.127***	-0.127***	-0.108***	-0.108***	-0.108***
	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)
Math Test 8° grade	0.096***	0.096***	0.096***	-0.238***	-0.239***	-0.238***	-0.013	-0.013	-0.013
	(-0.03)	(-0.03)	(-0.03)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)	(-0.02)
constant	0.172***	0.166***	0.166***	0.046	0.045	0.043	0.237***	0.239***	0.240***
	(-0.04)	(-0.04)	(-0.04)	(-0.04)	(-0.04)	(-0.04)	(-0.04)	(-0.04)	(-0.04)
R-sqr	0.261	0.261	0.261	0.218	0.218	0.218	0.19	0.19	0.19
F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BIC	320156.1	320159.3	320171.1	294432.7	294444.5	294452.5	298275.5	298286.9	298297.9
AIC	320027.9	320021.3	320023.3	294304.6	294306.5	294304.6	298147.3	298148.9	298150

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

	Math Performance			Dropout Intention		
	Model 7	Model 8	Model 9	Model 7	Model 8	Model 9
Girls	-0.354***	-0.354***	-0.354***	-0.097***	-0.097***	-0.097***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Rank	0.018	0.161*	0.057	-0.105*	-0.087	-0.08
	(-0.04)	(-0.07)	(-0.12)	(-0.05)	(-0.08)	(-0.14)
Rank^2	0.248***	-0.159	0.389	0.00	-0.052	-0.088
	(-0.04)	(-0.17)	(-0.55)	(-0.04)	(-0.18)	(-0.59)
Rank^3		0.284*	-0.613		0.036	0.096
		(-0.12)	(-0.87)		(-0.12)	(-0.91)
Rank^4			0.458			-0.031
			(-0.44)			(-0.45)
Non-natives	-0.028**	-0.028**	-0.028**	-0.099***	-0.099***	-0.099***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parents with secondary education	0.048	0.048	0.048	-0.095**	-0.095**	-0.095**
	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parents with tertiary education	0.132***	0.132***	0.132***	-0.149***	-0.149***	-0.149***
	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Parental Job II	-0.015	-0.016	-0.016	-0.008	-0.008	-0.008
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	-0.019	-0.019	-0.019	0.02	0.02	0.02
	(-0.01)	(-0.01)	(-0.01)	(-0.02)	(-0.02)	(-0.02)
Parental Job IV	0.042**	0.042**	0.042**	-0.02	-0.02	-0.02
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	0.054***	0.054***	0.054***	-0.02	-0.02	-0.02
	(-0.01)	(-0.01)	(-0.01)	(-0.02)	(-0.02)	(-0.02)
Regular enrollment	0.237***	0.236***	0.236***	-0.209***	-0.209***	-0.209***
	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	0.294***	0.294***	0.294***	-0.222***	-0.222***	-0.222***
	(-0.01)	(-0.01)	(-0.01)	(-0.02)	(-0.02)	(-0.02)
Math Test 8° grade	0.552***	0.552***	0.552***	-0.067**	-0.067**	-0.067**
	(-0.03)	(-0.03)	(-0.03)	(-0.02)	(-0.02)	(-0.02)
constant	-0.028	-0.034	-0.032	0.339***	0.338***	0.338***
	(-0.03)	(-0.03)	(-0.03)	(-0.04)	(-0.04)	(-0.04)
R-sqr	0.518	0.518	0.518	0.189	0.189	0.189
F	0.000	0.000	0.000	0.000	0.000	0.000
BIC	285044.648	285048.591	285058.97	307806.022	307817.775	307829.629
AIC	284916.477	284910.561	284911.08	307677.852	307679.745	307681.74

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

	Expected university enrollment			Actual dropout			Academic track choice		
	Model 7	Model 8	Model 9	Model 7	Model 8	Model 9	Model 7	Model 8	Model 9
Girls	0.133***	0.133***	0.133***	0.014***	0.014***	0.014***	0.194***	0.194***	0.194***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00	(0.00)	(0.00)	(0.00)	(0.00)
Rank	0.047	0.065	0.166*	0.080***	0.159***	0.128**	-0.002	0.02	0.034
	(-0.03)	(-0.04)	(-0.07)	(-0.01)	(-0.02	(-0.04)	(-0.02)	(-0.04)	(-0.07)
Rank^2	0.054*	0.002	-0.531	-0.155***	-0.394***	-0.224	0.084***	0.022	-0.053
	(-0.02)	(-0.10)	(-0.33)	(-0.01)	(-0.06	(-0.18)	(-0.02)	(-0.10)	(-0.31)
Rank^3		0.037	0.91		0.173***	-0.116		0.044	0.167
		(-0.07)	(-0.51)		(-0.04)	(-0.29)		(-0.07)	(-0.48)
Rank^4			-0.446			0.151			-0.063
			(-0.26)			(-0.15)			(-0.24)
Non-natives	0.012*	0.012*	0.012*	-0.074***	-0.075***	-0.075***	-0.037***	-0.037***	-0.037***
	(-0.01)	(-0.01)	(-0.01)	(0.00)	(0.00)	(0.00)	(-0.01)	(-0.01)	(-0.01)
Parents with secondary education	0.073***	0.073***	0.074***	0.018***	0.017***	0.017***	0.099***	0.099***	0.099***
	(-0.02)	(-0.02)	(-0.02)	(0.00)	(0.00)	(0.00)	(-0.01)	(-0.01)	(-0.01)
Parents with tertiary education	0.200***	0.200***	0.201***	0.040***	0.040***	0.040***	0.254***	0.254***	0.254***
	(-0.02)	(-0.02)	(-0.02)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Parental Job II	-0.002	-0.002	-0.002	0.015***	0.015***	0.015***	0.012	0.012	0.012
	(-0.01)	(-0.01)	(-0.01)	(0.00)	(0.00)	(0.00)	(-0.01)	(-0.01)	(-0.01)
Parental Job III	-0.015	-0.015	-0.015	0.002	0.002	0.002	-0.018*	-0.018*	-0.018*
	(-0.01)	(-0.01)	(-0.01)	(0.00)	(0.00)	(0.00)	(-0.01)	(-0.01)	(-0.01)
Parental Job IV	0.060***	0.060***	0.060***	0.040***	0.040***	0.040***	0.076***	0.076***	0.076***
	(-0.01)	(-0.01)	(-0.01)	(0.00)	(0.00)	(0.00)	(-0.01)	(-0.01)	(-0.01)
Parental Job V	0.065***	0.065***	0.065***	0.040***	0.040***	0.040***	0.093***	0.093***	0.093***
	(-0.01)	(-0.01)	(-0.01)	(0.00)	(0.00)	(0.00)	(-0.01)	(-0.01)	(-0.01)
Regular enrollment	0.113***	0.113***	0.113***	0.151***	0.151***	0.151***	0.108***	0.108***	0.108***
	(-0.01)	(-0.01)	(-0.01)	(0.00)	(0.00)	(0.00)	(-0.01)	(-0.01)	(-0.01)
Early enrollment	0.123***	0.123***	0.123***	0.143***	0.143***	0.143***	0.117***	0.117***	0.117***
	(-0.01)	(-0.01)	(-0.01)	(0.00)	(0.00)	(0.00)	(-0.01)	(-0.01)	(-0.01)
Math Test 8° grade	0.058***	0.058***	0.058***	-0.042***	-0.042***	-0.042***	0.048***	0.048***	0.048***
	(-0.01)	(-0.01)	(-0.01)	(0.00)	(0.00)	(0.00)	(-0.01)	(-0.01)	(-0.01)
constant	0.217***	0.217***	0.215***	0.262***	0.260***	0.260***	0.149***	0.148***	0.148***
	(-0.02)	(-0.02)	(-0.02)	(-0.01)	(-0.01)	(-0.01)	(-0.02)	(-0.02)	(-0.02)
R-sqr	0.317	0.317	0.317	0.411	0.411	0.411	0.423	0.423	0.423
F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BIC	149772.9	149784.4	149792.6	333630.9	333619.4	333630.9	126849.1	126860.4	126872.2
AIC	149644.7	149646.4	149644.7	333490.1	333467.8	333468.5	126721	126722.4	126724.3

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Section 4: Descriptive table VIII: Statistics

Take note that the design I relies on 143.420 students and design II relies on 264.172 students. In addition, Note that the variable actual relies on all students not matched. Hence, the sample size for this variable is larger. For the design I, the sample size is 373.096. For the design II, the sample size is 337.653. For the other outcomes, I point out that I use IPW.

Design I	**N: 143.420**			
	MEAN	**SD**	**MIN**	**MAX**
Intrinsic Motivation	0	1	-3	1
Extrinsic Motivation	0	1	-1	2
Math Confidence	0	1	-1	2
Consciousness	0	1	-2	2
Agreeableness	0	1	-3	1
Neuroticism	0	1	-1	2
Self-stigma	0	1	-1	3
Math Perfomance 10° grade	0	1	-2	3
Dropout Intention	0	1	-1	3
Rank	0	0	0	1
Math Standardized Test - 8° grade	0	1	-1	10
	%			
Academic track	56.9			
Not academic track	55.9			
Expected university enrollment	57.1			
Not expected university enrollment	56.1			
No Actual Dropout	54.11			
Actual dropout	45.89			
Boys	50.3			
Girls	49.7			
Early Enrollment	9.17			
Regular Enrollment	80.55			
Late Enrollment	10.28			
Native	90.07			
Non-native	9.93			
Primary Education	1.69			
Secondary Education	78.1			
Tertiary Education	20.21			
Parental Job I	3.71			
Parental Job II	40.36			
Parental Job III	13.96			
Parental Job IV	32.14			
Parental Job V	9.84			

Design II	N: 264.172			
	MEAN	SD	MIN	MAX
Math Performance 10° grade	0	1	-3	4
Test based Hierarchy	0.5	0.3	0	1
Mark based Hierarchy	0.4	0.3	0	1
Math Standardized Test - 5° grade	0	1	-5.1	4.3
	%			
Academic track choice	57.0			
Not academic track choice	43.0			
No Actual Dropout	85.39			
Actual dropout	16.41			
Boys	50.89			
Girls	49.11			
Early Enrollment	8.77			
Regular Enrollment	82.66			
Late Enrollment	8.77			
Native	90.18			
Non-native	9.82			
Primary Education	1.15			
Secondary Education	59.62			
Tertiary Education	17.91			
Parental Job I	4.71			
Parental Job II	37.3			
Parental Job III	13.77			
Parental Job IV	33.34			
Parental Job V	10.87			

Appendix Chapter III

Note that in all models references are native, male, speaking Dutch, and with a late enrolment. You can find synthetic and extensive tables for the main analysis of best friends effect and non-reciprocal ties. Hansen Test is reported with the disclaimer Equation Exactly Identified (EEI). Indeed, Hansen test is informative when more variables instrument the same endogenous regressor.

Section 1: Best Friends, synthetic tables for OLS, OLS and Classroom FE, and IV regressions

	Smoking I			
	Model I	Model II	Model III	Model IV
	Ols	Ols + FE	IV - A	IV - B-I
Smoking Clique	0.53***	0.38***	0.66***	0.69***
	(0.00)	(0.00)	(0.00)	(0.07)
R^2 adj	0.15	0.13	0.04	0.05
			First stage	
Indirect smoking clique			0.49***	0.49***
			(0.03)	(0.00)
Hansen Test			EEI	EEI
Kleibergen Test (reported p-value)			0.00	0.01
F - Statistic				
Indirect smoking clique			184	45
Indirect girls clique				112
Indirect ethnic clique				19
Indirect socioeconomic clique				3

	Smoking II			
	Model I	Model II	Model III	Model IV
	Ols	Ols + FE	IV - A	IV - B-I
Smoking Clique	0.47***	0.41***	0.74***	0.72***
	(0.00)	(0.04)	(0.08)	(0.08)
R^2 adj	0.11	0.13	0.04	0.03
			First stage	
Indirect smoking clique			0.48***	0.47***
			(0.04)	(0.04)
Hansen Test			EEI	EEI
Kleibergen Test (reported p-value)			0.00	0.01
F - Statistic				
Indirect smoking clique			132	32
Indirect girls clique				65
Indirect ethnic clique				19
Indirect socioeconomic clique				4

Drinking I	Model I	Model II	Model III	Model IV
	Ols	Ols + FE	IV - A	IV - B-I
Drinking Clique	0.44***	0.24***	0.65***	0.62***
	(0.03)	(0.04)	(0.07)	(0.08)
R^2 adj	0.13	0.13	0.02	0.02
			First stage	
Indirect drinking clique			0.36***	0.35***
			(0.02)	(0.02)
Hansen Test			EEI	EEI
Kleibergen Test (reported p-value)			0.00	0.01
F - Statistic			174	
Indirect drinking clique				49
Indirect girls clique				111
Indirect ethnic clique				18
Indirect socioeconomic clique				5

Drinking II	Model I	Model II	Model III	Model IV
	Ols	Ols + FE	IV - A	IV - B-I
Drinking Clique	0.33***	0.21***	0.44***	0.45***
	(0.03)	(0.04)	(0.08)	(0.08)
R^2 adj	0.13	0.05	0.03	0.02
			First stage	
Indirect drinking clique			0.39	0.38***
			(0.03)	(0.03)
Hansen Test			EEI	EEI
Kleibergen Test (reported p-value)			0.00	0.01
F - Statistic				
Indirect drinking clique			151	43
Indirect girls clique				68
Indirect ethnic clique				19
Indirect socioeconomic clique				5

Section 2: Non-reciprocal ties, synthetic tables for OLS, OLS and Classroom FE, and IVs

	Smoking I			
	Model I	Model II	Model III	Model IV
	Ols	Ols + FE	IV - A	IV - B-I
Smoking Clique	0.20***	0.08*	0.25***	0.26***
	(0.03)	(0.03)	(0.06)	(0.06)
R^2 adj	0.06	0.02	0.05	0.04
			First stage	
Indirect smoking clique			0.52***	0.52***
			(0.04)	(0.04)
Hansen Test			EEI	EEI
Kleibergen Test (reported p-value)			0.00	0.01
F - Statistic				
Indirect smoking clique			140	35
Indirect girls clique				109
Indirect ethnic clique				24
Indirect socioeconomic clique				6

	Smoking II			
	Model I	Model II	Model III	Model IV
	Ols	Ols + FE	IV - A	IV - B-I
Smoking Clique	0.20***	0.14**	0.49***	0.52***
	(0.03)	(0.04)	(0.09)	(0.10)
R^2 adj	0.06	0.03	0.01	0.07
			First stage	
Indirect smoking clique			0.50***	0.49***
			(0.05)	(0.05)
Hansen Test			EEI	EEI
Kleibergen Test (reported p-value)			0.00	0.01
F - Statistic				
Indirect smoking clique			83	20
Indirect girls clique				79
Indirect ethnic clique				17
Indirect socioeconomic clique				4

Drinking I				
	Model I	Model II	Model III	Model IV
	Ols	Ols + FE	IV - A	IV - B-I
Drinking Clique	0.20***	0.05	0.40***	0.29***
	(0.03)	(0.03)	(0.07)	(0.10)
R^2 adj	0.09	0.12	0.07	0.04
			First stage	
Indirect drinking clique			0.41***	0.41***
			(0.03)	(0.03)
Hansen Test			EEI	EEI
Kleibergen Test (reported p-value)			0.00	0.01
F - Statistic				
Indirect drinking clique			154	57
Indirect girls clique				119
Indirect ethnic clique				27
Indirect socioeconomic clique				6

Drinking II				
	Model I	Model II	Model III	Model IV
	Ols	Ols + FE	IV - A	IV - B-I
Drinking Clique	0.15***	0.06	0.41***	0.39***
	(0.03)	(0.04)	(0.08)	(0.09)
R^2 adj	0.12	0.14	0.07	0.07
			First stage	
Indirect drinking clique			0.42***	0.42***
			(0.03)	(0.04)
Hansen Test			EEI	EEI
Kleibergen Test (reported p-value)			0.00	0.01
F - Statistic				
Indirect drinking clique			116	45
Indirect girls clique				81
Indirect ethnic clique				17
Indirect socioeconomic clique				4

Section 3: Best friends for OLS, OLS and Classroom FE, and IV regressions

Smoking I				
	Model I	Model II	Model III	Model IV
	Ols	Ols + FE	IV - A	IV - B-I
Clique Smoking	0.531***	0.381***	0.663***	0.694***
	(-0.03)	(-0.05)	(-0.07)	(-0.08)
Individual Characteristics				
Girls	0.012	0.012	0.015	0.052
	(-0.04)	(-0.04)	(-0.04)	(-0.05)
Ability	-0.007**	-0.007**	-0.005*	-0.006*
	(0.00)	(0.00)	(0.00)	(0.00)
BMI	0.003	0.003	0.004	0.003
	(0.00)	(0.00)	(0.00)	(0.00)
Non-natives	-0.046**	-0.028	-0.034	-0.048*
	(-0.02)	(-0.02)	(-0.02)	(-0.02)
Socioeconomic Index	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Regular	-0.049	-0.046	-0.034	-0.04
	(-0.03)	(-0.04)	(-0.04)	(-0.04)
Early	-0.085**	-0.082*	-0.06	-0.063
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
Clique Characteristics				
Clique Girls	-0.012	-0.011	-0.013	-0.074
	(-0.04)	(-0.04)	(-0.04)	(-0.06)
Clique non-natives	0.023	0.052	0.05	-0.044
	(-0.02)	(-0.03)	(-0.03)	(-0.19
Clique socioeconomic Index	0.000	0.001	0.001	-0.018
	(0.00)	(0.00)	(0.00)	(-0.01)
Clique Ability	-0.002	-0.001	0.003	0.003
	(0.00)	(0.00)	(0.00)	(0.00)
Clique BMI	0.000	-0.001	0.000	-0.002
	(0.00)	(0.00)	(-0.01)	(-0.01)
Clique Late enrolled	-0.014	-0.017	0.006	0.014
	(-0.02)	(-0.03)	(-0.03)	(-0.03)
R-sqr	0.156	0.199	0.04	0.052
BIC	3154.717	2972.639	2867.796	3142
AIC	3063.147	2887.173	2783.37	3057.611

* $p<0.05$, ** $p<0.01$, *** $p<0.001$
(Standard errors in parentheses)
Source: CILS4EU

Smoking II				
	Model I	Model II	Model III	Model IV
	Ols	Ols + FE	IV - A	IV - B-I
Clique Smoking	0.468***	0.406***	0.736***	0.724***
	(-0.04)	(-0.04)	(-0.08)	(-0.09)
Individual Characteristics				
Girls	0.004	-0.001	-0.013	0.065
	(-0.04)	(-0.04)	(-0.05)	(-0.06)
Ability	-0.006*	-0.004	-0.004	-0.003
	(0.00)	(0.00)	(0.00)	(0.00)
BMI	0.003	0.002	0.003	0.002
	(0.00)	(0.00)	(0.00)	(0.00)
Non-natives	-0.047*	-0.045*	-0.054*	-0.043
	(-0.02)	(-0.02)	(-0.02)	(-0.03)
Socioeconomic Index	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Regular	-0.024	-0.019	-0.029	-0.029
	(-0.05)	(-0.05)	(-0.05)	(-0.05)
Early	-0.066	-0.06	-0.061	-0.056
	(-0.04)	(-0.04)	(-0.04)	(-0.05)
Clique Characteristics				
Clique Girls	0.004	-0.002	0.017	-0.093
	(-0.05)	(-0.05)	(-0.06)	(-0.07)
Clique non-natives	0.04	0.022	0.023	0.43
	(-0.03)	(-0.04)	(-0.04)	(-0.25)
Clique socioeconomic Index	0.000	0.001	0.000	0.01
	(0.00)	(0.00)	(0.00)	(-0.01)
Clique Ability	-0.006	-0.001	0.002	0.007
	(0.00)	(0.00)	(0.00)	(-0.01)
Clique BMI	-0.004	-0.007	-0.006	-0.008
	(0.00)	(0.00)	(-0.01)	(-0.01)
Clique Late enrolled	-0.006	0.000	0.018	0.038
	(-0.03)	(-0.03)	(-0.03)	(-0.04)
R-sqr	0.114	0.206	0.036	0.038
BIC	2871.208	2567.156	2422.685	2598.697
AIC	2782.945	2484.798	2341.437	2517.494

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: CILS4EU

Drinking I				
	Model I	Model II	Model III	Model IV
	Ols	Ols + FE	IV - A	IV - B-I
Clique Drinking	0.445***	0.243***	0.652***	0.629***
	(-0.03)	(-0.05)	(-0.07)	(-0.08)
Individual Characteristics				
Girls	0.03	0.034	0.023	-0.029
	(-0.03)	(-0.03)	(-0.03)	(-0.05)
Ability	-0.004	-0.003	-0.002	-0.002
	(0.00)	(0.00)	(0.00)	(0.00)
BMI	0.004	0.004	0.003	0.003
	(0.00)	(0.00)	(0.00)	(0.00)
Non-natives	-0.123***	-0.091***	-0.088***	-0.084***
	(-0.02)	(-0.02)	(-0.02)	(-0.02)
Socioeconomic Index	0.00	0.001*	0.002*	0.002*
	(0.00)	(0.00)	(0.00)	(0.00)
Regular	-0.062	-0.062	-0.064	-0.059
	(-0.03)	(-0.03)	(-0.04)	(-0.04)
Early	-0.133***	-0.128***	-0.124**	-0.121**
	(-0.03)	(-0.03)	(-0.04)	(-0.04)
Clique Characteristics				
Clique Girls	-0.037	-0.063	-0.034	0.044
	(-0.03)	(-0.03)	(-0.04)	(-0.06)
Clique non-natives	-0.007	0.071	0.102*	0.164
	(-0.03	(-0.04)	(-0.04)	(-0.20)
Clique socioeconomic Index	-0.001	0.000	-0.001	0.009
	(0.00)	(0.00)	(0.00)	(-0.02)
Clique Ability	-0.002	0.002	0.005	0.005
	(0.00)	(0.00)	(0.00)	(0.00)
Clique BMI	-0.003	-0.004	-0.001	0.000
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Clique Late enrolled	-0.028	-0.025	0.019	0.012
	(-0.02)	(-0.03)	(-0.03)	(-0.03)
R-sqr	0.137	0.195	-0.014	0.026
BIC	4389.638	4148.463	3987.503	4016.551
AIC	4298.067	4062.997	3903.087	3932.166

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: CILS4EU

Drinking II				
	Model I Ols	Model II Ols + FE	Model III IV - A	Model IV IV - B-I
Clique Drinking	0.331***	0.210***	0.442***	0.453***
	(-0.03)	(-0.04)	(-0.08)	(-0.08)
Individual Characteristics				
Girls	0.018	0.011	-0.006	-0.001
	(-0.03)	(-0.03)	(-0.03)	(-0.06)
Ability	-0.006*	-0.005*	-0.005	-0.005
	(0.00)	(0.00)	(0.00)	(0.00)
BMI	0.000	0.000	0.001	0.000
	(0.00)	(0.00)	(0.00)	(0.00)
Non-natives	-0.162***	-0.136***	-0.133***	-0.133***
	(-0.02)	(-0.02)	(-0.03)	(-0.03)
Socioeconomic Index	0.002*	0.002	0.001	0.001
	(0.00)	(0.00)	(0.00)	(0.00)
Regular	0.041	0.046	0.051	0.057
	(-0.04)	(-0.04)	(-0.04)	(-0.04)
Early	-0.059	-0.054	-0.047	-0.04
	(-0.04)	(-0.04)	(-0.04)	(-0.04)
Clique Characteristics				
Clique Girls	0.013	-0.004	0.038	0.029
	(-0.04)	(-0.04)	(-0.04)	(-0.09)
Clique non-natives	-0.090*	-0.053	-0.065	0.058
	(-0.04)	(-0.05)	(-0.05)	(-0.22)
Clique socioeconomic Index	0.001	-0.001	-0.002	-0.002
	(0.00)	(0.00)	(0.00)	(-0.01)
Clique Ability	0.000	0.000	-0.001	0.000
	(0.00)	(0.00)	(-0.01)	(-0.01)
Clique BMI	-0.005	-0.004	0.001	-0.001
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Clique Late enrolled	-0.029	-0.038	-0.018	-0.009
	(-0.03)	(-0.03)	(-0.03)	(-0.04)
R-sqr	0.134	0.22	0.034	0.025
BIC	3080.562	2789.588	2588.077	2600.517
AIC	2992.299	2707.23	2506.846	2519.331

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: CILS4EU

Section 4: Non-reciprocal ties for OLS, OLS and Classroom FE, and IV regressions

Smoking I				
	Model I Ols	Model II Ols + FE	Model III IV - A	Model IV IV - B-I
Clique Smoking	0.198***	0.078*	0.251***	0.261***
	(-0.03)	(-0.04)	(-0.06)	(-0.07)
Individual Characteristics				
Girls	0.057	0.054	0.057	0.034
	(-0.03)	(-0.03)	(-0.03)	(-0.04)
Ability	-0.007*	-0.005	-0.006	-0.007
	(0.00)	(0.00)	(0.00)	(0.00)
BMI	0.003	0.002	0.005	0.006
	(0.00)	(0.00)	(0.00)	(0.00)
Non-natives	-0.085***	-0.048*	-0.088***	-0.070*
	(-0.02)	(-0.02)	(-0.02)	(-0.03)
Socioeconomic Index	0.000	0.000	0.000	0.000
	(0.00)	(0.00)	(0.00)	(0,00)
Regular	-0.069	-0.084	-0.102	-0.119*
	(-0.05)	(-0.05)	(-0.05)	(-0.06)
Early	-0.126**	-0.130**	-0.156**	-0.178**
	(-0.04)	(-0.05)	(-0.05)	(-0.06)
Clique Characteristics				
Clique Girls	-0.057	-0.078*	-0.055	-0.019
	(-0.03)	(-0.03)	(-0.03)	(-0.06)
Clique non-natives	0.047	0.058	0.052	-0.038
	(-0.03)	(-0.04)	(-0.04)	(-0.09)
Clique socioeconomic Index	0.001	0.001	0.001	0.005
	(0.00)	(0.00)	(0.00)	(-0.01)
Clique Ability	-0.003	-0.003	-0.001	-0.003
	(0.00)	(0.00)	(0.00)	(0.00)
Clique BMI	0.006	0.005	0.003	0.006
	(0.00)	(0.00)	(0.00)	(-0.01)
Clique Late enrolled	-0.01	-0.004	-0.01	-0.026
	(-0.02)	(-0.02)	(-0.02)	(-0.02)
R-sqr	0.067	0.187	0.058	0.043
BIC	1893.022	1635.741	1608.697	1612.617
AIC	1810.994	1559.23	1529.916	1534.018

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: CILS4EU

Smoking II				
	Model I Ols	Model II Ols + FE	Model III IV - A	Model IV IV - B-I
Clique Smoking	0.202***	0.142**	0.491***	0.528***
	(-0.03)	(-0.04)	(-0.09)	(-0.11)
Individual Characteristics				
Girls	-0.009	-0.001	-0.003	-0.047
	(-0.03)	(-0.03)	(-0.03)	(-0.05)
Ability	-0.007*	-0.002	-0.004	-0.007
	(0.00)	(0.00)	(0.00)	(0.00)
BMI	0.003	0.003	0.003	0.004
	(0.00)	(0.00)	(0.00)	(-0.01)
Non-natives	-0.094**	-0.086**	-0.090**	-0.096*
	(-0.03)	(-0.03	(-0.03)	(-0.05)
Socioeconomic Index	0.001	0.001)	0.000	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)
Regular	-0.016	-0.021)	-0.071	-0.087
	(-0.06)	(-0.07)	(-0.07)	(-0.07)
Early	-0.095	-0.092	-0.135*	-0.162*
	(-0.06)	(-0.07)	(-0.07)	(-0.07)
Clique Characteristics				
Clique Girls	0.029	-0.014	0.022	0.068
	(-0.04)	(-0.04	(-0.04)	(-0.06)
Clique non-natives	0.088**	0.085	0.113*	0.12
	(-0.03)	(-0.04	(-0.04)	(-0.14)
Clique socioeconomic Index	0.001	0.002	0.001	0.011
	(0.00)	(0.00	(0.00)	(-0.01)
Clique Ability	0.000	0.002	0.006	0.002
	(0.00)	(0.00	(-0.01)	(-0.01)
Clique BMI	-0.003	-0.001	-0.006	-0.005
	(0.00)	(0.00	(-0.01)	(-0.01)
Clique Late enrolled	-0.007	-0.002	0.01	-0.004
	(-0.02)	(-0.02	(-0.03)	(-0.03)
R-sqr	0.061	0.24	0.001	0.078
BIC	1615.157	1296.86	1409.11	1486.102
AIC	1536.601	1223.663	1333.875	1411.015

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: CILS4EU

	Drinking I			
	Model I	Model II	Model III	Model IV
	Ols	Ols + FE	IV - A	IV - B-I
Clique drinking	0.201***	0.055	0.401***	0.295**
	(-0.03)	(-0.03)	(-0.07)	(-0.10)
Individual Characteristics				
Girls	0.053	0.048	0.029	0.018
	(-0.03)	(-0.03)	(-0.03)	(-0.04)
Ability	-0.005	-0.004	-0.005	-0.007
	(0.00)	(0.00)	(0.00)	(0.00)
BMI	0.002	0.003	0.002	0.004
	(0.00)	(0.00)	(0.00)	(0.00)
Non-natives	-0.205***	-0.133***	-0.177***	-0.140***
	(-0.03)	(-0.03)	(-0.03)	(-0.04)
Socioeconomic Index	0.001	0.002*	0.001	0.001
	(0.00)	(0.00)	(0.00)	(0.00)
Regular	-0.110**	-0.130**	-0.136**	-0.132**
	(-0.04)	(-0.05)	(-0.05)	(-0.05)
Early	-0.190***	-0.195***	-0.205***	-0.214***
	(-0.04)	(-0.05)	(-0.04)	(-0.05)
Clique Characteristics				
Clique Girls	-0.052	-0.082*	-0.042	-0.035
	(-0.03)	(-0.04)	(-0.03)	(-0.05)
Clique non-natives	0.016	0.110**	0.084*	-0.141
	(-0.04)	(-0.04)	(-0.04)	(-0.16
Clique socioeconomic Index	0.000	0.001	0.001	0.007
	(0.00)	(0.00)	(0.00)	(-0.01)
Clique Ability	0.001	0.002	0.001	-0.005
	(0.00)	(-0.01)	(0.00)	(-0.01)
Clique BMI	0.005	0.005	0.005	0.012
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Clique Late enrolled	-0.008	0.003	0.014	-0.022
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
R-sqr	0.094	0.229	0.075	0.049
BIC	2451.488	2153.082	2024.882	2041.472
AIC	2369.461	2076.571	1946.07	1962.84

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: CILS4EU

Drinking II				
	Model I	Model II	Model III	Model IV
	Ols	Ols + FE	IV - A	IV - B-I
Clique drinking	0.153***	0.064	0.414***	0.393***
	(-0.04)	(-0.04)	(-0.08)	(-0.10)
Individual Characteristics				
Girls	0.029	0.019	0.004	0.021
	(-0.04)	(-0.04)	(-0.04)	(-0.05)
Ability	-0.005	-0.003	-0.002	-0.002
	(0.00)	(0.00)	(0.00)	(0.00)
BMI	0.00	0.00	0.00	0.00
	(0.00)	(-0.01)	(0.00)	(0.00)
Non-natives	-0.246***	-0.194***	-0.216***	-0.202***
	(-0.03)	(-0.03)	(-0.03)	(-0.04)
Socioeconomic Index	0.003**	0.002*	0.002*	0.002
	(0.00)	(0.00)	(0.00)	(0.00)
Regular	-0.015	-0.035	-0.052	-0.048
	(-0.05)	(-0.06	(-0.05)	(-0.05)
Early	-0.094	-0.11	-0.121*	-0.118*
	(-0.05)	(-0.06)	(-0.05)	(-0.05)
Clique Characteristics				
Clique Girls	0.016	-0.006	0.031	0.005
	(-0.04)	(-0.04)	(-0.04)	(-0.06)
Clique non-natives	-0.059	-0.032	-0.013	-0.078
	(-0.04)	(-0.05)	(-0.04)	(-0.17)
Clique socioeconomic Index	0.001	0.000	0.001	0.000
	(0.00)	(0.00)	(0.00)	(-0.01)
Clique Ability	-0.005	-0.011*	-0.003	-0.003
	(0.00)	(0.00)	(0.00)	(-0.01)
Clique BMI	0.004	0.004	0.005	0.007
	(-0.01)	(-0.01)	(-0.01)	(-0.01)
Clique Late enrolled	-0.03	-0.044	0.000	-0.005
	(-0.03)	(-0.03)	(-0.03)	(-0.03)
R-sqr	0.125	0.277	0.077	0.078
BIC	1652.766	1366.864	1403.06	1388.198
AIC	1574.21	1293.666	1327.771	1313.057

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: CILS4EU

Section 5: Descriptive Statistics

Best friends, Time I

	%			
Smokers	22.33			
Non-smokers	77.67			
Drinkers	43.84			
Non-drinkers	56.16			
Girls	48.1			
Boys	51.9			
Natives	63.08			
Non-natives	35.92			
Late-enrolled	6.47			
Regular-enrolled	40.54			
Early-enrolled	52.99			
	MEAN	SD	MIN	MAX
Socioeconomic Index	47.6	13.0	13.0	78.2
BMI	20.3	2.9	12.4	43.0
Ability	19.6	3.8	2	27
Clique smoking	0.25	0.36	0	1
Clique drinking	0.60	0.41	0	1
Clique ability	19.5	2.8	4	27
Clique BMI	20.2	1.9	13.8	37.9
Clique Late enrolled	0.1	0.1	0	1
Clique Girls	0.5	0.4	0	1
Clique non-natives	0.4	0.3	0	1
Clique Socioeconomic Index	47.8	8.6	17.7	78.2
Indirect Smoking clique	0.2	0.3	0	1
Indirect drinking clique	0.5	0.4	0	1
Indirect girls clique	0.5	0.4	0	1
Indirect ethnic clique	0.7	0.3	0	1
Indirect socioeconomic clique	0.5	0.3	0	1

Best friends, Time II

	%			
Smokers	25,76			
Non-smokers	74.24			
Drinkers	70.73			
Non-drinkers	29.27			
Girls	47.68			
Boys	52.32			
Natives	64.52			
Non-natives	35.48			
Late-enrolled	5.42			
Regular-enrolled	40.08			
Early-enrolled	54.69			
	MEAN	SD	MIN	MAX
Socioeconomic Index	47.6	12.9	16	78.2
BMI	20.2	3.0	14.0	43.0
Ability	19.8	3.8	2	27
Clique smoking	0.24	0.36	0	1
Clique drinking	0.60	0.41	0	1
Clique ability	19.6	2.8	4	27
Clique BMI	20.2	1.9	14.5	37.9
Clique Late enrolled	0.1	0.1	0	1
Clique Girls	0.5	0.5	0	1
Clique non-natives	0.4	0.3	0	1
Clique Socioeconomic Index	47.9	8.6	20	78.2
Indirect Smoking clique	0.2	0.3	0	1
Indirect drinking clique	0.6	0.3	0	1
Indirect girls clique	0.5	0.4	0	1
Indirect ethnic clique	0.8	0.3	0	1
Indirect socioeconomic clique	0.5	0.3	0	1

Non-reciprocal ties, Time I

	%			
Smokers	22.32			
Non-smokers	77.68			
Drinkers	56.68			
Non-drinkers	43.32			
Girls	53.65			
Boys	46.35			
Natives	63.13			
Non-natives	36.87			
Late-enrolled	6.68			
Regular-enrolled	41.5			
Early-enrolled	51.83			
	MEAN	SD	MIN	MAX
Socioeconomic Index	47.4	13.1	16.0	78.2
BMI	20.3	3.0	12.4	43.0
Ability	19.6	3.8	2	27
Clique smoking	0.25	0.36	0	1
Clique drinking	0.60	0.47	0	1
Clique ability	19.4	3.4	6	27
Clique BMI	20.2	2.4	14.5	36.3
Clique Late enrolled	0.1	0.2	0	1
Clique Girls	0.5	0.5	0	1
Clique non-natives	0.4	0.4	0	1
Clique Socioeconomic Index	48.5	10.9	20.0	78.2
Indirect Smoking clique	0.2	0.3	0	1
Indirect drinking clique	0.6	0.4	0	1
Indirect girls clique	0.4	0.4	0	1
Indirect ethnic clique	0.7	0.4	0	1
Indirect socioeconomic clique	0.6	0.4	0	1

Non-reciprocal ties, Time II

	%			
Smokers	19.8			
Non-smokers	80.3			
Drinkers	54.5			
Non-drinkers	45.5			
Girls	53.1			
Boys	46.9			
Natives	64.5			
Non-natives	35.5			
Late-enrolled	5.5			
Regular-enrolled	41.7			
Early-enrolled	52.9			
	MEAN	SD	MIN	MAX
Socioeconomic Index	47.4	13.1	16.0	78.2
BMI	20.3	3.0	14.1	43.0
Ability	19.8	3.7	2	27
Clique smoking	0.24	0.36	0	1
Clique drinking	0,60	0.41	0	1
Clique ability	19.5	3.4	5	27
Clique BMI	20.2	2.4	14.5	36.3
Clique Late enrolled	0.1	0.2	0	1
Clique Girls	0.5	0.5	0	1
Clique non-natives	0.4	0.4	0	1
Clique Socioeconomic Index	48.5	11.2	20.0	78.2
Indirect Smoking clique	0.2	0.3	0	1
Indirect drinking clique	0.6	0.4	0	1
Indirect girls clique	0.4	0.4	0	1
Indirect ethnic clique	0.8	0.4	0	1
Indirect socioeconomic clique	0.5	0.4	0	1

Appendix Chapter IV

Note that in all models references are native, male, speaking Italian, and with a standard enrolment. I report all models used for the main analysis and for graphical visualization in the chapter IV. I do not report the variable language to achieve a better layout. Language is only a control variable and does not show interesting patterns for my empirical chapter. Available upon requests.

Table 1: Ethnic share, School Fixed Effects

	Bullied	Bullying	Social Integration	Extrinsic Motivation
	Model I	Model I	Model I	Model I
Ethnic Share	0.02***	0.03***	0.00	0.02**
	(0.01)	(0.01)	(0.01)	(0.01)
Individual Characteristics				
Girls	-0.12***	-0.27***	0.07***	-0.16***
	(0.01)	(0.01)	(0.01)	(0.01)
2-Generation	-0.09***	-0.12**	-0.01	-0.01
	(0.02)	(0.04)	(0.03)	(0.06)
1-Genearation	-0.05**	-0.11***	-0.03	0.00
	(0.02)	(0.02)	(0.03)	(0.03)
Early Enrollment	-0.01	-0.13	0.06	0.08
	(0.06)	(0.13)	(0.07)	(0.04)
Late Enrollment	-0.01	0.04	-0.1	0.03
	(0.04)	(0.04)	(0.07)	(0.04)
Italian teacher's mark	-0.07***	-0.08***	0.04***	-0.11***
	(0.01)	(0.01)	(0.01)	(0.01)
Math teacher's mark	-0.07***	-0.01	0.07***	-0.08***
	(0.01)	(0.01)	(0.01)	(0.01)
ESCS	0.00	0.00	0.02**	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Classroom Characteristics				
Share of girls	-0.11***	-0.07*	0.06	-0.07*
	(0.03)	(0.03)	(0.03)	(0.03)
Share of Late enrolled	0.09	0.05	0.00	0.18
	(0.10)	(0.12)	(0.11)	(0.09)
Average classroom ESCS	-0.02	-0.05**	0.04**	-0.13***
	(0.01)	(0.02)	(0.01)	(0.02)
Average in language test	0.06***	0.05**	-0.01	0.09***
	(0.02)	(0.02)	(0.02)	(0.02)
Average in math test	0.02	-0.03	-0.02	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
Class-size	0.00	-0.00***	-0.01***	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.64***	0.76***	-0.51***	0.65***
	(0.07)	(0.08)	(0.07)	(0.09)
R-sqr	0.04	0.05	0.03	0.06
F-Statistic	0.000	0.000	0.000	0.000
BIC	420427.1	396824.7	400228.4	446596.2
AIC	420121.3	396518.9	399922.6	446290.4
Obs.	197444	197444	197444	197444
School FE	YES	YES	YES	YES

* $p<0.05$, ** $p<0.01$, *** $p<0.001$
(Standard errors in parentheses)
Source: INVALSI

Table 2: Shannon Equitability Index, School Fixed Effects

	Bullied	**Bullying**	**Social Integration**	**Extrinsic Motivation**
	Model II	Model II	Model II	Model II
Diversity Index	0.01**	0.02***	-0.01	0.01
	(0.00)	(0.00)	(0.00)	(0.01)
Individual Characteristics				
Girls	-0.12***	-0.27***	0.07***	-0.16***
	(0.01)	(0.01)	(0.01)	(0.01)
2-Generation	-0.08***	-0.11**	-0.01	-0.01
	(0.02)	(0.04)	(0.03)	(0.06)
1-Genearation	-0.05*	-0.10***	-0.03	0.01
	(0.02)	(0.02)	(0.03)	(0.03)
Early Enrollment	-0.01	-0.13	0.06	0.08
	(0.06)	(0.13)	(0.07)	(0.04)
Late Enrollment	-0.01	0.04	-0.1	0.03
	(0.04)	(0.04)	(0.07)	(0.04)
Italian teacher's mark	-0.07***	-0.08***	0.04***	-0.11***
	(0.01)	(0.01)	(0.01)	(0.01)
Math teacher's mark	-0.07***	-0.01	0.07***	-0.08***
	(0.01)	(0.01)	(0.01)	(0.01)
ESCS	0.00	0.00	0.02**	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Classroom Characteristics				
Share of girls	-0.11***	-0.07*	0.06	-0.07*
	(0.03)	(0.03)	(0.03)	(0.03)
Share of Late enrolled	0.12	0.07	0.01	0.21*
	(0.10)	(0.12)	(0.10)	(0.09)
Average classroom ESCS	-0.03	-0.05**	0.04**	-0.13***
	(0.01)	(0.02)	(0.01)	(0.02)
Average in language test	0.05***	0.05**	-0.01	0.09***
	(0.02)	(0.02)	(0.01)	(0.02)
Average in math test	0.02	-0.03	-0.02	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
Class-size	0.00	-0.00***	-0.01***	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.63***	0.74***	-0.50***	0.64***
	(0.07)	(0.08)	(0.07)	(0.09)
R-sqr	0.04	0.05	0.03	0.06
F-Statistic	0.000	0.000	0.000	0.000
BIC	420437.6	396815.9	400225	446607
AIC	420131.8	396510.1	399919.2	446301.2
Obs.	197444	197444	197444	197444
School FE	YES	YES	YES	YES

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Table 3: Joint inclusion of Ethnic share and Diversity Index				
	Bullied	**Bullying**	**Social Integration**	**Extrinsic Motivation**
	Model III	Model III	Model III	Model III
Ethnic Share	0.02**	0.02**	0.00	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)
Diversity Index	0.01	0.02**	-0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Individual Characteristics				
Girls	-0.12***	-0.27***	0.07***	-0.16***
	(0.01)	(0.01)	(0.01)	(0.01)
2-Generation	-0.09***	-0.11**	-0.01	-0.01
	(0.02)	(0.04)	(0.03)	(0.06)
1-Genearation	-0.05**	-0.10***	-0.03	0.00
	(0.02)	(0.02)	(0.03)	(0.03)
Early Enrollment	-0.01	-0.13	0.06	0.08
	(0.06)	(0.13)	(0.07)	(0.04)
Late Enrollment	-0.01	0.04	-0.1	0.03
	(0.04)	(0.04)	(0.07)	(0.04)
Italian teacher's mark	-0.07***	-0.08***	0.04***	-0.11***
	(0.01)	(0.01)	(0.01)	(0.01)
Math teacher's mark	-0.07***	-0.01	0.07***	-0.08***
	(0.01)	(0.01)	(0.01)	(0.01)
ESCS	0.00	0.00	0.02**	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Classroom Characteristics				
Share of girls	-0.11***	-0.07*	0.06	-0.07*
	(0.03)	(0.03)	(0.03)	(0.03)
Share of Late enrolled	0.08	0.03	0.01	0.18
	(0.10)	(0.12)	(0.11)	(0.09)
Average classroom ESCS	-0.02	-0.05**	0.04**	-0.13***
	(0.01)	(0.02)	(0.01)	(0.02)
Average in language test	0.06***	0.05**	-0.01	0.09***
	(0.02)	(0.02)	(0.02)	(0.02)
Average in math test	0.02	-0.03	-0.02	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
Class-size	0.00	-0.00***	-0.01***	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.63***	0.74***	-0.50***	0.65***
	(0.07)	(0.08)	(0.07)	(0.09)
R-sqr	0.04	0.05	0.03	0.06
F-Statistic	0.000	0.000	0.000	0.000
BIC	420435.5	396816	400236.6	446607.6
AIC	420119.5	396500.1	399920.6	446291.6
Obs.	197444	197444	197444	197444
School FE	YES	YES	YES	YES

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Table 4: Ethnic share interacted with Diversity Index in tertiles				
	Bullied	**Bullying**	**Social Integration**	**Extrinsic Motivation**
	Model IV	Model IV	Model IV	Model IV
Ethnic Share	0.06***	0.05***	-0.01	0.02
	(0.01)	(0.01)	(0.01)	(0.01)
Medium	-0.01	0.00	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
High	-0.01	0.01	-0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Medium # Ethnic Share	-0.06***	-0.02	0.02	0.01
	(0.02)	(0.01)	(0.01)	(0.02)
High # Ethnic Share	-0.05***	-0.04**	0.02	0
	(0.01)	(0.01)	(0.01)	(0.02)
Individual Characteristics				
Girls	-0.12***	-0.27***	0.07***	-0.16***
	(0.01)	(0.01)	(0.01)	(0.01)
2-Generation	-0.09***	-0.12**	-0.01	-0.01
	(0.02)	(0.04)	(0.03)	(0.06)
1-Genearation	-0.05**	-0.11***	-0.03	0
	(0.02)	(0.02)	(0.03)	(0.03)
Early Enrollment	-0.01	-0.13	0.06	0.08
	(0.06)	(0.13)	(0.07)	(0.04)
Late Enrollment	-0.01	0.04	-0.1	0.03
	(0.04)	(0.04)	(0.07)	(0.04)
Italian teacher's mark	-0.07***	-0.08***	0.04***	-0.11***
	(0.01)	(0.01)	(0.01)	(0.01)
Math teacher's mark	-0.07***	-0.01	0.07***	-0.08***
	(0.01)	(0.01)	(0.01)	(0.01)
ESCS	0.00	0.00	0.02**	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Classroom Characteristics				
Share of girls	-0.11***	-0.07*	0.06	-0.07*
	(0.03)	(0.03)	(0.03)	(0.03)
Share of Late enrolled	0.07	0.03	0.01	0.18
	(0.10)	(0.12)	(0.11)	(0.09)
Average classroom ESCS	-0.02	-0.05**	0.04**	-0.13***
	(0.01)	(0.02)	(0.01)	(0.02)
Average in language test	0.05***	0.05**	-0.01	0.09***
	(0.02)	(0.02)	(0.02)	(0.02)
Average in math test	0.02	-0.03	-0.02	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
Class-size	0.00	-0.00***	-0.01***	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.66***	0.75***	-0.50***	0.64***
	(0.07)	(0.08)	(0.07)	(0.09)
R-sqr	0.04	0.05	0.03	0.06
F-Statistic	0.000	0.000	0.000	0.000
BIC	420439.3	396846.8	400264.3	446641.8
AIC	420092.8	396500.2	399917.8	446295.2
Obs.	197444	197444	197444	197444
School FE	YES	YES	YES	YES

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Table 5: Squared specification with Joint inclusion of Ethnic share and Diversity Index, School Fixed Models				
	Bullied	**Bullying**	**Social Integration**	**Extrinsic Motivation**
	Model V	Model V	Model V	Model V
Ethnic Share	0.03***	0.03***	0.00	0.02**
	(0.01)	(0.01)	(0.01)	(0.01)
Ethnic Share # Ethnic Share	-0.01*	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Individual Characteristics				
Girls	-0.12***	-0.27***	0.07***	-0.16***
	(0.01)	(0.01)	(0.01)	(0.01)
2-Generation	-0.09***	-0.12**	-0.01	-0.01
	(0.02)	(0.04)	(0.03)	(0.06)
1-Genearation	-0.05**	-0.11***	-0.03	0
	(0.02)	(0.02)	(0.03)	(0.03)
Early Enrollment	-0.01	-0.13	0.06	0.08
	(0.06)	(0.13)	(0.07)	(0.04)
Late Enrollment	-0.01	0.04	-0.1	0.03
	(0.04)	(0.04)	(0.07)	(0.04)
Italian teacher's mark	-0.07***	-0.08***	0.04***	-0.11***
	(0.01)	(0.01)	(0.01)	(0.01)
Math teacher's mark	-0.07***	-0.01	0.07***	-0.08***
	(0.01)	(0.01)	(0.01)	(0.01)
ESCS	0.00	0.00	0.02**	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Classroom Characteristics				
Share of girls	-0.11***	-0.07*	0.06	-0.07*
	(0.03)	(0.03)	(0.03)	(0.03)
Share of Late enrolled	0.08	0.05	0.00	0.18
	(0.10)	(0.12)	(0.11)	(0.09)
Average classroom ESCS	-0.02	-0.05**	0.04**	-0.13***
	(0.01)	(0.02)	(0.01)	(0.02)
Average in language test	0.06***	0.05**	-0.01	0.09***
	(0.02)	(0.02)	(0.02)	(0.02)
Average in math test	0.02	-0.03	-0.02	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
Class-size	-0.00*	-0.00***	-0.01***	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.65***	0.76***	-0.51***	0.66***
	(0.07)	(0.08)	(0.07)	(0.09)
R-sqr	0.04	0.05	0.03	0.06
F-Statistic	0.000	0.000	0.000	0.000
BIC	420431.3	396835.8	400238.6	446607.2
AIC	420115.3	396519.8	399922.6	446291.2
Obs.	197444	197444	197444	197444
School FE	YES	YES	YES	YES

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Table 6: Asymmetric effect of Ethnic share over Ethnic Status				
	Bullied	**Bullying**	**Social Integration**	**Extrinsic Motivation**
	Model VI	Model VI	Model VI	Model VI
Ethnic Share	0.03***	0.02***	0.00	0.02**
	(0.01)	(0.01)	(0.01)	(0.01)
2-Generation	-0.09***	-0.13**	-0.03	-0.04
	(0.02)	(0.05)	(0.04)	(0.08)
1-Genearation	-0.04	-0.11***	-0.04	0.00
	(0.02)	(0.02)	(0.03)	(0.03)
2-Generation # Ethnic Share	-0.01	0.03	0.04	0.06
	(0.01)	(0.03)	(0.02)	(0.05)
1-Genearation # Ethnic Share	-0.03*	0	0.02	0.02
	(0.01)	(0.01)	(0.01)	(0.01)
Individual Characteristics				
Girls	-0.12***	-0.27***	0.07***	-0.16***
	(0.01)	(0.01)	(0.01)	(0.01)
Early Enrollment	-0.01	-0.13	0.05	0.07
	(0.06)	(0.13)	(0.07)	(0.04)
Late Enrollment	-0.01	0.04	-0.1	0.03
	(0.04)	(0.04)	(0.07)	(0.04)
Italian teacher's mark	-0.07***	-0.08***	0.04***	-0.11***
	(0.01)	(0.01)	(0.01)	(0.01)
Math teacher's mark	-0.07***	-0.01	0.07***	-0.08***
	(0.01)	(0.01)	(0.01)	(0.01)
ESCS	0.00	0.00	0.02**	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Classroom Characteristics				
Share of girls	-0.11***	-0.07**	0.06	-0.07*
	(0.03)	(0.03)	(0.03)	(0.03)
Share of Late enrolled	0.09	0.05	0	0.18*
	(0.10)	(0.12)	(0.11)	(0.09)
Average classroom ESCS	-0.02	-0.05**	0.04**	-0.13***
	(0.01)	(0.02)	(0.01)	(0.02)
Average in language test	0.06***	0.05**	-0.01	0.09***
	(0.02)	(0.02)	(0.01)	(0.02)
Average in math test	0.02	-0.03	-0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.02)
Class-size	0.00	-0.00***	-0.01***	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.64***	0.76***	-0.51***	0.65***
	(0.07)	(0.08)	(0.07)	(0.09)
R-sqr	0.04	0.05	0.03	0.06
F-Statistic	0.000	0.000	0.000	0.000
BIC	420446.3	396838.8	400234.4	446589.6
AIC	420120.1	396512.7	399908.2	446263.5
Obs.	197444	197444	197444	197444
School FE	YES	YES	YES	YES

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Table 7: Asymmetric effect of Shannon Equitability Index over Ethnic Status

	Bullied	**Bullying**	**Social Integration**	**Extrinsic Motivation**
	Model VII	Model VII	Model VII	Model VII
Diversity Index	0.02***	0.02***	-0.01*	0.01
	(0.00)	(0.01)	(0.00)	(0.01)
2-Generation	-0.08***	-0.11**	-0.02	-0.01
	(0.02)	(0.04)	(0.03)	(0.07)
1-Genearation	-0.04*	-0.09***	-0.03	0.01
	(0.02)	(0.02)	(0.03)	(0.02)
2-Generation # Diversity Index	-0.03	0.01	0.04*	0.03
	(0.02)	(0.03)	(0.02)	(0.05)
1-Genearation # Diversity Index	-0.04**	-0.02*	0.02	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Individual Characteristics				
Girls	-0.12***	-0.27***	0.07***	-0.16***
	(0.01)	(0.01)	(0.01)	(0.01)
Early Enrollment	-0.01	-0.13	0.06	0.07
	(0.06)	(0.13)	(0.07)	(0.04)
Late Enrollment	-0.01	0.04	-0.1	0.03
	(0.04)	(0.04)	(0.07)	(0.04)
Italian teacher's mark	-0.07***	-0.08***	0.05***	-0.11***
	(0.01)	(0.01)	(0.01)	(0.01)
Math teacher's mark	-0.07***	-0.01	0.07***	-0.08***
	(0.01)	(0.01)	(0.01)	(0.01)
ESCS	0.00	0.00	0.02**	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Classroom Characteristics				
Share of girls	-0.11***	-0.07*	0.06	-0.07*
	(0.03)	(0.03)	(0.03)	(0.03)
Share of Late enrolled	0.12	0.07	0.01	0.21*
	(0.10)	(0.12)	(0.10)	(0.09)
Average classroom ESCS	-0.03	-0.05**	0.04**	-0.13***
	(0.01)	(0.02)	(0.01)	(0.02)
Average in language test	0.05***	0.05**	-0.01	0.09***
	(0.02)	(0.02)	(0.01)	(0.02)
Average in math test	0.02	-0.03	-0.02	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
Class-size	0.00	-0.00***	-0.01***	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.63***	0.74***	-0.50***	0.64***
	(0.07)	(0.08)	(0.07)	(0.09)
R-sqr	0.04	0.05	0.03	0.06
F-Statistic	0.000	0.000	0.000	0.000
BIC	420446.5	396833.8	400224.9	446619
AIC	420120.3	396507.6	399898.7	446292.9
Obs.	197444	197444	197444	197444
School FE	YES	YES	YES	YES

* p<0.05, ** p<0.01, *** p<0.001

(Standard errors in parentheses)

Source: INVALSI

Table 1: Ethnic share, School Fixed Effects		
	Performance in Language	**Performance in Mathematics**
	Model I	Model I
Ethnic Share	-0.02*	-0.02
	(0.01)	(0.01)
Individual Characteristics		
Girls	0.07***	-0.14***
	(0.01)	(0.01)
2-Generation	-0.12**	0.04
	(0.04)	(0.09)
1-Genearation	-0.19***	-0.07
	(0.03)	(0.04)
Early Enrollment	-0.09	-0.14
	(0.08)	(0.11)
Late Enrollment	-0.23***	-0.22**
	(0.06)	(0.07)
ESCS	0.22***	0.20***
	(0.01)	(0.01)
Classroom Characteristics		
Share of girls	0.04	-0.01
	(0.05)	(0.05)
Share of Late enrolled	-0.25	-0.07
	(0.15)	(0.19)
Average classroom ESCS	0.07**	0.07**
	(0.03)	(0.03)
Class-size	-0.00**	0.00
	(0.00)	(0.00)
Constant	0.21***	0.32***
	(0.04)	(0.04)
R-sqr	0.06	0.06
F-Statistic	0.000	0.000
BIC	472680.6	470019.8
AIC	472415.5	469754.8
Obs.	197444	197444
School FE	YES	YES

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Table 2: Shannon Equitability Index, School Fixed Effects		
	Performance in Language	**Performance in Mathematics**
	Model II	Model II
Diversity Index	-0.03***	-0.01
	(0.01)	(0.01)
Individual Characteristics		
Girls	0.07***	-0.14***
	(0.01)	(0.01)
2-Generation	-0.12**	0.04
	(0.04)	(0.09)
1-Genearation	-0.19***	-0.07*
	(0.03)	(0.03)
Early Enrollment	-0.09	-0.14
	(0.08)	(0.11)
Late Enrollment	-0.23***	-0.22**
	(0.06)	(0.07)
ESCS	0.22***	0.20***
	(0.01)	(0.01)
Classroom Characteristics		
Share of girls	0.03	-0.01
	(0.05)	(0.05)
Share of Late enrolled	-0.22	-0.09
	(0.15)	(0.18)
Average classroom ESCS	0.07**	0.07**
	(0.03)	(0.03)
Class-size	-0.01**	-0.00*
	(0.00)	(0.00)
Constant	0.23***	0.33***
	(0.04)	(0.04)
R-sqr	0.06	0.06
F-Statistic	0.000	0.000
BIC	472643.8	470021.4
AIC	472378.8	469756.4
Obs.	197444	197444
School FE	YES	YES

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Table 3: Joint inclusion of Ethnic share and Diversity Index

	Performance in Language	Performance in Mathematics
	Model III	Model III
Ethnic Share	0.00	-0.01
	(0.01)	(0.01)
Diversity Index	-0.03**	-0.01
	(0.01)	(0.01)
Individual Characteristics		
Girls	0.07***	-0.14***
	(0.01)	(0.01)
2-Generation	-0.12**	0.04
	(0.04)	(0.09)
1-Genearation	-0.19***	-0.07
	(0.03)	(0.04)
Early Enrollment	-0.09	-0.14
	(0.08)	(0.11)
Late Enrollment	-0.23***	-0.22**
	(0.06)	(0.07)
ESCS	0.22***	0.20***
	(0.01)	(0.01)
Classroom Characteristics		
Share of girls	0.04	-0.01
	(0.05)	(0.05)
Share of Late enrolled	-0.22	-0.06
	(0.15)	(0.18)
Average classroom ESCS	0.07**	0.07**
	(0.03)	(0.03)
Class-size	-0.01**	0.00
	(0.00)	(0.00)
Constant	0.22***	0.32***
	(0.04)	(0.04)
R-sqr	0.06	0.06
F-Statistic	0.000	0.000
BIC	472655.9	470029.6
AIC	472380.6	469754.3
Obs.	197444	197444
School FE	YES	YES

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Table 4: Ethnic share interacted with Diversity Index in tertiles

	Performance in Language	**Performance in Mathematics**
	Model IV	Model IV
Ethnic Share	-0.03	-0.03
	(0.02)	(0.02)
Medium	0.00	0.00
	(0.02)	(0.02)
High	-0.03	-0.01
	(0.02)	(0.02)
Medium # Ethnic Share	0.04	0.04
	(0.02)	(0.02)
High # Ethnic Share	0.02	0.02
	(0.03)	(0.02)
Individual Characteristics		
Girls	0.07***	-0.14***
	(0.01)	(0.01)
2-Generation	-0.12**	0.04
	(0.04)	(0.09)
1-Genearation	-0.19***	-0.07
	(0.03)	(0.04)
Early Enrollment	-0.09	-0.14
	(0.08)	(0.11)
Late Enrollment	-0.23***	-0.22**
	(0.06)	(0.07)
ESCS	0.22***	0.20***
	(0.01)	(0.01)
Classroom Characteristics		
Share of girls	0.04	-0.01
	(0.05)	(0.05)
Share of Late enrolled	-0.22	-0.05
	(0.15)	(0.19)
Average classroom ESCS	0.07**	0.07*
	(0.03)	(0.03)
Class-size	-0.01**	0
	(0.00)	(0.00)
Constant	0.22***	0.32***
	(0.04)	(0.04)
R-sqr	0.06	0.06
F-Statistic	0.000	0.000
BIC	472695.7	470051.6
AIC	472389.9	469745.8
Obs.	197444	197444
School FE	YES	YES

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Table 5: Squared specification with Joint inclusion of Ethnic share and Diversity Index, School Fixed Models

	Performance in Language	**Performance in Mathematics**
	Model IV	Model V
Ethnic Share	-0.02	-0.01
	(0.01)	(0.01)
Ethnic Share # Ethnic Share	0.00	0.00
	(0.00)	(0.01)
Individual Characteristics		
Girls	0.07***	-0.14***
	(0.01)	(0.01)
2-Generation	-0.12**	0.04
	(0.04)	(0.09)
1-Genearation	-0.19***	-0.07
	(0.03)	(0.04)
Early Enrollment	-0.09	-0.14
	(0.08)	(0.11)
Late Enrollment	-0.23***	-0.22**
	(0.06)	(0.07)
ESCS	0.22***	0.20***
	(0.01)	(0.01)
Classroom Characteristics		
Share of girls	0.04	-0.01
	(0.05)	(0.05)
Share of Late enrolled	-0.25	-0.07
	(0.15)	(0.19)
Average classroom ESCS	0.07**	0.07**
	(0.03)	(0.03)
Class-size	-0.00**	0.00
	(0.00)	(0.00)
Constant	0.21***	0.32***
	(0.04)	(0.04)
R-sqr	0.06	0.06
F-Statistic	0.000	0.000
BIC	472692.1	470030.4
AIC	472416.9	469755.2
Obs.	197444	197444
School FE	YES	YES

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Table 6: Asymmetric effect of Ethnic share over Ethnic Status		
	Performance in Language	**Performance in Mathematics**
	Model VI	Model VI
Ethnic Share	-0.01	0.00
	(0.01)	(0.01)
2-Generation	-0.09	0.09
	(0.05)	(0.11)
1-Genearation	-0.17***	-0.05
	(0.03)	(0.03)
2-Generation # Ethnic Share	-0.06*	-0.11
	(0.03)	(0.06)
1-Genearation # Ethnic Share	-0.04**	-0.06***
	(0.01)	(0.02)
Individual Characteristics		
Girls	0.07***	-0.14***
	(0.01)	(0.01)
Early Enrollment	-0.09	-0.14
	(0.07)	(0.11)
Late Enrollment	-0.23***	-0.22**
	(0.06)	(0.07)
ESCS	0.22***	0.20***
	(0.01)	(0.01)
Classroom Characteristics		
Share of girls	0.04	-0.01
	(0.05)	(0.05)
Share of Late enrolled	-0.25	-0.08
	(0.15)	(0.18)
Average classroom ESCS	0.07**	0.07**
	(0.03)	(0.03)
Class-size	-0.01**	0.00
	(0.00)	(0.00)
Constant	0.22***	0.33***
	(0.04)	(0.04)
R-sqr	0.06	0.06
F-Statistic	0.000	0.000
BIC	472676.1	469947.5
AIC	472390.7	469662.1
Obs.	197444	197444
School FE	YES	YES

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Source: INVALSI

Table 7: Asymmetric effect of Shannon Equitability Index over Ethnic Status

	Performance in Language	Performance in Mathematics
	Model VII	Model VII
Diversity Index	-0.03***	-0.01
	(0.01)	(0.01)
2-Generation	-0.13**	0.04
	(0.04)	(0.09)
1-Genearation	-0.19***	-0.08**
	(0.03)	(0.03)
2-Generation # Diversity Index	0.03	-0.02
	(0.03)	(0.07)
1-Genearation # Diversity Index	0.01	0.01
	(0.02)	(0.02)
Individual Characteristics		
Girls	0.07***	-0.14***
	(0.01)	(0.01)
Early Enrollment	-0.09	-0.14
	(0.08)	(0.11)
Late Enrollment	-0.23***	-0.22**
	(0.06)	(0.07)
ESCS	0.22***	0.20***
	(0.01)	(0.01)
Classroom Characteristics		
Share of girls	0.03	-0.01
	(0.05)	(0.05)
Share of Late enrolled	-0.22	-0.09
	(0.15)	(0.18)
Average classroom ESCS	0.07**	0.07**
	(0.03)	(0.03)
Class-size	-0.01**	-0.00*
	(0.00)	(0.00)
Constant	0.22***	0.33***
	(0.04)	(0.04)
R-sqr	0.06	0.06
F-Statistic	0.000	0.000
BIC	472662.2	470042.7
AIC	472376.7	469757.3
Obs.	197444	197444
School FE	YES	YES

* $p<0.05$, ** $p<0.01$, *** $p<0.001$
(Standard errors in parentheses)
Source: INVALSI

Table I: Descriptive Statistics

VARIABLES	SHARE	MEAN	SD	MIN	MAX
Bullied		0.0	0.7	-0.8	3.3
Bullying		0.0	0.7	-0.7	4.4
Social Integration		0.0	0.7	-3.3	1.0
Extrinsic Motivation		0.0	0.8	-1.2	1.6
Ethnic Share		-0.1	0.8	-0.8	7.5
Shannon Equitability Index		-0.1	0.8	-0.9	16.2
Performance in Language		0.1	1.0	-3.4	2.2
Performance in Math		0.1	1.0	-2.9	2.0
Individual Characteristics					
Boys	49.9				
Girls	50.1				
Natives	91.6				
2-Generation	5.6				
1-Genearation	2.8				
Language (see focus)					
Regular Enrollment	96.9				
Early Enrollment	1.3				
Late Enrollment	1.8				
Italian teacher's mark		7.9	1.0	1.0	10.0
Math teacher's mark		7.9	1.1	1.0	10.0
ESCS		0.1	0.9	-2.8	2.3
Classroom characteristics					
Average in language test		7.8	0.5	5.3	10.0
Average in math test"		7.9	0.5	5.5	10.0
Share of Late enrolled		0.0	0.0	0.0	1.0
Average classroom ESCS		0.0	0.5	-2.2	1.9
Share of girls		0.5	0.1	0.0	1.0
Class-size		19.3	4.5	1.0	34.0

Table II: Focus Language

	Freq	%
Italian	189205	91.65
Arabic	2646	1.28
Albanese	2500	1.21
Romanian	2444	1.18
English	1472	0.71
Spanish	1365	0.66
Chinese	950	0.46
French	475	0.23
Hindi	324	0.16
German	253	0.12
Portuguese	211	0.10
Croatian	127	0.06
Ladin	78	0.04
Greek	53	0.03
Slovenian	67	0.03
Other	4273	2.07
Total	206443	100

LIST OF FIGURES

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

LIST OF TABLES

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV