Università
Ca'Foscari
Venezia

Master's Degree programme – Second Cycle
(*D.M. 270/2004*)
in Models and Methods in Economics and Management

–

Final Thesis

Ca' Foscari
Dorsoduro 3246
30123 Venezia

Peer Effects of Immigrant Children in Education:
A cross-country analysis

Supervisor
Ch. Prof. Elena Francesca Meschi

Graduand
Alberto Marcato
Matricula 830265

Academic Year
2014 / 2015

Abstract

The thesis studies whether the presence of non-native students in classrooms produces externalities affecting math outcomes of 8th grade native pupils. This analysis is performed over 9 European countries using data from the 2011 TIMMS database. First, the basic theoretical concepts of peer effect mechanisms are described, then the immigrant peer effects is estimated through several regression strategies, such as OLS school fixed-effect, quantile, piecewise, and separate regressions. In order to address the probable heterogeneous and nonlinear nature of peer effects, I exploit the random variations in non-native class shares within schools, conditional on observable covariates. The main findings support the nonlinear and heterogeneous effects hypothesis, in particular the study evidences a shifting trend of peer effects along the immigration concentration distribution.

Contents

Introduction

Over the last years, European countries have faced a dramatic growth of immigration, which has provoked an incredible anxiety for its potential threats and problems. However, while most of the studies concern the effects of migration on the labor market network of the host countries, research on the impact that non-native children could have on the local educational systems are still narrow, as well as presenting unclear and mixed results. Furthermore education, as fundamental determinant of human capital, is crucial for the economy of any State. For those reasons, I want to contribute to the sparse literature on the possible impact that non-native students might have on their native peers. In this regard, I exploit a very rich cross-country database (TIMSS), which has the great convenience of collecting data at the class-level, thus allowing to measure the peer effects at their most appropriate level of analysis.

Starting from a theoretical perspective, I describe the potential reasons for the possible peer effects of immigrant students, illustrating also some stylized models on how these effects may work. In addition, I present possible sources of bias. Indeed, assessing the immigrant impact on the educational performance of native pupils is really complex, because self-selection mechanisms lead students to be non-randomly allocated in classes and schools. By virtue of these sorting problems, analyses that do not adequately solve the endogenous problems can not distinguish correctly whether results are derived from a proper causal relationship or rather from a spurious one.

From the empirical point of view, I implement my model exploiting the presumed exogenous variations of the concentrations of non-natives in classrooms within the same school, and further controlling for some alleged

important covariates. Consequently, the basing assumption requires that within the schools, immigrant students are randomly assigned in classes, i.e. not correlated to any relevant unobservable characteristics. Firstly, I estimate the average marginal effect of a linear regression with school fixed-effects, conditional to a large set of covariates at both individual and class level. The estimated non-native peer effects are highly insignificant, in accordance with several previous studies. Moreover, I also investigate potential non-linearity and heterogeneity, through diverse regression strategies. The resulting immigrant peer effects have a very complex nature, with both heterogeneous and non-linear components. In particular, non-native students might have positive impacts on the math attainments of native students, whether class-concentration of immigrants is quite small. On the contrary, if the concentration is really small or large, the immigrant peer effects are negative. Furthermore, the peer effects even vary in accordance with the characteristics of students and class. For example, while the lowest skilled students spur non-native peer effects, the most highly skilled able students are almost unaffected.

The thesis's results thus suggest that the anxiety for immigration negative effects in education are excessive, given the small magnitude of these externalities. Moreover, the results suggest that immigration might even be beneficial for native students. In conclusion, the thesis provides interesting contributions to the narrow literature. It attempts to better describe the complex nature of non-native peer effects by proposing and analyzing theoretical mechanisms of peer interactions. Furthermore, from the empirical perspective it exploits multiple strategies to produce additional evidence for this complex phenomenon.

This thesis is organized as follows: Chapter 1 describes the theoretical and conceptual framework of both immigration phenomenon and peer effects, with a brief literature review. Chapter 2 displays the data and the descriptive statistics. Chapter 3 presents the empirical procedure adopted. Chapter 4 reports the results and the robustness checks. Chapter 5 concludes.

1. Conceptual Framework

1.1 Immigration phenomenon

Nowadays, many States are facing a worrying and growing phenomenon: *Immigration*. The migrations are really varied and complex, as they are guided by many factors. Anyway, considering size and impact, wage differences in the unskilled labor market are causing the most serious flows. Instead, migration of highly-qualified workers, known as "brain drain", still represents a narrow range of the total phenomenon (Čiarnienė-Kumpikaitė, 2008).

Migration is not a new phenomenon, people have always moved to seek better opportunities to live, but recently it reached an unprecedented level. Europe is dealing with the greatest influx of non-European immigrants in history, and more troubling: it could still rise. Indeed, flows do not seem to be going to disappear shortly and permanently (*New York Times*, 2015).

Although the latest crises in the Middle-East and Africa have led a plethora of people to escape towards developed countries in search of refuge and fortune, these flows can not be attributed solely to wars, conflicts, persecutions, and pauperizations. As reported by Dustmann-Okatenko (2014), the relationship between income and propensity to migrate is described by an *"inverted-U"*. The current migration boom is not derived from a substantial impoverishment of undeveloped countries, but probably from the fact that life in these countries does not match the ambitions of the inhabitants, who are influenced by the western lifestyle. Hence, it is a relative dissatisfaction, rather than an absolute, that spurs emigration. Internet, social media, and email, characterizing the ongoing *digital age*, have produced network connections, which enable migrants, former-migrants, and natives

to share news and knowledge. Therefore, information on the persisting wealth gaps between countries, as well as available and simple tips and guides to reach the developed countries, have stimulated the willingness to emigrate of disadvantage populations. Moreover, although still being far from closing the gap with Europe, in recent decades the per capita income of the underdeveloped countries has grown. This, together with modern accessibility to information, has provided people with the resources and knowledge necessary to emigrate.

Eventually, there may come a time when the wealth gap between developed and undeveloped countries will be filled and people no longer crave to migrate. However, that is unlikely to happen for many years, so in the short term, as poor countries get richer, more people will emigrate (*Telegraph*, 2015).

Migration is an important topic of study, as it provides many issues that public policy needs to address. Although not widely analyzed, a really interesting subject are the potential externalities that increasing share of immigrant children in classes (or schools) might have in the educational attainments of natives.[1] In spite of the importance for educational policies, the research on immigrant peer effect has not produced a vast literature yet. The current results are neither clear nor univocal. Indeed, they present ambiguous and sometimes contradictory evidence.[2]

1.2 Peer Effects

The study of the immigrant concentration impact on educational attainments is largely based on the relatively large literature on peer effects. Social

[1]In this thesis, the effects will be estimated exclusively on natives, in order to weaken the *reflection problem* — Manski, 1993. The correct estimation of the influential effect of one group on another is complex and extremely difficult, whether the two groups are nested, because it causes a problem of simultaneity: it is not easy to discern whether the subgroup reflects a variation in its belonging group, or whether, instead, the subgroup is the cause of the variation in the group.

[2]For example, Contini (2013), Gould et al. (2009) Jansen-Rasmussen (2011), Ballatore et al. (2014), and Tonello (2015) find (weak) negatively influences on natives' achievements. While, Geay et al. (2013), Cortes (2006), and Ohinata-Van Ours (2013) present no sizable effects. Instead, Hunt (2012) obtains positive externalities. For further details, see *Section 1.6*.

interactions, such as peer effects, are universally recognized as common and fundamental determinants in human decision-making mechanism. Furthermore, as economic agents, even students share and consume goods. The particular good that students consume is education, which has the peculiar characteristic to be in turn affected by pupils' attitude. Indeed, students can harm the teaching quality, ruining the regular lessons schedule through disrupting behavior. Therefore, the whole class academic results can be influenced by the behavior of just some pupil students (Lazear, 2001).

In addition, like all individuals, children are neither perfectly rational nor possess absolute knowledge, so their decisions can not be perfect. By virtue of that, they decide by comparing and imitating the beliefs and behaviors of the physically and socially close individuals.[3]

« The propensity of an individual to behave in some way varies with the prevalence of that behaviour in some reference group containing the individual. »

— Manski (1993), p.1.

Children grow up mimicking, adapting at the environment in which they live. Nevertheless, children could not completely mimic parents or teachers, because they are not enabled to act as adults. Indeed, most of the situations suppose and require diverse behaviors for children and adult. Hence, the peer group likely has an incredible influence on students; influence that can be even stronger than the *parental nurture* (Harris, 1998).

According to Conlisk (1980), mimic behaviors are stimulated by lack of information, thus youngsters may be more susceptible to imitate than adults. In fact, generally children possess less knowledge and experience, thus they might be more likely pushed to imitate their reference group (peers). Several studies provide evidence for peer effects. In particular, Goux and

[3]Conlisk (1980) describes that in an incomplete information framework, where take a decision is costly, the optimal action is to imitate the behavior of believed-better informed individuals. Bernheim (1994) demonstrates that in non-cooperative games, where popularity is included in the utility function, individuals tend to conform their behavior, even in case of heterogeneous preferences. Case (1991) shows how, *ceteribus paribus*, an individual demand for a certain good is positively influenced by the peer group mean demand. Grinbaltt-Keloariju-Ikaheimo (2008) show that automobile purchase decisions are strongly affected by the type of cars owned by the neighbors.

Maurin (2007) assess it in an educational context. Through an instrumental variable approach, they find positive peer effects in educational attainments.

The positive influence of group behavior consists in a common result in the peer effects studies. However, these effects are highly complex and unclear, in fact, they generally show nonlinear and heterogeneous effects among different groups.[4]

While effects of aggregate variations on individual demand, driven by price or market fluctuations, are generally accepted and considered as crucial endogenous social effect determinants, peer effects' influence has remained controversial for long time, on account of mixed and ambiguous results obtained in studies of this phenomenon. Peer effect is a compound and vague object, which requires a precise definition. The multiple channels, through which it operates, may drive to accept several different interpretations. In accordance with Sacerdote (2011), in this thesis peer effects concern any externality produced from any behaviors, backgrounds, or results of peer group that in turn influence the behavior or performance of individuals. This broader definition of peer effects includes both endogenous and exogenous effects, as defined by Manski (1993). Manski identifies three different cases to justify why individuals belonging to the same group act in a similar way.

Endogenous effects denote whenever the behavior of an individual varies accordingly to the behavior of its group. This effect could lead to a social multiplier.

Exogenous effects (or contextual effects) wherein individuals modify their attitude in response to peers' background qualities.

Correlated effects whenever same-group individuals behave similarly because they possess common unobserved characteristics (or face similar institutional environments), which are not included in the model.

For the thesis purpose, it is not particularly important to distinguish whether peer effects exerted by the immigrant students derive from backgrounds *(exogenous effect)* or attainments *(endogenous effect)*.

[4] *See section 1.5*, for more details.

On the other hand, a crucial issue in the peer effect literature consists in addressing the potential bias caused by correlated effects. Indeed, unobserved correlated effects may drive to erroneously attribute these effects to other analyzed variables, as peer effects, thus biasing the estimation. For example, extraordinary teachers or additional resources drive non-native concentration effects to be upwardly estimated, if those are correlated with high proportions of immigrant students.

Coleman (1968) is one of the first researchers who assesses the importance of peers in educational attainments. He describes peers' influence and classroom composition as fundamental determinants for children cognitive development. Therefore, as members of the peer group, non-native students may influence native pupils' performance. Moreover, peer effects are generally concordant in sign, thus an higher share of non-natives in schools might entail negative externalities. This because typically immigrants tend to perform worse than natives, as even reported in the TIMMS and PISA surveys.[5]

1.3 Why Immigrants Perform Worse

The peer effects have long influenced parents about the potential leverage that bad acquaintances may have on their children. This apprehension has a relevant importance on the identification of immigrant peer effects, because native parents might decide to transfer their children in schools with fewer immigrants, in order to avoid the negative influence that those could have on their offspring.[6] This *native flight* thus might exacerbate negative non-natives' spillovers. The fact that immigrant students generally perform worse than natives drives even public opinions and policy makers to react, deeming that

[5]PISA 2012 results document that immigrant students are reducing their gap with the natives, indeed in Canada, Ireland, and New Zealand they score equally at the natives in mathematics. Moreover, in Australia, Hungary, and Macao-China the non-native students are even better. Nevertheless, excluding these exceptions, immigrants are still performing averagely worse than the natives. — *OECD (2014), PISA 2012 Results.*

[6]See for example Rangvid (2007), Gould et al. (2009), Rangvid (2010), Tonello (2015), and Brunello-Rocco (2013).

immigrants could harm the educational attainments of native students.[7]

The diffuse anxiety is that an excessive concentration of immigrants in classrooms can cause negative externalities, because their supposed bad influence or limited proficiency in local language could absorb the teacher time and resources, reducing the quality of teaching.[8]

The non-natives students are disadvantaged both because they often start the school with a very limited knowledge of the local language[9] (sometimes no knowledge at all) and because they usually have underprivileged socioeconomic background.[10] Although socioeconomic status is one of the most influential determinant of differences between immigrants and natives on educational achievements, it is not the only reason. Indeed, the gap between natives and immigrants remains relevant in most of the countries, even conditional to socioeconomic characteristics.[11]

Low proficiency in the host language might push non-native students to cluster themselves in ethnic groups, exacerbating so the negative effects. *Acting white* is a particular kind of minority self-clustering. It concerns the pressure among black students to do not achieve good results for not appearing as they are betraying their origins posing as "white".[12] On the other hand, a large share of individuals of the same minority may also have positive spillovers. Second generation immigrants may be stimulated to study harder in order to distinguish themselves in the labor market from the newcomer immigrants (Hunt, 2012). However, these particular externalities could work if countries have a long immigration tradition, like the USA.

In conclusion, non-native students likely cause negative externalities,

[7]Ballatore et al. (2015) and Brunello-Rocco (2013) report the Italian example, where the Education Minister decreed a threshold for the maximum concentration of immigrant pupils, and encouraged desegregation policies.

[8]Georgiades et al (2013) found how the immigrants and minority students exhibited more emotional and behavioral problems that their native peers, resulting so more disrupting.

[9]According to Tonello (2015), differences in languages can lead to exacerbate the gap when the non-native is rejected, or they self-cluster in a foreign group.

[10]See Ammermueller (2007), Schnepft (2007), and Colding et al. (2009).

[11]Tonello (2015), OECD (2014) PISA 2012 Results.

[12]Fryer-Torelli (2010) analyze this phenomenon and its economic effects. Noting that this harming peer interaction is not exclusively of the African-Americans, but types of "acting white" exist in other ethnic subcultures across the world.

either because they require more time and special attention from the teachers — causing disruption; or because immigrant pupils need special services — reducing so the school expenditures useful for native pupils performance. In contrast with the mix and conflicting literature on the effects of immigrant concentration in school, the determinants of the lower performance of non-native students are more clear and unequivocal. The common accepted reasons of the test score gap are home environment and student characteristics.[13] Differences in parents' education and family situation are still important, albeit less influential (Ammermueller, 2007). However, the gap decreases with test score level: better immigrant students achieve similar results of their native counterparts.

1.4 Models of Peer Effects

The thesis focuses on the influence of peer groups on individual students outcomes, therefore of particular interest is the design of how these social interactions work. One of the most commonly used approach for addressing the peer effects is the *linear-in-means model*.[14]

$$y_i = \alpha + \beta \bar{y}_{(-i)} + \gamma x_i + \delta \bar{x}_{(-i)} + \varepsilon_i \tag{1.1}$$

where y_i represents the student-i's achievement, $\bar{y}_{(-i)}$ the mean achievement of student-i's peers, x_i and $\bar{x}_{(-i)}$ the characteristics of student-i and her peers.

Simplicity is the main reason for the popularity of linear-in-means approach, although unfortunately the model presents also two important drawbacks. Firstly, it is not interesting in a political perspective. Indeed, it does not allow to analyze the optimal peers allocation of students, as consisting in a *zero-sum game*; i.e. it assigns null net-effect for the reassignment of a student. In other words, the loss of a class due to the quit of a good peer is perfectly

[13] Ammermueller (2007) and Jansen-Rasmussen (2011) find that the language spoken in the family is a fundamental reason of the educational gap, follow by a comfortable space for studying. Other studies on the topics are: Schnepft (2007), Colding et al. (2009), and Schneeweis (2011).

[14] See Hoxby (2000), Hoxby-Weingarth (2005), Sacerdote (2011).

compensated by the gain achieved in the class where she has moved. Secondly, the linear-in-means model constrains the peer effects to work solely through the mean. Hence, mean-preserving variations do not matter. Therefore, this model does not enable nonlinearities nor heterogeneous effects, unless with some expansions.

However, peer effects generally present both non linearities – the effects increase (reduce) along the explanatory variable distribution – and heterogeneous effects. Hence, peer effects tend to vary by group, both for non-native class concentration level and for other covariates' values. Nevertheless, the linear-in-means models still represent a useful starting point in empirical research (Sacerdote, 2011).

Evidence of heterogeneous peer effects is widespread. In particular, studies supporting tracking policies generally point out that both most-able and least-able students can benefit from having peers of the same ability-level, while both could benefit less, or even be harmed, by different level-students.[15] For example, Jonsson-Mood (2008), analyzing Swedish secondary schools, show that medium-achievers are harmed by a large proportion of high-achievers, because it provokes a loss of self-confidence in the former, thus reducing their willingness to attend university.

In regard to ethnic groups, several studies reveal how peers' ethnicities influence individual students' achievements. Hoxby (2000) finds that peer's effects are stronger within individuals of the same ethnicity. Hanushek et al. (2009) notice that high shares of black students in a school have negative externalities on all the students, albeit these effects are stronger for the black students, thus emphasizing the idea of stronger peer effects within groups than across. Moreover, they find that concentration of black students influences black high-achiever students more than what it does for black low-achievers.

On the other hand, even researches against tracking system stress on nonlinearity and heterogeneity. Burke-Sass (2013) show that the best students benefits more from low-achievers than high-achievers. Vigdor-Nechyba (2007) find positive externalities from the increase in classroom heterogeneity.

In conclusion, almost the totality of the studies display evidence of non-

[15] Hoxby-Weingarth (2005), Duflo et al. (2008), Imberman et al. (2012), Hunt (2012).

linearities, it is thus important in peer effects estimation to take account of their potential nonlinearity effects. Models commonly adopted that allow nonlinear and heterogeneity are *bad apple*, *integration*, and *boutique* model.[16]

In the *bad apple model*, through disruption, the worst students provide negative externalities, affecting the entire classroom. Undisciplined behavior or extra-teaching needs harm teaching quality, causing the reduction of class educational attainments. Hence, assuming that non-natives are more likely disruptive than natives, the model predicts that rises in immigrant concentration ruin educational outcomes. These effects, however, are marginally decreasing in absolute value, because in this pattern one disrupting student is enough to provoke negative externalities. Therefore whether the probability of non-disruption is high, the substitution effect of switching a low-likely disruptive student with a high-likely one is bigger than when the probability of disruption is already high.

The *integration model* is also based on the disruption mechanism, but in this model, natives are willing to use efforts in order to integrate immigrant pupils. Therefore the gap between natives and non-natives tends to disappear, erasing the potential negative externalities of immigration. However, integration processes are costly, so whenever immigrant concentration exceeds a *critical mass*, natives are no more willing to help and integrate non-native pupils, in fact they could even start to reject and segregate them. Hence, this model predicts negligible negative effects for low immigrant concentrations, but which become extremely strong with very high concentrations, due to the marginally increasing externalities.

The *boutique model* predicts that any student benefits most from similar peers. Low-achievers profit more when surrounded by low-achiever peers than by high-achievers. Hence, according to the boutique model, non-natives score better if their peers are also immigrants. By virtue of that, tracking policies not only benefit high-achievers, but they are the best solution for any type of student. Higher cooperation and teaching customization are the typical proposed justifications for this model (Duflo et al., 2010).

Although highly simplified, those models provide satisfactory theoretical

[16]Lazear (1999), Lazear (2001), Sacerdote (2011), Tonello (2015).

bases about peer effect mechanism, providing thus an useful guide in the empirical implementation.

1.5 Endogeneity and identification issues

From the empirical point of view, the crucial challenge for a correct estimation of peer effects is their *identification*, which is made complex by *sorting mechanisms*, *omitted variable bias*, and *reflection problem*.

The allocation of immigrant children in the school is plausibly non-random, either because of self-selection or because of Manski (1993)'s correlated effects (i.e. neighborhood common characteristics). Non-randomness allocation worsens the identification of immigration impact on educational performance, because it is hard to understand whether the correlation is driven by causal relationships or by different characteristics in the groups. In fact, as a response to a large immigrant concentration in school, native families may choose to transfer their children in a school with a lower share of non-native pupils, as they are concerned about the possible immigrant bad influence on their offspring – independently of whether these worries are real or not.

Furthermore, common poor socioeconomic situations of non-natives induce them to concentrate in areas mainly populated by underprivileged natives, which would have achieved low educational attainments anyway.[17]

Additionally, in case of cross-country analyses, immigrants tend to settle in the most developed countries, where even the schooling systems are more efficient, that drives to underestimate the negative impact of immigrants (Brunello–Rocco, 2013). Finally, even the class composition within school may be endogenous: Ammermueller-Pischke (2009) and Ballatore et al. (2014) assess that both for political regulations and/or for school managers' personal choices, ethnic composition of classes could be non-random.

The general approaches in literature to solve the non-random composi-

[17]Cortes (2006) documents how 1st generation immigrants are more likely to live in enclave schools than 2nd generation ones. Even Jansen-Rasmussen (2011) and Brunello-Rocco (2013) report how immigrants tend to concentrate in less wealthy neighborhood or area.

tion of groups consist in relying in some exogenous variations, otherwise in controlling for fixed effects at school or class level.[18]

The omitted variable bias derives substantially from problem of missing data. As suggested by Todd-Wolpin (2003), modeling the child cognitive development as the production process of a firm, the educational attainments can be seen as a cumulative dynamic process of child endowed ability, family and school time-variables. Ideally, it will be necessary to possess complete data on all the inputs. However, it requires to identify all the relevant inputs for the child educational production function. This problem is hard to solve with the available incomplete data sets, also because some crucial inputs could not be observed, such as endowed latent ability.[19]

To sidestep the missing data problem, common techniques consist in designing an estimation model that overcomes the potential omitted variable bias or to use proxy variables for the significant unobservable ones. Notwithstanding, even these techniques have some weaknesses. To get over data missing problem and omitted variable bias, models require assumptions that are often implausible, because based on highly restrictive conditions.

On the other hand, obtaining unbiased estimates with the proxy variable approach is even more troublesome. In fact, although proxies can alleviate omitted variable bias, they require strong conditions, which if dissatisfied lead to the *bad control* case, worsening the bias. Indeed, if proxies, besides controlling for the latent variable, are correlated to other relevant explanatory variables, the estimates of both the latent variable and the correlated regressors may be seriously biased.[20]

[18] Unfortunately, exogenous variations are rare, literature presents just few historical cases. For example, Gould et al. (2009) exploit the random allocation in Israel of migrants from the former Soviet Union, after its collapse. Imberman-Kugler-Sacerdote (2012), instead, use the random allocation of refugees after the Katrina tornado. Angrist-Lang (2004) rely on the randomness allocation adopted by the METCO desegregation program.

[19] A largely adopted model is the *value added model*, which supposes that lagged outcomes are sufficient statistics for precedent missing values of the variables. According to Todd-Wolpin (2003), however, mixed and contrasting results characterizing the literature of peer effects, are determined, to a certain extent, by different and incorrect identifications of the real crucial inputs of the educational production function, entailing misspecified and biased estimations.

[20] For further details see Angrist-Pischke, 2008.

1.6 Literature Review

The current rising of the immigration flows and the well-established relevance of peer effects have encouraged a growing concern for the immigration, even in its impact on education systems. The literature, however, is still limited and ambiguous. The existing studies have found evidence for both significant (both positive and negative) and insignificant influence of non-native peers. The adoption of different models and data may have provoked such conflicting results (Todd-Wolpin, 2003). Actually, more recent studies select various approaches in virtue of the available data and because peer effects do not permit an easy identification, given their puzzling and complex nature.

One main challenge derives from the students' distribution in classes and schools, which, in reason of the sorting processes, is widely plausibly non-random. Indeed, self-sorting mechanisms and neighborhood characteristics likely drive the composition of classes and schools to be heterogeneous. Student unobserved features are plausibly correlated with non-native concentration, thus the estimation of the immigration share impact will be biased.

By virtue of the non-random allocation of non-native students, assessing their causal relation with native pupils' attainments becomes more tricky. Generally speaking, the causal relation describes the impact that an individual would have experienced if he had been treated, whereas he was not. In order to assess this counterfactual effect, natural sources of exogenous variation are preferred. Unfortunately, these are extremely rare in educational framework. In fact, very few studies exploit natural exogenous shocks in the peer group composition, among these Gould et al. (2009) and Angrist-Lang (2004), although none of these concerns European countries.[21]

Gould et al. (2009) use the mass migration of Jews in Israel generated from the collapse of the Soviet Union in the early 90s. The study shows

[21]Indeed, most of these researches concern the USA, besides in reason of the long tradition of immigration, these are more concerned the reasons behind the ethnicity school attainment gap. For example, Angrist-Lang (2004) do not analyze the immigration impact, although they assess the impact of the introduction of black pupils into school with "only whites", analyzing the data from the METCO program in the Boston area. They examine the within school variation across classes, finding insignificant impact on white-students' performances.

negative externalities of immigrant concentration on the probability of natives to be promoted. Their identifying assumption relies on the *as good as random* allocation of non-native pupils across grades within school. Conditional on the total number of enrolled immigrants in a given school, the immigrant share should depends exclusively on exogenous demographic factors.

However, the scarcity in nature of exogenous variations drives researchers to practice other approaches in order to address the problem of endogeneity. The main approaches are Instrumental Variables (IV), aggregation at large level, school fixed effects (SFE), or compounds of them.

Jansen-Rasmussen (2011) report significant negative externalities of immigrant concentration on native pupils' school performances. They exploit the variations in the immigration county proportion as instrument for the school concentration of non-natives. The paper supposes that parents could change children schools only in the same county, because bounded by housing and job conditions.

Similar to Jansen and Rasmussen, Hunt (2012) aggregates data of the USA at state level and uses the decennial variation in immigration as instrument. She finds that immigration has a positive effect on minorities, in particular among blacks, because increasing immigration implies higher supply in unskilled labor market, thus encouraging "non-white" students to improve their school performances, in order to elude the resulting higher competition.

Brunello and Rocco (2013) assess the immigrant impact on natives' school attainments by aggregating data at country level and exploiting the variations of non-native concentration in schools. Their results are negative, albeit weak, non-native externalities.

Chaparro et al. (2013) use a *Difference-in-Difference* approach, which relies on the assumption that after selecting on observables, the remaining correlation between immigration share and unobservable relevant features is due exclusively to time invariant components. They observe a negative effect on the Spain grade-retention rate, although only if the non-native proportion exceeds the threshold of 15%.

On the other hand, school-fixed-effects models address the endogenous sorting problem by exploiting the variations within schools either across adjacent

cohorts (Hoxby, 2000; Tonello, 2015) or between classes (Ammermueller-Pischke, 2006; Ohinata-Van Ours, 2013; Contini, 2013). Therefore, the basing assumption requires that the compositions of classes (or cohorts) within school are as good as random.

Analyzing Italian census data (INVALSI), Tonello (2015) detects a small negative impact of immigrant concentration on Italian students outcomes. The estimation model exploits variations across adjacent cohorts within schools, which are assumed to be exogenous. His results support the idea of peer interactions, following an integration mechanism. Indeed, including the quadratic term of the immigrant school share, he uncovers marginal increasing negative effects.

Ohinata-Van Ours (2013) use cross-sectional data from an international survey (PIRLS), assuming as good as random differences of immigrant class concentrations within school. They find no-sizable effect of non-native students on educational attainments of Dutch pupils.

2. Data and Descriptive Statistics

2.1 Data

The data used in this study are the 2011 Trends in International Mathematics and Science Study (TIMSS), a project of the International Association for the Evaluation of Educational Achievement (IEA), which aims to measure the educational performances of students all over the world. TIMSS is one of the largest international student assessment database, besides, in contrast with others such as PISA, it is very suitable to analyze peer effects. Indeed, TIMSS pays particular attention to student's background and instructional practices, and it further organizes data at class level. Performed on a regular 4-year basis, TIMSS measures mathematics and science performances for 4th and 8th grades students. The thesis focuses on mathematics, because it consists in a crucial requisite to access at scientific studies, and thus to acquire advanced competences and skills, fundamental requirements for high level and high remunerated occupations.[1]

Questionnaires surveyed at students, parents, teachers, and principals provide a database with rich and extensive information on home and school background. Unfortunately, the immigration status is indicated exclusively for the 8th grade students. As consequence, I analyze solely this grade.

TIMSS 2011 reports the performances of 8^{th}grade students from 45 countries and 14 benchmarking participants. I restrict the analysis on all the European countries, by virtue of the current growing concern in Europe for the immigration phenomenon. The countries analyzed are therefore: England, Finland, Hungary, Italy, Norway, Romania, Slovenia, and Sweden.

[1]Martin et al. (2012) TIMSS 2011 International Results in Mathematics.

TIMSS data are collected by a rigorous and complex survey design, consisting in a stratified two-stage cluster sample. The primary sample units are schools, which are stratified according to important common demographic or geographical characteristics. Stratification is a simple and effective method to improve the reliability of survey estimates, because it ensures the correct proportional representation of each different group within-country. In fact, sub-country groups presumably possess peculiar characteristics, which without stratification they might be underestimated, or even neglected completely.

Additionally, being drawing in proportion to their size, schools are more likely selected whether they have more classes of the targeted grade, guaranteeing so that each student has the same likelihood to be chosen. Moreover, TIMSS ensures an adequate sample size and correct representation for any subpopulation, selecting from the same strata two replacement schools for each sampled one. Therefore, whenever a school refuses to participate, it will be replaced by a likely analogous one.

Furthermore at the second stage, entire classes are picked out with equal probability through systematic random sampling. In contrast with other procedures that elect pupils across all the students belonging to the same grade or possessing a certain age, drawing intact classes permits a better analysis of peer effects. Indeed, these are plausibly most strong at class level. By the way, at the second stage there are not replacement, if a class does not participate it is simply ruled out.

Even though classes should be analyzed entirely, TIMSS enables principals to exclude some students in accordance with some precise conditions. Among these, students with functional or intellectual disabilities or that have severe lack in the native language could be excluded. Conventionally, should be excluded students that have received less than one year of instruction in the language test. Anyway, the omitted students must be less than the 5% of the national target population.[2]

In principle, this complex and rigorous procedure should provide samples of students with equal selection probability. However, in practice, disproportionate sampling across strata and non-participating units vary consistently

[2]Martins et al. (2012) Methods and procedures in TIMSS and PIRLSS 2011.

the students' probability of selection, therefore, TIMSS provides a sampling weight system. The students' weights are essentially the inverse of their probability of being sampled.

The students' assessment scale method used in TIMSS consists in plausible values. Plausible values are imputed scores, which are derived from the distribution of the estimated student ability, based on an Item Response Theory model, the Rasch model. The Rasch model has become extremely popular in the education large-scale surveys, because this procedure improves the estimation of the student ability. Indeed, a simple analysis of the correct answers, which weights items as equally difficult, leads to biased results, especially if students face tests that are necessarily diverse. The test scores are further standardized at the international mean of 500, with a standard deviation of 100.

Table A1 provides of the entire sample the information about the number of observations, mean, and standard deviation for each analyzed variable. I exploit a large number of variables at individual, class, and school level. The individual statistics consists in Math score, Immigrant status, Gender, Book-index, PARED, Quarter of birth, and Repetition dummy. *Book* is an ascending index that assumes values from 1 to 5, whether respectively the student's family possesses from 0 to 10 books, 11 to 25, 26 to 100, 101 to 200, and more than 200. For relieving the burden of the computation, I recoded this index variable in three dummies: *Few* whether the books owned are less than 25, *Some* when between 25 and 100, *Many* if more than 100. Also PARED is another ascending (1-5) index, describing the parents' highest level of education, in accordance with the ISCED.[3] Derived from PARED, I

[3]UNESCO (2006) *International Standard Classification of Education: ISCED 1997.*

Level 1 Finished Primary school, or lower.

Level 2 Finished Lower-Secondary school.

Level 3 Finished Upper-Secondary school.

Level 4 (or 5b) Finished Post-Secondary Education but not University.

Level 5 Finished University or higher.

Maintained by the United Nations Educational, Scientific and Cultural Organization (UNESCO), the International Standard Classification of Education (ISCED) is a system of standards that allows to design standards, which indicate the levels of education achieved by individuals, thus useful for comparing individuals education within and across countries.

construct the *Parent graduated* dummy variable, which assumes value 1 when at least one parent is graduated, 0 otherwise.

Although TIMSS database does not indicate whether a student has failed a grade, it at least reports the children' age. Hence, I repute "repeaters" those students that have entered the eight grade later than at the regular age, independently if they have really failed a grade. However, I do not consider older immigrants as repeaters, because many non-native students are held back in the earlier grades for language deficiency, therefore they could bias the estimation. In fact, whether older non-natives are considered as repeaters, immigrant peer effects are going to be partly captured by the control variable of the repeaters class share, given the serious correlation between older and non-native pupils (Contini, 2013).

In regard to the definition of *immigrant*, I prefer to adopt the most common *ius soli* determination, thus, unconditionally to the citizenship of the parents, everyone born in the country is considered native.

The variables at class-level are instead Teacher Gender, Teacher Years of Experience, Teacher Master Education, and Class size. *Teacher Years of Experience* is a dummy that indicates whether teacher has taught for more than 10 years. *Teacher Master Education* reports whether teacher has attained a Master Degree (or higher). Unfortunately, TIMSS does not measure the real number of students in the classes, however being an important determinant of student achievements, I approximate that with the number of pupils that have participated at the test.[4]

Finally, the school covariates are: City size, City income, Material, Room, and Staff availability. *City size* is an increasing (1-3) index measuring how many people live in the school area, it assumes value 1 whether there are less than 15 000 inhabitants, 2 if between 15 000 and 100 000, 3 if more than 500 000. *City income* is also an (1-3) index, which describes the average

[4]Although this proxy approach might conduct to biased results, the thesis purpose needs the class size measure, because it is a crucial determinant of students' performance, and it is indispensable to assess the non-native class concentration. Anyway, this approach leads to erroneously impute peer effects at participating students exclusively. Therefore, whether non-participation is correlated with some relevant variables, for example non-native status, the estimation results are biased. Unfortunately, the data do not enable to obviate this serious problem, therefore I suggest particular caution in the results interpretation.

income of students' family: 1 stands for low average income, 3 for high. The *resources indexes* describes the availability of materials, rooms, or competent staffs. As the city variables, these indexes are increasing: 1 whether resources are few, 3 if they are many. For a better understanding, each school-level variable is decompose into three dummy variables: Low, Medium, High.

2.2 Descriptive statistics

As reported in Appendix (Table A2), student non-participation and missing values of relevant variables constitute a delicate issue for the empirical analysis. Indeed, student non-participation shrinks the number of observed pupils per class. Therefore, for limiting the attenuation bias due to an excessive proportion of non-participating children, I analyze exclusively the classrooms with at least 10 participating students. This first constrain discards the 4% of the total student sample.

On the other hand, the missing values of relevant characteristics do not look a so troublesome problem. The most serious case is the parent's education, which reaches the dramatic 25.56% of missing values, on the contrary, for the other variables the issue is not so alarming. The exceptions are Norway and Sweden. While Norway has no information on the city average income, Sweden presents more than 10% of missing observations for the school characteristics. Anyway, the missing values are not negligible, because the reduction of the observations decreases the effectiveness of both statistical tests and estimators. By virtue of this problem, I implement several regressions to investigate the potential bias and robustness. Being the variable most affected by missing values, PARED is ruled out and substituted by the number of books possessed, which is correlated with parents' education, and it is even considered a very appreciable predictor of family background.[5]

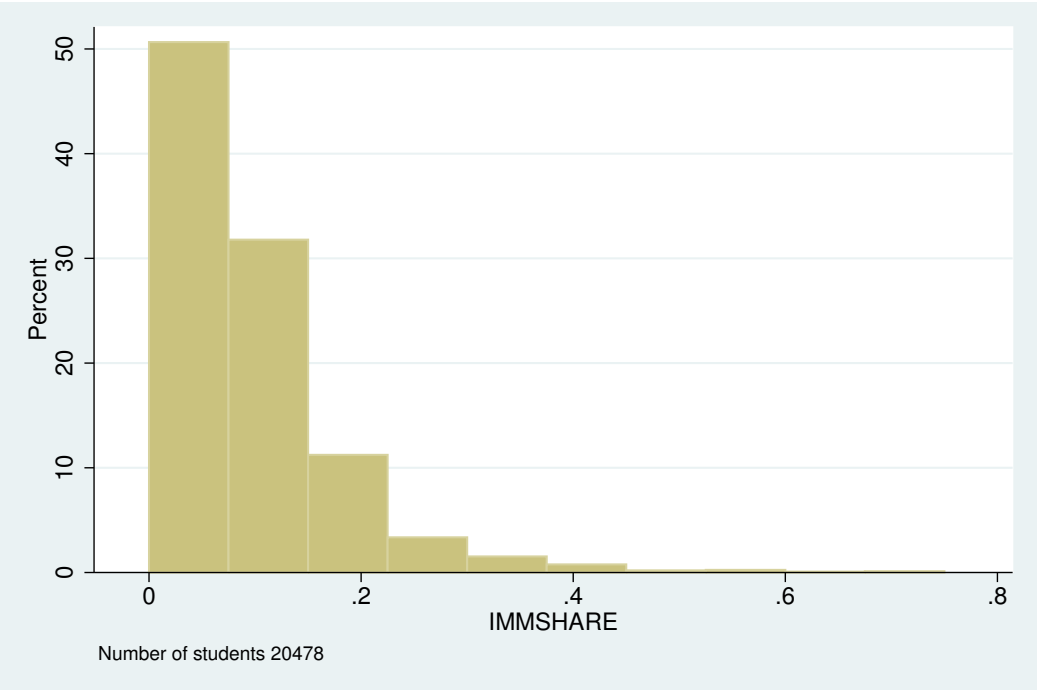
In addition, school fixed effect models require multiple classes per school, albeit many schools have only one class. Therefore, for completeness, in the 2nd subsample of Table A2, I display the sample size of the only schools

[5]See for example Ammermueller–Pischke (2009).

possessing at least two surveyed classes.

The absence of multiple classes provokes the discard of the 35.1% of students, 38.4% of classes, and 53.8% of schools from the original sample. This loss of information is extremely troublesome, especially in some countries, such Italy, whose sample becomes incredibly small. Short sample decreases the standard deviations, thus even the power of significance tests declines, leading so to an increase of *false positive* cases.[6]

Figure 2.1: Share of students for immigrant concentration in classes



The histogram is built at student-level. To obtain a better insight of the right tail, I rule out the students facing classes without immigrants, these students constitutes almost the 50%, they are 20 376 on a total of 40 854.

Tables A3 and A4 present mean and standard deviation for each variable respectively at individual level, and at both class and school level. For the sake

[6]By definition is impossible to observe the error terms, therefore the estimation of their standard deviation is obtained with $\sigma/\sqrt{n}$. Where σ represent the standard deviation of the residuals, and n the sample size. Intuitively, more data imply less variation, therefore results more precise.

For further details, especially in hierarchical data structure, see Hox (2010).

Table 2.1: Weighted Mean values comparison between Native and Immigrant students. *(Standard errors in parentheses)*

Table 2.1.A

	ENG	FIN	HUN	ITA	LTU
Score	510.9 (79.7) 493.3 (98.1)	517.7 (59.6) 484.4 (67.5)	507.7 (84.6) 443.1 (124.4)	500.4 (69.6) 474.7 (71.0)	506.4 (74.8) 447.9 (89.6)
BOOK *Index(1-5)*	3.00 (1.29) 2.55 (1.35)	3.23 (1.14) 2.71 (1.37)	3.24 (1.29) 2.57 (1.30)	3.07 (1.25) 2.52 (1.20)	2.81 (1.14) 2.78 (1.43)
Girl	0.48 (0.50) 0.48 (0.50)	0.49 (0.50) 0.46 (0.50)	0.49 (0.50) 0.38 (0.49)	0.49 (0.50) 0.44 (0.50)	0.49 (0.50) 0.33 (0.48)
Quarter *Index(1-3)*	1.99 (0.82) 1.97 (0.80)	1.98 (0.80) 1.90 (0.82)	1.99 (0.81) 1.99 (0.79)	1.98 (0.81) 2.06 (0.85)	1.95 (0.80) 1.93 (0.89)
PARED *Index(1-5)*	3.74 (1.11) 3.79 (1.32)	3.95 (1.01) 3.85 (1.24)	3.53 (1.00) 3.48 (1.09)	3.32 (1.13) 3.43 (1.15)	3.85 (0.88) 3.97 (1.00)
Classize	23.70 (5.25) 21.82 (5.70)	18.55 (3.40) 18.56 (3.91)	21.29 (5.26) 19.15 (6.05)	20.20 (4.17) 19.82 (3.96)	19.18 (3.93) 18.36 (4.89)

Table 2.1.B

	NOR	ROM	SVN	SWE	TOT
Score	476.9 (60.6) 451.6 (68.7)	458.8 (98.1) 436.1 (122.5)	507.2 (66.9) 453.2 (65.0)	487.7 (62.5) 457.2 (72.6)	495.9 (77.3) 464.8 (84.2)
BOOK *Index(1-5)*	3.40 (1.21) 2.84 (1.33)	2.50 (1.20) 2.40 (1.17)	2.93 (1.14) 2.49 (1.29)	3.29 (1.26) 2.58 (1.40)	3.05 (1.25) 2.61 (1.33)
Girl	0.49 (0.50) 0.44 (0.50)	0.49 (0.50) 0.30 (0.50)	0.49 (0.50) 0.43 (0.49)	0.48 (0.50) 0.45 (0.50)	0.49 (0.50) 0.44 (0.50)
Quarter *Index(1-3)*	1.96 (0.81) 2.03 (0.81)	1.91 (0.76) 2.23 (0.76)	2.00 (0.80) 2.02 (0.80)	1.93 (0.81) 1.97 (0.82)	1.96 (0.81) 2.00 (0.82)
PARED *Index(1-5)*	4.47 (0.82) 4.09 (1.22)	3.40 (0.99) 3.68 (0.90)	4.08 (0.80) 3.66 (0.99)	4.15 (0.95) 4.15 (1.16)	3.77 (1.02) 3.81 (1.19)
Classize	23.02 (5.40) 23.32 (5.72)	22.60 5.42) 22.46 (5.61)	20.26 (3.97) 20.32 (4.06)	22.36 (4.97) 21.87 (5.38)	21.24 (5.00) 21.14 (5.35)

In black the native values, in blue the immigrant ones. In parenthesis the standard deviations. In the TOT column, the sample includes all the analyzed countries. The mean values are obtained using HOUWGT: a set of weights included in TIMMS database, which ensures that the weighted sample corresponds to the actual sample size in each country.

of brevity, I do not report the values of the peer variables, because these are simply generated averaging the individual-level variables at class-level, thus their mean values are unchanged, instead logically their standard deviations are shrunk. Although peer effects might exist at different levels, they are plausibly stronger at class-level, because teachings takes place in classroom and further pupils generally spend more time with their class-peers than with other individuals, increasing so the *"time-exposure to contagion"*.

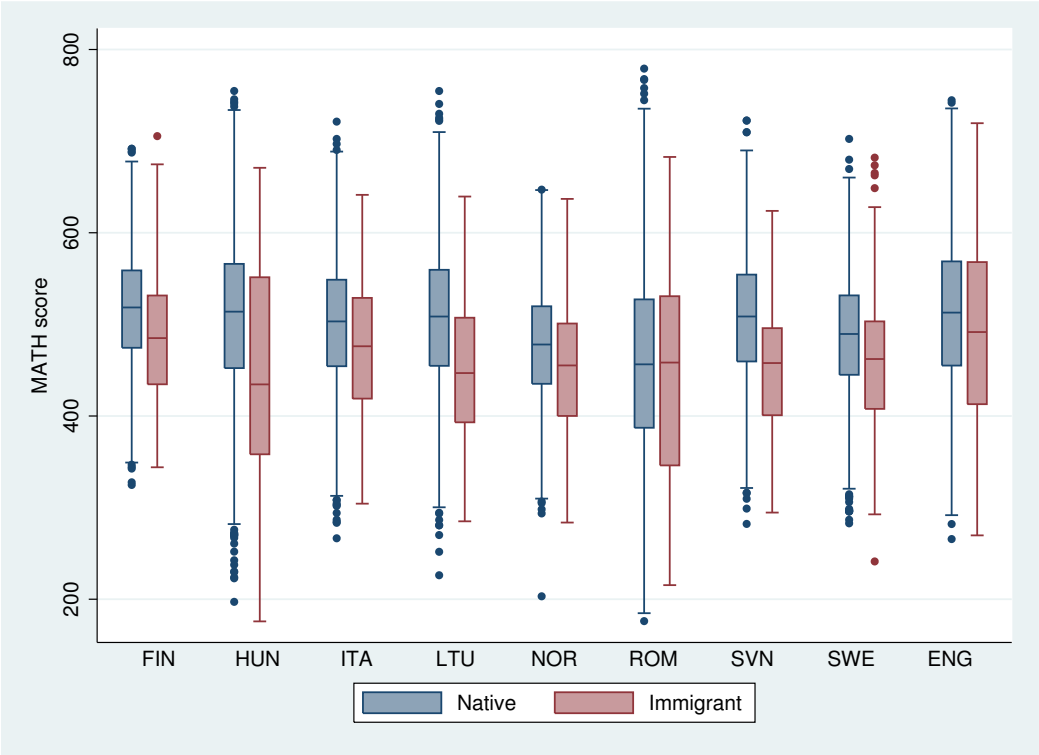
By the way, to obtain a further insight in the crucial variable of my study, the non-native class concentration, Figure 2.1 displays its distribution for the whole sample. The distribution possesses a really long right tail, thus showing that most of the classes have a low proportion of immigrant students.

In regard to the differences between native and non-native students, Table 2.1 presents their separate individual variables measures, demonstrating thus that natives achieve math scores significantly better than immigrants. On the other hand, *Gender* and *Quarter of birth* variables do not show any statistically relevant difference. Average class size instead seems smaller for immigrant students, albeit this gap is likely negligible.

The most remarkable outcome is the contrasting results between *Book index* and *PARED*: while natives own more books than immigrants, immigrant parents result better schooled.

Further analyses about immigrants performing worse in math are provided in Figures 2.2 and A1, which compare math scores distributions of native and immigrant pupils.

Figure 2.2: MATH scores distribution: Natives vs Immigrants



3. Empirical framework

3.1 Model selection

In order to investigate the immigrant peer effects is first necessary to specify the education production function. As reported in *Section 1.3*, a model widely used in literature is the *linear-in-means model*, with peer effects assessed at classroom level. Adapting equation (1.1) to include the influence of class and school characteristics, the model becomes:

$$y_{ics} = \alpha + \delta \bar{y}_{(-i)cs} + \beta X_{ics} + \gamma \bar{X}_{(-i)cs} + \phi S_s + \varrho C_{cs} + \mu_s + \mu_{cs} + \varepsilon_{ics} \quad (3.1)$$

where y and X are respectively the test score and the other individual characteristics, while the subscripts i, c, s denote the individual i, of class c, in school s. The $(-i)$ describes the peer characteristics obtained averaging the individual characteristics at class level ruling out the individual i's values.[1] Moreover, S and C are the school and class level characteristics, while μ_s, μ_{cs}, and ε_{ics} are the error terms at school, class, and individual level. In particular, the μ identify the *correlated effects*, while γ and δ determine the *exogenous* and *endogenous effects* – using the terminology defined by Manski (1993).[2]

However, this model incurs in simultaneity problem, which prohibits a direct estimation. Indeed, the peers' average outcome ($\bar{X}_{(-i)cs}$) affects by γ the result of student i (y_{ics}), which in turn influences the performances of all her peers j, where $j = \{1, \ldots, (1-i), (i+1), \ldots, N_{cs}\}$. By virtue of the symmetry of equation (3.1), each student j is influenced by ($\bar{X}_{(-j)cs}$), of which

[1] The independent variable under analysis *(proportion of immigrants per class)* is therefore embodied in the averaged peers' characteristics per class ($\bar{X}_{(-i)cs}$).

[2] *See Section 1.1.*

student i is part of. This process makes almost impossible to distinguish, without strong assumptions, exogenous and endogenous effects, hindering so the direct and correct identification of the structural equation.[3] Anyway, a strategy commonly used in literature is to indistinctly combined endogenous and exogenous effects, deriving thus a reduced model where all the effects of immigrant peers are mixed. It hence:

$$y_{ics} = \alpha + \beta X_{ics} + \lambda \bar{X}_{(-i)cs} + \phi S_s + \varrho C_{cs} + \mu_s + \mu_{cs} + \varepsilon_{ics} \tag{3.2}$$

where the parameter λ embodies both the endogenous and exogenous effects.

On the other hand, even in the reduced structural equation, correlated effects persist. Therefore, they should be controlled, otherwise estimation results might be misleading, driving to identify causal relations, whether they are instead spurious. Indeed, correlated effects cause serious and severe problem for the identification of the peer effects, because they break the assumption of random composition of students in classes and school.

As described in *Section 1.4*, the literature suggests several identification strategies to address this problem, each one, however, requires particular assumptions. I address the endogenous problem exploiting the within schools variations across classes. The strategy reflects the idea that school composition is hardly random, because school-selection mechanisms drive that to be correlated with students' unobservable characteristics. Accordingly to specific country policy, students are either free to choose their school or else to enroll in the closest school. In both cases, however, the random composition is implausible. In the latter case, the neighborhood common characteristics spur schoolmates to posses similar family background and socio-economic status, such as unobserved significant characteristics. On the other hand, whether parents may decide their children' school, they may do that accordingly to quality or other school characteristics, driving thus at endogenous student allocations. As explained in *Chapter 1*, for the native flight phenomenon.

By virtue of those reasons, within school transformation is a very useful method, because ruling out the school level variables, it also erases the

[3]See the "reflection model" described by Manski (1993).

unobserved school selection determinants, which represent the most probable sources of endogeneity.

Anyway, even this strategy is based on a strong assumption, i.e. the random allocation of students across classes within schools. In other words, schools do not have to group students according to some significant student characteristics. Hence, the sorting of students in classes must be random. Formally, it is assumed that the class-specific error term (μ_{cs}) is identically and independently normally distributed, and further uncorrelated with the regressors, like a *White Noise*. The within transformed structural equation, which consists in my final estimation model, is:

$$y_{ics} - \bar{y}_s = \beta(X_{ics} - \bar{X}_s) + \lambda(\bar{X}_{(-i)cs} - \bar{X}_s) + \varrho(C_{cs} - \bar{C}_s) + \eta_{ics} \tag{3.3}$$

$$\text{where} \qquad \eta_{ics} = (\mu_{cs} - \bar{\mu}_{cs(s)}) + (\varepsilon_{ics} - \bar{\varepsilon}_{ics(s)})$$

However, random class composition within school is not necessary assured. Indeed, even after within transformation correlated effects might persist. For example, at 8th grade, schools may have already tracked students for ability. Moreover, in class with high proportion of non-native students, principals may assign better teachers or resources, or even enroll less students, in order to relieve the supposed negative impact of immigrant pupils. Therefore, in these cases, the within school transformation fails to completely address the correlated effect issue, leading thus the estimates of peer effects to be biased.

3.2 Potential class selection within schools

The analysis by school fixed effects addresses adequately the probable endogenous composition of schools, such as *native flight*. On the other hand, it does not consider the possibility that a sort of endogeneity persists within school in the student class composition. Therefore, it is interesting to study the possible differences in resources allocations between classes.[4]

[4]Furthermore, although schools differences in resources do not consist in a problem in the final estimation with the school fixed effects, in the preliminary estimations these differences can explain some misleading results.

As described *Chapter 1*, the causes for different allocations of resources could be multiple, such as worried principals or educational organizations might provide schools or classes with more resources proportionally to the share of immigrants, in order to relieve their presumed negative impact.

Ballatore et al. (2015) advise that worried about the potentially adverse immigrant externalities principals could deliberately manage non-native student allocations within school. They may avoid to assign many immigrants students in large classes, preventing thus to mix the negatives externalities of high concentration of both students and immigrants. As consequence, whether large class size really produces significant negative effects on the student performance, the deliberate allocation of non-native students will bias the results. Indeed, detrimental effects of immigrant peers may so be mitigated by the positive externalities of small class sizes.

Similar to Ammermueller-Pischke (2009), I investigate if immigrants are allocated according to particular deliberate choices, performing a *Pearson Chi^2 test*, which verifies whether, within school, classes are composed unconditionally to the citizenship of the students.[5] The test assesses whether the observed number of immigrant students in a class are statistically different from the predicted number in case of random allocation. In other words, the predicted values implies that the share of non-native students is proportional to the total number of immigrant students enrolled in the school. The test is so formally implemented:

$$\mathcal{P}_s = \sum_c \sum_k \frac{(n_{sck} - \hat{n}_{sck})^2}{\hat{n}_{sck}}$$

where

$$\hat{n}_{sck} = \frac{\left(\sum_k n_{sck}\right)\left(\sum_c n_{sck}\right)}{\sum_k \sum_c n_{sck}} \qquad \text{and} \qquad \begin{array}{l} c = \{0, 1, \ldots, C_s\} \\ k = \{0, 1\} \end{array}$$

the parameters n_{sck} and $\hat{n}_{sck}$ are respectively the observed and the predicted number of students in school s, class c, and with the immigrant status k, where

[5] Undoubtedly, the test works exclusively for schools with multiple classes. Therefore, these tests do not analyze all the schools in the database.

0 indicate native and 1 non-native students. In Appendix, Figure A2 shows the distribution of the school p-values across countries. All the countries, except Italy, present p-values lower than 0.05 for less than the 10% of schools. Furthermore, assuming that school decisions are mutually uncorrelated, it is possible to generate the test at country level. The country level statistic $\mathcal{P}$ is obtained adding all the $\mathcal{P}_s$ statistics within-country. Under the null hypothesis of random allocation, $\mathcal{P} \sim \mathcal{X}^2$ with $\sum_s (C_s - 1)$ degrees of freedom, where C_s are the number of sample classes in school-s. These country-level p-values are reported in Column (5) in Table 3.1.

Another potential cause of within-school endogeneity might be non-random differences in teachers' ability. As important determinant for students attainments, teacher quality should be randomly assigned to classes. However, generally skilled teachers are matched with skilled pupils, which thus may suggest non-random assignments. Therefore, since teacher and students abilities tend to co-vary, the immigrant peer effects might result upwardly biased (Vigdor-Nechyba 2007, Burke-Sass 2013). Anyway, this matching system of ability between teachers and students might be quite difficult, because teacher quality is latent and it changes accordingly to the typology of students. Moreover, whether teachers may either decide autonomously their class, or if principals decide for them, in both cases it is probable that the teacher assignment is correlated to the observable qualities of the teacher such as experience or highest education. Therefore, controlling for teacher observable characteristics should be a satisfactory check of the potential teacher and students ability match.[6]

In order to investigate the existence of endogenous allocations of teachers and resources among schools, I regress the immigrant class proportion on school and class characteristics. Then I test whether these covariates depends on immigrant concentration, the results are in Table 3.1.

Table 3.1 displays the p-values of the tests, which verify the independence of the allocation of resources among classes and schools. With the exception of

[6]Notwithstanding, Burke-Sass (2013) state that teacher real effectiveness is weakly correlated to observed teacher characteristics. However, by virtue of what I say about the latency and relatively of teacher ability, and without panel data to implement a "teacher fixed effects" model as they suggest, I consider acceptable to check for statistics observable.

Table 3.1: Independence of immigrants allocation with class and school characteristics

	(1) Class	(2) School	(3) Class+School	(4) +SFE	(5) Person Chi^2
ENG	0.207	0.028**	0.059*	0.446	0.906
FIN	0.305	0.038**	0.127	0.290	0.969
HUN	0.143	0.150	0.289	0.120	0.998
ITA	0.366	0.413	0.669	0.000***	0.0103**
LTU	0.223	0.873	0.630	0.344	1
NOR[1]	0.004***	0.068*	0.003***	0.346	0.235
ROM	0.657	0.177	0.443	0.810	1
SVN	0.230	0.550	0.264	0.000***	0.895
SWE	0.393	0.006***	0.028**	0.000***	0.004***
TOT	0.000***	0.001***	0.000***	0.018**	0.999

The first three columns report the p-values of the Wald test computed on the regression of the class immigration concentration on the respective regressors. The 4th column carreis out the regression on the class covariates adopting the school fixed effects. Finally, last column displays the p-values for the Pearson Chi^2 test, testing whether within school immigrant is randomly allocated in classes. All the tests are computed using class level data.

[1] Norway data set does not possess information on the *city income* explanatory school variable, so their school covariates in (2) and (3) are assessed ruling out that variable.

the last column, the outcomes show the p-values of the Wald test F-statistics, which test the joint significance of all the regressors. In the first column, the regressions are exclusively on class characteristics. The assumption of the independence between immigrant concentration and class resources allocation is not rejected for any country, except Norway. Column 2, instead, implements the regressions on school covariates. England, Finland, and Sweden do not satisfy the independence assumption, with a significance of the 5%. As discussed in the *Sections 1.4 - 1.5*, correlation of school level covariates and non-native concentration is a common belief in literature. Anyway, rejecting the assumption of independence between immigrant concentration and school covariates is not troublesome, whether school-fixed effects are included.

While Column 3 investigates both class and school covariates, Column 4 adds the school fixed effects (SFE). In the regressions with SFE, the independence assumption seems rejected for Italy, Slovenia, and Sweden. In spite of the reduction of the sample size, however, the results might be erroneously rejected. Indeed, the within-transformations shrink the number of observations valid to the estimation. Therefore, the restricted sample sizes decreasing the power of the tests, may drive to potential *false negative*. That could be happened for Italy and Slovenia, which are the two countries most affected by the sample size reduction. The same reasoning should be valid also for the Pearson Chi^2 tests in Column 5. In conclusion, Wald and Pearson Chi^2 tests seem to support the exogeneity between non-native class concentration and class resources, that at least for the six of the nine countries analyzed.[7]

Anyway, although wide adopted in literature, this strategy to investigate the random allocation among class and schools presents a critical drawback. Technically, significance tests indeed can only reject or not-reject a null-hypothesis. A rejection with the significance level of 5% implies that the probability for this hypothesis to be true is less than the 5%. On the contrary, whether the p-value of the test is beyond the significance level, it does not provide any information, it simply fails to reject it. Therefore, if a test fails to reject the null-hypothesis it has not confirmed the randomness.

[7]By virtue of this partially confirmation, I implement further regressions ruling out the countries that have failed these tests, more detail in *Section 4.3*.

4. Results

4.1 Baseline Results

The results of the regressions of native student achievements on non-native class are reported in Table 4.1. For further investigate the non-native peer impact issue, I exploit 9 different regression equations. These regressions are implemented by including iteratively additional covariates, showing so the threads of misleading results whether endogeneity is not properly erased. Column (1) reports naive OLS regression of native student test scores on class share of immigrants. As expected, this specification yields an apparently negative and significant non-native peer effects. However, whenever individual (2), class (3), and school (4) covariates are included in the model, the apparent negative effect fades, becoming insignificant. Hence, these results support how described in *Chapter 1* about the spurious relations due to non-random allocation of immigrants.

In Columns (5), (7), (8), and (9) I include school fixed effects for addressing the potential endogeneity due to school selection and thus removing its spurious effect. As collateral consequence, the differences of school-level variables become ineffectual to estimation, hence, they are ruled out from the model. Additionally from column (6) peer covariates are included. Peer covariates interestingly turn into positive the immigrant impact, which even gets slightly significant in column (8), where variables measuring parents' education are excluded for increasing the sample size and improving the estimation precision. However, the apparent positive impact of non-native peers may be determined by the strong correlation between immigrant students and the few book possessed variable.

Table 4.1: Regressions for math test score on immigrant concentration

(Standard deviation in parentheses)

	(1)	(2)	(3)	(4)	(5)
IMMSHARE	–56.92***	–33.37**	–27.15*	–21.62	13.16
	(19.66)	(15.82)	(15.35)	(18.09)	(31.50)
Girl		–5.147***	–5.038***	–4.518***	–5.542***
		(1.020)	(1.054)	(1.161)	(1.465)
few_Book		–41.54***	–39.74***	–37.76***	–35.70***
		(1.473)	(1.414)	(1.441)	(1.809)
many_Book		19.18***	18.21***	17.08***	11.74***
		(1.149)	(1.134)	(1.218)	(1.394)
Parent graduate		27.39***	25.05***	23.49***	16.71***
		(1.274)	(1.242)	(1.277)	(1.418)
Repeaters		–50.11***	–47.30***	–45.19***	–40.09***
		(3.180)	(2.982)	(2.979)	(2.899)
1^{st}quarter_birth		1.453	1.589	0.890	1.631
		(1.219)	(1.216)	(1.369)	(1.571)
3^{rd}quarter_birth		2.663**	2.582**	3.696***	3.246**
		(1.152)	(1.174)	(1.294)	(1.503)
T.female			–3.469	–3.173	4.547
			(2.702)	(3.071)	(5.059)
T.10ys_teaching			2.878	–0.0680	6.053
			(2.428)	(2.623)	(6.816)
T.mast_graduate			6.793**	4.833	1.735
			(3.461)	(3.809)	(5.189)
Class size			2.316***	2.158***	3.901***
			(0.243)	(0.288)	(0.681)
School covariates				✓	
Observations	38,583	28,979	27,603	23,558	27,603
No. of schools					1,288

Weighted least squares regressions for math test scores on a set of covariates. Standard errors are robust to clustering at school level. The weights are the students' sampling probability (HOUWGT) where possible, where it was not I used the schools' sampling probability (SCHWGT) – School fixed effect models require school common weight. The student level covariates are student's sex, index of books at home, parent graduate, repeater, and month quarter of birth. The class level covariates are teacher's sex, teacher master graduate, 10 years of experience in teaching, and class size. The peer are share of girls, of quarters of birth, and of repeaters per class. Finally, the school covariates are indexes of city average income and city size, and the index for the disposal of material, room, and staff. The different groups (individual, classroom, school, and peer variables) are separated by lines. T precedes the teacher statistics. Each regression includes country dummies. Country dummies and school level variables estimates are omitted due to problem of available space.

Table 4.1: *Continue*

	(6)	(7)	(8)	(9)
IMMSHARE	–2.681	34.04	45.94*	22.50
	(16.55)	(29.10)	(26.44)	(27.82)
Girl	–4.880***	–5.270***	–6.116***	–5.915***
	(1.096)	(1.619)	(1.424)	(1.537)
few_Book	–32.95***	–37.59***	–37.77***	–35.47***
	(1.250)	(1.904)	(1.652)	(1.521)
many_Book	13.50***	11.74***	15.00***	14.32***
	(1.123)	(1.487)	(1.410)	(1.273)
Parent graduate	18.68***	17.72***		
	(1.097)	(1.492)		
Repeaters	–43.17***	–41.13***	–42.14***	–44.25***
	(2.604)	(2.997)	(2.808)	(2.816)
1^{st}quarter_birth	1.082	2.103	2.441	2.290
	(1.371)	(1.685)	(1.535)	(1.613)
3^{rd}quarter_birth	3.326**	4.367***	3.807**	4.105***
	(1.291)	(1.678)	(1.498)	(1.578)
T.female	–3.007	2.364	3.646	4.353
	(2.893)	(4.380)	(4.250)	(5.047)
T.10ys_teaching	–0.546	1.414	3.271	4.472
	(2.448)	(5.708)	(5.487)	(6.299)
T.mast_graduate	1.604	–1.250	–0.235	3.378
	(3.512)	(4.668)	(4.609)	(5.437)
Class size	1.238***	2.391***	3.248***	4.613***
	(0.260)	(0.597)	(0.559)	(0.654)
Peers girl	8.183	9.226	–3.497	3.755
	(7.784)	(14.90)	(15.17)	(18.55)
Peers few book	–55.31***	–67.80***	–80.77***	
	(9.489)	(18.98)	(18.90)	
Peers many book	17.83**	12.00	30.37*	
	(8.879)	(15.84)	(16.27)	
Peers parent graduate	26.19***	36.87***		
	(6.203)	(9.532)		
Peers repeaters	–28.73*	–37.10	–41.58*	–87.15***
	(15.49)	(23.23)	(25.05)	(24.72)
Peers 1^{st}quar_birth	–15.82	14.21	15.87	12.66
	(15.63)	(16.76)	(16.79)	(19.22)
Peers 3^{rd}quar_birth	15.36	30.35*	23.47	29.21
	(13.60)	(17.87)	(17.56)	(19.87)
School covariates	✓			
Observations	23,558	27,603	36,469	36,469
No. of schools		1,288	1,288	1,288

Robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1

As described in Table 2.1, non-natives statistically report lower average measure of the book possession index. Therefore, this correlation likely biases the peer effect estimate upwardly. Indeed, part of the impact of non-native concentration may be absorbed by the peer book low possession variable. In fact, as in Column (9) the peer book variables are excluded the non-native peer effects turn again to be insignificant, supporting the idea of spurious relation.[1]

By the way, these results are based on a country-pooled model, which requires that immigrant peer effects are constant across countries. In order to provide more credibility to this assumption, I include in my preferred baseline specification (Column 9) the interaction terms between immigrant class concentration and each country dummy variable, in accordance with the strategy adopted by Brunello-Rocco (2013). Afterward, I perform Wald tests on the interaction terms, assessing whether there are country-specific peer effects. All the tests lead to non-reject the null-hypothesis of insignificant country-specif peer effects. Besides, the p-value of the joint significance test is 0.7383.[2]

Analyzing the covariates, individual characteristics affect student's outcomes in concordance with literature findings. Girl, grade-repeating, and disadvantaged socio-economic students (measured by the number of book possessed), meanly obtain worse results in the math tests. Instead, class-level covariates appear insignificant for math attainments. The sole exception is the class size, which, in contrast with Ballatore et al. (2014), seems to have a positive impact, i.e. large classes apparently improve student's performances. In addition, peer covariates also look insignificant, with the exception of the negative effect produced by grade-repeating peers. Although, the class percentage of girls and quarter of birth, in literature, are generally considered

[1]The problem of correlation between regressors becomes troublesome, because I am exploiting proxy variables. Hence, as briefly reported in *Section 1.5*, whether a proxy variable is correlated with both the dependent variable and other regressors the estimation could be seriously biased.

[2]As should be well-known, a p-value that is above the critical value does not lead to accept the null-hypothesis of non-significance, but it simply fails to reject that. Anyway, the non-rejection of the null-hypothesis permits, at least slightly, to strengthen the belief that immigrant peer effects are not country-specific.

quite important, the insufficient within-school variations in the data might they look negligible in student attainments (Rangvid, 2007). However, I am not going to further investigate this aspect, as not in my thesis purpose.

A lastly clarification is for student birth month, which results are not directly interpretable. Even though age is an important determinant for student achievement, it is only in relation with the school cohort age, and not in absolute term (Crawford et al., 2010). Therefore, the diverse starting school age differs between countries makes the pooled results inapprehensible. Indeed, whether in England the older pupils in a cohort are the ones born in September, in Italy are the ones born in January. Hence, while in England the third quarter students represent the oldest pupils, in Italy they consist in the youngest.

In conclusion, to further analyze the possible differences among countries, I estimate the non-native peer effect for each country separately. Results are reported in Table 4.2. For available space problem, I display only the results for my preferred model (Table 4.1, column (9)), i.e. the one with SFE, and conditional to individual, class, and peer covariates, with the exclusion of parents' education and peer book possession variables. The country-specific regressions confirm average insignificant effect of non-native class concentration.

Table 4.2: Estimated peer immigrant effect per country

	obs.	*no.schools*	IMMSHARE	*(std.dev.)*
England	2 989	*107*	-3.110	*(39.20)*
Finland	3 819	*143*	43.39	*(59.63)*
Hungary	4 893	*143*	48.72	*(57.67)*
Italy	3 293	*178*	-6.516	*(120.2)*
Lithuania	4 522	*129*	-19.34	*(72.39)*
Norway	3 339	*128*	3.526	*(34.10)*
Romania	5 140	*142*	-153.1	*(149.4)*
Slovenia	4 055	*176*	32.27	*(65.64)*
Sweden	4 419	*142*	35.81	*(24.95)*

Weighted least squares regression with school fixed effects for math test scores on immigrant concentration in class, controlled for a set of covariates. Standard errors are robust to clustering at the school level. The weights are the schools' sampling probability (SCHWGT). The student level covariates are student's sex, index of books at home, parents' graduated, repeater, and month quarter of birth. The class level covariates are teacher's sex, teacher's master graduated, 10 years of experience in teaching, and class size. The peer are share of girls, of quarters of birth, and of repeaters per class.

4.2 Nonlinear and Heterogeneous effects

In section 1.4, I have pointed out that *linear-in-means* model is likely inadequate to structure the peer effect mechanism, plausibly because it is characterized by nonlinear and heterogeneous components. I exploit several approaches in order to test the presence of nonlinearities. A first procedure consists in a parametric approach, similar to the one adopted by Tonello (2015), which outcomes are reported in Table 4.3. I include in the regression a quadratic term of the immigrant class-concentration variable, in Column (1), and then even the cubic term, Column (2).[3] Although results are not statistically significant, both the regressions' highest term shares negative estimates. Despite this is not enough to declare nonlinear components, negative coefficients could be due to marginally decreasing effects, at least in part of the distribution.

In the interest of further insight on possible nonlinearities, implementing other approaches should be worthwhile. Hence, I divide the distribution of non-native class-share into dummies, regressing then student's performances on these, in order to investigate the potential heterogeneous effects at different level of non-native concentration.[4] I adopt two different partitions: first grouping by quartiles, then dividing by deliberate cut-points; the results are reported in Table 4.4.[5] The cut-point partition is the most meaningful, because it allows comparisons among countries and a better insight in potential non-linear peer effects. Indeed, immigrant peer impact is plausibly an "absolute" effect, which, in other words, does not depend on the national distribution

[3]The table reports even the estimates of separate regressions run for males and female students.

[4]The interpretation of partitioned regressions is pretty different from the continuous ones. The regression in Table 4.1 describes marginal effects of rising in immigrant classroom concentration at the mean, while the partitioned estimates describe the mean impact to attend class with different concentration of non-native in comparison with a zero-immigrant classroom.

[5]The quartile partition is conducted at student-level on classes that possess non-native students. The classes with exclusively native students are taken as the reference group. Instead, for the cut-point partition I split the distribution into 6 groups, choosing as non-native concentration thresholds 5%, 8%, 15%, and 35%. Again the reference group is composed by the classes with no immigrant.

Table 4.3: Parametric approach at non-linear immigrant peer effect.
(Standard deviation in parentheses)

	ALL		MALE		FEMALE	
	(1)	(2)	(1)	(2)	(1)	(2)
IMMshare	52.1	12.0	47.0	14.2	61.4	–47.3
	(55.3)	*(83.6)*	*(64.5)*	*(95.0)*	*(60.1)*	*(97.5)*
IMMshare^2	–110	245	–65.5	208	–196	886
	(160)	*(522)*	*(180)*	*(525)*	*(177)*	*(691)*
IMMshare^3		–553		–401		–1 965*
		(808)		*(746)*		*(1 058)*

Weighted least squares regression with school fixed effects for math test scores on immigrant concentration in class, controlled for a set of covariates. Standard errors are robust to clustering at the school level. The weights are the schools' sampling probability (SCHWGT). The regression (1-4) are based on regression (9) of Table 3.2.B. Regression (1) introduces a quadratic term of the immigrant concentration at class level. In (2) a cubic term is added. The regressions are computed for the whole sample, and separately for the student's gender.

of immigrant class share. Nevertheless, this type of partition easily incurs in precision problems, because each group is not equally represented in the sample. Therefore, whether the sample is small or whether observations are not sufficiently distributed throughout the entire domain, estimates may lack of precision, driving thus to potentially misleading and biased results.

On the other hand, a percentile partition oversteps precision problems, because dividing the sample at percentile assures equal sample size for each group. However, this method drives groups to cover extremely large or short area of the variable domain, puzzling the direct comprehension of the estimates.

By the way, the results of the regression on the quartile partition seems support the nonlinear immigrant peer effects. Indeed, in the first two quartiles peer effects are negative, while in the second two they are positive, albeit only the 2nd and 4th quartiles present significant estimates. Surprisingly, these outcomes seem completely opposite to those commonly obtained in literature, in which, immigrant peer effects are weakly negative, and increasing in absolute value. However, these apparent contradictory results could be caused by the strong unskewned distribution of non-native class concentration. As showed in Figure 2.1, the distribution presents a really long right tail, driving the quartiles to be concentrated in low levels of immigrant proportion, hence the last quartile contains classes with non-native percentages very far apart, however for different rates, peer effects may diverge.[6]

On the other hand, the "cut-points partition" allows to create groups embodying less divergent concentrations of non-native students. I deliberately choose the cut-point percentages at 5%, 8%, 15%, and 35%.[7] This regression partly confirms the previous results, besides it presents a new interesting finding: whether immigrant concentration becomes quite high (over 35%), peer effects turn again to be significantly negative. This shifting trend is not observable studying non-native impact on quartiles, because students

[6]In fact, as described in Table 4.4, the quartiles' delimiters are 4.8%, 7.4%, and 11.8%. Therefore the 4th quartile embodies classes very different in term of immigrant proportion (from 11.8% to 75%).

[7]These partition is the one that presents the most significant results, in spite of the handful diverse groupings carried out, which however have rather robust results.

Table 4.4: Nonlinear non-native peer effects:
Regression on categorical variables of immigrant class share.

Table 4.4.A

	Partition by QUARTILES			
	Q1	Q2	Q3	Q4
	–1.22	–11.00**	7.93	13.8**
	(5.58)	*(5.02)*	*(5.51)*	*(6.04)*
N.students	5 631	4 742	5 032	5 073
N.classes	221	232	230	247
N.schools	188	196	199	218

Table 4.4.B

	Partition by CUT-POINTS				
	C1	C2	C3	C4	C5
	-1.34	-9.22*	8.67	17.6**	-24.6**
	(5.63)	*(4.79)*	*(5.49)*	*(7.50)*	*(10.6)*
N.students	5 631	5 936	5 315	2 989	607
N.classes	221	278	248	145	34
N.schools	188	229	212	131	33

Weighted least squares regression with school fixed effects, and usual set of covariates, as in regression (9) Table 3.2.B. Independent variable is the categorical variable built by a partition of the continuous varible IMMSHARE. Cut-point partition splits the distribution at 5%, 8%, 15%, and 35%. While quartiles have the following maximum percentage of immigrant: Q1 4,8%, Q2 7,4%, Q3 11,8%, Q4 75%. The number of students in classes without immigrant students are 20 376. Detailed information in text. (Standard deviations in parentheses)

belonging to high immigrant concentrated class are really scarce.[8] Hence, in the quartile study, these peer effects are overwhelmed by the ones produced in the less concentrated classes, within the same quartile. Summing up, although the parametric approach does not reach a significant outcomes, the dummies partitions assess that the immigrant peer effects are decreasingly negative whether the non-native concentration is quite low (approximately under 10%), beyond that they turn to be increasingly positive, but finally whether immigrant concentration are very high (approximately over 35%) the non-native peer effects re-becomes negative.[9]

By the way, the cut-point analyses by countries achieves almost entirely inconsistent estimates, probably caused by the lack in precision due to the shrunk sample sizes. Instead, per countries quartile analyses are incomparable, because they have quartile delimiters at different percentage of non-native concentration, further peer effects are sensible to the absolute number of immigrant pupils.[10]

Anyway, the partition analyses support the idea of non-linear relationship between student's test scores and proportion of non-native peers. In fact, they present evidence of shifting peer effects along the explanatory variable domain. A piecewise regression may therefore improve the estimation of the diverse marginal peer effects. Indeed, through break-points, a piecewise regression partitions the domain of the explanatory variable into sub-domains, allowing the relationship between response variable and independent variable to varies across the sub-domains, leading so to a more flexible and richer characterization of the phenomenon. Although, in principle, there are several methods to estimate the best break-points, in a model with handful covariates, fixed effects, and little valid observations within groups, these methods are likely going to achieve analogously inconsistent and biased results. Hence, I preferred to avoid the hard and long computational effort and simply assumed

[8]Students with more than the 35% of non-native peers represent just the 0.74% of the total sample. See Table 4.4.

[9]Curiously, 30% is the maximum threshold of immigrants recommended by the Italian Ministry of Education, Circular Letter n. 2, 8/1/2010 (MIUR, 2012).

[10]For the sake of brevity, I have not reported the per-country outcomes of the partition regressions.

break-points at the same point of the "cut-partition", i.e. at 5%, 8%, 15% and 35% of immigrant classroom share.

Table 4.5: Pieciwise linear regression.
(Standard deviation in parentheses)

	S1	S2	S3	S4	S5
IMMSHARE	–140.8	165.9	354.2**	–168.7***	–65.0
	(96.3)	*(234.6)*	*(143.2)*	*(59.6)*	*(178.9)*

Weighted piecewise least squares regression of student's test scores in math on class-concentration of immigrant peers. Condition on school fixed effects, and the usual set of covariates, as in regression (9) Table 4.1. The break-points are assumed at 5%, 8%, 15%, and 35%. Detailed information in text.

In addition to nonlinearity, peer effects are even plausibly heterogeneous. Hence, non-native peers probably cause different externalities for different group of students. I investigate that, running separate regressions for different typology of pupils.[11] Firstly, I test whether immigrant peer effects are diverse for student's gender. Table 4.5 reports the regressions for both males and females. Moreover, in *Panel B*, I exploit the finding of non-linear effects to analyze whether low, medium, and high proportion of non-native peers influence equally male and female pupils. Accordingly to the cut-point partition findings, I decided to construct three dummies identifying low (<8%), medium (between 8 and 35%), and high concentrated (>35%) classes.[12]

[11] Although several studies preferred to analyze heterogeneity, interacting the variables analyzed. I preferred to run separate regressions, because by interacting the other predictor coefficients are force to remain constant across the different variable type. Hence, given the handful number of predictors included in the model and their plausible correlation, a pooled regression likely obtain misleading results, driven by the steadiness of the covariates' coefficients. Indeed, it is highly plausible that different characterization of some variables changes not only the effect of immigration, but even other variables' impact. For example, whether classes are small, the class proportion of girl students in the class probably have a negative impact on the average class attainments, because girls perform worse in mathematics and ability peer effects are generally positive. On the contrary, in large classes this adverse effect could be overwhelmed by the girls positive effect on the reduction of the lesson disruption. Anyway, separate regressions have a drawback. They do not enable

Table 4.6: Gender Heterogeneous effects.

(Standard deviation in parentheses)

	FEMALE	MALE
Panel A: Immigrant class-concentration continuous variable		
IMMSHARE	17.92	29.11
	(36.79)	*(32.88)*
Panel B: Categorical variable of Immigrant share.		
Low Immshare	–5.16	–6.12
	(4.45)	*(4.51)*
Medium Immshare	10.30*	11.18**
	(5.49)	*(5.47)*
High Immshare	–17.62	–25.57
	(11.90)	*(17.25)*

Weighted least squares regression with school fixed effects, and the usual set of covariates, as in regression (9) Table 4.1. In *Panel A*, I exploit two separate regression for female and male students on the continuous IMMSHARE variable. In *Panel B*, the studied independent variable is the categorical variable built by a partition of the continuous variable IMMSHARE. The categorical variable is divided in Low when the immigrant concentration is $\leq 8\%$, Medium when $\leq 35\%$, and High when $> 35\%$.

Table 4.6 investigates potential differences between girl and male students, its results do not provide any evidence of heterogeneous immigrant-peer effects in gender.

Furthermore, given the crucial importance, another possible source of heterogeneity that I would like to analyze is the class size. Analogously to Table 4.6, Table 4.7 studies the immigrant peer effects in different class sizes. To get a better insight I group classes according to their quartile into 3 dummy variables: SMALL, MEDIUM, and LARGE. The first quartile of the distribution represents the Small classes, while the fourth identifies the Large class, Medium classes stay in between. The outcomes apparently suggest that medium classes are the only ones that benefits by non-native students. On the other hand, differences in non-native class-proportions among size of class do not look significant.

Finally, in Table 4.8, I analyze the potential heterogeneous effects for student's ability. I divide students by test score quartiles per each country, defining unskilled students the ones in the first quartile, competent the ones in the top quartile, and ordinary the in-betweeners. The most remarkable result is that whether immigrant class-concentration is higher the negative non-native peer effects are spurred by the less-able students.

According to the previous procedure, I test also the heterogeneous effect for: student's socio-economic status (measured by number of books possessed), and teacher's experience. Particularly interesting results are that socio-economically advantaged students apparently are not corrupted by non-native pupils, even when their concentration is quite high, in contrast to less advantaged students. In addition, the teacher's experience seems to alleviate the detrimental effects of high proportion of non-native pupils. The results are reported in Appendix, Tables (A5-A6).

---

precise verification on the significance of the difference regression estimates.

[12]Table 4.4.B shows that students suffer different immigrant spillover effects according to the number of non-native peers. Initially the impact is positive, then it becomes negative, and for very high share of non-native the effect turns again into negative. Therefore, the thresholds of these dummy variables are whether the immigrant effects change sign.

Table 4.7: Class size Heterogeneous effects.
(Standard deviation in parentheses)

	SMALL	MEDIUM	LARGE
Panel A: Immigrant class-concentration continuous variable			
IMMSHARE	–12.35	41.48	–85.14*
	(69.60)	*(54.32)*	*(49.28)*
Panel B: Categorical variable of Immigrant share.			
Low Immshare	–13.41	–1.13	–11.21
	(8.73)	*(6.86)*	*(9.23)*
Medium Immshare	8.52	5.69	–10.38
	(9.62)	*(7.01)*	*(12.33)*
High Immshare	21.95	–5.29	–16.56
	(25.39)	*(17.93)*	*(10.48)*

Weighted least squares regression with school fixed effects, and the usual set of covariates, as in regression (9) Table 4.1. In *Panel A*, I exploit three separate regression for class size type on the continuous IMMSHARE variable. In *Panel B*, the studied independent variable is the categorical variable built by a partition of the continuous variable IMMSHARE. The categorical variable is divided in Low when the immigrant concentration is $\leq 8\%$, Medium when $\leq 35\%$, and High when $> 35\%$.

Table 4.8: Student's ability: Heterogeneous effects.
(Standard deviation in parentheses)

	UNSKILLED	ORDINARY	COMPETENT
Panel A: Immigrant class-concentration continuous variable			
IMMSHARE	–41.55	6.68	–15.69
	(29.08)	*(13.60)*	*(22.17)*
Panel B: Categorical variable of Immigrant share.			
Low Immshare	–4.92	–1.40	–3.67
	(4.05)	*(1.71)*	*(2.52)*
Medium Immshare	0.08	3.22	1.53
	(3.55)	*(1.97)*	*(2.93)*
High Immshare	–45.69***	–18.56**	–4.23
	(13.45)	*(9.16)*	*(19.52)*

Weighted least squares regression with school fixed effects, and the usual set of covariates, as in regression (9) Table 4.1. In *Panel A*, I exploit three separate regression for student' test scores, on the continuous IMMSHARE variable. In *Panel B*, the studied independent variable is the categorical variable built by a partition of the continuous variable IMMSHARE. The categorical variable is divided in Low when the immigrant concentration is $\leq 8\%$, Medium when $\leq 35\%$, and High when $> 35\%$. The division in Unskilled, Ordinary, and Competent is computed by quartiles over country. Ordinary identifies the students at both 2nd and 3rd quartiles.

Quantile Regression

The separate regressions for investigating heterogeneity in immigrant spillovers in students' ability have presented some interesting results. Therefore, in order to get a better overview, I exploit a *Quantile Regression*, which provides a richer characterization of the phenomenon, because it permits to study the impact of regressors at every location of the distribution of the dependent variable, not just at the mean. Therefore, quantile regression is highly suitable for studying heteroskedastic data. However, a quantile regression with clustered data and fixed effects is very complex and so extremely hard and long to elaborate, thus, I opt to implement the Canay's simplified quantile regression procedure. Canay (2010) develops a simple two-step method for quantile estimations, which allows to estimate quantile regressions with fixed effects. The method consists of running a classic quantile regression adding the estimates for fixed effects, obtained by standard OLS regression with fixed effects. Canay demonstrates that the mechanism produces consistent estimates whether the fixed effects are assumed to be *"location shift variables"*, in other words, fixed effects do not change across quantiles, i.e. they affect each quantile in the same way.[13]

In Table 4.8, I report the outcomes of the quantile regression so implemented. The outcomes suggest that while unskilled students improve their average results whether immigrant class-proportion rises, best students are not significantly affected by that. Unfortunately, the cluster option in the estimate of the immigrant spillovers across the students' ability distribution does not allow to obtain simultaneous quantile regression. Therefore, I cannot correctly test the differences between the estimated coefficients between diverse quantiles. Nevertheless, I am mainly interested in assessing the size and the sign of the immigrant peer effects, and not in testing whether similar coefficients are statistically diverse. Figure 4.1 plots immigrant peer effects across quantiles, showing the presence of heterogeneity and how the spillover effects decrease progressively with the students' ability.

[13]Even though the assumption that school fixed effects are constant across quantiles is pretty strong, the outcomes could still provide an additional insight on the heterogeneity of the immigrants spillover effects along students' ability.

Table 4.9: Quantile Regressions

(Standard deviation in parentheses)

	Student's Ability				
	p(10)	p(25)	p(50)	p(75)	p(90)
IMMSHARE	68.82***	40.41***	29.47***	1.89	–7.96
	(6.61)	*(6.35)*	*(4.83)*	*(5.93)*	*(8.13)*

The quantile regressions is implemented at the following percentiles: 10, 25, 50, 75, 90. To sidestep the problem of the model with school fixed effect, I first estimate it through by a OLS regression with SFE, and then I add the estimated fixed effect in the quantile regression. The graph below plots the results of this quantile regressions, in comparison with the OLS estimate (the dashed line). The confidential intervals are underestimated, because they are constructed without considering the clustering, that is omitted for a more easy computation.

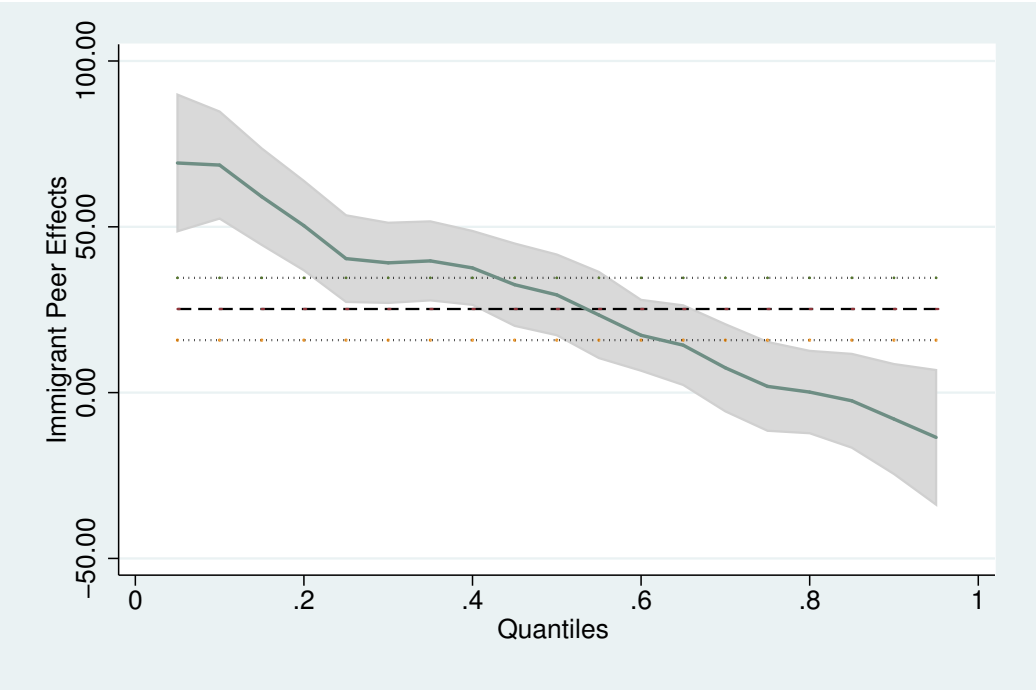


Figure 4.1: Quantile Regression Plot

4.3 Robustness Check

The previous sections' results rely on a series of assumptions. In this section, I discuss the robustness of the baseline regression (Column 9, Table 4.1) about the non-randomnesses class formation within-school and the definition of non-native students.

Firstly, I restrict the sample to schools that possess at least two classrooms. Secondly, concerning the random allocation of resources and non-native students across class within school, I discard the schools that failed at diverse level of significance the Pearson Chi^2 tests.[14] Finally, I have also estimate the peer effects in the only country that have passed the test of random allocation of resources, after controlling for school fixed effects. Hence, Sweden, Italy, Slovenia, and Norway were been discarded from the estimation. All these robustness checks do not lead to any significant and appreciable results.

I also analyzed the immigrant peer effects adopting different definition of "immigrant student". I firstly adopted the "ius sanguinis", so a child with both the parents foreign-born is considered an immigrant. Further, I exploit the difference between first-generation and second-generation immigrant. Even here the results do not add any further information, supporting the idea of Harris (1998) that peer influence is more important than "parental nurture" in children cognitive development.

Finally, according to the idea that State industrial development is a crucial determinants in the selection-country decision of immigrants, as final robustness check, I analyze separately the non-native peer effects for the most developed countries, consisting in the plus Norway, and for the non-EU15 countries. This separation is even justified by the drastic difference in the average share of non-native students, indeed, while for the developed countries the mean is 7,31%, for the non-EU15 the average immigration share is only the 2,21%.

Although, the peer impacts are still highly insignificant for both the groups, the outcomes confirm the shifting non-linear trend of the non-native

[14]The test are described in Section 3.2., the levels of significance adopt are 5%, 10%, 25%, 50%.

peer effects. Positive for medium immigrant class-share as in the EU15 group, negative for small share as in the non-EU15, results are reported in Appendix, Table A7.

Attenuation bias

The complex model have not enabled to remove all the possible sources of bias, which afflict the subject studied. In particular, although the database used (TIMMS) is achieved through a rigorous, well-organized survey, it does not constitute census data, thus it is more prone to be affected by sampling problems and measurement errors. In particular, as expressed in Section 2.1, TIMMS permits to exclude some students according to precise conditions, one is the host language deficiencies, hence typically non-native students. Therefore, the estimated immigrant classroom proportion may be underestimated. Moreover, Ammermueller-Pischke (2009) demonstrate that, especially in studies with school fixed effects, the measurement errors likely drive to underestimate peer effects.

Furthermore, the hierarchical structure of data leads to another concerning problem. Studying the impact of non-native class share to the student test scores implies the analysis of variable from different level, because test scores are at individual level, while non-native share is at classroom level. Therefore, the immigrant concentration variable is disaggregated at the individual level, driving few data class level to become many more values. This mechanism provokes a fictional increase in the sample size, and thus the precision rises, undervaluing the standard errors estimates. Consequently, these issues suggest that the results should be interpreted with appropriate caution.

Conclusions

The thesis aims to provide evidence on potential impacts of non-native students on native peers' educational performance. For this purpose, I exploit the 2011 international survey database TIMMS, which reports the math test scores of 8th grade students from 9 European countries. TIMMS data possesses also wide additional information on student characteristics, allowing to efficiently address the probable endogenous class allocation of students. In addition, the class-level information of TIMSS enables a better understanding of the phenomenon. Indeed, peer effects are plausibly stronger at class-level, rather than at school-level (or even country-level).

From the empirical point of view, the identification strategy relies on the random idiosyncratic variations in non-native class shares within schools, assuming that these changes are uncorrelated with relevant unobserved characteristics. Therefore as baseline strategy a school-fixed effect OLS regression is run on the country-pooled data. The main results are that non-native students do not averagely affect the math performance of native pupils.

However, as seen in *Chapter 1*, literature and theoretical models suggest heterogeneous and nonlinear nature of peer effects. Therefore, the present study investigates these possibilities with several strategies, such as partitions, quantiles, separate and piecewise regressions. The procedures support the heterogeneous results, in particular for ability.[15]

The most interesting outcome is the unusual non-linear estimated peer effects. My analyses, although not allowing to correct point out the optimal immigrant class concentration, disclose that while moving along the level of non-native share the peer effects are initially negative, then become positive

[15]Further details in Chapter 4.

and finally return negative. Nevertheless, the effects remain weak. Indeed, with a medium score of 500 points, the biggest impact is of almost 25 points in the dummy partition (Table 4.4). While the effects are less than 5 points impact for percentage changes in the immigrant share when looking at the marginal effects computed in the piecewise regressions (Table 4.5).

In regard to the theoretical considerations of *Chapter 1*, none of the stylized models singularly predicts correctly the particular and complex tendency of peer effects. Despite this, by emerging and hiding each other mutually along the distribution of the immigrant class concentration, the models may concurrently contribute to constitute the peer effects.

In the low immigrant concentration case, all the three described models suggest small insignificant negative effects. While *bad apple* evokes an increase in disruption propensity, *boutique model* justifies that with a rise in classroom heterogeneity. However, as non-native proportion rises, the impact due to rising disruption decreases and so does heterogeneity, given the presence already in the class of a fair amount of immigrants. Moreover, the moderate number of non-natives might lead to a slightly slower teaching pace, thus possibly improving the results of non-top-class students. Hence, if benefits overcome the potential disadvantages for top-class students, the overall peer effects result positive.[16] Finally, at large non-native share, *integration model* emerges. Indeed, if immigrants become numerous, the integration processes become quite costly, driving natives to ostracize and segregate non-native pupils. This deterioration of the school environment harms everybody's education attainments.

On the other hand, besides the *compound effect* of peer models, other possible explanations could justify the results obtained in this study. For example, pretty anxious about immigration, principals and public politicians might react at high share of non-native students (Ballatore et al.,2014). Plausibly, they may ignore immigrant effects, when the non-native share is low. But as it rises, they might increase the resources available for those

[16]In section 4.3, the quantile regression reports that peer effects are stronger in the bottom quantiles of the students' ability distribution, thus supporting the possibility of positive immigrant peer effects.

classes, which will thus improve students performance. Yet, if non-natives reach high levels of concentration, the impact of the additional resources could be overtaken by the negative immigrant peer effects. Otherwise, class within a school, which has higher concentration of non-native pupils can be symptomatic of a very bad or under-endowed class.[17]

To sum up, findings presented constitute a substantial contribution to the limited literature of non-native peer effects, in particular for European countries. Although this study can not estimate a precise threshold, the results suggest that immigrant shares below an undefined threshold would improve the native performances. Thus, the growing anxiety about the potential negative spillovers seems ill-founded, especially seen their really weak impact. However, further researches is necessary to support the findings, because the randomness within school is not completely guaranteed, and moreover the sample, being a survey database, is plausible serious affected by measurement errors. By virtue of these potential threats, diverse data-sets and different methods should be used, in order to not only support or reject my results, but also to disentangle other interesting unsolved issues. For example, the measurement of the endogenous or exogenous components of immigrant peer effects could be interesting. In addition, a study verifying whether immigrants have always the same impact or not and whether immigrants and natives share a similar language or culture could be of interest to improve the knowledge on the topic. Furthermore, also some further analyses on the test score impact in different subjects and grades could provide better insight.

In conclusion, the immigration effects on education is a subject with yet multiple open questions and still little evidence. Therefore, seen the relevance that immigration is going to have in the future for all the European countries, further research is required.

[17]Different provision of resources per class within school is not so unlikely. Indeed, it can be quite normal to observe this different treatment in schools practicing tracking systems.

Bibliography

[1] Ammermueller, A. (2007) *Poor background or low returns? Why immigrant students in Germany perform so poorly in the Programme for International Student Assessment.* Education Economics, 15(2), 215-230.

[2] Ammermueller, A., Pischke, J.S. (2009) *Peer effects in European primary schools: Evidence from PIRLS.* J. Labor Econ. 27(3): 315-348.

[3] Angrist, J.D., Lang, K. (2004) *Does school integration generate peer effects? Evidence from Boston's Metco program.* Am. Econ. Rev. 94, 1613-1634.

[4] Angrist, J.D., Pischke, J.S. (2008) Mostly Harmless Econometrics: An empiricist's companion. Princeston University Press, Princeston.

[5] Fryer Jr, R.G., Torelli, P. (2010) *An Economic Analysis of "Acting White".* Journal of Public Economics 94. 380–396.

[6] Ballatore, M.D., Fort, M., Ichino, A. (2014) *The Tower of Babel in the classroom: immigrant and natives in Italian schools.* IZA Discussion Paper, No. 8732.

[7] Bernheim, B. D. (1994) *A theory of conformity.* J. Polit. Econ., 841-877.

[8] Brunello, G., Rocco, L. (2013) *The effect of immigration on the school performance of natives: Cross Country evidence using PISA test scores.* Econ. Educ. Rev. 32(C): 234-246.

[9] Burke, M.A., Sass, T.R. (2013) *Classroom peer effects and student achievement.* Journal of Labor Economics, 31(1), 51-82.

[10] Cameron, A.C., P.K. Trivedi (2009) Microeconometrics using Stata. College Station, TX, Stata Press.

[11] Canay, I.A. (2011) *A simple approach to quantile regression for panel data.* The Econometrics Journal, 14(3), 368-386.

[12] Card, D. (2013) *Peer effects of immigrant children on academic performance of native speakers: Introduction.* Econ. J. 123(570): F279-F280.

[13] Case, A. (1991) *Spatial Patterns in Household Demand.* Econometrica, 59, 953-965.

[14] Čiarnienė, R., Kumpikaitė, V., (2008) *The Impact of Globalization on Migration Processes.* Social Research, Siauliai, 13(3), 42-48.

[15] Chaparro, F.P., Gonzalez, D.S., Rodriguez, R.S. (2013) *The Impact of Immigrant Concentration in Schools on Grade Retention in Spain: a Difference-in-Difference Approach.* MPRA Paper, No. 46888.

[16] Clarke, P., Crawford, C., Steele, F., Vignoles, A. F. (2010). *The choice between fixed and random effects models: some considerations for educational research. IZA Discussion Paper No. 5287*

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

[18] Colding, B., Husted, L., Hummelgaard, H. (2009) *Educational progression of second-generation immigrants and immigrant children.* Economics of Education Review, 28(4), 434-443.

[19] Conlisk, J. (1980) *Costly optimizers versus cheap imitators* J. Econ. Behav. Organ., 1(3), 275-293.

[20] Contini, D. (2013) *Immigrant background peer effects in Italian schools.* Soc. Sci. Res. 42(4): 1122-1142.

[21] Cortes, K.E. (2006) *The effects of age at arrival and enclave schools on the academic performance of immigrant children.* Econ. Educ. Rev., 25(2): 121-132.

[22] Crawford, C., Dearden, L., Meghir, C. (2010) *When you are born matters: the impact of date of birth on educational outcomes in England.* IFS working papers No. 10, 06.

[23] Epple, D., Romano, R.E. (2011) *Peer effects in education: a survey of theory and evidence.* In: Benhabib J., Bisin A., Jackson M. (eds) Handbook of social economics, vol 1B, ch. 4. North-Holland, Amsterdam.

[24] Dustmann, C., Okatenko, A. (2014) *Out-migration, wealth constraints, and the quality of local amenities.* Journal of Development Economics, 110, 52-63.

[25] Duflo, E., Dupas, P., Kremer, M. (2008) *Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya.* (No. w14475). National Bureau of Economic Research.

[26] Foy, P., Arora, A., Stanco G.M. (2013) TIMSS 2011 User Guide for the International Database. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

[27] Geay, C., McNally, S., Telhaj, S. (2013) *Non-native Speakers of English in the Classroom: What Are the Effects on Pupil Performance?* Economic Journal, 123(570), F281-F307.

[28] Georgiades, K., Boyle, M.H., Fife, K.A. (2013) *Emotional and behavioral problems among adolescent students: the role of immigrant, racial/ethnic congruence and belongingness in schools.* Youth Adolesc. 42(9): 1473-1492.

[29] Gould, E.D., Lavy, V., Paserman, M.D. (2009) *Does immigration affects the long-term educational outcomes of natives? Quasi-experimental evidence.* Econ. J. 119(540): 1243-1269.

[30] Goux, D., Maurin, E. (2007) *Close neighbours matter: Neighbourhood effects on early performance at school.* The Economic Journal, 117(523), 1193-1215.

[31] Grinblatt, M., Keloharju, M., Ikäheimo, S. (2008) *Social influence and consumption: Evidence from the automobile purchases of neighbors.* Review of Economics and Statistics, 90(4), 735-753.

[32] Hanushek, E.A., Kain, J.F., Rivkin, S.G. (2002) *New evidence about Brown v. Board of Education: The complex effects of school racial composition on achievement.* NBER Working Paper, No. 8741.

[33] Harris, J.R. (1998) The Nurture Assumption: Why children turn out the way they do. 2nd edition (2009) New York: Free Press.

[34] Hox, J. (2010) Multilevel analysis: Techniques and applications. Routledge.

[35] Hoxby, C.M. (2000) *Peer effects in the classroom: learning from gender and race variation.* NBER Working Paper, No. 7867.

[36] Hoxby, C.M., Weingarth, G. (2005) *Taking race out of the equation: school reassignment and the structure of peer effects.* Unpublished Working Paper.

[37] Hunt, J. (2012) *The impact of immigration on the educational attainment of natives* NBER Working Paper, No. 18047.

[38] Imberbam, S.A., Kugler, A., Sacerdote, B.I. (2012) *Katrina's children: evidence on the structure of peer effects from hurricane evacuees.* Am. Econ. Rev. 102(5): 2048-2082.

[39] Jansen, P., Rasmussen, A.W. (2011) *The effect of ethnic concentration schools on native and immigrant children's reading and math skills.* Econ. Educ. Rev. 30(6): 1503-1515.

[40] Jonsson, J.O., Mood, C. (2008) *Choice by contrast in Swedish schools: How peers' achievement affects educational choice.* Soc. Forces 87(2).

[41] Lazear, E.P. (1999) *Culture and language.* J. Public Econ. 107(S6)

[42] Lazear, E.P. (2001) *Educational production* Q.J. Econ. 116(3): 551-583.

[43] Manski, C.F. (1993) *Identification of endogenous social effects: the reflection problem.* Rev. Econ. Stud. 60(3): 531-542.

[44] Martin, M.O., Mullis, I.V.S. (2012). Methods and procedures in TIMSS and PIRLS 2011. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

[45] Martin, M.O., Mullis, I.V.S., Foy, P., Arora, A. (2012). TIMSS 2011 International Results in Mathematics. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

[46] MIUR (2012). *Gli alunni con cittadinanza non italiana nel sistema scolastico italiano.* Annual report. Italian Ministry of Education Statistical Office.

[47] Nelson, F. (2015, September 3) "Prepare yourselves: The Great Migration will be with us for decades." *The Telegraph.* [18]

[48] Nordland, R. (2015, November 1) "A Mass Migration Crisis, and It May Yet Get Worse." *The New York Time*, p. A6.

[49] OECD (2005) PISA Data Analysis Manual, 2nd edition (2009), OECD, Paris.

[50] OECD (2013) PISA Results 2012. OECD, Paris.

[51] Ohinata, A., Van Ours, J.C. (2012) *Spillover effects of studying with immigrant students in the same classroom: Evidence from Quantile regression analysis.*[19]

[52] Ohinata, A., Van Ours, J.C. (2013) *How immigrant children affect the academic achievement of native Dutch children.* Econ. J. 123(579): F308-331.

[53] Rangvid, B.S. (2007) *School composition effects in Denmark: quantile regression evidence from PISA 2000.* Empirical Economics, 33(2), 359-388.

[54] Rangvid, B.S. (2010) *School choice, universal vouchers and native flight from local schools.* European Sociological Review, 26(3),319–335.

[18] Retrieved from: *www.telegraph.co.uk/news/uknews/immigration*
[19] Retrieved from: $http://cream.conference-services.net$

[55] Sacerdote, B. (2011) *Peer Effects in Education: How might they work, How big they are and How much do we know thus far?* In: Hanushek, E.A., Machin, S., Woessman, L., (eds) Handbook of the economics of education, vol. 3, ch. 4. North-Holland, Amsterdam.

[56] Schneeweis, N. (2011) *Educational institutions and the integration of migrants.* J. Pop. Econ., 24(4), 1281-1308.

[57] Schnepf, S.V. (2007) *Immigrants' edcational disadvantage: an examination across ten countries and three surveys.* J. Pop. Econ., vol. 20: 527-545.

[58] Silaghi, F.R. (2011) *Immigration and peer effects: evidence from primary education in Spain.* Master thesis CEMFI.

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

[60] Tonello, M. (2015) *Peer effects of non-native students on natives' educational outcomes: mechanism and evidence.* Empirical Economics, 1-32.

[61] UNESCO (2006) *International Standard Classification of Education: ISCED 1997.*

[62] Vigdor, J.L., Nechyba, T.S. (2007) *Peer effects in North Carolina public schools.* In: Woessmann, L., Peterson, P.E., (eds) Schools and the Equal Opportunity Problem. The MIT Press, pp. 73-101.

[63] Zavodny, M. (1999) *Determinants of recent immigrants' locational choices.* International Migration Review, 1014-1030.

Appendix

Table A1: Summary statistics

	obs.	mean	sd
Math score	40854	495.2	*(77.1)*
Immigrant	40573	0.05	*(0.22)*
Female stud.	40849	0.49	*(0.50)*
Book *(Index)*	40508	3.04	*(1.25)*
Book *[0-25)*	40508	0.35	*(0.48)*
Book *[26-100)*	40508	0.30	*(0.46)*
Book *[>100)*	40508	0.34	*(0.47)*
PARED *(Index)*	30410	3.78	*(1.03)*
Quarter *(Index)*	40665	1.97	*(0.81)*
Repeatears	40847	0.05	*(0.22)*
Teacher stats			
Female teach.	39367	0.33	*(0.46)*
10ys of teach.	38943	0.75	*(0.43)*
Master grad.	39260	0.24	*(0.42)*
School stats			
Citysize*(Index)*	38435	1.84	*(0.81)*
Citysize*(small)*	38435	0.42	*(0.49)*
Citysize*(med.)*	38435	0.32	*(0.47)*
Citysize*(high)*	38435	0.26	*(0.44)*

	obs.	mean	sd
Cityincome*(Index)*	34727	1.75	*(0.55)*
Cityincome*(small)*	34727	0.30	*(0.46)*
Cityincome*(med.)*	34727	0.64	*(0.48)*
Cityincome*(high)*	34727	0.05	*(0.23)*
Class size	40854	21.22	*(5.01)*
School resources			
Material*(Index)*	38365	2.53	*(0.57)*
Material*(low)*	38365	0.04	*(0.19)*
Material*(med.)*	38365	0.40	*(0.49)*
Material*(high)*	38365	0.56	*(0.50)*
Room*(Index)*	38263	2.37	*(0.62)*
Room*(low)*	38263	0.08	*(0.27)*
Room*(med.)*	38263	0.48	*(0.50)*
Room*(high)*	38263	0.44	*(0.50)*
Staff*(Index)*	38075	2.33	*(0.56)*
Staff*(low)*	38075	0.05	*(0.21)*
Staff*(med.)*	38075	0.57	*(0.49)*
Staff*(high)*	38075	0.38	*(0.48)*

The stastics are computed on the whole sample — dropped the classes with less than 10 students. As suggested by the TIMMS user guide, the weights used are the SENWGT, because they treat each country equally, independently on the selected units or the real number of elegible students. Using those weights, each country has a weighted sample size of 500.

Table A2: Sample characteristics and Percentages of missing values

	ENG	FIN	HUN	ITA	LTU	NOR	ROM	SVN	SWE
No. of schools	118	145	146	197	141	134	147	186	153
– *1st subsample*	**117**	**145**	**143**	**195**	**129**	**129**	**143**	**181**	**151**
– *2nd subsample*	*48*	*83*	*102*	*7*	*112*	*36*	*100*	*39*	*105*
No. of classes	176	258	251	205	258	170	248	225	266
– *1st subsample*	**165**	**228**	**247**	**202**	**241**	**165**	**243**	**220**	**256**
– *2nd subsample*	*96*	*166*	*206*	*14*	*224*	*72*	*200*	*78*	*210*
No. of students	3842	4266	5178	3979	4747	3862	5523	4415	5573
– *1st subsample*	**3762**	**4149**	**5152**	**3953**	**4631**	**3828**	**4375**	**4375**	**5520**
– *2nd subsample*	*2201*	*3038*	*4356*	*299*	*4389*	*1762*	*4607*	*1689*	*4503*
Missing values (%)									
Student stats									
Immshare	0.80	0.94	0.19	0.53	0.43	0.50	0.73	0.69	1.50
Student female	0	0	0	0	0	0	0	0	0.09
Books	0.72	1.11	0.21	0.61	0.41	1.07	0.75	0.85	1.87
Semester	0	0.02	0	0	0.02	0.03	0	0.02	0.09
Repeater	0	0.02	0	0	0.02	0	0	0	0.09
PARED	43.57	28.73	11.55	10.14	20.77	39.50	16.58	25.56	44.09
Teacher stats									
Female	7.71	2.92	2.19	4.65	0.45	1.62	0.88	1.76	10.34
Master grad.	9.20	3.49	2.19	6.65	0	1.62	0.88	1.76	9.78
Yrs. teaching	10.77	2.92	2.19	7.44	0	3.03	4.14	2.17	9.78
School stats									
Citysize	7.74	7.25	3.63	2.96	4.56	2.19	1.48	4.69	17.07
CityIncome	6.67	8.24	3.63	3.82	3.95	100	0.73	5.28	16.56
Materials	7.34	8.22	3.34	1.32	6.69	1.59	1.46	4.89	17.81
Rooms	6.03	9.23	3.61	1.42	7.69	2.72	0.47	5.14	18.62
Staff	6.03	9.06	6.02	2.07	6.69	2.35	1.15	6.90	18.46

The 1st subsampling consists to discard the classes attended by less than 10 students. While, the 2nd subsampling the schools having less than 2 classes surveyed. The 1st subsample is highlighted because it represents the reference sample in the following analyses; for exception of the school fixed effect analyses, which adopt the 2nd subsample.

Table A3: Weighted Mean values of individual statistics

(Standard deviation in parentheses)

Table A3.A

	ENG	FIN	HUN	ITA⋆	LTU
Math Score	509.0 *(81.7)*	516.4 *(60.2)*	505.9 *(86.6)*	498.7 *(70.0)*	505.1 *(75.6)*
Immigrants	0.09 *(0.29)*	0.03 *(0.18)*	0.03 *(0.16)*	0.07 *(0.25)*	0.01 *(0.12)*
Girl	0.48 *(0.50)*	0.49 *(0.50)*	0.49 *(0.50)*	0.49 *(0.50)*	0.49 *(0.50)*
BOOK	2.95 *(1.30)*	3.30 *(1.16)*	3.22 *(1.29)*	3.03 *(1.25)*	2.81 *(1.14)*
– *Book* [0-25)	0.38 *(0.49)*	0.24 *(0.43)*	0.31 *(0.46)*	0.37 *(0.48)*	0.41 *(0.49)*
– *Book [25-100)*	0.28 *(0.45)*	0.35 *(0.48)*	0.27 *(0.45)*	0.29 *(0.45)*	0.34 *(0.47)*
– *Book [>100)*	0.33 *(0.47)*	0.41 *(0.49)*	0.41 *(0.49)*	0.34 *(0.47)*	0.24 *(0.43)*
1^{st}Semester	0.50 *(0.50)*	0.51 *(0.50)*	0.49 *(0.50)*	0.50 *(0.50)*	0.52 *(0.50)*
Repeater	0.00 *(0.03)*	0.04 *(0.19)*	0.12 *(0.32)*	0.07 *(0.26)*	0.08 *(0.27)*
PARED	3.75 *(1.05)*	3.94 *(1.01)*	3.53 *(0.99)*	3.32 *(1.13)*	3.85 *(0.88)*
– *Parent Grad.*	0.32 *(0.47)*	0.42 *(0.49)*	0.26 *(0.44)*	0.24 *(0.43)*	0.26 *(0.44)*

Table A3.B

	NOR⋆	ROM	SVN⋆	SWE	MEAN
Math Score	474.6 *(61.7)*	458.4 *(98.5)*	504.9 *(67.7)*	484.3 *(64.3)*	495.2 *(77.1)*
Immigrants	0.08 *(0.28)*	0.01 *(0.09)*	0.04 *(0.20)*	0.08 *(0.28)*	0.05 *(0.22)*
Girl	0.49 *(0.50)*	0.48 *(0.50)*	0.49 *(0.50)*	0.48 *(0.50)*	0.49 *(0.50)*
BOOK	3.35 *(1.23)*	2.51 *(1.20)*	2.91 *(1.14)*	3.23 *(1.29)*	3.04 *(1.25)*
– *Book* [0-25)	0.25 *(0.43)*	0.54 *(0.27)*	0.37 *(0.48)*	0.31 *(0.46)*	0.35 *(0.48)*
– *Book [25-100)*	0.31 *(0.46)*	0.26 *(0.44)*	0.36 *(0.48)*	0.28 *(0.45)*	0.30 *(0.46)*
– *Book [>100)*	0.45 *(0.50)*	0.19 *(0.39)*	0.27 *(0.44)*	0.42 *(0.49)*	0.34 *(0.47)*
1^{st}Semester	0.53 *(0.50)*	0.55 *(0.50)*	0.50 *(0.50)*	0.53 *(0.50)*	0.51 *(0.50)*
Repeater	0.01 *(0.09)*	0.11 *(0.32)*	0.04 *(0.19)*	0.02 *(0.15)*	0.05 *(0.21)*
PARED	4.44 *(0.87)*	3.40 *(0.99)*	4.06 *(0.81)*	4.15 *(0.97)*	3.78 *(1.03)*
– *Parent Grad.*	0.62 *(0.48)*	0.20 *(0.40)*	0.31 *(0.46)*	0.47 *(0.50)*	0.33 *(0.47)*

BOOK and **PARED** are indexes, which take values respectively of (1-5) and of (1-5).

⋆ In these countries, the standard cohort for the 8^{th}graders is the 1997, instead of the 1996.

Table A4: Mean values for the class and school statistics.

Table A4.A

	ENG	FIN	HUN	ITA	LTU
Class stats					
Teacher sex	0.50 (0.45)	0.47 (0.45)	0.16 (0.33)	0.21 (0.41)	0.04 (0.16)
10ys teaching	0.48 *(0.47)*	0.65 *(0.41)*	0.88 *(0.30)*	0.80 *(0.40)*	0.92 *(0.25)*
Teacher Mast.	0.39 *(0.43)*	0.79 *(0.36)*	0.20 *(0.36)*	0.26 *(0.43)*	0.30 *(0.39)*
Class size	22.90 *(5.13)*	18.28 *(3.07)*	20.83 *(4.96)*	19.51 *(4.21)*	18.97 *(3.76)*
School stats					
City size	2.35 *(0.73)*	1.92 *(0.76)*	1.83 *(0.84)*	1.72 *(0.75)*	1.87 *(0.89)*
– *Small*	0.15 *(0.36)*	0.33 *(0.47)*	0.44 *(0.50)*	0.45 *(0.50)*	0.46 *(0.50)*
– *Medium*	0.35 *(0.48)*	0.42 *(0.49)*	0.28 *(0.45)*	0.36 *(0.48)*	0.20 *(0.40)*
– *Large*	0.49 *(0.50)*	0.25 *(0.43)*	0.28 *(0.45)*	0.18 *(0.38)*	0.33 *(0.47)*
City income	1.81 *(0.68)*	1.98 *(0.30)*	1.53 *(0.53)*	1.89 *(0.47)*	1.55 *(0.52)*
– *Low*	0.34 *(0.48)*	0.05 *(0.22)*	0.48 *(0.50)*	0.17 *(0.38)*	0.46 *(0.50)*
– *Medium*	0.51 *(0.50)*	0.91 *(0.29)*	0.50 *(0.50)*	0.77 *(0.42)*	0.53 *(0.50)*
– *High*	0.15 *(0.36)*	0.04 *(0.19)*	0.01 *(0.12)*	0.06 *(0.24)*	0.01 *(0.09)*
Material	2.56 *(0.52)*	2.58 *(0.52)*	2.46 *(0.66)*	2.43 *(0.54)*	2.53 *(0.55)*
– *Low*	0.01 *(0.10)*	0.01 *(0.12)*	0.09 *(0.29)*	0.02 *(0.14)*	0.02 *(0.16)*
– *Medium*	0.43 *(0.50)*	0.39 *(0.49)*	0.35 *(0.48)*	0.53 *(0.50)*	0.42 *(0.50)*
– *High*	0.56 *(0.50)*	0.59 *(0.49)*	0.56 *(0.50)*	0.45 *(0.50)*	0.55 *(0.50)*
Room	2.24 *(0.63)*	2.25 *(0.57)*	2.36 *(0.65)*	2.27 *(0.64)*	2.46 *(0.61)*
– *Low*	0.10 *(0.30)*	0.07 *(0.25)*	0.09 *(0.29)*	0.10 *(0.31)*	0.06 *(0.24)*
– *Medium*	0.55 *(0.50)*	0.62 *(0.49)*	0.45 *(0.50)*	0.53 *(0.50)*	0.42 *(0.50)*
– *High*	0.35 *(0.48)*	0.31 *(0.47)*	0.45 *(0.50)*	0.37 *(0.48)*	0.52 *(0.50)*
Staff	2.52 *(0.50)*	2.27 *(0.48)*	2.26 *(0.58)*	1.98 *(0.51)*	2.45 *(0.60)*
– *Low*	0	0.01 *(0.12)*	0.07 *(0.26)*	0.14 *(0.34)*	0.06 *(0.23)*
– *Medium*	0.48 *(0.50)*	0.70 *(0.46)*	0.59 *(0.49)*	0.74 *(0.44)*	0.44 *(0.50)*
– *High*	0.52 *(0.50)*	0.29 *(0.45)*	0.33 *(0.47)*	0.12 *(0.33)*	0.50 *(0.50)*

Table A4: Mean values for the class and school statistics.

Table A4.B

	NOR	ROM	SVN	SWE	MEAN
Class stats					
Teacher sex	0.59 (0.46)	0.41 (0.44)	0.18 (0.29)	0.46 (0.41)	0.32 (0.42)
10ys teaching	0.57 *(0.48)*	0.90 *(0.26)*	0.82 *(0.37)*	0.70 *(0.42)*	0.76 *(0.40)*
Teacher Mast.	0.01 *(0.09)*	0.19 *(0.33)*	0.01 *(0.05)*	0.02 *(0.12)*	0.23 *(0.39)*
Class size	22.87 *(5.55)*	22.44 *(5.08)*	19.51 *(4.15)*	21.75 *(4.72)*	20.66 *(4.81)*
School stats					
City size	1.94 *(0.75)*	1.70 *(0.87)*	1.41 *(0.68)*	1.85 *(0.76)*	1.81 *(0.81)*
– *Small*	0.31 *(0.46)*	0.57 *(0.50)*	0.70 *(0.46)*	0.37 *(0.48)*	0.44 *(0.50)*
– *Medium*	0.44 *(0.50)*	0.16 *(0.37)*	0.19 *(0.40)*	0.40 *(0.49)*	0.31 *(0.46)*
– *Large*	0.25 *(0.44)*	0.27 *(0.44)*	0.11 *(0.31)*	0.23 *(0.42)*	0.25 *(0.43)*
City income	.	1.62 *(0.57)*	1.62 *(0.51)*	1.91 *(0.58)*	1.74 *(0.55)*
– *Low*	.	0.42 *(0.50)*	0.39 *(0.49)*	0.22 *(0.41)*	0.31 *(0.46)*
– *Medium*	.	0.53 *(0.50)*	0.60 *(0.49)*	0.65 *(0.48)*	0.63 *(0.48)*
– *High*	.	0.04 *(0.20)*	0.01 *(0.11)*	0.13 *(0.34)*	0.05 *(0.22)*
Material	2.70 *(0.48)*	2.08 *(0.63)*	2.85 *(0.36)*	2.48 *(0.58)*	2.52 *(0.57)*
– *Low*	0.01 *(0.09)*	0.16 *(0.36)*	0	0.04 *(0.20)*	0.04 *(0.20)*
– *Medium*	0.28 *(0.45)*	0.60 *(0.49)*	0.15 *(0.36)*	0.43 *(0.50)*	0.40 *(0.49)*
– *High*	0.71 *(0.46)*	0.24 *(0.43)*	0.85 *(0.36)*	0.52 *(0.50)*	0.56 *(0.50)*
Room	2.43 *(0.53)*	2.27 *(0.72)*	2.50 *(0.56)*	2.57 *(0.62)*	2.37 *(0.62)*
– *Low*	0.01 *(0.13)*	0.15 *(0.36)*	0.03 *(0.17)*	0.07 *(0.25)*	0.08 *(0.27)*
– *Medium*	0.54 *(0.50)*	0.41 *(0.49)*	0.44 *(0.50)*	0.30 *(0.46)*	0.47 *(0.50)*
– *High*	0.45 *(0.50)*	0.43 *(0.50)*	0.53 *(0.50)*	0.64 *(0.48)*	0.45 *(0.50)*
Staff	2.29 *(0.46)*	2.24 *(0.60)*	2.58 *(0.53)*	2.43 *(0.54)*	2.32 *(0.56)*
– *Low*	0	0.08 *(0.28)*	0.01 *(0.11)*	0.02 *(0.16)*	0.05 *(0.22)*
– *Medium*	0.71 *(0.46)*	0.59 *(0.49)*	0.40 *(0.49)*	0.52 *(0.50)*	0.58 *(0.49)*
– *High*	0.29 *(0.46)*	0.33 *(0.47)*	0.59 *(0.49)*	0.45 *(0.50)*	0.37 *(0.48)*

Table A5: Student's socio-economic: Heterogeneous effects.
(Standard deviation in parentheses)

	FEW	SOME	MANY
Panel A: Immigrant class-concentration continuous variable			
IMMSHARE	–14.76	41.58	51.00
	(50.15)	*(32.44)*	*(32.31)*
Panel B: Categorical variable of Immigrant share.			
Low Immshare	–14.23**	0.43	–1.52
	(6.27)	*(4.50)*	*(4.74)*
Medium Immshare	12.64*	10.26*	13.93***
	(6.96)	*(5.33)*	*(4.51)*
High Immshare	–37.81**	–36.56**	–14.97
	(16.15)	*(15.69)*	*(16.85)*

Weighted least squares regression with school fixed effects, and the usual set of covariates, as in regression (9) Table 4.1. In *Panel A*, I exploit three separate regression for student' socio-economic status, on the continuous IMMSHARE variable. In *Panel B*, the studied independent variable is the categorical variable built by a partition of the continuous variable IMMSHARE. The categorical variable is divided in Low when the immigrant concentration is $\leq 8\%$, Medium when $\leq 35\%$, and High when $> 35\%$. The division in Few, Some, and Many books follows how described in Section 2.1, Few when books possessed are less than 25, Some whether between 25 and 100, Many if over 100.

Table A6: Teacher's experience: Heterogeneous effects.

(Standard deviation in parentheses)

	BEGINNER	EXPERT
Panel A: Immigrant class-concentration continuous variable		
IMMSHARE	11.70	30.19
	(40.63)	*(43.66)*
Panel B: Categorical variable of Immigrant share.		
Low Immshare	–5.35	–10.80
	(4.40)	*(6.23)*
Medium Immshare	8.10	18.87**
	(6.04)	*(10.73)*
High Immshare	–20.91***	–37.71
	(6.02)	*(26.36)*

Weighted least squares regression with school fixed effects, and the usual set of covariates, as in regression (9) Table 4.1. In *Panel A*, I exploit two separate regression whether students face beginner and expert teacher on the continuous IMMSHARE variable. In *Panel B*, the studied independent variable is the categorical variable built by a partition of the continuous variable IMMSHARE. The categorical variable is divided in Low when the immigrant concentration is $\leq 8\%$, Medium when $\leq 35\%$, and High when $> 35\%$. A teacher is considered "expert" when she has experienced more than 10 years of teaching.

Table A7: Separate Baseline Regressions of EU15 and non-EU15.
(Standard deviation in parentheses)

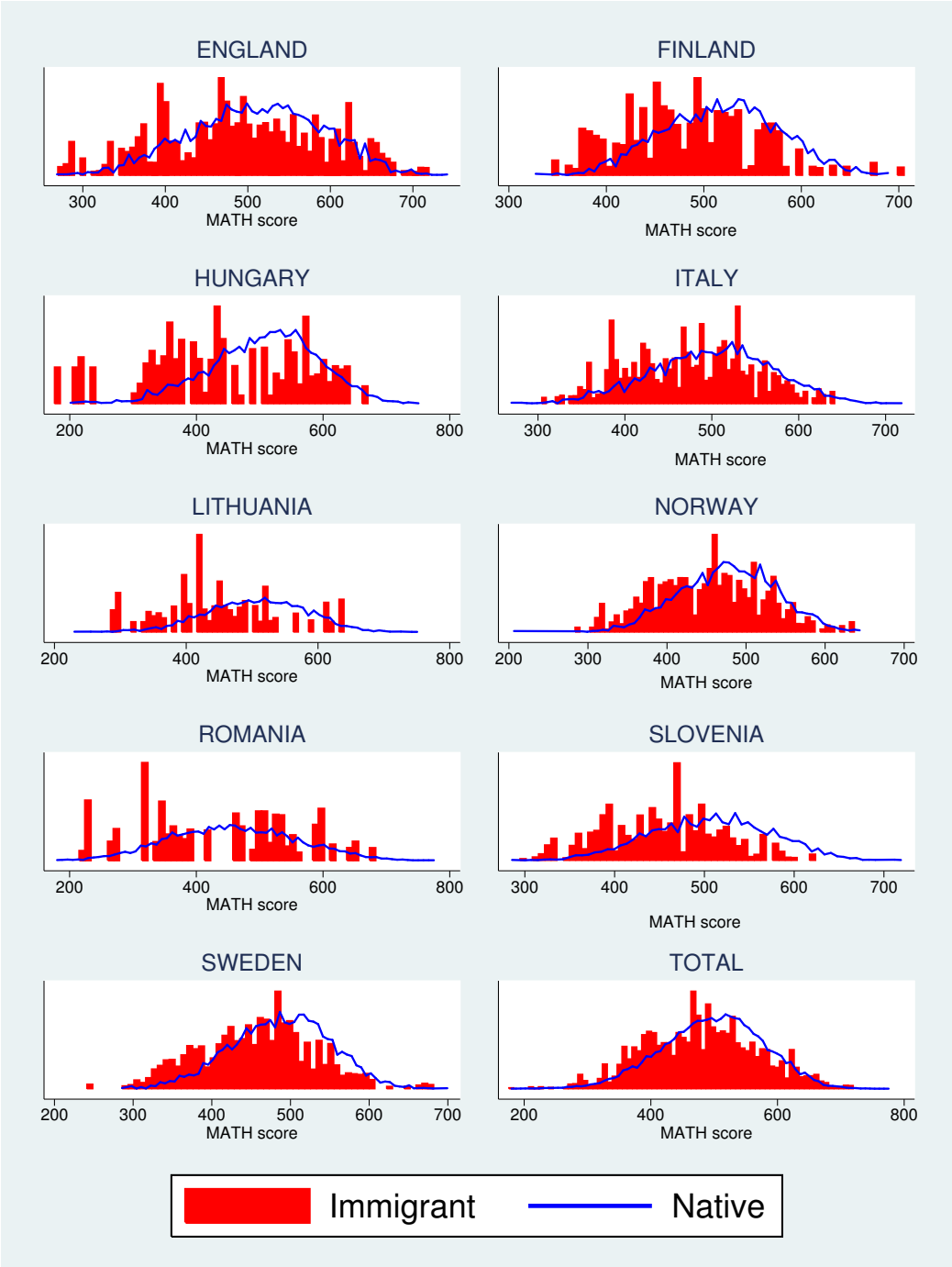
	EU15⋆	non-EU15
Peer effects	36.24 *(24.23)*	–20.77 *(66.06)*
Average non-natives	7.31% *(0.09)*	2.21% *(0.04)*
Observation	17 859	18 610

Weighted least squares regression with school fixed effects, and the usual set of covariates, as in regression (9) Table 4.1. I exploit two separate regression for EU15 and non-EU15 countries. Tthe studied independent variable is the usual continuous variable IMMSHARE. The dependent variable is the students' test scores in math. *Average non-native* indicates the average immigrant class-concentration in percentage.

The EU15 countries are England, Finland, Italy, Norway⋆, and Sweden. Instead, the non-EU15 countries are Hungary, Lithuania, Romania, and Slovenia.

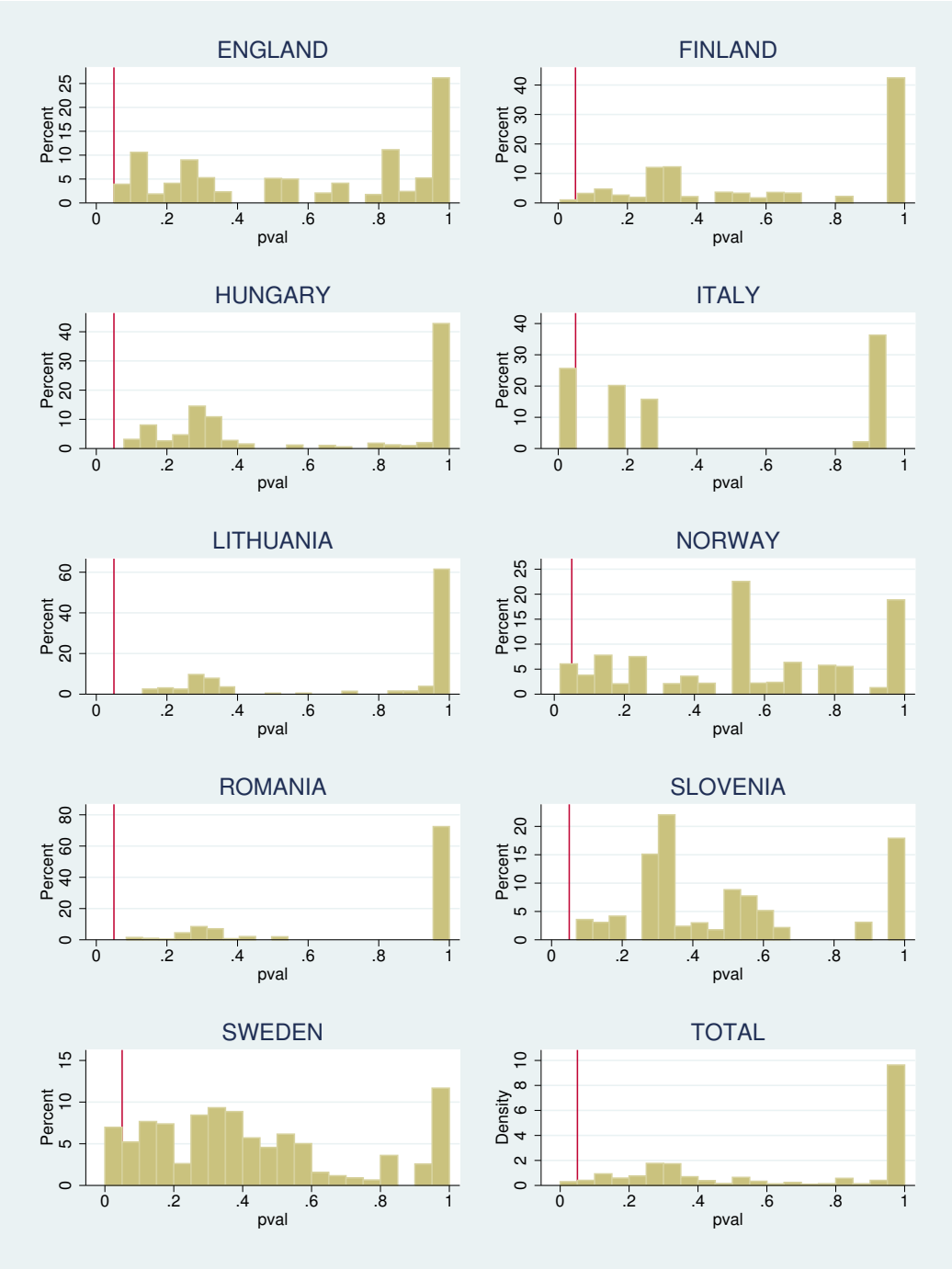
⋆ Norway should not belong to EU15, however I have included in that, because it is more similar to the EU15 countries. Further detaiil in *Section 4.3*.

Figure A2: Comparison MATH scores distribution: Natives vs Immigrants



The histograms describe the test scores distribution for native and immigrant students. The y-axis denotes the frequencies. Histograms are built using the HOUWGT, student weight system that ensures that the weighted sample corresponds to the actual sample size in each country.

Figure A3: Pearson's Chi^2 p-values distribution



The histograms describe the p-values of the Pearson's Chi^2, which test the random allocation of non-native students per class within school. scores distribution for native and immigrant students. The vertical red line indicates the p-value of the 5%. For further detail see *Section 3.2*.