The Tower of Babel in the Classroom?

Immigrants and Natives in Italian Schools*

Rosario Ballatore
Bank of Italy

Margherita Fort
University of Bologna

Andrea Ichino
University of Bologna

This version: October 25, 2013
work in progress

Abstract

We exploit rules of class formation to identify and estimate the causal effect of increasing the *number of immigrants* in a classroom keeping class size constant. Since principals use class size to neutralize the possible detrimental effects of immigrants' inflows, this composition effect is the relevant policy parameter but it has been neglected by the literature so far. Our empirical analysis is guided by a model of the educational production function that allows for two types of students, natives and immigrants, who have a non-negative probability of generating disruption or positive externalities across or within groups. We show that the pure composition effect is substantial and negative at age 7 (-12% in language and -7% in mathematics), but vanishes when children grow up. The reasons of this vanishing of effects, even at constant class size, remains to be explored.

JEL Classification: C36, I20, I24, J15
Keywords: Education, Immigration, Integration, Instrumental Variable

*We would like to thank INVALSI and in particular Paolo Sestito, Patrizia Falzetti and Roberto Ricci for giving us access to the many sources of restricted data that have been used in this paper. We received useful comments from seminar participants at the European University Institute, the University of Bologna, the II International Workshop on Applied Economics of Education and the European Economic Association of Labour Economists, the Einaudi Institute for Economics and Finance. We also would like to thank Massimo Pietroni, Enrico Rettore, Paolo Sestito, Marco Tonello, Lucia Rizzica and Marco Paccagnella for their useful suggestions. We acknowledge financial support from MIUR- PRIN 2009 project 2009MAATFS_001. Margherita Fort is affiliated with IZA. Andrea Ichino is a research fellow of CEPR, CesIFO and IZA. The views here expressed are those of the authors and do not necessarily reflect those of the Bank of Italy.

1 Introduction

The integration of non-native children in classrooms is a problem that many countries are facing. Anecdotal evidence of class disruption involving immigrants often generates concerns in the public opinion. As a result of these concerns, educational authorities implement policies that are typically not based on reliable evidence of the real dimension of the problem. An example is the rule introduced by the Italian Ministry of Education, according to which no class should have more than 30% of immigrants: the reason why this threshold was chosen is unclear and certainly not based on experimental evidence. On the other hand, in partial excuse of what educational authorities can do, there is admittedly little evidence on the causal effect of an immigrants' inflow on the performance or natives and immigrants in a classroom. It is indeed difficult to find exogenous sources of variation in *class composition* that leave unchanged other factors entering the educational production function, such as class size and/or parental background.

In this paper, we estimate the causal effect of the *numbers* of immigrants and natives in a class on the performance of native students distinguishing between effects of *size* and *composition*. Identification is achieved exploiting discontinuities generated by the enrolment thresholds at which classes are split because of institutional rules of class formation. These discontinuities generate exogenous sources of variation in class size in schools with and without immigrants: while in the former both the number of natives and the number of immigrants in class is affected, in the latter only the number of natives changes discontinuously. In the same spirit of Nannicini and Gagliarducci (forthcoming), by contrasting the change in performance at the threshold in schools with and without immigrants, we can estimate the effect of the *numbers* of immigrants and natives in a class on the performance of native students, as well as the *pure composition change*, namely the effect of increasing the number of immigrants keeping class size constant.

We interpret the evidence using Lazear (2001) model of the educational production function in which the absolute numbers of students in a class determines class performance, if each student has some positive probability of generating disruption. We extend this model in two ways. First, by allowing for two types of students, natives and immigrants, each with its own probability of disruption. Second, by introducing the possibility that the behaviour of a student generates positive externalities, not only disruption, across or within groups. For example, natives may ask questions that are irrelevant or even disruptive for their learning process, but the teacher's answers to these questions may instead be very useful to immigrants.

We find that the effect of increasing the number of immigrants reducing at the same time the number of natives (a pure composition change with constant class size) is negative and amounts to a reduction of nearly 12% in the national language standardized test score and a reduction of 7% in the corresponding mathematics test score at age 7 (grade 2) but these effects vanish as children age.

The paper is organized as follows. After a brief review of the related literature (Section 2), we present the theoretical model that justifies our empirical analysis (Section 3). We replicate in Section 4 the estimates obtained with conventional approaches by the existing literature and we use them as a starting point to discuss identification issues in Section 5, where we also show how we address them. We then illustrate our main findings in Section 7. Details on the institutional setting, the data and the estimation are presented in the Appendix. Section 8 concludes.

2 Related Literature

Our paper bridges the literatures on segregation and on peer effects in the educational context, focusing specifically on how the presence of immigrants affects natives' performance. Both literatures face similar identification problems in exploring the inter- or intra- group interaction between subjects. First, the difficulty of finding credible exogenous sources of variability in the exposure to minorities, mainly because there is endogenous sorting across cities, schools and even classes within schools. Parents not only choose the school for their children on the basis of their income or preferences - which is likely to affect students' outcomes directly - but may influence the allocation of children to classes as well. Moreover, school policies concerning class formation may encourage allocation of students based on their ability or socio-economic background. Second, these difficulties are increased by the well known *reflection problem* (Manski, 1993).

Different identification strategies (aggregation, within school variability, fixed effects, instrumental variables) have been explored in the literature and here we focus on the most recent contributions.

Ammermuller and Pischke (2009), Contini (2011) and Ohinata and Van Ours (2011) address the problem of the endogenous sorting of immigrants between schools by exploring the variability in the share of immigrants students within schools between classes of a given grade while Hoxby (2000), Bossavie (2011), Tonello (2012) exploit the variability in ethnic composition between adjacent cohorts within the same schools. The first approach rests on the assumption that, once school fixed effects are controlled for, the allocation of immigrants between classes is as good as random; the second approach argues that the variability between subsequent cohorts is random when the data are aggregated at the school-cohort level. Ammermuller and Pischke (2009), Contini (2011) and Ohinata and Van Ours (2011) find a weak negative effect of immigrants concentration on average natives' performances, but also show that this effect becomes larger for students with low family background. Our evidence suggests that there may be substantial flaws in this approach: we show that immigrant pupils are more likely to be allocated in classes in which natives have a worse family background. Hoxby (2000), Bossavie (2011) and Tonello (2012) find a negative weak inter-race peer effect in language test scores (and no effect in mathematics performance), while the intra-race peer effect is found to be negative and stronger (i.e. high share of immigrants affects mainly the other immigrants' language test scores, while there are no effects in mathematics).

Using a detailed individual dataset for Texas students Hanushek, Kain, and Rivkin (2009) address the endogeneity in the exposure to black minority exploring the variability given by the movement of students between classes or schools but can also include controls for individual fixed effects. They find small negative (not significant in some cases) effects of black concentration on white performance and a sizeable reduction for black schoolmates.

Angrist and Lang (2004) use the quasi-experimental variation in exposure to minorities provided by the METCO de-segregation program and find no significant effect on white students and a negative (though small) effect on other minority students. Gould, Lavy, and Paserman (2009) use the mass immigration from Soviet Union that occurred in Israel during the 1990s to identify the long run causal effect of having immigrants as classmates, finding a negative effect of immigration on the probability of passing the high-school matriculation exam. Negative effects on students' performance are also found by Jensen and Rasmussen (2011), who use the immigrants concentration at larger geographical areas as an instrumental variable to deal with the endogeneity of the share of immigrant at school. Card and Rothstein (2007) overcome the endogenous sorting of students among schools by looking at the relationship between the black/white test score gap and the degree of segregation at the city level, rather than at the individual level. They find a negative effect of segregation at the school and neighbourhood level on the achievement gap, with the latter being stronger than the former. Along the same line, Brunello and Rocco (2011) aggregate the data at country level and use the within country variation in the share of immigrants at school finding small negative effects. On the contrary Hunt (2012), using cohort variation across US states and years, reports positive effects (though small) of immigrants' concentration on the probability that natives complete high-school.

All these contributions provide informative - but not conclusive - evidence on the effects of the presence of immigrants in a school. The majority of them report small (and in some cases not significant) negative effects of immigrant concentration on natives performance. But some positive estimates emerge as well in this literature. Most of these studies, however, exploit a source of variation in class composition that does not leave unchanged other factors that enter in the educational production function, such as class size, which is found to be important in influencing the performance of students (among others: Angrist and Lavy, 1999; Pellizzari, De Giorgi, and Woolston, 2012). Specifically, we believe that the ethnic composition of a class is not independent of class size. For instance, a principal could minimize the negative externalities coming from the presence of immigrants placing them in smaller classes. Furthermore, principals could also add extra classes when the immigrants enrolment increases. When this correlation is neglected the negative effect of immigrants' concentration in classes could be seriously underestimated.

Our contribution to this literature is twofold. First, we use a credible source of exogenous variability in the number of natives and immigrants in classes using the institutional rules of class formation in the same spirit of Angrist and Lavy (1999) but within the framework recently developed by Nannicini and Gagliarducci (forthcoming). Second, to our knowledge, we are first in this literature to focus on a policy-relevant parameter, the *pure immigrants-natives composition effect* (pure composition effect in what follows), which is the effect of a unitary increase in the number of immigrants keeping constant class size (i.e. reducing at the same time the number of natives). Class size is the most accessible instrument that principals have to neutralize the negative causal effect of an immigrants' inflow. Therefore, to evaluate the consequences of this inflow, class size must be controlled for and this is exactly what we are able to do in our empirical analysis. Note that this is possible precisely because we exploit *exogenous* variability in class size by adapting to our setting the Difference-in-Discontinuities Design of Nannicini and Gagliarducci (forthcoming) and the quasi-experimental approach of Angrist and Lavy (1999).

Differently than in the previous literature that relies on more conventional strategies, our results point in the direction of a sizeable reduction in natives performance when one immigrant is added to a class and the number of natives is correspondingly reduced, holding class size constant. While this finding confirms and measures precisely what is typically believed by the public opinion, we also find that this effect becomes insignificant in the final grade of primary school. We think that this is evidence that, after a sufficient amount of time, the school system and more generally the Italian society succeed in reducing the negative consequences of immigrants inflows in classes, through channels that are different from changes in class size. Whether these channels result from conscious choices of school operators and policy makers or are just a natural by-product of the educational process and of the permanence in the Italian society, is a question that we plan to explore in future research.

3 Why immigrants (and natives) may affect classroom's performance

In this Section we illustrate the theoretical framework that guides our empirical analysis.

We start from the model of the educational production function proposed by Lazear (2001) , but we adapt it for the possibility that a larger fraction of natives or immigrants may have not only a "detrimental" but also a "constructive" effect on the educational outcomes of immigrants and/or natives . The data will speak on which effect prevails.

The main idea of Lazear's model is that the time devoted by teachers to students in a classroom is a public good. A "private" use of this time (i.e. a student that asks or requires specific attention) creates negative externalities that spill over the entire class, affecting the performance of all the other pupils. Consider a class with C students. If no student asks for specific attention, all students benefit fully and equally from the time of the teacher. Let $P \in [0,1]$ be the probability that a student *does not require* specific attention by the teacher at the expenses of others. Then, the likelihood that the event of disruption does not occur is a function of the number of students in the class: P^C. Define $\bar{V}$ as the maximum performance of a student (as for example measured by a test) if the teacher could devote full attention to her. The actual performance is a fraction of maximum performance:

$$V = \bar{V}P^{C} < \bar{V} \tag{1}$$

The equation above shows that the performance of students is strictly linked with the size of the class C. Even small deviations from $P = 1$ (i.e. even rare episodes of disruption) can generate large performance losses when class size is large. For example, if $P = 0.98$ in a class of 25 students, $V = 0.6\bar{V}$. Therefore, a 2% individual probability of demanding specific teacher's attention decreases by 40% potential performance of the average student in the classroom.

We extend the above framework in two directions. First, we generalize the model allowing for the possibility of episodes in which the behaviour of a single student has a positive externality on classmates: for example a student may be asking interesting questions that allow the teacher to clarify points that are unclear to all students. Second, we adapt the framework to the possibility that students are of two types (natives and immigrants) and, thus, class composition, in terms of these two groups, matters for performance.

In order to introduce the first kind of extension, consider the following modified educational production function:

$$V = \bar{V}P^{\phi C} < \bar{V} \tag{2}$$

where $\phi \in \mathbb{R}\backslash\{0\}$. A negative $\phi < 0$, captures the situation of a constructive behaviour of students, so that the effect of class size on performance is reversed and becomes positive. If $\phi > 0$ we are back in Lazear's world in which behaviour is "mis-behaviour" and class size can only reduce performance. Therefore, under this extension, the effect of class size may change depending on ϕ for given P. In this way we can accommodate the possibility that a larger number of natives might affect positively the performance of immigrants (or vice versa). This happens, for example, if immigrants' proficiency in Italian benefits from a larger size of Italians asking questions that for them may be a waste of time but are useful to immigrants.

Before introducing the possibility of heterogeneous types of students, however, it is interesting to see how the optimal class size that would be chosen by a principal is affected by the introduction of the parameter ϕ in Lazear's model. The principal solves the problem:

$$\underset{C}{\mathrm{Max}}\, \Pi = P^{\phi C} - \frac{W}{C} \tag{3}$$

where W is the wage of a teacher and the rental cost of the capital she uses. The first order condition is:

$$\phi P^{\phi C} p + \frac{W}{C^2} = 0 \tag{4}$$

where $p = Ln(P)$, which implies that the optimal class size is:

$$C^* = f(\underset{+}{P}, \underset{+}{W}, \underset{-}{\phi}) \tag{5}$$

Therefore, optimal class size increases with the good behaviour of students (higher P) and with the cost of providing the educational public good (W). As ϕ increases (for example because the quality and usefulness of student's questions in class decreases) the optimal class size is smaller.

It is crucial to note for our purposes, as shown by Lazear, that the positive relationship between C and P at the optimum hides, in observational data, the negative causal effect of C on V keeping P constant. In other words, without an exogenous variation of class size C, independent of P, it is impossible to estimate the causal effect of class size on students' performance.

We now relax the hypothesis of student's homogeneity within a class. Assume that students are divided in two groups with different probabilities of disruption: N natives (with $P = P_n$ and $\bar{V} = \bar{V}_n$) and I immigrants (with $P = P_i$ and $\bar{V} = \bar{V}_i$). The test scores of the average native and immigrant in a class are:

$$V_n = \bar{V}_n P_n^{\phi_{nn} N} P_i^{\phi_{in} I} \tag{6}$$

$$V_i = \bar{V}_i P_n^{\phi_{ni} N} P_i^{\phi_{ii} I} \tag{7}$$

where $\phi_{rq} \in \{-\infty, +\infty\}$, with $r, q \in \{n, i\}$, captures the possibility that teachers' attention asked by group r affects group q differently than group r. For example, a question asked in class by a native may be a waste of time for other natives but beneficial for immigrants, in which case:

$$\phi_{nn} > 0 \text{ and } \phi_{ni} < 0.$$

Using small letters for logs and $h \in \{n, i\}$ for student's etnicity, the performance of a generic student can be written as:

$$v_h = \bar{v}_h + p_n \phi_{nh} N + p_i \phi_{ih} I \tag{8}$$

which implies that the effects of class size when natives (immigrants) are increased keeping immigrants (natives) constant are, respectively:

$$\begin{aligned} \beta_h &= \frac{\partial v_h}{\partial N} = p_n \phi_{nh} \\ \gamma_h &= \frac{\partial v_h}{\partial I} = p_i \phi_{ih} \end{aligned} \tag{9}$$

From these group specific parameters we can derive the effect of a composition change:

$$\delta_h = \left(\frac{dv_h}{dI} \right)_{C=\bar{C}} = \gamma_h - \beta_h \tag{10}$$

which is the effect of increasing the number of immigrants keeping class size constant (i.e. reducing at the same time the number of natives).

Consider for instance the effects on natives' performance. Conventional wisdom posits that immigrants are more in need of specific attention ($p_i < p_n < 0$) and that the effects of attention requests are more damaging for natives when they originate from immigrants ($\phi_{in} \geq \phi_{nn}$>0). In this case, our model would predict that $\gamma_n < \beta_n < 0$ and $\delta_n < 0$. In words, this configuration of parameters implies that the effect of an increase of class size on natives' performance is negative and stronger when it occurs because of an increase in the number of immigrants and that substituting one native with an immigrant, keeping class size constant, reduces natives' performance.

Other configurations of the structural parameters, different from the conventional wisdom, are plausible as well. We thus move to the empirical analysis in order to establish which is the relevant configuration supported by our data. We first replicate, in Section 4, the estimates based on identification strategies previously employed in the literature. We then illustrate our approach in Section 5 and 6, while Section 7 presents the main findings.

4 Conventional Evidence

We consider first the population regressions that constitute the empirical counter-part of the above theoretical model. Equation (8) describes the performance (in logs) of a representative student as a function of the number of natives and immigrants in the class and of his/her ethnicity (native or immigrant). Let s denote a student, $h \in \{n, i\}$ denote the student's ethnicity, j denote schools and k denote classes. Let $\alpha_h = \bar{v}_h$. Then, the following two equations (11) and (12)

$$v_{ihjk} = \alpha_h + \beta_h N_{jk} + \gamma_h I_{jk} + \epsilon_{ihjk} \tag{11}$$

$$v_{ihjk} = \alpha_h + \beta_h C_{jk} + \delta_h I_{jk} + \epsilon_{ihjk} \tag{12}$$

are the empirical counterparts of equation (8), in which the terms ϵ_{ihjk} capture other unobserved determinants of students' performance and $C_{jk} = N_{jk} + I_{jk}$ is class size. Equation (11) makes explicit the role of the number of natives and of the number of immigrants, while equation (12) highlights the effects of class size and class composition. These two equations are equivalent for our purposes. The problem is to find a set of reasonable conditions under which the population parameters β_h, γ_h and $\delta_h = \gamma_h - \beta_h$ are identified in these equations.

We start by applying conventional identification approaches to the Italian data described in Tables 1-2 and in the Appendix 9.1. For these analysis, as typically done in the literature, the unit of observation is the class, not the individual student.

Table 3 reports OLS estimates of the parameters β_h, γ_h and δ_h in the equations (11) and (12) augmented with schools fixed effect and using language test scores (Italian) as the measure v_{ihjk} of students' performance. This table replicates the approach used by Ammermuller and Pischke (2009), Contini (2011) and Ohinata and Van Ours (2011), i.e. it exploits the within-school variation across classes for identification. The first four columns of the table report results for the second grade, while the fifth grade results are in the remaining columns. For each grade, in the two columns on the left the dependent variable is the performance of natives while in the two columns on the right the focus is on immigrants. For each grade and ethnicity, the table reports estimates with and without controls for (average) individual and family characteristics of the students in the class.[1] Table 4 replicates the analogous estimates for the case in which the measure v_{ihjk} of students' performance is the mathematics test score.

Several interesting results emerge from these tables. First, note from the outset that these estimates are sensitive to the inclusion or exclusion of observable controls: this suggests that the allocation of students across classes is not random in our setting, a fact that is explicitly

[1] These variables include the averages of a series of dummies indicating if the mother's (or the father's) highest educational degree is at most lower secondary education, if the parents are unemployed and if the students went to nursery or kindergarten.

confirmed in Table 5. This table shows, for schools where there is at least one immigrant in the relevant grade, the relationship between the number of immigrants in a class and natives' background characteristics[2]: the estimated coefficients are statistically distinct from zero (at the 1% significance level) when school fixed effects are not included, but remain statistically significant also when school fixed effects are included instead.[3] This is evidence that, within a school and a grade, immigrants are dis-proportionately allocated to classes in which natives have a less favourable background. As a consequence, the OLS estimates of Tables 3 and 4 do not have a causal interpretation. In other words, school fixed effects solve the problem of the non-random allocation of immigrants across schools, but we find strong evidence that also within schools immigrants are not allocated randomly across classes.

Nevertheless, the correlations estimated in Tables 3 and 4 are worth attention because they are in line with the common wisdom. First, an increase in the number of natives is essentially uncorrelated with both test scores, for all grades and ethnicities, once observables are controlled for. Second, an increase in the number of immigrants is instead correlated with a reduction of test scores for both natives an immigrants and for both language and mathematics in grade 2. Third, this negative correlation between test scores and the number of immigrants in grade 2 is stronger for language than for mathematics. But, fourth, it is considerably smaller in grade 5 becoming essentially nil when the performance of immigrants is considered. As a result, keeping class size constant, the substitution of one native with one immigrant is associated with a decline in the performance of both ethnicities in both subjects, that fades away in the fifth grade particularly for immigrants.

To establish whether these associations, that largely correspond to common wisdom, are the outcome of causal relationships, we need to find a more convincing identification strategy.

[2]We include the following controls: the fraction of natives with low educated mothers, with low educated fathers, with unemployed mothers and with unemployed fathers; for grade 5, also the fraction of natives with few books at home, not available for grade 2.

[3]The estimates reported in Table 5 use data from the Language sample (i.e. the sample of students with no missing information in the language test score, see Appendix 9.1), using only observations for classes in schools with at least one immigrant in the relevant grade. Estimates for the Mathematics sample (i.e. the sample of students with no missing information in mathematics test score, see Appendix 9.1) are similar. For both samples, we get the same results also if we include all schools, i.e. also schools without immigrants. These additional results are not reported for brevity but are available from the authors upon request.

5 Identification

We now focus on our most innovative contribution, namely the identification of the pure composition effect, i.e. the effect on students' performance of increasing the number of immigrants in a class keeping class size constant (thus decreasing at the same time the number of natives). This parameter is interesting from a policy perspective because class size is an endogenous choice variable that principals may potentially use to neutralize the effects of immigrants. To take decisions, policy makers need an estimate of the effects of immigrants inflows net of the endogenous reactions of principals.

Our strategy relies on the identification of the causal effects of class size on student's performance in the presence and in the absence of immigrants. With the identification assumptions and the correction factors that will be explained below, we show that the class size effects in the presence or absence of immigrants can be used to identify the effects of increasing natives (immigrants) keeping immigrants (natives) constant, and thus the pure composition effect at constant class size. Our identification strategy rests on what we call the "Stable Native Treatment Effect" assumption, which mirrors, for the case of *Difference-in-Discontinuities*, the well known *parallel trend* hypothesis of Difference-in-Differences strategies.

For reasons that will become clear below, we will be able to identify only the effects on natives' test score, not on immigrants' test scores: we therefore drop the subscript h and focus on the performance of natives only.

Consider equation (12), which we can rewrite as:

$$v_{ijk} = \alpha + \gamma C_{jk} - (\gamma - \beta) N_{jk} + \epsilon_{ijk} \tag{13}$$

Using this equation the overall effect of class size in schools with immigrants is

$$\left[\frac{dv_{ijk}}{dC_{jk}}\right]_{I>0} = \gamma - (\gamma - \beta_{I>0}) \left[\frac{dN_{jk}}{dC_{jk}}\right]_{I>0} \tag{14}$$

where $\beta_{I>0}$ is the effect of increasing natives on natives performance in this type of schools. In schools without immigrants the overall effect of class size is simply

$$\left[\frac{dv_{ijk}}{dC_{jk}}\right]_{I=0} = \beta_{I=0} \tag{15}$$

where, therefore, $\beta_{I=0}$ is the effect of increasing natives on natives performance when there are only natives inside the school.

Building on previous work by Nannicini and Gagliarducci (forthcoming), the pure composition effect can be identified on the basis of the following assumption, which, incidentally, is implicit in the theoretical model of Section 3:

Assumption 1 *SNTE: Stable Native Treatement Effect.*
The effect on natives performance of increasing the number of natives is the same in schools with and without immigrants:

$$\beta_{I>0} = \beta_{I=0} = \beta \tag{16}$$

As already mentioned, note that the SNTE assumption is equivalent to the *parallel trends* assumption in the popular Difference-in-Differences (Angrist and Pischke, 2008, sec. 5.2) identification approach and in the more more-novel Difference-in-Discontinuities (Nannicini and Gagliarducci, forthcoming) approach.

Under the SNTE assumption, we can relate the parameters of interest to class size effects in the two types of schools, using equations (14) and (15),

$$\delta = \frac{\left[\frac{dv_{ijk}}{dC_{jk}}\right]_{I>0} - \left[\frac{dv_{ijk}}{dC_{jk}}\right]_{I=0}}{\left(1 - \left[\frac{dN_{jk}}{dC_{jk}}\right]_{I>0}\right)} \tag{17}$$

$$\beta = \left[\frac{dv_{ijk}}{dC_{jk}}\right]_{I=0} \tag{18}$$

$$\gamma = \frac{\left[\frac{dv_{ijk}}{dC_{jk}}\right]_{I>0} - \left[\frac{dv_{ijk}}{dC_{jk}}\right]_{I=0}}{\left(1 - \left[\frac{dN_{jk}}{dC_{jk}}\right]_{I>0}\right)} + \left[\frac{dv_{ijk}}{dC_{jk}}\right]_{I=0} \tag{19}$$

Equation 17 shows that the identification of the pure composition effect requires:

i) the identification of the causal effects of class size on student performance in the presence or absence of immigrants;

ii) the SNTE assumption and,

iii) the estimation of a correction factor that translates the difference between the effects of class size into the difference between the effects of changes in the numbers of immigrants and natives keeping (respectively) natives and immigrants constant.

Items ii) and iii) lead to the novel contribution of this paper, while the identification of class size effects (item i) exploits the well established identification method proposed by Angrist and Lavy (1999) for Israel, based on the discontinuities generated by the enrollment thresholds at which classes are split because of institutional rules of class formation, which we apply to the case of Italy. It is crucial, for our purposes, to note that these discontinuities generate exogenous sources of variation of class size in schools with and without immigrants: while in the former type of school both the number of natives and the number of immigrants is affected, in the latter only the number of natives changes discontinuously. With the help of the SNTE assumption, the effects of class size in the two types of schools give the identification of β, γ and δ

Note, also, that one could in principle identify the causal effects of class size in schools with and without immigrants using other methods, different from the one proposed by Angrist and Lavy (1999) (for example, within-school variation across classes or across cohorts), but would then need anyway the novel part (ii) and (iii) or our identification strategy to estimate the pure composition effect. With appropriate data to implement alternative identification strategies of class size effects (which are not at our disposal), one would be able to construct an overidentification test.

6 Estimation

Consider the regression:

$$v_{ijk} = \omega + \Psi_0 * 1_{I>0} + \Psi_1 C_{jk} + \Psi_2 C_{jk} * 1_{I>0} + \eta_{ijk} \tag{20}$$

where $1_{I>0}$ is a dummy for schools with immigrants and the other variables are defined as above.

Our identification strategy requires to first obtain, following Angrist and Lavy (1999), causal estimates of the effect of class size on student performance in schools without immigrants

$$\Psi_1 = \left[\frac{dv_{ijk}}{dC_{jk}}\right]_{I=0} \tag{21}$$

and the differential class size effect in schools with immigrants

$$\Psi_2 = \left[\frac{dv_{ijk}}{dC_{jk}}\right]_{I>0} - \left[\frac{dv_{ijk}}{dC_{jk}}\right]_{I=0} \tag{22}$$

Since this is the less novel part of the paper, the details of this step of the estimation process are described in the appendix Section 9.2.

Given the estimates of Ψ_1 and Ψ_2 obtained with equation (20) and the rules of class formation, using equations (17), (18) and (19), we can write the parameters β, γ and δ, in which we are interested, as:

$$\beta = \Psi_1 \tag{23}$$

$$\gamma = \frac{\Psi_2}{\left(1 - \left[\frac{dN_{jk}}{dC_{jk}}\right]_{I>0}\right)} + \Psi_1 \tag{24}$$

$$\delta = \frac{\Psi_2}{\left(1 - \left[\frac{dN_{jk}}{dC_{jk}}\right]_{I>0}\right)} \tag{25}$$

However, to obtain these parameters, in addition to estimates of Ψ_1 and Ψ_2, we also need an estimate of the correction factor $\left(1 - \left[\frac{dN_{jk}}{dC_{jk}}\right]_{I>0}\right)$. We get this estimate using the following auxiliary regression, restricted to schools in which both immigrants and natives are present:

$$N_{jk} * 1_{I>0} = \theta_1 + \theta_2 C_{jk} * 1_{I>0} + u_{jk} \tag{26}$$

where

$$\theta_2 = \left[\frac{dN_{jk}}{dC_{jk}}\right]_{I>0} \tag{27}$$

and where, as in equation (20), the rules of class formation are used to construct an instrument for $C_{jk} * 1_{I>0}$.[4]

7 New evidence

Table 6 reports estimates based on the novel identification strategy described in the previous sections, in which the dependent variables are the language and mathematics test scores of native students in the second and fith grade. It should be immediately noted that differently than in the case of the OLS estimates of Tables 3 and 4, these estimates are fairly robust to the inclusion or exclusion of individual and family characteristics into the specification of the relevant equations. This is reassuring because it suggests that our identification strategy

[4]Equation (27) could be also estimated simply with school fixed effects if one were willing to exploit within-school across-cohorts variation for the identification of class size effects.

is robust with respect to the non random allocation of students across classrooms within a school.

The estimates of the first two columns of Table 6 suggest that an increase in the number of natives has a small negative effect on the natives test score ($\beta = \Psi_1 \approx$ - 1% in the first row). This is also the effect of increasing class size in a school without immigrants. An increase in the number of immigrants has instead a sizeable and statistically significant negative effects on the same dependent variable ($\gamma \approx$ - 13% in the second row). As a result, the pure composition effect of substituting one native with an immigrant is negative, sizeable and significant as well ($\delta \approx$ - 12% in the third row).

The estimates of the fourth row indicate that the effect of an increase in class size on the language performance of natives is 3% more positive in schools with immigrants than in schools without them. This may seem puzzling at first site, but it is reasonable in the light of our estimate of the correction factor $\left(1 - \left[\frac{dN_{jk}}{dC_{jk}}\right]_{I>0}\right)$ in the fifth row. This estimate is negative and significant which implies that $\left[\frac{dN_{jk}}{dC_{jk}}\right]_{I>0} > 1$; in words this means that, in the sample, increases of class size in schools with immigrants are typically associated with compositional changes in favour of natives, whose number increases more than the increase in class size. Therefore, when class size increases in schools with immigrants this is typically associated with a decline in the number of immigrants per class, which has a positive effect on natives language test scores. This finding suggests that principals try to compensate increases of class size with decrease in the number of immigrants or, vice-versa; when immigrants increase, they try to reduce class size to compensate for the negative effects of immigrants on natives performance.

Interestingly, all these effects become essentially insignificant in the fifth grade, which is reasonable under the assumption that the school system is capable, after a sufficient amount of time to minimize the problems caused by immigration in classrooms.

The pattern for the results concerning mathematics test scores in the last four columns of Table 6 are qualitatively similar but smaller in size, which is reasonable under the assumption that the disruption caused by immigrants in a classroom is more sizeable during Italian language lectures than during mathematics lectures.

Leaving to the Appendix 9.3 a discussion of the details concerning the estimation of first stages and of the auxiliary regression (26), which are reported in Tables 7 and 8, here we prefer to focus on the validity of the crucial identification assumption (i.e. the SNTE), which states that the effect of an increase in the number of natives on natives performance is the same in schools with or without immigrants. Unlike in standard Difference-in-Differences approaches, where the parallel trend assumption is testable provided that data on a sufficient time-span are available, the SNTE is not directly testable in our setting We thus propose a robustness check aimed at indirectly assessing the validity of this assumption. We claim that the SNTE is more likely to hold in a sample where

i) the possibility of sorting betweeen different neighbourhoods is limited, which happens in small cities;

ii) the probability of enrolling immigrants, based on observed characteristics at the municipality level, including enrollment, is similar in schools with and without immigrants;

iii) the probability of enrolling immigrants is not too high if, at such high probabilities, there are too few schools without immigrants in the sample.

In practice, we restrict our analysis to a sub-sample of small municipalities (i.e. municipalities with less than 100,000 inhabitants) where the probability of enrolling immigrants is below the sample median (≈ 0.90). As shown in Figure 2, above this threshold there are very few schools without immigrants in the sample. To compute the probability of enrolling immigrants we use a probit model in which the unit of observation is a single school and the dependent variable is a dummy indicating the presence of immigrants enrolled in that school. We specify the probability as a function of several municipal characteristics in terms of immigrants concentration and population. Details of this estimation are reported in the Appendix 9.3.3. Figure 2 shows the empirical probability distribution function of the predicted probability of enrolling immigrants by school type in the sub-sample of small cities. The red and the blue bars represent respectively the density of schools with and without immigrants for a given probability level, while the vertical line denotes the median of the probability of enrolling immigrants. As we move towards lower probability levels the density of schools without immigrants increases with respect to the density of schools with immigrants, but there is a wide range of values where we have common-support.

Since the SNTE is arguably more likely to hold in this restricted sub-sample, finding similar results here and in our main sample would support our identification assumption. Results are reported in Table 9, and are qualitatively the same as those presented above in Table 6, although slightly smaller in size. We believe that this evidence supports our assumption that the effect of an increase of natives is the same independently of the presence of immigrants (SNTE). In particular the composition effect of substituting one native with one immigrants continues to be in the order of - 9% on the natives' language test score in the second grade, and in the order of -5% on the natives' mathematics test score in the same grade. Both negative effects vanish in the fifth grade. The fact that this happens also in this restricted sample where the SNTE assumption is more likely to hold more tightly, confirms that a longer permanence in the school system or the process of integration in the Italian society or both are somehow capable of reducing the detrimental effect of immigrants inflows on the performance of natives.

As already mentioned, the policy relevance of these results is twofold. On the one hand we are able to estimate the causal effect of an immigrants' inflow controlling for the endogenous changes of class size that principals implement in order to neutralize the detrimental components of such causal effect. Hence, the causal parameter that we estimate is the most relevant policy parameter in comparison to those estimated so far in the literature. On the other hand, we also show that with the passage of time, the school system and the society (at least the Italian one) are somehow capable of reducing the problems caused by immigrants in classes thanks to channels that are different from changes in class size. We cannot say, at the current state of our analysis, whether these happens because of a conscious choices of school operators, parents and policy makers or it is just a natural by-product of the on-going educational and integration process.

8 Conclusions

We propose a strategy to identify class size and *class composition* effects, in terms of native and immigrants, on students performance. This strategy exploits rules of class formation in Italy, in the spirit of Angrist and Lavy (1999), and identification assumptions similar to those of the Difference in Discontinuities framework proposed by Nannicini and Gagliarducci (forthcoming). By combining these two approaches, we obtain estimates of the effect of changes in the number of native students holding immigrants constant in a class, and estimates of the effect of changes in the number of immigrants holding natives constant. We then use these two estimates to assess the class composition effect, for given class size, a parameter on which there is no clean evidence in the literature, as far as we know. We believe that the estimates of this composition effect are the most important contribution of our paper. Principals do use class size to neutralize the detrimental effects of immigrants' inflows. Therefore, only controlling effectively for class size, as we do thanks to our novel identification strategy, it is possible to measure the real dimension of such possible detrimental effects.

We interpret our evidence in the light of the Lazear (2001)'s model of the educational production function in which the absolute numbers of students in a class determines performance, if each student has some positive probability of generating disruption. We extend this model in two ways. First, by allowing for two types of students, natives and immigrants, each with its own probability of disruption. Second, by introducing the possibility that the behaviour of a student generates positive externalities, not only disruption, across or within groups. This model is particularly suited to interpret our estimates because it shows why class size is endogenous with respect to immigrants inflows and therefore needs to be controlled for, as we do in our empirical analysis.

Our results suggest that the pure composition effect is sizeable in grade 2 ranging between -12% in language and -9% in mathematics, but vanishes in grade 5. Class size has opposite effects on natives' performance in language ($\approx$ +1% at age 7) and mathematics ($\approx$ -3% at age 7) and also these effects vanish as children age. The effect of class size is typically larger in absolute magnitude in schools with immigrants rather than in schools without immigrants. For language test score, our estimates suggest that the negative effect of adding one additional native student (nearly -3%) is much smaller (in absolute terms) than the negative effect of adding an immigrant student (about -16%), so that the composition effect is negative and significant (almost -13%). The magnitude of these estimates is larger (in absolute terms) than the ones obtained by OLS and are comparable in size to the effect of having a parent who has completed, at most, a lower secondary school curriculum.

We think it is remarkable that these effects tend to vanish as children age, because it indicates that the school system is somehow capable of implementing educational strategies aimed at neutralizing the negative effects of immigrant inflows in classrooms. These strategies do not consist only in the reduction of class size, because we estimate a negative composition effect also when class size is kept constant. Therefore, principals must have other tools to neutralize these composition effects at constant class size or the vanishing of the effects must be a natural by-product of the on-going educational process. Next in our research agenda is a study of what these tools may be.

9 Appendix

9.1 The data

The individual data on test scores used in this paper were provided by the Italian National Institute for the Evaluation of the Education System (INVALSI in what follows). These data are collected yearly through a standardized testing procedure that assesses both language (Italian) and mathematical skills of pupils in 2nd and 5th grade (primary school). We use the 2009-2010 edition, the first in which all schools and students of the selected grades are required to take part in the assessment .[5] The testing procedure and its implementation are described in details in the annual reports of the INVALSI. The test scores in language and mathematics are available both for natives and immigrants and are measured as the fraction of correct answers over the total number of questions.[6]

In addition to the test score the dataset contains a large number of individual socio-economic variables which allow us to control properly for factors correlated to school performance. Part of this information is available for all the grades, and is provided by the schools' administration employees who fill out a record-sheet with personal and family characteristics of each student. Examples of these variables are: gender, citizenship[7], place of birth (Italy, European Union, Not-European Union, other), language spoken at home (Italian, dialect, other), the attendance of nursery or kindergarten, the highest educational level achieved by parents and their occupational status. Other information is self-reported and is collected through a questionnaire administered only to students enrolled in the 5th grade. These variables include data on students' habits and home possessions like the number of books at home, which is commonly used in the international literature as a valid proxy of family socio-economic background. Finally, the dataset contains a unique student/class/grade identifier which can be used to relate each student to a specific class in a particular grade of a given school.

We link this individual dataset with schools' administrative information coming from the Italian Ministry of Education and Research (MIUR). We measure class size and the number of natives and immigrants in a class with this administrative data instead of using individual information. This procedure has an important advantage: it allows us to avoid the bias induced by absenteeism on the day of the test, if we were using the number of students taking the test to construct measures of class size and ethnicity size.

To measure the total number of children enrolled in a given grade we use the *long-run* enrolment, i.e. the highest level of enrolment reached for each school-grade in the last 5 years. We do so because schools may adjust upward the number of teachers as a response to increases in enrollment levels but, later, downward adjustments may be more difficult. In practice, we replace the level of enrollment in grade 2 with the corresponding enrollment in

[5]In previous waves of the survey the participation of the schools to the test was on a voluntary basis and only a limited number of schools was sampled on a regional basis. Within these randomly selected schools, only a randomized sample of students took the test.

[6]Our empirical model is in logarithms. Since nothing prevents students to score 0 on the test we add a positive constant (0.000001) to each score.

[7]We have only the distinction between natives and non natives and no information on the country of origin.

grade 5 if the latter is higher. This correction is retrospective and feasible only from grade 5 to grade 2, not vice-versa. Another administrative information used in this study is the number of disabled children in each school.

Our empirical strategy (see Section 5) is based on the comparison between schools with and without immigrants. We define a school *with immigrants* if there is at least one immigrant enrolled in the relevant grade.

Differently from other studies that aggregate the data at the class (or school) level we prefer to keep the unit of observation at an individual level, so that we are able to check the sensitivity of our findings to the inclusion of individual control variables.[8]

Note that since language and mathematics tests were held on different days and students may have missed none, one of the two or both, we decided to have a distinct data set for each outcome of interest in order to maximize the number of observations used in the empirical analysis.

Our datasets cover on average 7,500 primary schools, 80 per cent of which have at least one immigrant enrolled. Table 1 reports descriptive statistics averaged at the school level for both the language (Panel A) and the mathematics sample (Panel B). There are no apparent differences in the average class size, the enrollment level and the average number of natives and immigrants per class between the two samples, as expected. The average school has a total of 140 pupils enrolled in grade 2 and 290 in grade 5. On average there are 19 students overall and about 1.5 immigrants per classes in both grades. The table also reports the mean test scores for both ethnicities.

Immigrants tend to perform worse than natives in reading and mathematics, but the gap between ethnic groups is more sizeable in language. Scores in language are generally higher than scores in mathematics for natives, while immigrants perform relatively better in mathematics. The gap between natives and immigrants in reading tends to narrow across grades but remains relatively more stable in mathematics. Finally, the dispersion in the score distribution for both reading and mathematics is lower among natives with respect to immigrants: natives, as expected, are more homogeneous.

The average individual characteristics for both ethnicities are reported in Table 2. We use this information to construct control variables in our regressions. Our sample covers, on average, more than 430,000 natives and about 43,000 immigrants students in both grades. The table shows that on average natives students have a better family background than immigrants. The share of low educated mothers or fathers, however, seems to be very similar (and in some case higher) for natives than for immigrants. This can be explained looking at the pattern of missing records reported for both ethnicities. In general the share of missing values for all the variables is higher for immigrants than for natives. This under-reporting, which we believe is not at random, may bias the comparison of family background between ethnicities. In all regression, we will include dummy variables to control for the pattern of item non-response.

9.2 Rules for class size formation in Italy

There is a large consensus in Italy about the importance of class formation in schools. Families, principals and teachers tend to agree on the fact that classes should be as homogeneous

[8]We use data aggregated at class level in Section 4 when we estimate the association between the number of immigrants and the characteristics of natives in the class, or in the case of the OLS estimates in Tables 3 and 4.

as possible and that congestion should be avoided. These preferences are partially reflected in the Italian legislation that establishes a minimum and a maximum of students per class.[9] These limits are set separately for each educational level, but the differences are only minimal. In particular, in primary schools a class should be formed with a minimum of 10 students and a maximum of 25, in the lower secondaries these limits are 15 and 25 students, while for high schools, the law foresees a maximum of 25 students per class. Since the process of class formation starts at the end of every academic year, when the number of children enrolled in the incoming year is only provisional, the law provides a 10% tolerance in the upper and lower limits to accomodate possible gaps between the expected and the actual number of classes, as a function of the final enrollment rates. Moreover, there are additional exceptions for particular cases. First, the maximum number of students in class is set to 20 if disabled pupils are present. Second, municipalities in mountain and relative remote areas are allowed to have classes with less than 10 students.

Neglecting, for the time being, all these additional exceptions provided by the law, and setting at 28 the maximum number of students in class (25 + 10% of tollerance), we can express the theoretical function of class size derived by the rule as follows:

$$\bar{C}_j = \frac{Z_j}{\text{Int}\left(\frac{Z_j - 1}{28}\right) + 1} \tag{28}$$

where Z_j is the total enrollment in school j at the beginning of the year[10] and $\text{Int}(x)$ is the largest integer smaller or equal to x. This equation highlights that theoretical class size is a function of school enrollment and displays discontinuities at every multiple of 28. For instance the predicted class size increases linearly until total enrollment in the grade reaches the 28 limit, and in this enrollment range there will be one class. When there are 29 children enrolled, two classes of 14.5 students on average should be formed. This *class-splitting* occurs every time the enrollement level reaches a multiple of 28. As discussed in Section 5, we use this institutional rule as a source of exogenous variation in class size to identify our causal parameters of interest, in the same spirit of Angrist and Lavy (1999).

We expect compliance with the rule of class formation not to be perfect because exceptions are formally tolerated and the margin of 10% on the maximum number of students per class is prescribed. Our intuition is confirmed when we contrast theoretical and actual class size in grade 2 and 5 (see Figure 1), using data aggregated by enrollment level. The x-axis reports the number of students enrolled in the relevant grade and the y-axis reports the average class size. The dashed red line shows the theoretical class size based on a critical threshold of 28 and its multiples, while the solid black line corresponds to the actual average class size for a given level of enrolment. The graphs focus on all the enrolment levels covered by our data and suggest that the average class size approximates fairlyl well its predicted theoretical function, even if compliance is not perfect.

In particular, as expected, compliance appears to be stronger at low levels of enrollment, where the constraint on the minimum number of students per class matters. It is also noteworthy that, in any case, the actual class size reaches the maximum level predicted by the rules at each discontinuity point. Moreover, the figures show that the black solid line lies

[9]The official rules for class formation in Italy are contained the DL n. 331/1998 and the DPR n. 81/2009 of the Ministry of Education and Research.

[10]We use long-run enrollment as explained in Section 9.1.

always below the dashed red line of the theoretical class size (with few exceptions at very low levels of enrollment). This can be explained by the fact that some schools anticipate the condition of class-splitting to cope with the negative effect of class congestion when problematic students are present.

9.3 Econometric details

This section presents details on: i) the first stage estimates using the rules for class formation discussed above; ii) the estimates of the correction factor (26) and iii) the estimates of the probability that a school enrolls immigrants, which we need to restrict the analysis to schools in which the SNTE identifying assumption is more likely to be satisfied.

9.3.1 Instrumental variables and first stage estimates

In the same spirit of Angrist and Lavy (1999), we exploit the institutional rules of class formation described in Section 9.2 as a source of exogenous variability in average class size. More specifically, we use the discontinuous function of the enrollment (i.e. the predicted class size $\bar{C}$ defined in equation (28)) to instrument the actual class size in schools with and without immigrants.

We extend the set-up of Angrist and Lavy (1999) to the case of two endogenous variables and two instruments, where we obtain the second instrument interacting the predicted class size with the dummy for the presence of immigrants in the relevant grade ($\bar{C}_{ijk} * 1_{I>0}$): this is equivalent as stating that the rule of class formation applies in both schools. Note that differences across schools with and without immigrants are also accounted for by our design (see Section 5) as in Nannicini and Gagliarducci (forthcoming).

Indeed, one remarkable difference with the *Maimonides rule* study of Angrist and Lavy (1999) is that we use the exogenous discontinuity generated by the rule within each type of schools (i.e. with and without immigrants). Note, in fact, that equation (20) can be traced back to the standard class size model by estimating separately two equations for the schools with and without immigrants (with one endogenous and one instrumental variable per equation) and then taking the difference between the estimated parameters to obtain the Ψ_2 parameter of the pooled regression. Intuitively, while the IV strategy in Angrist and Lavy (1999) can be framed in a standard regression discontinuity design, our extended framework can be seen as exploiting the comparison between the two discontinuities given by the theoretical rule within each group of schools.

In details, the first stages of equation (20) take the form:

$$C_{ijk} = \pi_{00} + \pi_0 1_{I>0} + \pi_1 \bar{C}_{ijk} + \pi_2 \bar{C}_{ijk} * 1_{I>0} + \pi_3 Z_{ijk} + \pi_4 X_{ijk} + \mu_p + \upsilon_{ijk} \quad (29)$$

$$C_{ijk} * 1_{I>0} = \rho_{00} + \rho_0 1_{I>0} + \rho_1 \bar{C}_{ijk} + \rho_2 \bar{C}_{ijk} * 1_{I>0} + \rho_3 Z_{ijk} + \rho_4 X_{ijk} + \mu_p + \nu_{ijk} \quad (30)$$

where the dependent variables are the actual class size and its interaction with the dummy for the presence of immigrants and the instruments are the theoretical class size defined by the rules of class formation and its interaction with same dummy for the presence of immigrants. We add to these regressions the vectors of schools and individuals' controls (Z_{jk} and X_{ijk}) and the provincial level fixed effects (μ_p) already included in equation (20). Specifically, among school level covariates, we include a second order polynomial of the enrollment to control for the continuous relationship between class size and enrollment itself.[11] Consistently with equation (20), we estimate the first stage regressions using individuals as units of observation, even if the variability exploited by our instruments is at school level. Of course, we cluster standard errors.

Our identification strategy works as long as the instruments affect the test scores of natives only through their effect on actual class size. Failure of the exclusion restriction may occur, for example, because of the *manipulation* of the rule by the parents, which can place their children in schools with lower average class size, i.e. in schools with an enrolment level slightly on the "right" of each discontinuities. This is implausible because the enrollment is not observed by families at the time of application and mobility across school districts in Italy is very low. Since our identification framework differs from the simple case of one endogenous variable and one instrument as in Angrist and Lavy (1999) and relies on the novel Difference-in-Discontinuities strategy proposed by Nannicini and Gagliarducci (forthcoming), it is useful to report formally the effects identified in (29) and (30). Let $c^* = Int(\frac{Z_j - 1}{28})$, where Z_j is the total enrolment in school j at the beginning of the year and $Int(x)$ is the largest integer smaller or equal to x.

$$\begin{aligned}
\pi_1 = \rho_2 &= E[C|I=0,\bar{C}^-] - E[C|I=0,\bar{C}^+] \\
&= lim_{\bar{C}\uparrow c^*} E[C|I=0,\bar{C}] - lim_{\bar{C}\downarrow c^*} E[C|I=0,\bar{C}] \\
&\equiv E[N|I=0,\bar{C}^-] - E[N|I=0,\bar{C}^+] \\
&= lim_{\bar{C}\uparrow c^*} E[N|I=0,\bar{C}] - lim_{\bar{C}\downarrow c^*} E[N|I=0,\bar{C}] \qquad (31) \\
\pi_2 = \rho_1 &= E[C|I>0,\bar{C}^-] - E[C|I>0,\bar{C}^+] - E[C|I=0,\bar{C}^-] - E[C|I=0,\bar{C}^+] \\
&= lim_{\bar{C}\uparrow c^*} E[C|I>0,\bar{C}] - lim_{\bar{C}\downarrow c^*} E[C|I>0,\bar{C}] \\
&\quad -(lim_{\bar{C}\uparrow c^*} E[C|I=0,\bar{C}] - lim_{\bar{C}\downarrow c^*} E[C|I=0,\bar{C}]) \qquad (32)
\end{aligned}$$

From these two equations we can easily obtain the effect of the theoretical rule on actual class size in schools with immigrants: $(\pi_1 + \pi_2) = (\rho_1 + \rho_2)$.

Table 7 shows the results of the first stage estimates for each grade and estimation sample (language and mathematics). Panel A of the table displays the results without including any individual level control, while panel B reports the first stage estimates of the regressions where these controls are included. Importantly, in every specification the predicted class size and its interaction are strong instruments for the actual class size. Indeed, the F-statistics are always greater than the critical threshold indicated by Staiger and Stock (1997) , underlying that weak identification is not a problem in our set-up.

Focusing, for instance, on the estimated parameters for the 2nd grade (columns 1-2 of panel A) it is easy to see that the association between predicted and actual class size is higher in schools without immigrants. In these schools an increase of one theoretical student is associated with an increase of 0.246 students in actual class size, while in schools with immigrants this effect is approximately 0.12. These results remain unchanged when we add individual controls to the regressions (panel B) or if we look at the estimates for the 5th grade (colums 3-4, 7-8). This evidence suggests a different degree of compliance with the

[11] We checked the robustness of our results to other choices for the polynomial in enrolment. Results, not reported here for brevity, are available form the author upon request.

rules of class formation in the schools with immigrants. We believe that the explanation for such behaviour relies on the fact that principals of these schools face pressing incentives to anticipate the moment of class splitting (i.e. they do not consider the margin of flexibility allowed by the Italian rule), in order to compensate for the presence of immigrants with lower average class size in the schools. This evidence is analogous to what Lavy (1995) shows for the case of schools with low socio-economic background in his study. Table 10 reports the first stage estimates in a sample of schools located in small municipalities with a similar (low) probability of enrolling immigrants. Results are qualitatively unchanged, even if point estimates differ slightly in the magnitude.

9.3.2 Auxiliary regression for the correction factor

In this section we illustrate how we estimate the correction factor $\left[1 - \left(\frac{\partial N_{jk}}{\partial C_{jk}}\right)_{I>0}\right]$ introduced in Section 5 to allow us to link the effects Ψ_1 and Ψ_2 of class size on students performance (see equation (20)) to the causal parameters of interest β (for the number of natives), γ (for the number of immigrants) and δ (the composition effect).

We estimate $\left[\frac{\partial N_{jk}}{\partial C_{jk}}\right]_{I>0}$ with equation (26), which we report here for convenience

$$N_{jk} * 1_{I>0} = \theta_1 + \theta_2 C_{jk} * 1_{I>0} + u_{jk}$$

where

$$\theta_2 = \left[\frac{\partial N_{jk}}{\partial C_{jk}}\right]_{I>0}$$

The units of observation are classes in schools with immigrants and rules of class formation are exploited to construct an instrument for $C_{jk} * 1_I > 0$. We add to the regression the standard controls included in equation (20), for school characteristics, average individual characteristics in the class and province fixed effects.

Table 8 reports the results of the auxiliary regression where the dependent variable is the average number of natives per class. In the first two columns we report the estimates for grade 2 without including any individual level control. The estimate of $\left[\frac{dN_{jk}}{dC_{jk}}\right]_{I>0}$ is positive and greater than one ($\hat{\theta}_2 = 1.234$) which means that an increase of overall class size is associated with a more than proportional increase in N. This is possible if, when class size increases, principals tend to reduce the fraction of immigrants in the class. Or, viceversa, if principals reduce class size when the fraction of immigrants in the class increases. Because of these reactions of principals, it follows that the estimated correction factor is negative and statistically significant.

This is consistent with the evidence presented in Section 9.3.1 where the estimates of the "jump" in class size in schools with immigrants are lower than the ones observed in schools where only natives are enrolled. The first stage reported in column 2 shows the validity of the instrument as a predictor of class size. Results remain unchanged when we add controls to the regression (columns 3-4) and when grade 5 is considered (columns 5-8). Table 11 reports the auxiliary regression estimates considering a sample of schools located in small municipalities with a similar (low) probability of enrolling immigrants. Again the results are qualitatively unchanged.

9.3.3 Estimates of the probability that a school enrolls immigrants

One may argue that schools with and without immigrants are fundamentally different, which would make our identification strategy not valid. If schools with and without immigrants are intrinsically different and impossible to compare, then the effect of a change in the number of natives cannot be the same in the two types of schools as required by the Stable Native Treatment Effect assumption at the core of our identification approach (see Section 5). This critique becomes much less plausible, however, if our results do not change when one restricts the attention to schools that have a similar probability to enroll immigrants, for a given total enrollment level.

We thus estimate this probability by relating the presence of immigrants in a school with several characteristics of municipalities and schools. More specifically, we run the following probit regression using a dataset averaged at the school/grade level.

$$1(I > 0)_j = \alpha + \theta_1 imm_m + \theta_2 under_m + \theta_2 pop_m + \theta_{3k} X'_{mk} + \theta_4 Z_j + \mu_p + \xi_j m \qquad (33)$$

where $1(I > 0)_j$ is a dummy indicating the presence of at least one immigrant enrolled in school j. We include as covariates the following municipal characteristics such: the share of the immigrant population at the end of 2009 (imm_m), the share of under-aged immigrants over the total foreign population in 2009 ($under_m$) and the total population (pop_m). We also add to the specification a dummy indicating if the municipality is the capital of the Province, a dummy for a coastline location of the municipality, a series of dummies for each decile of the share of females in the local population and of under-aged or elderly subjects (all these dummy variables are grouped in the vector $X'_m k$).[12] In addition to these municipal variables, we add to the regression the total enrollment of the school (Z_j) and the provincial fixed effects (μ_p). Finally, $\xi_j m$ is a normally distributed error term.

We estimate equation (33) pooling data of 2nd and 5th grade and keeping separated the language and the mathematics samples. As explained in section 7 estimation is carried out on a sample of schools in cities with less than 100,000 inhabitants (small municipalities), where the probability of sorting across neighborhoods is limited.

As expected, these estimates show that the share of immigrants in the municipality has a positive and a sizeable effect in increasing the probability of enrolling immigrants[13]. Schools located in the Province's capital and the ones located in the coastline cites have higher probability of enrolling immigrants. This probability increases with the share of the elderly population, which is plausible given the fact that a considerable share of immigrants in Italy (particularly women with children) are involved in the care of elderly natives and in family assistance (house cleaning, child care etc.). Finally, school's size (in terms of enrollment) is positively correlated with presence of immigrants.

Figure 2 plots the empirical distribution function of the estimated probabilities enrolling immigrants, for the subset of schools in small cities.

[12] Data on municipal characteristics come from the Italian National Statistics Institute (ISTAT) and it was possible to link them with the school dataset thanks to a town identifier kindly provided by the INVALSI for this purpose.

[13] These results are not reported for brevity, but are available upon request.

References

Ammermuller, A. and J-S Pischke. 2009. "Peer Effects in European Primary Schools: Evidence from PIRLS." *Journal of Labour Economics* 27 (3):315–348.

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

Angrist, J.D. and V. Lavy. 1999. "Using Maimonides Rule To Estimate The Effect of Class Size on Scholastic Achievement." *The Quarterly Journal of Economics* 114 (2):533–575.

Angrist, J.D. and J-S Pischke. 2008. *Mostly Harmless Econometrics.* Pricenton University Press.

Bossavie, L. 2011. "Does Immigration Affect the School Performance of Natives? Evidences From Microdata." Mimeo, EUI.

Brunello, G. and L. Rocco. 2011. "The Effect of Immigration on the School Performance of Natives: Cross Country Evidence Using PISA Test Scores." Discussion Papers 5479, IZA.

Card, D.E. and J. Rothstein. 2007. "Racial Segregation and the Black-White Test Score Gap." *Journal of Public Economics* 91 (11-12):2158–2184.

Contini, D. 2011. "Immigrant Background Peer Effect in Italian Schools." Working paper, IRVAPP.

Gould, Eric. D., V. Lavy, and M.Daniele Paserman. 2009. "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence." *Economic Journal* 119 (540):1243–1269.

Hanushek, E.A, J.F Kain, and S.G Rivkin. 2009. "New Evidence About Brown V. Board of Education: The Complex Effects of School Racial Composition on Achievement." *Journal of Labour Economics* 27 (3).

Hoxby. 2000. "Peer Effect in the Classroom: Learning from Gender and Race Variation." Working paper 7876, NBER.

Hunt, J. 2012. "The Impact of Immigration on Educational Attainment of Natives." Working paper 18047, NBER.

Jensen, P. and A.W. Rasmussen. 2011. "The Effect of Immigrant Concentration in Schools on Native and Immigrant Children's Reading and Math Skils." *Economics of Education Review* :1503–1515.

Lavy, V. 1995. "Endogenous School Resources and Cognitive Achievement in Primary Schools in Israel." Falk Institute Discussion Paper 95.03, Hebrew University.

Lazear, E.P. 2001. "Educational Production." *Quarterly Journal of Economics* CXVI:777–803.

Manski, C.F. 1993. “Identification of Endogenous Social Effects: The Reflection Problem.” *Review of Economic Studies* 60 (3):531–542.

Nannicini, T. and S. Gagliarducci. forthcoming. “Do Better Politicians Perform Better? Disentangling Incentives from Selection.” *Journal of the European Economic Association* .

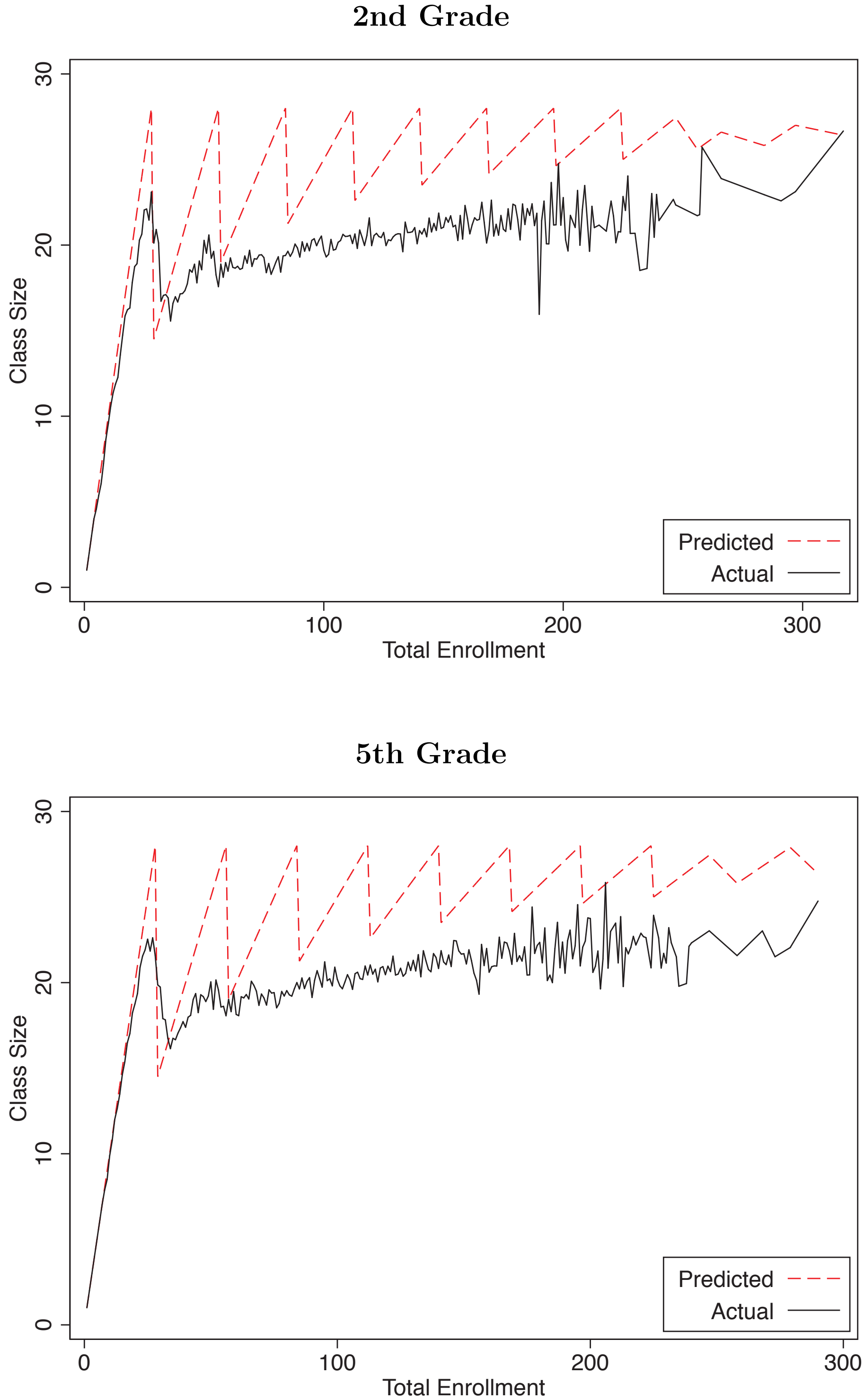
Ohinata, A. and J.C Van Ours. 2011. “How Immigrant Children Affect the Academic Achievement of Native Duch Children.” Discussion Papers 6212, IZA.

Pellizzari, M., G. De Giorgi, and W.G. Woolston. 2012. “Class Size and Class Heterogeneity.” *Journal of the European Economic Association* .

Staiger, D. and J.H. Stock. 1997. “Instrumental Variables Regression with Weak Instruments.” *Econometrica* 65 (3):557–586.

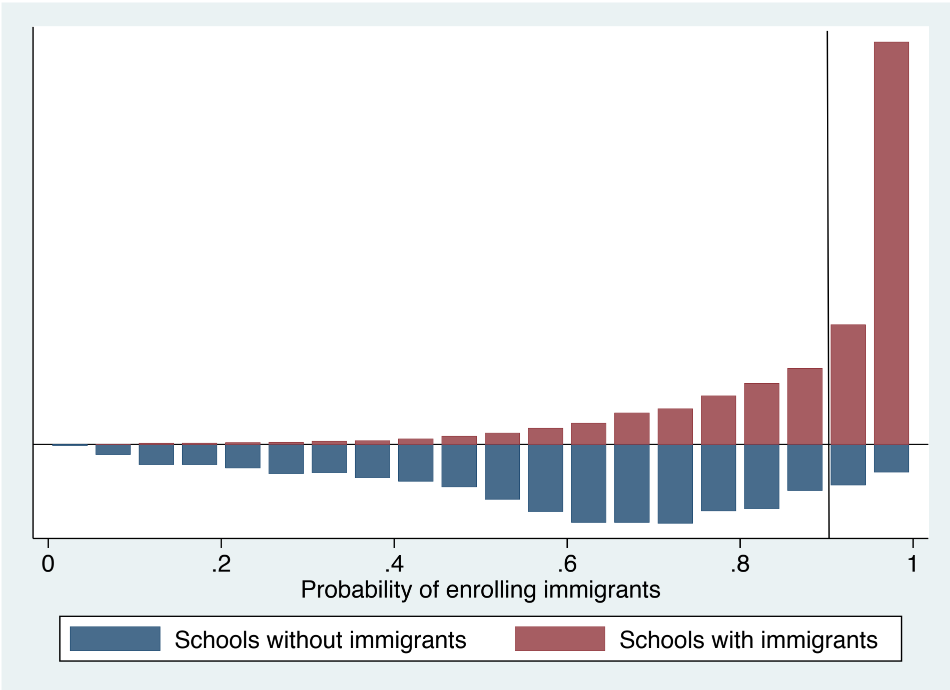
Tonello, M. 2012. “Social Interaction Between Native and Non-Native Students: Mechanisms and Evidences.” Working paper 65, Università Cattolica del Sacro Cuore.

Figure 1: Actual and theoretical class size in Italy.



Source: Invalsi data for academic year 2009-10. The dashed red line and the solid black line are respectively the predicted class size according to the rule of class formation in equation (28) and the actual average class size by level of enrollment.

Figure 2: Empirical probability distribution function of the estimated probability of enrolling immigrants in small municipalities by type of school.



Source: Invalsi data for academic year 2009-10. The red and the blue bars represent respectively the density of schools with and without immigrants for a given probability level. The vertical black line is the median of these estimated probabilies.

Table 1: Descriptive statistics

	Mean	S.D.	p10	p25	p50	p75	p90
	Panel A. Language Sample						
2nd Grade:	*7508 schools, 5960 with immigrants; 29179 classes*						
Class size	19	4.2	14	17	20	22	24
Number of natives	18	4.3	12	15	18	20	23
Number of immigrants	1.5	1.5	0	.26	1	2.1	3.5
Enrollment (*Long run)*	78	46	23	39	73	110	140
Test score language, Natives	.68	.11	.55	.61	.67	.73	.81
Test score language, Immigrants	.56	.17	.36	.45	.54	.66	.81
5th Grade:	*7493 schools, 6007 with immigrants; 29443 classes*						
Class size	19	4.3	14	17	20	24	35
Number of natives	18	4.3	13	15	18	23	33
Number of immigrants	1.5	1.5	0	.27	1	3.5	13
Enrollment	74	45	21	36	68	135	290
Test score language, Natives	.71	.09	.62	.67	.71	.81	1
Test score language, Immigrants	.62	.14	.46	.54	.61	.81	1
	Panel B. Math Sample						
2nd Grade:	*7503 schools, 5958 with immigrants; 29148 classes*						
Class size	19	4.2	14	17	20	22	24
Number of natives	18	4.3	12	15	18	20	23
Number of immigrants	1.5	1.5	0	.26	1	2.2	3.5
Enrollment (*Long run)*	78	46	23	39	73	110	140
Test score math, Natives	.64	.13	.51	.56	.62	.7	.84
Test score math, Immigrants	.57	.16	.4	.47	.54	.66	.81
5th Grade:	*7491 schools, 6005 with immigrants; 29430 classes*						
Class size	19	4.3	14	17	20	24	35
Number of natives	18	4.3	13	15	18	23	33
Number of immigrants	1.5	1.5	0	.27	1	3.5	13
Enrollment	74	45	21	36	68	135	290
Test score math, Natives	.66	.11	.55	.6	.65	.81	1
Test score math, Immigrants	.6	.14	.44	.51	.58	.8	1

Source: Invalsi data for academic year 2009-10. Test scores are measured as the fraction of correct answers. Immigrants are defined as pupils who do not have an Italian citizenship or who report to speak a foreign language at home.

Table 2: Average individual charateristics by ethnicity

	2nd Grade		5th Grade	
	Natives	Immigrants	Natives	Immigrants
	Panel A. Language Sample			
Share with low educated mother	.27	.29	.31	.28
	(.20)	*(.32)*	*(.19)*	*(.32)*
Share with low educated father	.34	.28	.36	.26
	(.21)	*(.35)*	*(.20)*	*(.35)*
Share with unemployed mother	.52	.72	.53	.69
	(.19)	*(.26)*	*(.18)*	*(.24)*
Share with unemployed father	.24	.37	.23	.37
	(.21)	*(.31)*	*(.20)*	(37)
Share with kindergarten attendance	.87	.74	.86	.67
	(.12)	*(.18)*	*(.13)*	*(.21)*
Share with nursery attendance	.18	.19	.15	.19
	(.42)	*(.43)*	*(.40)*	*(.40)*
Share of male	.51	.51	.51	.51
Observations	434,586	43,763	439,512	43,481
	Panel B. Math Sample			
Share with low educated mother	.27	.29	.31	.28
	(.20)	*(.33)*	*(.19)*	*(.32)*
Share with low educated father	.33	.28	.36	.26
	(.21)	*(.35)*	*(.20)*	*(.35)*
Share with unemployed mother	.52	.72	.53	.69
	(.19)	*(.26)*	*(.18)*	*(.24)*
Share with unemployed father	.24	.37	.23	.37
	(.21)	*(.30)*	*(.20)*	*(.30)*
Share with kindergarten attendance	.87	.74	.86	.67
	(.12)	*(.18)*	*(.13)*	*(.21)*
Share with nursery attendance	.18	.19	.15	.18
	(.42)	*(.43)*	*(.40)*	*(.40)*
Share of male	.51	.51	.51	.51
Observations	430,610	43,773	433,148	43,362

Source: Invalsi data for academic year 2009-10. Variable definitions are the following: Low educated (mother, father)= 1 if the maximum level of education achieved is at most the lower secondary diploma; Unemployed (mother, father)= 1 if not in employement; Kindergarten attendance= 1 if the student has previously attended the kindergarten; Nursery attendance: 1 if the student has previously attended nursery. All the dummy variables set to 0 if the record contains missing values. The share of missing values for each variables is reported in parentheses.

Table 3: OLS estimates of the effect of the number of natives and immigrants on language test scores

	2nd Grade				5th Grade			
	Test Score language natives	Test Score language natives	Test Score language immigrants	Test Score language immigrants	Test Score language natives	Test Score language natives	Test Score language immigrants	Test Score language immigrants
N. natives: $\hat{\beta}$	0.002***	-0.001	0.000	-0.002	0.003***	0.000	0.003	0.001
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
N. immigrants: $\hat{\gamma}$	-0.008***	-0.005**	-0.014***	-0.010**	-0.005**	-0.004**	-0.002	0.001
	(0.001)	(0.001)	(0.004)	(0.004)	(0.001)	(0.002)	(0.004)	(0.004)
Composition: $\hat{\delta}$	-0.010***	-0.004***	-0.015***	-0.008**	-0.008***	-0.004***	-0.005	0.000
	(0.001)	(0.001)	(0.004)	(0.004)	(0.001)	(0.001)	(0.003)	(0.004)
Constant	-0.512***	-0.378***	-0.746***	-1.079***	-0.424***	-0.322***	-0.624***	-0.769***
	(0.012)	(0.077)	(0.042)	(0.190)	(0.012)	(0.046)	(0.046)	(0.112)
School fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Class Specific controls	NO	YES	NO	YES	NO	YES	NO	YES
Mean score	0.67		0.54		0.71		0.61	
(S.D.)	0.12		0.20		0.10		0.16	
(S.D. Log score)	0.32		0.67		0.25		0.52	
Observations	25,750	25,750	17,124	17,124	26,033	26,033	17,244	17,244
R-squared	0.492	0.421	0.418	0.421	0.399	0.399	0.395	0.399

Source: Invalsi data for academic year 2009-10. The unit of observation is the class. The dependent variable is the class average log test score in language. Individual and family background controls include: the share of mothers (or fathers) whose highest educational degree is at most a lower secondary diploma, the share of mothers (or fathers) unemployed, the share of children that have attended kindergarten or nursery and the share of males. We incude as well a set of variables measuring the shares of students in each class for which observations on control variables are missing. Robust standard errors in parentheses. Standard errors are corrected for within school correlation between classes.

Table 4: OLS estimates of the effect of the number of natives and immigrants on math test scores

	2nd Grade				**5th Grade**			
	Test Score math natives	Test Score math natives	Test Score math immigrants	Test Score math immigrants	Test Score in math natives	Test Score math natives	Test Score math immigrants	Test Score math immigrants
N. natives: $\hat{\beta}$	-0.001*	-0.002***	-0.001	-0.002	0.002***	0.000	0.002**	0.000
	(0.001)	(0.001)	(0.001)	(0.002)	(0.000)	(0.000)	(0.001)	(0.001)
N. immigrants: $\hat{\gamma}$	-0.006***	-0.004***	-0.007***	-0.006**	-0.005***	-0.004***	-0.004*	-0.001
	(0.001)	(0.001)	(0.003)	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)
Composition: $\hat{\delta}$	-0.005***	-0.002***	-0.006**	-0.003	-0.007***	-0.004***	-0.006***	-0.002
	(0.001)	(0.001)	(0.003)	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)
Constant	-0.503***	-0.374***	-0.634***	-0.784***	-0.500***	-0.425***	-0.630***	-0.734***
	(0.010)	(0.066)	(0.025)	(0.106)	(0.008)	(0.035)	(0.020)	(0.070)
School fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Class Specific controls	NO	YES	NO	YES	NO	YES	NO	YES
Mean score	0.63		0.55		0.65		0.58	
(S.D.)	0.13		0.17		0.12		0.16	
(S.D. Log score)	0.29		0.43		0.23		0.32	
Observations	25,728	25,728	17,109	17,109	26,018	26,018	17,224	17,224
R-squared	0.501	0.507	0.495	0.498	0.467	0.479	0.482	0.487

Source: Invalsi data for academic year 2009-10. The unit of observation is the class. The dependent variable is the class average log test score in math. Individual and family background controls include: the share of mothers (or fathers) whose highest educational degree is at most a lower secondary diploma, the share of mothers (or fathers) unemployed, the share of children that have attended kindergarten or nursery and the share of male. We incude as well a set of variables measuring the shares of students in each class for which observations on control variables are missing. Robust standard errors in parentheses. Robust standard errors in parentheses. Standard errors are corrected for within school correlation between classes.

Table 5: Natives background and number of immigrants in a class

	2nd Grade		5th Grade	
	Number of immigrants per class	Number of immigrants per class	Number of immigrants per class	Number of immigrants per class
% natives with low educated mother	0.290**	0.471***	0.322***	0.537***
	(0.122)	(0.114)	(0.109)	(0.110)
% natives with low educated father	0.372***	0.380***	-0.023	0.011
	(0.111)	(0.108)	(0.101)	(0.104)
% natives with unemployed mother	-2.711***	-0.463***	-2.660***	-0.399***
	(0.097)	(0.105)	(0.086)	(0.099)
% natives with unemployed father	1.849***	0.759***	1.680***	0.683***
	(0.103)	(0.122)	(0.095)	(0.108)
Constant	2.679***	1.741***	2.811***	1.792***
	(0.049)	(0.058)	(0.049)	(0.058)
School fixed effects	NO	YES	NO	YES
Observations	25,750	25,750	26,033	26,033
R-squared	0.056	0.561	0.061	0.548

Source: Invalsi data for academic year 2009-10. The dependent variable is the number of immigrants per class. Regressions are based on the "language sample" comprising subject who were not abent on the date of the language test. We incude as well province fixed effects and a set of dummies indicating whether individual and family control variables are missing for some observations, in which case the correspondent variable is set to zero. Robust standard errors in parentheses. Standard errors are corrected for within school correlation between classes.

Table 6: Estimates of the effects of immigrants and natives on natives' performance, using the SNTE assumption and rules for class formation as exogenous source of variation

	Language test score				**Math test score**			
	2nd Grade		**5th Grade**		**2nd Grade**		**5th Grade**	
Natives' increase: $\hat{\beta} = \hat{\psi}_1$	-0.007	-0.010**	-0.003	-0.004	-0.014***	-0.015***	-0.007**	-0.008***
	(0.005)	(0.005)	(0.003)	(0.003)	(0.005)	(0.005)	(0.003)	(0.003)
Immigrants' increase: $\hat{\gamma}$	-0.134***	-0.131***	-0.023	-0.024	-0.091***	-0.090***	0.017	0.020
	(0.027)	(0.029)	(0.028)	(0.030)	(0.024)	(0.026)	(0.033)	(0.034)
Pure Composition Change: $\hat{\delta} = \hat{\gamma} - \hat{\beta}$	-0.128***	-0.121***	-0.021	-0.020	-0.077***	-0.074***	0.024	0.028
	(0.026)	(0.028)	(0.028)	(0.030)	(0.022)	(0.024)	(0.033)	(0.035)
Differential class size effect effect: $\hat{\psi}_2$	0.030***	0.026***	0.003	0.003	0.016***	0.017***	-0.004	-0.004
	(0.006)	(0.006)	(0.004)	(0.004)	(0.006)	(0.006)	(0.004)	(0.005)
Correction factor: $(1 - \frac{\hat{\partial} N}{\partial C}\|_{I>0})$	-0.234***	-0.213***	-0.155*	-0.146**	-0.250***	-0.227***	-0.145*	-0.137*
	(0.076)	(0.079)	(0.081)	(0.073)	(0.083)	(0.083)	(0.078)	(0.071)
Observations	434,586	434,586	439,512	439,512	430,610	430,610	433,148	433,148
Individual controls:	NO	YES	NO	YES	NO	YES	NO	YES

Source: Invalsi data for academic year 2009-10. The dependent variable is the log of test score in language (math). All regressions include a dummy indicating whether there are immigrants at the school/grade, the number of disabled students in the school/grade, a dummy for the presence of disabled students in the school/grade, a 2nd order polynomial of long run enrollment (2nd grade) and actual enrollment (5th grade). Regressions with controls include a set of family and individual covariates. These controls are: a dummy equal to 1 if the mother (or father) have attended, at most, a lower secondary school, a dummy indicating whether the mother (or the father) is unemployed, a dummy for kindergarten (and nursery) attendance and a dummy equal to 1 if the student is male. We incude as well province fixed effects and a set of dummies indicating whether individual and family control variables are missing for some observations, in which case the correspondent variable is set to zero. Robust standard errors are in parentheses. Standard errors are corrected for within school correlation between classes. Standard erros of γ and δ do not take into account the variability in the estimates of $\frac{\Delta N}{\Delta C}_{|I>0}$.

Table 7: First stage estimates

	Language sample				Math sample			
	2nd Grade		5th Grade		2nd Grade		5th Grade	
	(C)	$(C * 1_{I>0})$	(C)	$(C * 1_{I>0})$	(C)	$(C * 1_{I>0})$	(C)	$(C * 1_{I>0})$
Panel A. Without controls								
$\bar{C}$	0.246***	-0.101***	0.356***	-0.116***	0.245***	-0.101***	0.355***	-0.116***
	(0.029)	(0.007)	(0.026)	(0.007)	(0.029)	(0.007)	(0.026)	(0.007)
$\bar{C} * 1_{I>0}$	-0.128***	0.224***	-0.238***	0.228***	-0.128***	0.224***	-0.237***	0.228***
	(0.030)	(0.014)	(0.027)	(0.013)	(0.030)	(0.014)	(0.027)	(0.013)
F statistic	59.33	186.9	112.1	229.2	58.56	189.5	111.6	228.4
Observations	434,586	434,586	439,512	439,512	430,610	430,610	433,148	433,148
Panel B. With controls								
$\bar{C}$	0.247***	-0.100***	0.355***	-0.117***	0.246***	-0.101***	0.354***	-0.117***
	(0.028)	(0.007)	(0.026)	(0.007)	(0.029)	(0.007)	(0.026)	(0.007)
$\bar{C} * 1_{I>0}$	-0.133***	0.220***	-0.238***	0.228***	-0.132***	0.221***	-0.238***	0.228***
	(0.030)	(0.014)	(0.027)	(0.013)	(0.030)	(0.014)	(0.027)	(0.013)
F statistic	59.61	180.9	114.5	232.0	59.00	183.0	113.8	231.6
Observations	434,586	434,586	439,512	439,512	430,610	430,610	433,148	433,148

Source: Invalsi data for academic year 2009-10. All regressions include a dummy indicating whether there are immigrants at the school/grade level, the number of disabled students in the school/grade, a dummy for the presence of disabled students in the school/grade, a 2nd order polynomial of long run enrollment (2nd grade) and actual enrollment (5th grade). Regressions with controls include a set of family and individual covariates. These controls are: a dummy equal to 1 if the mother (or father) have at most attended the lower secondary school, a dummy indicating whether the mother (or the father) is unemployed, a dummy for kindergarten (and nursery) attendance and a dummy equal to 1 if the student is male. We incude as well province fixed effects and a set of dummies indicating whether individual and family control variables are missing for some observations, in which case the correspondent variable is set to zero. Province fixed effects are included. Robust standard errors are in parentheses. Standard errors are corrected for within school correlation between classes.

Table 8: Instrumental variable estimates of $\frac{\partial N}{\partial C}_{|I>0}$

	2nd Grade				**5th Grade**			
	Num. (Natives)	FS (Class size)	Num. (Natives)	FS (Class size)	Num. (Natives)	FS (Class size)	Num. (Natives)	FS (Class size)
				Panel A. Language Sample				
C: $\frac{\partial N}{\partial C}_{\vert I>0}$	1.2338***		1.2130***		1.1551***		1.1456***	
	(0.076)		(0.079)		(0.081)		(0.073)	
$\bar{C}$		0.0913***		0.0843***		0.0707***		0.0763***
		(0.016)		(0.016)		(0.015)		(0.015)
F statistic		31.51		29.15		21.49		26.73
Observations	25,750	25,750	25,750	25,750	26,033	26,033	26,033	26,033
				Panel B. Math Sample				
C: $\frac{\partial N}{\partial C}_{\vert I>0}$	1.2500***		1.2271***		1.1451***		1.1368***	
	(0.082)		(0.083)		(0.078)		(0.071)	
$\bar{C}$		0.0869***		0.0823***		0.0725***		0.0779***
		(0.016)		(0.016)		(0.015)		(0.015)
F statistic		28.40		27.67		22.64		27.97
Observations	25,728	25,728	25,728	25,728	26,018	26,018	26,018	26,018
Individual controls	NO	NO	YES	YES	NO	NO	YES	YES

Source: Invalsi data for academic year 2009-10. Regression in a sample of school with immigrants. The unit of observation is a class. All regressions include a dummy for the presence of disabled students in the school/grade, the number of disabled student in the school/grade, a 2nd order polynomial of long run enrollment (2nd grade) and actual enrollment (5th grade). Regressions with individual controls include a full set of family background controls. These controls are: the share of mothers (or fathers) that have attended at most the lower secondary school, the share of mothers (or fathers) unemployed, the share of children that have attended kindergarten or nursery and the share of males. We incude as well a set of variables measuring the shares of students in each class for which observations on control variables are missing. Robust standard errors in parentheses. Robust standard errors are in parentheses. Standard errors are corrected for within school correlation between classes.

Table 9: Estimates of the effects of immigrants and natives on natives' performance, using rules for class formation as exogenous source of variation, on the restricted sample of schools in small municipalities with probability of enrolling immigrants below the sample median .

	Language test score				**Math test score**			
	2nd Grade		**5th Grade**		**2nd Grade**		**5th Grade**	
Natives' increase: $\hat{\beta} = \hat{\psi}_1$	-0.004	-0.008	0.004	0.002	-0.007	-0.009	-0.002	-0.005
	(0.007)	(0.007)	(0.003)	(0.003)	(0.007)	(0.007)	(0.004)	(0.004)
Immigrants' increase: $\hat{\gamma}$	-0.103***	-0.088***	-0.025	-0.024	-0.060***	-0.058***	0.014	0.020
	(0.027)	(0.026)	(0.017)	(0.018)	(0.022)	(0.022)	(0.028)	(0.030)
Pure Composition Change: $\hat{\delta} = \hat{\gamma} - \hat{\beta}$	-0.098***	-0.079***	-0.029*	-0.026	-0.053**	-0.049**	0.017	0.024
	(0.025)	(0.024)	(0.018)	(0.019)	(0.019)	(0.020)	(0.028)	(0.031)
Differential class size effect: $\hat{\psi}_2$	0.025***	0.021***	0.008*	0.006	0.016***	0.014**	-0.003	-0.004
	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)
Correction factor: $(1 - \frac{\partial N}{\partial C}\|_{I>0})$	-0.257*	-0.262*	-0.260	-0.246	-0.293*	-0.282*	-0.186	-0.172
	(0.137)	(0.153)	(0.199)	(0.182)	(0.145)	(0.143)	(0.129)	(0.119)
Observations	202,870	202,870	201,030	201,030	176,420	176,420	175,473	175,473
Individual controls:	NO	YES	NO	YES	NO	YES	NO	YES

Source: Invalsi data for academic year 2009-10. The probability of enrolling immigrants is estimated, with a separate regression, at the school level as a function of municipalities and schools characteristic. IV estimates are restricted to the subsample of schools with an estimated probability of enrolling immigrants less than or equal to the median. Small municipalities are those with less than 100,000 inhabitants. The dependent variable is the log of the test score in language (math). All regressions include a dummy indicating whether there are immigrants at the school/grade level, the number of disabled students in the school/grade, a dummy for the presence of disabled students in the school/grade, a 2nd order polynomial of long run enrollment (2nd grade) and actual enrollment (5th grade). Regressions with controls include a set of family and individual covariates. These controls are: a dummy equal to 1 if the mothers (or fathers) have attended at most the lower secondary school, a dummy indicating whether the mother (or the father) is unemployed, a dummy for kindergarten (and nursery) attendance and a dummy equal to 1 if the student is male. We incude as well province fixed effects and a set of dummies indicating whether individual and family control variables are missing for some observations, in which case the correspondent variable is set to zero. Robust standard errors are in parentheses. Standard errors are corrected for within school correlation between classes. Standard erros of γ and δ do not take into account the variability in the estimates of $\frac{\Delta N}{\Delta C}_{|I>0}$.

Table 10: First stage estimates for the restricted sample of schools in small municipalities with probability of enrolling immigrants below the sample median.

	Language sample				Math sample			
	2nd Grade		5th Grade		2nd Grade		5th Grade	
	(C)	$(C * 1_{I>0})$	(C)	$(C * 1_{I>0})$	(C)	$(C * 1_{I>0})$	(C)	$(C * 1_{I>0})$
Panel A. Without controls								
$\bar{C}$	0.181***	-0.140***	0.313***	-0.149***	0.173***	-0.139***	0.324***	-0.134***
	(0.034)	(0.010)	(0.030)	(0.011)	(0.034)	(0.011)	(0.030)	(0.011)
$\bar{C} * 1_{I>0}$	-0.060*	0.282***	-0.185***	0.278***	-0.044	0.290***	-0.182***	0.276***
	(0.035)	(0.018)	(0.031)	(0.017)	(0.036)	(0.019)	(0.032)	(0.018)
F statistic	27.67	178.49	65.04	190.02	26.74	169.23	69.21	162.02
Observations	202,870	202,870	201,030	201,030	176,420	176,420	175,473	175,473
Panel B. With controls								
$\bar{C}$	0.183***	-0.138***	0.312***	-0.150***	0.175***	-0.137***	0.322***	-0.137***
	(0.033)	(0.010)	(0.029)	(0.011)	(0.034)	(0.011)	(0.030)	(0.011)
$\bar{C} * 1_{I>0}$	-0.068**	0.277***	-0.187***	0.276***	-0.050	0.287***	-0.184***	0.276***
	(0.035)	(0.018)	(0.030)	(0.017)	(0.035)	(0.019)	(0.031)	(0.018)
F statistic	27.61	173.69	66.25	191.42	26.83	165.96	70.03	164.16
Observations	202,870	202,870	201,030	201,030	176,420	176,420	175,473	175,473

Source: Invalsi data for academic year 2009-10. The probability of enrolling immigrants is estimated, with a separate regression, at the school level as a function of municipalities and schools characteristic. IV estimates are restricted to the subsample of schools with an estimated probability of enrolling immigrants less than or equal to the median. Small municipalities are those with less than 100,000 inhabitants. All regressions include a dummy indicating whether there are immigrants at the school/grade level, the number of disabled students in the school/grade, a dummy for the presence of disabled students in the school/grade, a 2nd order polynomial of long run enrollment (2nd grade) and actual enrollment (5th grade). Regressions with controls include a set of family and individual covariates. These controls are: a dummy equal to 1 if the mother (or the father) has attended at most the lower secondary school, a dummy indicating whether the mother (or the father) is unemployed, a dummy for kindergarten (and nursery) attendance and a dummy equal to 1 if the student is male. We incude as well province fixed effects and a set of dummies indicating whether individual and family control variables are missing for some observations, in which case the correspondent variable is set to zero. Robust standard errors are in parentheses. Standard errors are corrected for within school correlation between classes.

Table 11: Instrumental variable estimates of $\frac{\partial N}{\partial C}_{|I>0}$ in the restricted sample of schools in small municipalities with probability of enrolling immigrants below the sample median.

	2nd Grade				**5th Grade**			
	Num. (Natives)	FS (Class size)	Num. (Natives)	FS (Class size)	Num. (Natives)	FS (Class size)	Num. (Natives)	FS (Class size)
				Panel A. Language Sample				
C: $\frac{\partial N}{\partial C}_{\|I>0}$	1.257***		1.262***		1.260***		1.246***	
	(0.137)		(0.153)		(0.199)		(0.182)	
$\bar{C}$		0.061***		0.053**		0.040*		0.042**
		(0.022)		(0.021)		(0.021)		(0.020)
F statistic		7.53		6.45		3.84		4.48
Observations	11,476	11,476	11,476	11,476	11,312	11,312	11,312	11,312
				Panel B. Math Sample				
C: $\frac{\partial N}{\partial C}_{\|I>0}$	1.293***		1.282***		1.186***		1.172***	
	(0.145)		(0.143)		(0.129)		(0.119)	
$\bar{C}$		0.064***		0.061***		0.054**		0.056***
		(0.023)		(0.022)		(0.022)		(0.021)
F statistic		7.53		7.71		6.39		7.43
Observations	9,871	9,871	9,871	9,871	9,864	9,864	9,864	9,864
Individual controls	NO	NO	YES	YES	NO	NO	YES	YES

Source: Invalsi data for academic year 2009-10. The probability of enrolling immigrants is estimated, with a separate regression, at the school level as a function of municipalities and schools characteristic. IV estimates is restricted to the subsample of schools with an estimated probability of enrolling immigrants less than or equal to the median. Small municipalities are those with less than 100,000 inhabitants. Regression is based on a sample of school with immigrants. The unit of observation is a class. All regressions include a dummy for the presence of disabled students in the school/grade level, the number of disabled in the school/grade, a 2nd order polynomial of long run enrollment (2nd grade) and actual enrollment (5th grade). Regressions with individual controls include a full set of family background controls. These controls are: the share of mothers (or father) that have attended at most the lower secondary school, the share of mothers (or fathers) unemployed, the share of children that have attended kindergarten or nursery and the share of males. We incude as well a set of variables measuring the shares of students in each class for which observations on control variables are missing. Robust standard errors in parentheses. Robust standard errors are in parentheses. Standard errors are corrected for within school correlation between classes.