The Tower of Babel in the Classroom.
Immigrants and Natives in Italian Schools

Rosario Maria Ballatore[‡] Margherita Fort[§] Andrea Ichino[¶]

July 4, 2017

Online Appendix

In this appendix we collect additional figures and tables which could not be included in the main text due to space constraints. The material is listed below with brief comments and it is organized as follows.

Section A clarifies formally how the interaction between February and September enrolment and the rules of class formation can generate the exogenous variation that we exploit for identification.

Section B contains supplementary material concerning the main samples and the analysis in which, following the INVALSI classification, immigrants are defined as children born in Italy or elsewhere from parents who are both non-Italian, while natives are children born in Italy or elsewhere, from at least one Italian parent.

Section C contains supplementary material concerning the samples and the analysis focussed on first first generation immigrants defined as children born outside Italy from parents who are both non-Italian. The complement is therefore composed by quasi-natives, i.e. natives and children born in Italy from non-Italian parents.

[‡]Bank of Italy
[§]University of Bologna, Economics, CESifo and IZA
[¶]European University Institute and Bologna, Economics, CEPR, CESifo and IZA

A Native class size and number of immigrants

To clarify how the interaction between February and September enrolment can generate exogenous differential variations in the number of immigrants and natives per class, suppose that in grade g of school s predicted average class size $\bar{C}^N_{sg}$ at the school level, based on February native pre-enrolment and rules of class formation, can take three equally likely values: $H > M > L = H/2$. The principal knows that if $\bar{C}^N_{sg} = H$ for a class in February, with probability π that class will be split in September because of late enrolment in the corresponding school, originating two small classes each one with (approximately) $L = H/2$ natives. In the other two cases, instead, there is no risk of splitting.

Each principal manages three otherwise similar schools with different predicted average class sizes and has to allocate a total of I immigrants who enrol in February or September. Let's also assume, again for simplicity, that each school has one class. Since there is a probability $1 - \pi$ (with $0 < \pi < 1$) that the class expected to be large in February, will remain large (late enrolment insufficient to cause splitting in September), the principal will not plan to put immigrants in that class to avoid possible disruption. In the other two classes, instead, predicted class size based on native enrolment is low enough that immigrants cause no disruption and can be randomly distributed. Therefore, the average number of immigrants in the three types of classes, as anticipated in February, is:

$$I_{sg} = \begin{cases} 0 & \text{if} \quad \bar{C}^N_{sg} = H \\ \frac{I}{2} & \text{if} \quad \bar{C}^N_{sg} = M \\ \frac{I}{2} & \text{if} \quad \bar{C}^N_{sg} = L \end{cases} \tag{A–1}$$

In September, however, the schools with the high predicted class size will split their class with probability π. Therefore, the allocation of immigrants based on the *final* number of natives per class C^N_{sg}, after late enrolment has occurred, is,

$$I_{sg} = \begin{cases} \approx 0 & \text{if} \quad C^N_{sg} \approx H \\ \frac{I}{2} & \text{if} \quad C^N_{sg} \approx M \\ \approx \frac{I}{2}\frac{1}{(1+2\pi)} & \text{if} \quad C^N_{sg} \approx L \end{cases} \tag{A–2}$$

where the size of the three types of classes is now approximately H, M or L because of late enrolment. The average number of immigrants per class remains approximately zero in high-sized classes and does not change in medium- sized classes, while in the remaining group it is an average of the $\frac{I}{2}$ immigrants allocated to each one of the originally small classes and of the 0 (or very few) immigrants in the classes that become small because of September splitting.[1] As a result of this allocation mechanism, since $\frac{1}{1+2\pi} < 1$, the average number of immigrants per class is a hump-shaped function of the average number of natives.

[1]Suppose that there are R principals, and therefore R classes of each type in February (given that each principal manages one class of each type). The number of high-sized classes that split is πR and they originate $2\pi R$ small classes. Therefore after splitting, the number of small classes is $R + 2\pi R$. Each one of the R originally small classes have $I/2$ immigrants, while the new $2\pi R$ small classes have ≈ 0 immigrants. Thus the final average number of immigrants in small classes is given by

$$\approx \frac{R\frac{I}{2} + 2\pi R0}{R + 2\pi R} \approx \frac{I}{2}\frac{1}{(1+2\pi)} < \frac{I}{2}.$$

The number of intermediate size classes remains R in September, each one with $I/2$ immigrants. High-sized classes have ≈ 0 immigrants and their final number is $R(1-\pi)$.

B Main samples and analysis

B.1 Math sample

- Table B.1 reports descriptives statistics on the math sample composed of children who were not absent on the day of the Math test. We construct the sample as follows. Starting from the universe of schools (16,780 in 7,561 institutions in grade 2; 16,789 in 7,549 institutions in grade 5) we operate the following sample restrictions. We retain all schools in institutions in which at least one immigrant applies for a given grade (14,546 schools of 5,966 different institutions in grade 2; 14,603 schools of 6,010 different institutions in grade 5). We then restrict the attention to schools that enroll between 10 and 75 native students (11,555 schools of 5,462 different institutions in grade 2; 11,638 schools of 5,486 different institutions in grade 5). We remove records of students with missing test scores in math and we collapsed the data at the school level. We drop schools in which at least one of the included covariate is missing (1695 schools and 744 institutions in grade 2; 1674 schools and 728 institutions in grade 5) and the few schools (3 in grade 2; 2 in grade 5) in which no native took the test. We restrict our attention to the schools that are grouped together with other schools in educational institutions managed by a single principal: as explained in the paper (see the Introduction and Section 3), our identification strategy cannot apply to "stand-alone" schools. This leaves us with 8,006 schools of 2,865 different institutions in grade 2 and 8,085 schools in 2,883 different institutions in grade 5.

  The average enrolment of natives per school-grade is 28.1 while for immigrants it is 3.76. As we note in the paper for the language sample also in this sample, immigrants tend to perform worse than natives in reading and math, but the gap between ethnic groups is more sizeable in language. Natives perform relatively better in Italian than in math and unsurprisingly the opposite happens for immigrants. The gap between natives and immigrants in reading tends to narrow across grades but remains relatively more stable in math. Finally, the dispersion in the score distribution for both Italian and math is lower among natives who are more homogeneous than immigrants. The fact that immigrants test scores are lower on average, has motivated the public opinion concern that immigrant inflows reduce native performance.

- Figure B.1 plots, for the math sample, the average number of natives per class (squares - left vertical axis) and of immigrants per class (circles - right vertical axis) for each level of theoretical class size based on native enrolment using data on the math sample.

The figure also plots fitted values of the two relationships (solid for immigrants and dashed for natives). Theoretical class size is calculated as a function of final enrolment of natives N_{sg} in school s and grade g, using the "Maimonides-type" rule in equation (2) in the paper. The figure shows that the average number of natives in a class is an increasing function of theoretical class size, as predicted by the conventional effects of a Maimonides-type rule.

- Figure B.2 plots, for the math sample, the "Maimonides-type" rule in equation (2) in the paper on the number of natives per classe (left panel) and the available space for immigrants. In the left panel the dashed line plots the theoretical class size C_{sg}^{N} as a function of the final enrolment of natives in each school. The dark dots describe how the actual number of natives per class changes as a function of their enrolment and the light dots describe instead the total actual class size, including immigrants, as a function of native enrolment. The right panel of Figure B.2 plot this vertical distance (the connected light dots, which represent the actual number of immigrants per class) as a function of native enrolment, suggesting that it is not constant. The same panels also plot the theoretically available space for immigrants (dashed line), defined as the maximum number of students in a class (25) minus the theoretical class size based on the number of natives C_{sg}^{N}. Note that there is a correspondence between the spikes of the space actually used for immigrants (i.e., their number per class) and the theoretically available space, not only in the 10-75 range of native enrolment. Moreover the used space for immigrants tends to be relatively higher than expected for intervals of native enrolment that generate medium size classes. This result is due to the interaction between early/late enrolment and rules of class formation on the allocations of immigrants across schools and is responsible for the difference in the shapes displayed in Figure B.1.

- In Figure B.3 and Figure B.4 we construct, for the math sample, graphs similar to those in Figure B.1 and in Figure B.2 using as outcomes not class size variables (total class size, number of natives or number of immigrants in class) but rather the average values of the control variables included in our regression, net of institution-grade fixed effects: if our identification strategy is internally valid, we should observe that covariates do not exhibit spikes nor hump shapes but evolve smoothly over native enrolment in schools or over corresponding predicted class size. Figure B.3 plots the within-group (instituion-grade) residuals of each covariate against the predicted class size based on natives enrolment and Maimonides rule: each graph shows remarkably flat lines in sharp contrast with the relationships highlighted in Figure B.1, where we could see that the number of natives in class is an increasing function of theoretical class size, while the number of immigrants in class is a humped-shaped function of theoretical class size. Figure B.4 plots the within-group residuals of each covariates against native enrolment to check if there is evidence of discontinuities as the one observes at the 25,50 and 75 cutoffs in class size measures in Figure B.2: we do not observe sharp jumps but there is a trend in the parental education that suggests that lower educated parents are more likely to have children attending small schools. This negative correlation is not a threat to the internal validity of our identification strategy and it is accounted for in the regression models we estimate. To sum up, the pattern we observe in both Figure B.3 and Figure B.4 is re-assuring and we are confident in the internal validity of our identification strategy.

- In Table B.2 we formally test, for the math sample, if our instrumental variables are correlated with the included covariates, net of institution-grade fixed effects and trends in enrolment. We do so estimating a system of equations where the dependent variables are within-group (institution-grade) residuals of each covariate and the regressors are the corresponding within-group residuals of quadratic trend in native enrolment and our 15 instrumental variables: as one would expect given the pattern in Figure B.3 and Figure B.4, most coefficients are not statistically different from zero but we also detect significant correlations with some background variables, notably parental education. Angrist et al. (forthcoming) clarify that this correlation between rules of class formation and observables is likely to be the result of cheating, mainly due to shirking. As documented in Section 5.6 in the paper, our results are robust to score manipulation.

B.2 Exact identification

- Table B.3 reports in each column a different first stage regression correspondent to the IV estimates of Table 7 in the paper. The estimates with two endogenous regressors confirm the intuitions illustrated above (Section A): principals tend to allocate less immigrants to classes as predicted class size increases but do allocate more immigrants to classes that have a medium predicted class size.

- Table B.4 reports in each column a different first stage regression correspondent to the IV estimates of Table 8 in the paper in the paper. Also in this case the estimates with two endogenous regressors confirm the intuitions illustrated above (Section A):

principals tend to allocate less immigrants to classes as predicted class size increases but do allocate more immigrants to classes that have a medium predicted class size.

B.3 Internal validity: robustness to score manipulation

- Table B.5 report the first stage estimates that correspond to the instrumental variable estimates of Table 12 in the paper. When we consider only the number of immigrants in class as endogenous and we treat the number of natives in class as a nuisance parameter, we do not face weak instruments problems. However our instruments for the number of immigrants are weak when also the number of natives is treated as endogenous. There is no indication of under-identification of the model parameters.

B.4 Internal validity: IV estimates with no school controls

- Table B.6 and Table B.7 report the IV estimates that correspond to the instrumental variable estimates of Tables 2 and 3 in the paper where we change the set of controls and we do not include school level variables. The main results of the paper are confirmed: in the pooled sample for both language and math test scores, estimates of Pure Ethnic Composition Effect (PEC) are negative, statistically significant and similar in size to those reported in the paper. The estimates of the PEC in grade 2 and grade 5 follow the same pattern observed in the paper.

- Table B.8 and Table B.9 report the first stage estimates that correspond to the instrumental variable estimates of Table B.6 and Table B.7, respectively. The most precise results pool data on the 2nd and 5th grade. In this pooled sample for both language and math test scores, we do not face weak instruments problems when we consider only the number of immigrants in class as endogenous and we treat the number of natives in class as a nuisance parameter. However our instruments for the number of immigrants are weak when also the number of natives is treated as endogenous. There is no indication of under-identification of the model parameters. Results by grade are more imprecise, as in the corresponding Tables in the paper, but fully consistent with those reported in the main pooled sample.

B.5 Comparison with fixed effect strategis: class samples

- Table B.10 reports descriptives statistics on the language sample used for the fixed effect regressions of Table B.12. In this sample we denote as immigrants first and second generation immigrants. We construct the sample as follows.

Starting from the universe of schools (16,828 in 7,561 institutions in grade 2; 16,803 in 7,549 institutions in grade 5) we operate the following sample restrictions. We retain all schools in institutions in which at least one immigrant applies for a given grade (14,580 schools of 5,966 different institutions in grade 2; 14,616 schools of 6,010 different institutions in grade 5). We then restrict the attention to schools that enroll between 10 and 75 native students (11,574 schools of 5,466 different institutions in grade 2; 11,639 schools of 5,484 different institutions in grade 5). We remove records of students with missing test scores in language and we collapsed the data at the class level. We drop classes in which at least one of the included covariate is missing (3,249 classes-; 1,955 schools and 1,136 institutions in grade 2; 3,133 classes, 1,923 schools and 1,121 institutions in grade 5) and the few classes (3 in grade 2; 9 in grade 5) in which no native took the test. We restrict our attention to the classes in schools that are grouped together with other schools in educational institutions managed by a single principal: as explained in the paper (see the Introduction and Section 3), our identification strategy cannot apply to "stand-alone" schools. This leaves us with 13,292 classes in 8,012 schools of 2,867 different institutions in grade 2 and with 13,449 classes of 8,082 schools in 2,880 different institutions in grade 5.

The average enrolment of natives per school-grade is about 28 while for immigrants it is approximately 4. As we note in the paper for the language sample also in this sample, immigrants tend to perform worse than natives in reading and math, but the gap between ethnic groups is more sizeable in language. Natives perform relatively better in Italian than in math and unsurprisingly the opposite happens for immigrants. The gap between natives and immigrants in reading tends to narrow across grades but remains relatively more stable in math. Finally, the dispersion in the score distribution for both Italian and math is lower among natives who are more homogeneous than immigrants. The fact that immigrants test scores are lower on average, has motivated the public opinion concern that immigrant inflows reduce native performance. There is no noticeable difference with the descriptive statistics reported in Table 1 in the paper.

- Table B.11 reports descriptives statistics on the math sample used for the fixed effect regressions of Table B.12. In this sample we denote as immigrants first and second generation immigrants. We construct the sample as follows.

  Starting from the universe of schools (16,780 in 7,561 institutions in grade 2; 16,789 in 7,549 institutions in grade 5) we operate the following sample restrictions. We retain all schools in institutions in which at least one immigrant applies for a given grade (14,546 schools of 5,966 different institutions in grade 2; 14,603 schools of 6,010 different

institutions in grade 5). We then restrict the attention to schools that enroll between 10 and 75 native students (11,555 schools of 5,462 different institutions in grade 2; 11,638 schools of 5,486 different institutions in grade 5). We remove records of students with missing test scores in math and we collapsed data at the class level. We drop classes in which at least one of the included covariate is missing (3,242 classes, 1,951 schools and 1,136 institutions in grade 2; 3,132 classes, 1,925 schools and 1,122 institutions in grade 5) and the few classes (3 in grade 2;6 in grade 5) in which no native took the test. We restrict our attention to the classes in schools that are grouped together with other schools in educational institutions managed by a single principal: as explained in the paper (see the Introduction and Section 3), our identification strategy cannot apply to "stand-alone" schools. This leaves us with 13,284 classes in 8,004 schools of 2,865 different institutions in grade 2 and with 13,449 classes of 8,082 schools in 2,881 different institutions in grade 5.
The average enrolment of natives per school-grade is about 28 while for immigrants it is approximately 4. As we note in the paper for the language sample also in this sample, immigrants tend to perform worse than natives in reading and math, but the gap between ethnic groups is more sizeable in language. Natives perform relatively better in Italian than in math and unsurprisingly the opposite happens for immigrants. The gap between natives and immigrants in reading tends to narrow across grades but remains relatively more stable in math. Finally, the dispersion in the score distribution for both Italian and math is lower among natives who are more homogeneous than immigrants. The fact that immigrants test scores are lower on average, has motivated the public opinion concern that immigrant inflows reduce native performance. There is no noticeable difference with the descriptive statistics reported in Table B.1.

- Table B.12 reports the results of school-grade fixed effects regressions on a class-level sample following Contini (2013) and Ohinata and Van Ours (2013). The key assumption to intepret this estimates as causal is that controlling for selection into schools the allocation of natives and immigrants to classes is as good as random in the sample. Compared to the main IV estimates we report in the paper in Table 2 and Table 3, the resulting estimates of δ from Table B.12 have the same sign but are significantly smaller. We interpret this as suggestive evidence that this strategy does not fully account for the non-random allocation of students with different ethnic background across classes within schools.

C Quasi-natives and first generation immigrants

C.1 Descriptives: language and math samples

- Table C.1 and Table C.2 report descriptives statistics on the language and math sample used for the OLS regression of Table C.3 and the IV regressions of Table 10 and Table 11 in the paper. In this sample immigrants are defined as first generation only. We describe the language sample construction below. We proceed in a similar way for the math sample. Starting from the universe of schools (16,828 in 7,561 institutions in grade 2; 16,803 in 7,549 institutions in grade 5) we operate the following sample restrictions. We retain all schools in institutions in which at least one *first generation* immigrant applies for a given grade (12,181 schools of 4,661 different institutions in grade 2; 12,908 schools of 4,988 different institutions in grade 5). We then restrict the attention to schools that enroll between 10 and 75 native students (9,712 schools of 4,296 different institutions in grade 2; 10,319 schools of 4,576 different institutions in grade 5). We drop records of students with missing test scores (10 schools and 1 institution in grade 2; 1 schools and no institutions in grade 5) and we collapse the data at the school level. We drop schools in which at least one of the included covariate is missing (1,440 schools and 603 institutions in grade 2; 1,490 schools and 609 institutions in grade 5). We restrict our attention to the schools that are grouped together with other schools in educational institutions managed by a single principal: as explained in the paper (see the Introduction and Section 3), our identification strategy cannot apply to "stand-alone" schools. This leaves us with 7,037 schools of 2,469 different institutions in grade 2 and 7,496 schools in 2,637 different institutions in grade 5 in the language sample. In the math sample, the analogous sample selection process leads to a final sample of 7,030 schools of 2,467 different institutions in grade 2 and 7,494 schools of 2,637 different institutions in grade 5. The descriptive statistics on average test scores of natives and first generation immigrants in Table C.1 and Table C.2 follow a similar pattern and we comment here on those reported in Table C.1. The average enrolment of natives per school-grade is 28.2 while for first generation immigrants it is 1.44 in grade 2 and 1.87 in grade 5, roughly less than half than the enrollment of first and second generation immigrants (see Table 1 in the paper). First generation immigrants tend to perform worse than natives in reading and math, but the gap between ethnic groups is more sizeable in language. Natives perform relatively better in Italian than in math and unsurprisingly the opposite happens for immigrants. The gap between natives and immigrants in reading tends to narrow across grades but remains relatively more stable in math. Finally, the dispersion in the score distribution for both Italian and math is lower among natives who are more homogeneous than immigrants. Test scores of first generation immigrants are fairly similar to those we observe in Table 1 when first and second generation immigrants are pooled.

- Table C.3 presents estimates of the counfounded ethnic composition effect based on a regression that includes institution×grade fixed effects on a sample where we use a more restrictive definition of immigrants and we focus on first generation immigrants. Here identification requires the assumption that students are randomly allocated across schools within a given instution and grade. Compared to the corresponding estimates of Table 4 in the paper, the results suggest a larger (in absolute values) counfounded ethnic composition effect in both language and math. When instead we compare these estimates with the one obtained relying on our instrumental variables identification strategy (Table 10 and 11 in the paper), we confirm that estimates that exploit only within institution variation across schools underestimate the true PEC.

C.2 IV: Language sample

- Figure C.1 and Figure C.2 replicate Figure 1 and Figure 2 in the paper for the language sample where we use the more restrictive definition of immigrants and focus on first generation immigrants: the pattern is similar to the one discussed in the paper.

- In Figure C.3 and Figure C.4 we construct graphs similar to those in Figure C.1 and Figure C.2 using as outcomes not class size variables (total class size, number of natives or number of immigrants in class) but rather the average values of the control variables included in our regression: if our identification strategy is internally valid when we focus on first generation immigrants, we should observe that covariates do not exhibit spikes nor hump shapes but evolve smoothly over native enrolment in schools or over corresponding predicted class size. Indeed, this is the pattern we observe in both Figure C.3 and Figure C.4.

- Table C.4 reports the first stage estimates that correspond to the instrumental variable estimates of Table 10 in the paper. The pattern of first stage estimates is fairly similar to the one observe in Table 5 in the paper.

- In Table C.5 we formally test if our instrumental variables are correlated with the included covariates, net of institution-grade fixed effects and trends in enrolment. We do so estimating a system of equations where the dependent variables are within-group (institution-grade) residuals of each covariate and the regressors are the corresponding within-group residuals of quadratic trend in native enrolment and our 15 instrumental variables: as one would expect given the pattern in Figure C.3 and Figure C.4, most coefficients are not statistically different from zero but we also detect significant correlations with some background variables, notably parental education and the share of missing values on control variables for quasi-natives. Quasi-natives include children born from at least one Italian parent (natives) and children born in Italy from not-Italian parents. Angrist et al. (forthcoming) clarify that the correlation we detect -as they do in their paper- between rules of class formation and observables is likely to be the result of cheating, mainly due to shirking.

C.3 IV: math sample

This Section contains additional figures and tables which could not be included in the main text due to space constraints. Here is the list with brief comments:

- Figure C.5 and Figure C.6 replicate Figure 1 and Figure 2 in the paper for the language sample where we use the more restrictive definition of immigrants and focus on first generation immigrants: the pattern is similar to the one discussed in the paper.

- In Figure C.7 and Figure C.8 we construct graphs similar to those in Figure C.5 and Figure C.6 using as outcomes not class size variables (total class size, number of natives or number of immigrants in class) but rather the average values of the control variates included in our regression: if our identification strategy is internally valid when we focus on first generation immigrants, we should observe that covariates do not exhibit spikes nor hump shapes but evolve smoothly over native enrolment in schools or over corresponding predicted class size. Indeed, this is the pattern we observe in both Figure C.7 and Figure C.8.

- Table C.6 reports the first stage estimates that correspond to the instrumental variable estimates of Table 11 in the paper. The pattern of first stage estimates is fairly similar to the one observe in Table 6 in the paper.

- In Table C.7 we formally test if our instrumental variables are correlated with the included covariates, net of institution-grade fixed effects and trends in enrolment. We do so estimating a system of equations where the dependent variables are within-group (institution-grade) residuals of each covariate and the regressors are the corresponding within-group residuals of quadratic trend in native enrolment and our 15 instrumental variables: as one would expect given the pattern in Figure C.7 and Figure C.8, most coefficients are not statistically different from zero but we also detect significant correlations with some background variables, notably parental education and the share of missing values on control variables for quasi-natives. Quasi-natives include children born from at least one Italian parent (natives) and children born in Italy from not-Italian parents. Angrist et al. (forthcoming) clarify that the correlation we detect -as they do in their paper- between rules of class formation and observables is likely to be the result of cheating, mainly due to shirking.

References

Angrist, J., Battistin, E., Vuri, D., forthcoming. In a small moment: class size and moral hazard in the Italian Mezzogiorno. American Economic Journal: Applied Economics.

Contini, D., 2013. Immigrant background peer effects in Italian schools. Social Science Research 42 (4), 1122–1142.

Ohinata, A., Van Ours, J., 2013. How Immigrant Children Affect the Academic Achievement of Native Dutch Children. Economic Journal 123 (560), F308–F331.

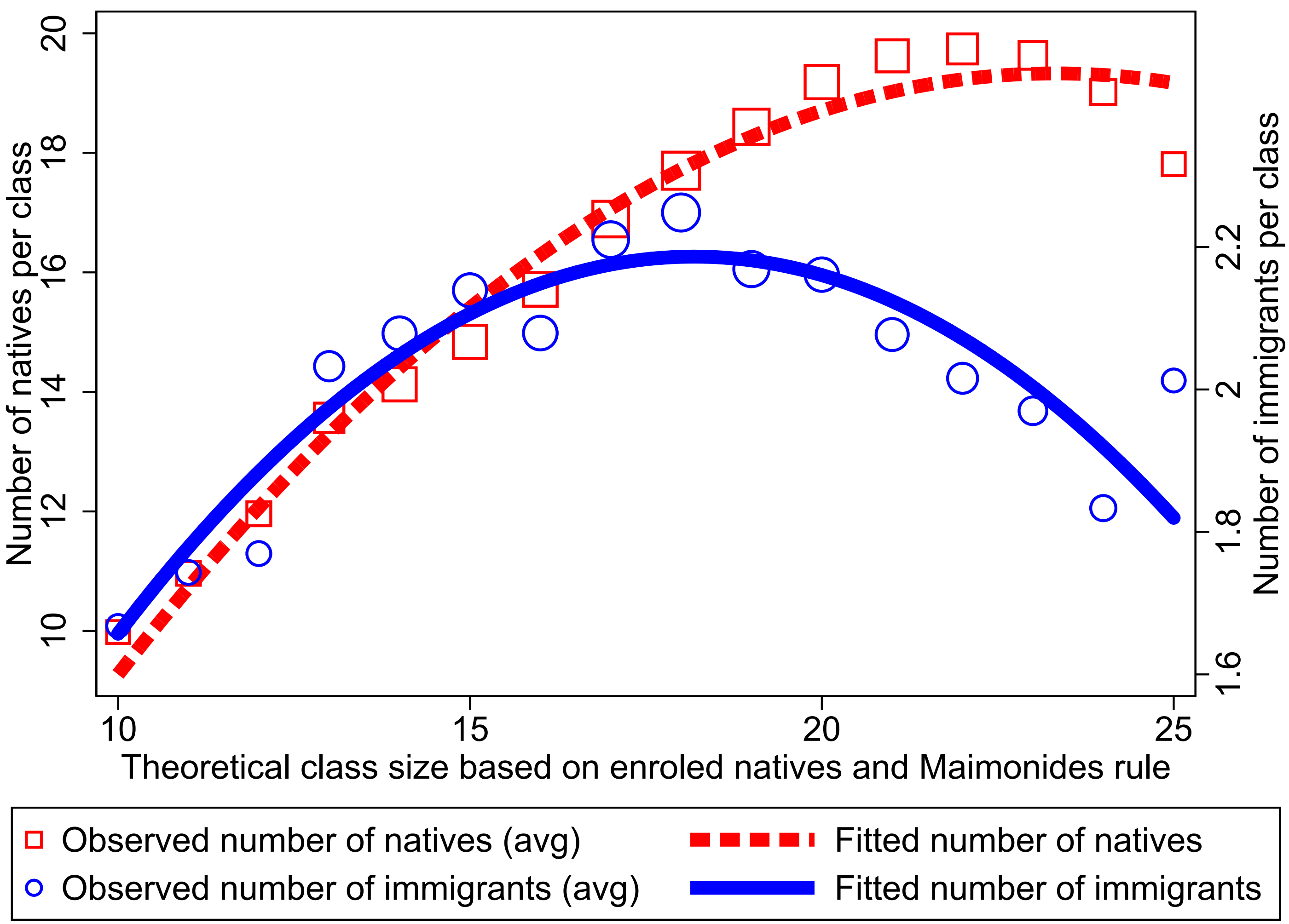
Quintano, C., Castellano, R., Longobardi, S., 2009. A Fuzzy Clustering Approach to Improve the Accuracy of Italian Student Data. An Experimental Procedure to Correct the Impact of the Outliers on Assessment Test Scores. Statistica & Applicazioni (2), 149–171.

Table B.1: Descriptive statistics for the math sample.

	2nd grade		5th grade	
	Mean	S.D.	Mean	S.D.
	Panel A. School characteristics			
Fraction of correct answers:				
language (natives)	0.67	0.11	0.71	0.08
math (natives)	0.62	0.12	0.65	0.10
language (immigrants)	0.54	0.18	0.61	0.14
math (immigrants)	0.55	0.16	0.58	0.15
Number of natives in class	16.5	3.83	16.7	3.81
Number of immigrants in class	2.07	2.04	2.06	1.99
Class size	18.6	3.92	18.7	4.01
Share (0-1) of natives in class with				
low educated father	0.44	0.20	0.47	0.20
low educated mother	0.35	0.19	0.39	0.20
employed father	0.96	0.07	0.96	0.07
employed mother	0.63	0.21	0.62	0.22
Share (0-1) of natives in class who				
attended kindergarten	0.99	0.04	0.99	0.04
are male	0.51	0.12	0.51	0.12
Cheating propensity (Quintano et al., 2009)	0.04	0.14	0.04	0.14
Cheating indicator (Angrist et al., forthcoming)	0.04	0.18	0.05	0.19
Enrolment (natives)	28.1	15.2	28.5	15.4
Enrolment (immigrants)	3.76	4.69	3.71	4.46
Average number of classes	1.68	0.77	1.68	0.78
Sample size (number of schools)	8,006		8,085	
	Panel C. Institution characteristics			
External monitored institutions	0.25	0.43	0.25	0.43
Average number of classes	4.76	1.71	4.79	1.7
Average number of schools	2.79	0.97	2.80	0.97
Sample size (number of institutions)	2,865		2,883	

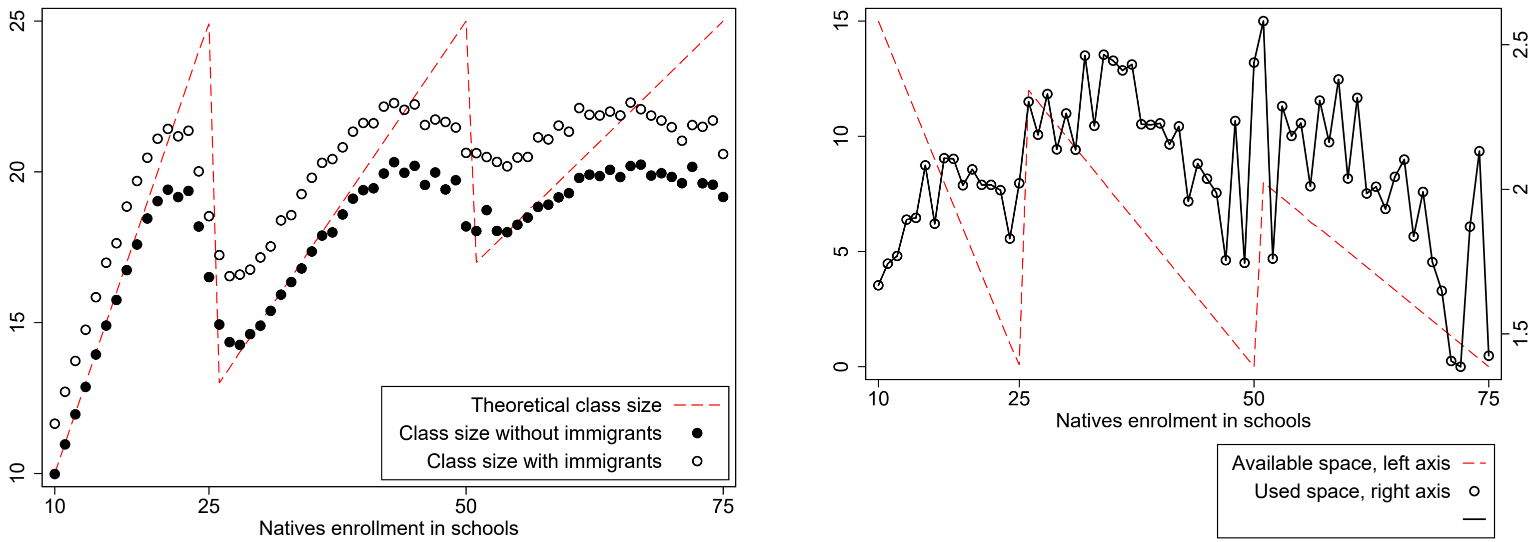
Notes: The unit of observation is a school in Panel A and an institution in Panel B. Institutions are groups of schools managed by the same principal. The family and individual background characteristics in Panel A are the school-average shares of natives in a class who have that specific characteristic over the total number of natives in the class. Missing values do not contribute to the computation of these shares. All regressions in the following tables include the school-average shares of missing values for each characteristic as an additional control. All these variables come from the school administration through the data file that we received from INVALSI, except for the cheating indicator that was computed by Angrist et al. (forthcoming) and kindly given to us by these authors.

Figure B.1: Number of natives and immigrants in a class as a function of theoretical class size based on native enrolment; math sample, 2nd and 5th grade.



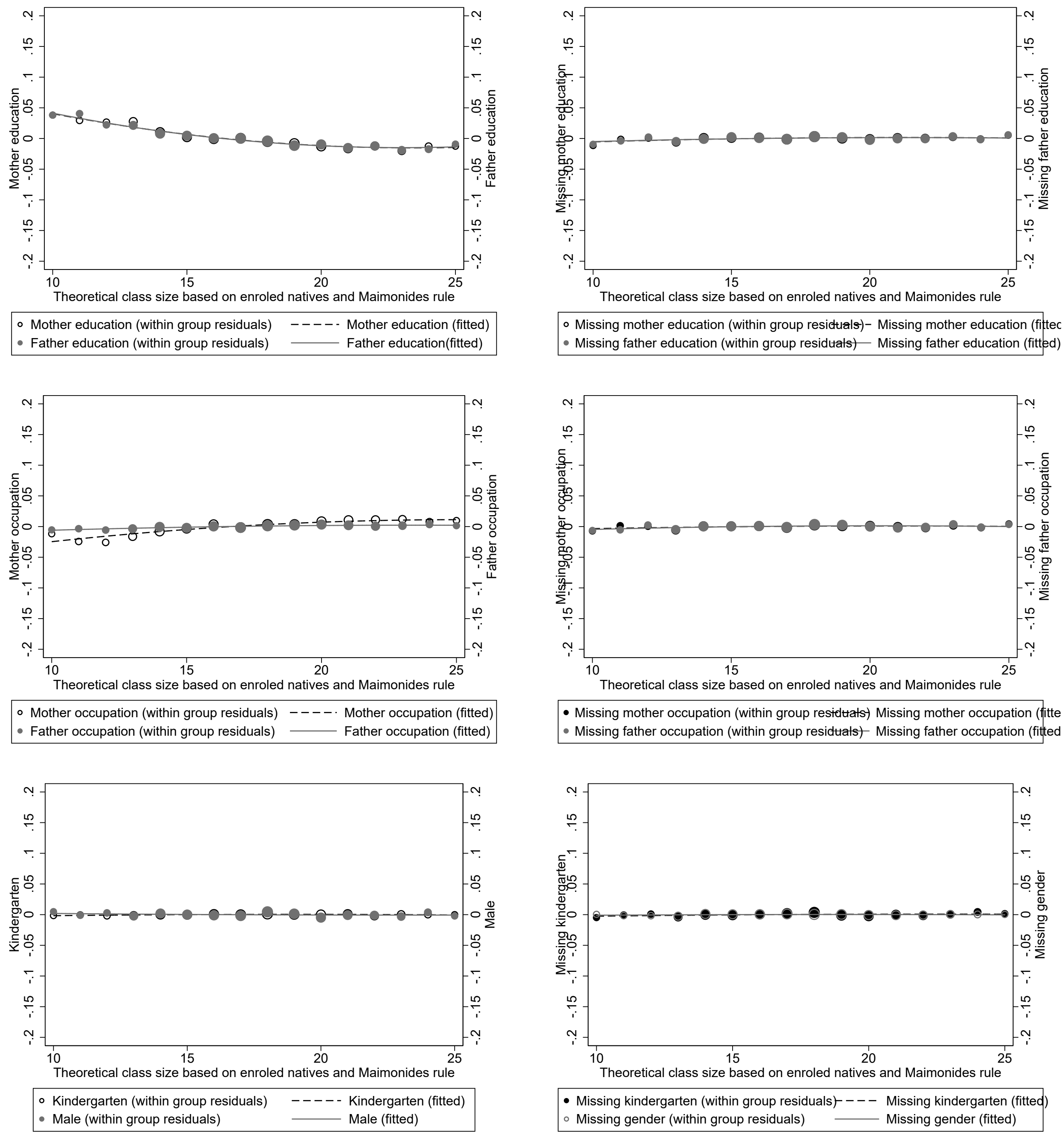
Notes: In this figure, squares (left vertical axis) indicate the average number of natives per class in schools with the correspondent theoretical class size based on enrolled natives (horizontal axis). The dashed line is a quadratic fit of these averages. Circles (right vertical axis) indicate the average number of immigrants per class in schools with the correspondent theoretical class size based on enrolled natives (horizontal axis). The continuous line is a quadratic fit of these averages. The size of squares and circles is proportional to the number of schools used to compute the averages that they represent. The quadratic fitted lines have been estimated with weights equal to the number of schools for each value of theoretical class size based on enrolled natives (horizontal axis).

Figure B.2: Number of natives and immigrants in a class and class size as a function of native enrolment; math sample, 2nd and 5th grades.



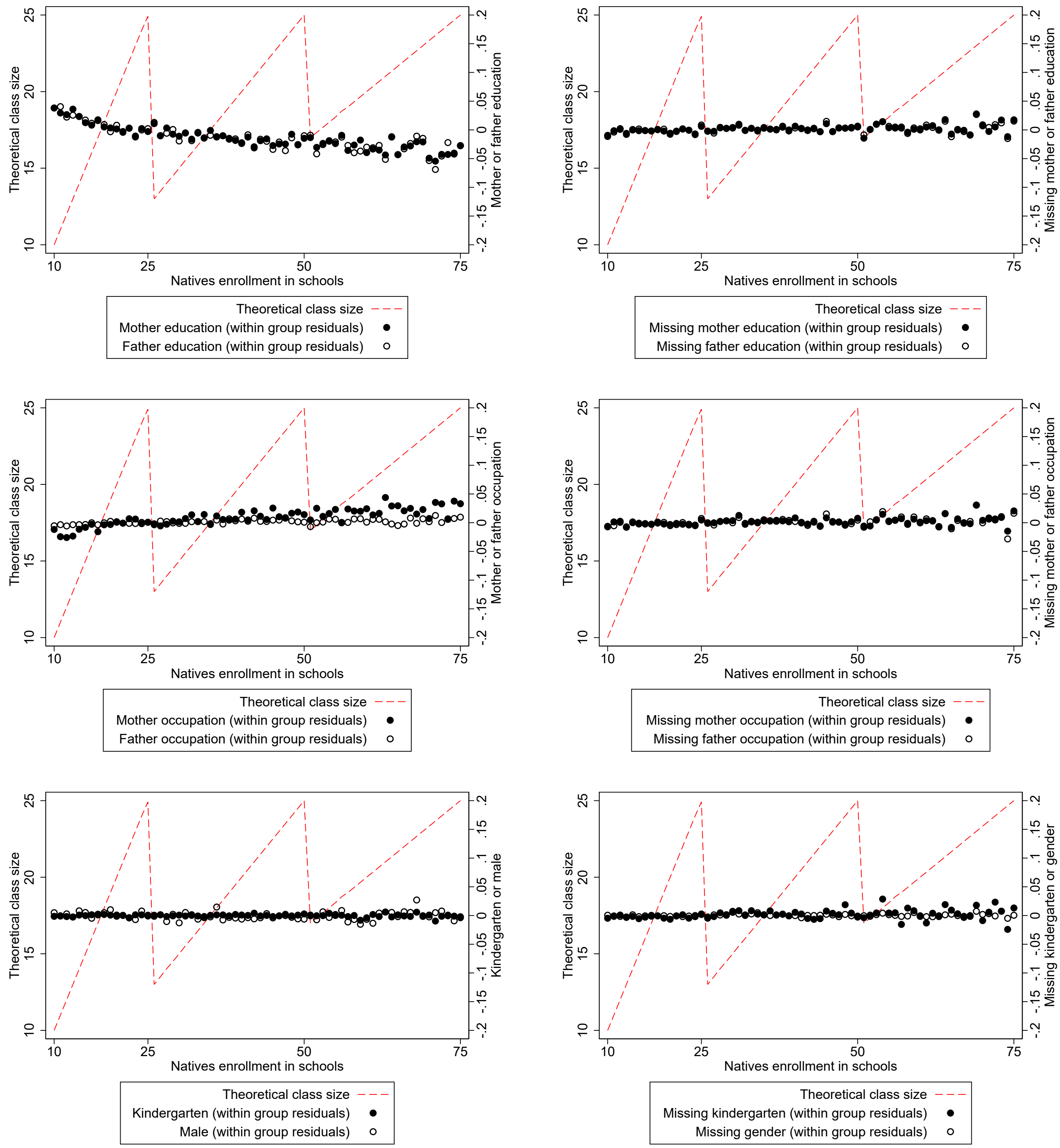
Notes: The left panels report the theoretical class size (dashed line), the class size without immigrants (dark dots) and the class size with immigrants (light dots) as a function of native enrollment in schools. In the right panels, the line connecing light dots represent the vertical distance between the light and dark dots of the left panels (the actual number of immigrants per class) as a function of native enrolment. The right panels also plot the theoretically available space for immigrants (dashed line), defined as the maximum number of students in a class (25) minus the theoretical class size based on the number of natives C_{sg}^{N}.

Figure B.3: Included covariates as a function of theoretical class size based on native enrolment; math sample, 2nd and 5th grade.



Notes: In these panels, rounds the average of the within-group (institution-grade) residual of each covariate included in the IV regression of the paper (Table 3) in schools at the correspondent theoretical class size based on enrolled natives (horizontal axis). Variable names are indicated on the left or right vertical axes according to the legend. The solid and dashed lines represent quadratic fits of these averages. The size of squares and circles is proportional to the number of schools used to compute the averages that they represent. The quadratic fitted lines have been estimated with weights equal to the number of schools for each value of theoretical class size based on enrolled natives (horizontal axis).

Figure B.4: Included covariates as a function of theoretical class size based on native enrolment; math sample, 2nd and 5th grade.



Notes: In these panels, we report theoretical class size (dashed line) and average the average of each covariate included in IV regression of the paper (Table 3) (dark or light dots) as a function of native enrollment in schools.

Table B.2: Effect of the instruments on covariates; math sample; pooled 2nd and 5th grade .

Outcomes	Education		Employed		Kindergarten	Male	Missing					
							Education		Employed		Kindergarten	Male
	Mother	Father	Mother	Father			Mother	Father	Mother	Father		
Regressors												
$1(10 \leq C_{sg}^{N} < 11)$	0.03***	0.01	-0.00	-0.00	-0.00	-0.00	-0.02***	-0.02***	-0.01*	-0.01**	-0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(11 \leq C_{sg}^{N} < 12)$	0.02*	0.02**	-0.02**	-0.00	-0.00	-0.01	-0.01	-0.01*	0.00	-0.01**	0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(12 \leq C_{sg}^{N} < 13)$	0.01	-0.00	-0.02***	-0.01	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(13 \leq C_{sg}^{N} < 14)$	0.02**	0.00	-0.02**	-0.00	-0.00	-0.00	-0.01**	-0.02***	-0.01**	-0.02***	-0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(14 \leq C_{sg}^{N} < 15)$	0.01	-0.00	-0.01	-0.00	-0.00	0.00	-0.00	-0.01	-0.00	-0.01	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(15 \leq C_{sg}^{N} < 16)$	-0.00	-0.01	-0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.01	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(16 \leq C_{sg}^{N} < 17)$	-0.01	-0.01	0.01	0.00	0.00	-0.01	-0.01	-0.01*	-0.01*	-0.01*	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(17 \leq C_{sg}^{N} < 18)$	0.00	-0.01	-0.01	-0.00	0.00	0.00	-0.01*	-0.01**	-0.01*	-0.01**	-0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(18 \leq C_{sg}^{N} < 19)$	-0.00	-0.01	0.00	0.00	0.00	0.01	-0.01	-0.01	-0.01	-0.01	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(19 \leq C_{sg}^{N} < 20)$	-0.01	-0.02**	0.00	0.00	0.00	0.00	-0.01*	-0.01*	-0.01	-0.01	-0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(20 \leq C_{sg}^{N} < 21)$	-0.01	-0.01	0.00	0.00	0.00	-0.00	-0.01**	-0.02***	-0.01	-0.01**	-0.01	-0.00**
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(21 \leq C_{sg}^{N} < 22)$	-0.01*	-0.02**	0.01	0.00	0.00	0.00	-0.01*	-0.01**	-0.01**	-0.02***	-0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(22 \leq C_{sg}^{N} < 23)$	-0.01	-0.02*	0.01	-0.00	-0.00	0.00	-0.01	-0.01*	-0.01**	-0.01	-0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(23 \leq C_{sg}^{N} < 24)$	-0.02**	-0.03***	0.00	0.00	0.00	-0.00	-0.01	-0.01*	-0.01	-0.01*	-0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(24 \leq C_{sg}^{N} < 25)$	-0.00	-0.01	0.00	0.00	0.00	0.01	-0.01**	-0.02***	-0.02**	-0.02***	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Observations							16,091					
Joint significance of instruments (p-value)	0.0002	0.0005	0.0071	0.1536	0.4155	0.7545	0.2424	0.1656	0.2791	0.2610	0.3900	0.3575

Notes: The table reports in each column the estimates of a system of equations (Seemingly Unrelated Regression Estimates) with one equation for each control variable included in the OLS and IV regression in Table 4 in the paper. The unit of observation is a school. The dependent variable in each column is the within-group (institution-grade) residual of the covariate indicated in the heading of each column, i.e. the observed value of the variable in the school minus the institution-grade average of the same variable. The controls are include the following set of within-group (institution-grade) residuals of family and individual covariates: the share of natives with mothers and fathers who attended, at most, a lower secondary school, the shares of natives with employed mothers and fathers, the share of natives who attended kindergarten and the share of male natives in class as well as the share of native students who report missing values in each of these variables. The instruments are a set of 15 dummies, one for each level of the theoretical number of natives in a class predicted by equation (**??**) according to the rules of class formation as a function of native enrolment at the school×grade level. The omitted category corresponds to a number of natives in a class equal to 25. There are no school-grades in which the number of natives in a class is less than 10. All regressions include the within-group residuals of a 2nd order polynomial of natives enrolment at the school×grade level. Robust standard errors clustered at the institution-grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%. The table also report the p-value for the test of joint significance of the instruments equation by equation.

Table B.3: First Stage for the number natives and immigrants (first and second generation); exact identification; language sample.

	Pooled 2nd & 5th grades			2nd grade			5th grade		
	Two endogenous		One endogenous	Two endogenous		One endogenous	Two endogenous		One endogenous
	N	I	I	N	I	I	N	I	I
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Predicted class size	0.52***	-0.02***	0.07***	0.50***	-0.02**	0.08***	0.54***	-0.02***	0.06***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Medium predicted class size	1.07***	0.14***		1.12***	0.10*		1.02***	0.18***	
	(0.06)	(0.04)		(0.09)	(0.05)		(0.08)	(0.05)	
Institution×grade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Polynomial in natives enrolment	✓	✓	✓	✓	✓	✓	✓	✓	
School level controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	16,100	16,100	16,100	8,014	8,014	8,014	8,086	8,086	8,086
F stat	1437.36	13.09	115.53	699.80	4.75	75.42	736.20	8.74	42.44
SW F stat	97.30	21.21	115.53	19.78	6.68	75.42	105.13	15.29	42.44
SW χ^2 p-value	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00

Notes: The table reports in each column a different first stage regression correspondent to the IV estimates of Table 8 in the paper. In the case of one endogenous variable, we use as instruments the theoretical number of natives in a class predicted by equation (2) in the paper according to the rules of class formation as a function of native enrolment at the school×grade level. In the case of two endogenous variables, we add to the predicted class size a dummy variable that takes the value one if the predicted class size falls in the medium range (i.e. between the median -17.5 native students - and the 75th percentile -20.5 native students-) and zero otherwise. All regressions include a 2nd order polynomial of natives enrolment at the school×grade level. The controls are aggregated at the school level and include the following set of family and individual covariates: the share of natives with mothers and fathers who attended, at most, a lower secondary school, the shares of natives with employed mothers and fathers, the share of natives who attended kindergarten and the share of male natives in the school. All regressions include also the share of native students who report missing values in each of these variables as well as institution×grade fixed effects. Robust standard errors clustered at the institution-grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%. The table reports also: i) i) the value of the F test of null hypothesis that the coefficients of the instruments are all zero in each first stage equation; and ii) the Sanderson-Windemeier first stage F statistic of each individual endogenous regressor (in the case of a model with one endogenous regressor this coincides with the F-test on excluded instruments) to test for weak identification ; and iii) the p-value of Sanderson-Windemeier χ^2 statistic of each individual endogenous regressor to test for under-identification.

Table B.4: First Stage for the number natives and immigrants (first and second generation); exact identification; math sample.

	Pooled 2nd & 5th grades			2nd grade			5th grade		
	Two endogenous		One endogenous	Two endogenous		One endogenous	Two endogenous		One endogenous
	N	I	I	N	I	I	N	I	I
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Predicted class size	0.52***	-0.02***	0.07***	0.50***	-0.02**	0.08***	0.54***	-0.02***	0.06***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Medium predicted class size	1.07***	0.14***		1.12***	0.10*		1.03***	0.18***	
	(0.06)	(0.04)		(0.09)	(0.05)		(0.08)	(0.05)	
Institution×grade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Polynomial in natives enrolment	✓	✓	✓	✓	✓	✓	✓	✓	
School level controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	16,091	16,091	16,091	8,006	8,006	8,006	8,085	8,085	8,085
F stat	1438.46	13.06	115.88	699.91	4.75	75.21	737.75	8.77	42.79
SW F stat	95.97	21.09	115.88	18.79	6.55	75.21	105.29	15.33	42.79
SW χ^2 p-value	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00

Notes: The table reports in each column a different first stage regression correspondent to the IV estimates of Table 9 in the paper. In the case of one endogenous variable, we use as instruments the theoretical number of natives in a class predicted by equation (2) according to the rules of class formation as a function of native enrolment at the school×grade level. In the case of two endogenous variables, we add to the predicted class size a dummy variable that takes the value one if the predicted class size falls in the medium range (i.e. between the median -17.5 native students - and the 75th percentile -20.5 native students-) and zero otherwise. All regressions include a 2nd order polynomial of natives enrolment at the school×grade level. The controls are aggregated at the school level and include the following set of family and individual covariates: the share of natives with mothers and fathers who attended, at most, a lower secondary school, the shares of natives with employed mothers and fathers, the share of natives who attended kindergarten and the share of male natives in the school. All regressions include also the share of native students who report missing values in each of these variables as well as institution×grade fixed effects. obust standard errors clustered at the institution-grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%. The table reports also: i) i) the value of the F test of null hypothesis that the coefficients of the instruments are all zero in each first stage equation; and ii) the Sanderson-Windemeier first stage F statistic of each individual endogenous regressor (in the case of a model with one endogenous regressor this coincides with the F-test on excluded instruments) to test for weak identification ; and iii) the p-value of Sanderson-Windemeier χ^2 statistic of each individual endogenous regressor to test for under-identification.

Table B.5: Robustness to cheating: first stage for the number of natives and immigrants (first and second generation); language and math samples; pooled 2nd and 5th grades

	Language			Math		
	Two endogenous		One endogenous	Two endogenous		One endogenous
	N	I	I	N	I	I
	(1)	(2)	(3)	(4)	(5)	(6)
$1(10 \leq C_{sg}^N < 11)$	-6.14***	-0.13	-1.36***	-6.08***	-0.11	-1.35***
	(0.31)	(0.15)	(0.16)	(0.31)	(0.15)	(0.16)
$1(11 \leq C_{sg}^N < 12)$	-5.32***	-0.11	-1.17***	-5.22***	-0.08	-1.14***
	(0.30)	(0.15)	(0.15)	(0.31)	(0.15)	(0.15)
$1(12 \leq C_{sg}^N < 13)$	-4.08***	-0.13	-0.94***	-3.98***	-0.14	-0.95***
	(0.30)	(0.15)	(0.15)	(0.30)	(0.15)	(0.15)
$1(13 \leq C_{sg}^N < 14)$	-2.95***	-0.01	-0.60***	-2.87***	0.02	-0.56***
	(0.30)	(0.13)	(0.12)	(0.30)	(0.13)	(0.12)
$1(14 \leq C_{sg}^N < 15)$	-2.47***	-0.07	-0.56***	-2.41***	-0.03	-0.51***
	(0.29)	(0.13)	(0.12)	(0.30)	(0.13)	(0.12)
$1(15 \leq C_{sg}^N < 16)$	-1.74***	-0.11	-0.46***	-1.61***	-0.08	-0.40***
	(0.29)	(0.13)	(0.12)	(0.30)	(0.13)	(0.12)
$1(16 \leq C_{sg}^N < 17)$	-0.91***	-0.05	-0.23**	-0.80***	-0.01	-0.18
	(0.29)	(0.12)	(0.12)	(0.29)	(0.12)	(0.12)
$1(17 \leq C_{sg}^N < 18)$	0.12	0.01	0.04	0.18	0.06	0.09
	(0.29)	(0.12)	(0.11)	(0.29)	(0.12)	(0.11)
$1(18 \leq C_{sg}^N < 19)$	0.80***	0.02	0.18*	0.89***	0.05	0.23**
	(0.29)	(0.12)	(0.11)	(0.29)	(0.12)	(0.11)
$1(19 \leq C_{sg}^N < 20)$	1.44***	-0.03	0.25**	1.55***	-0.00	0.31***
	(0.29)	(0.12)	(0.11)	(0.29)	(0.12)	(0.11)
$1(20 \leq C_{sg}^N < 21)$	2.13***	-0.15	0.27**	2.21***	-0.11	0.34***
	(0.29)	(0.12)	(0.11)	(0.30)	(0.12)	(0.11)
$1(21 \leq C_{sg}^N < 22)$	2.56***	-0.21*	0.30***	2.70***	-0.22*	0.33***
	(0.30)	(0.12)	(0.11)	(0.30)	(0.12)	(0.11)
$1(22 \leq C_{sg}^N < 23)$	2.37***	-0.29**	0.18	2.46***	-0.26**	0.24**
	(0.32)	(0.13)	(0.12)	(0.32)	(0.13)	(0.12)
$1(23 \leq C_{sg}^N < 24)$	2.27***	-0.31**	0.14	2.39***	-0.28**	0.20*
	(0.33)	(0.13)	(0.12)	(0.34)	(0.13)	(0.12)
$1(24 \leq C_{sg}^N < 25)$	1.62***	-0.37***	-0.05	1.71***	-0.35***	-0.01
	(0.34)	(0.13)	(0.12)	(0.34)	(0.14)	(0.12)
Institution×grade FE	✓	✓	✓	✓	✓	✓
Polynomial in natives enrolment	✓	✓	✓	✓	✓	
School level controls	✓	✓	✓	✓	✓	✓
Observations	14,888	14,888	14,888	14,798	14,798	14,798
F stat	352.81	2.82	17.47	359.91	3.09	17.79
SW F stat	22.99	2.71	17.47	25.02	2.99	17.79
SW χ^2 p-value	0.00	0.00	0.00	0.00	0.00	0.00

Notes: The table reports in each column a different first stage regression correspondent to the IV estimates in Table 12 in the paper for the language and math sub-samples in which for all classes in the school the cheating indicator computed by citeABV2017 signals no cheating. Institutions for which we do not have at least two schools that meet this criteria are also dropped. The unit of observation is a school. The dependent variable is the average number of natives (immigrants) per class in the school. The instruments are a set of 15 dummies, one for each level of the theoretical number of natives in a class predicted by equation (2) in the paper according to the rules of class formation as a function of native enrolment at the school×grade level. The omitted category corresponds to a number of natives in a class equal to 25. There are no schools in which the number of natives in a class is less than 10 in both grades. All regressions include a 2nd order polynomial of natives enrolment at the school×grade level. The controls are aggregated at the school level and include the following set of family and individual covariates: the share of natives with mothers and fathers who attended, at most, a lower secondary school, the shares of natives with employed mothers and fathers, the share of natives who attended kindergarten (and/or nursery) and the share of male natives in the class. All regressions include also the share of native students who report missing values in each of these variables as well as institution×grade fixed effects. Robust standard errors clustered at the institution×grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%.The table reports also: i) i) the value of the F test of null hypothesis that the coefficients of the instruments are all zero in each first stage equation; and ii) the Sanderson-Windemeier first stage F statistic of each individual endogenous regressor (in the case of a model with one endogenous regressor this coincides with the F-test on excluded instruments) to test for weak identification ; and iii) the p-value of Sanderson-Windemeier χ^2 statistic of each individual endogenous regressor to test for under-identification.

Table B.6: IV-FE estimates of the effect of the number of natives and immigrants (first and second generation) on language test scores of natives; language sample.

	Pooled 2nd & 5th		2nd grade		5th grade	
	Two endogenous	One endogenous	Two endogenous	One endogenous	Two endogenous	One endogenous
	(1)	(2)	(3)	(4)	(5)	(6)
Number of natives: $\hat{\beta}$	-0.0014***	-0.0014***	-0.0017**	-0.0016**	-0.0010*	-0.0010*
	(0.0005)	(0.0005)	(0.0008)	(0.0008)	(0.0006)	(0.0006)
Number of immigrants: $\hat{\gamma}$	-0.0189**	-0.0113***	-0.0170	-0.0106**	-0.0129*	-0.0109***
	(0.0080)	(0.0029)	(0.0134)	(0.0044)	(0.0077)	(0.0038)
Pure Ethnic Composition effect: $\hat{\delta}$	-0.0175**	-0.0099***	-0.0153	-0.0090*	-0.0119	-0.0099***
	(0.0078)	(0.0026)	(0.0129)	(0.0037)	(0.0075)	(0.0034)
Observations	16,100	16,100	8,014	8,014	8,086	8,086
Institution×grade FE	✓	✓	✓	✓	✓	✓
Polynomial in natives enrolment	✓	✓	✓	✓	✓	✓
School level controls	NO	NO	NO	NO	NO	NO
Hansen (p-value)	0.8477	0.7721	0.9106	0.9074	0.5328	0.5860
F stat (natives)	396.2750		190.4860		213.3212	
SW F stat (natives)	16.8477		3.8778		47.2156	
SW χ^2 p-value (natives)	0.00		0.00		0.00	
F stat (immigrants)	3.0590	18.0421	1.6401	9.6810	2.2354	8.9252
SW F stat (immigrants)	2.8152	18.0421	1.2305	9.6810	2.3301	8.9252
SW χ^2 p-value (immigrants)	0.00	0.00	0.29	0.00	0.00	0.00

Notes: The table reports in each column a different set of estimates for the language samples. The unit of observation is a school. The dependent variable is the average test scores in language for natives students (i.e. fraction of correct answers). The instruments are a set of 15 dummies, one for each level of the theoretical number of natives in a class predicted by equation (2) in the paper according to the rules of class formation as a function of native enrolment at the school×grade level. The omitted category corresponds to a number of natives in a class equal to 25. There are no school-grades in which the number of natives in a class is less than 10. All regressions include a 2nd order polynomial of natives enrolment at the school×grade level as well as institution×grade fixed effects. Robust standard errors clustered at the institution-grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%. The table reports also: i) the p-value of the Hansen test; ii) the value of the F test of null hypothesis that the coefficients of the instruments are all zero in each first stage equation; and iii) the Sanderson-Windemeier first stage F statistic of each individual endogenous regressor (in the case of a model with one endogenous regressor this coincides with the F-test on excluded instruments) to test for weak identification ; and iv) the p-value of Sanderson-Windemeier χ^2 statistic of each individual endogenous regressor to test for under-identification.

Table B.7: IV-FE estimates of the effect of the number of natives and immigrants (first and second generation) on math test scores of natives; math sample.

	Pooled 2nd & 5th		2nd grade		5th grade	
	Two	One	Two	One	Two	One
	endogenous		endogenous		endogenous	
	(1)	(2)	(3)	(4)	(5)	(6)
Number of natives: $\hat{\beta}$	-0.0015***	-0.0014***	-0.0019**	-0.0017**	-0.0010	-0.0010
	(0.0006)	(0.0006)	(0.0009)	(0.0009)	(0.0007)	(0.0007)
Number of immigrants: $\hat{\gamma}$	-0.0203**	-0.0122***	-0.0197	-0.0112**	-0.0111	-0.0121**
	(0.0091)	(0.0034)	(0.0142)	(0.0048)	(0.0093)	(0.0047)
Pure Ethnic Composition effect: $\hat{\delta}$	-0.0188**	-0.0108**	-0.0178	-0.0095**	-0.0101	-0.0111***
	(0.0088)	(0.0029)	(0.0137)	(0.0040)	(0.0092)	(0.0042)
Observations	16,091	16,091	8,006	8,006	8,085	8,085
Institution×grade FE	✓	✓	✓	✓	✓	✓
Polynomial in natives enrolment	✓	✓	✓	✓	✓	✓
School level controls	NO	NO	NO	NO	NO	NO
Hansen (p-value)	0.9615	0.9401	0.9671	0.9650	0.9336	0.9580
F stat (natives)	400.7382		194.3786		213.3349	
SW F stat (natives)	15.9277		3.7828		44.8846	
SW χ^2 p-value (natives)	0.00		0.00		0.00	
F stat (immigrants)	3.0359	17.8185	1.6533	9.6188	2.2105	8.7921
SW F stat (immigrants)	2.7741	17.8185	1.2243	9.6188	2.3010	8.7921
SW χ^2 p-value (immigrants)	0.00	0.00	0.29	0.00	0.00	0.00

Notes: The table reports in each column a different set of estimates for the math samples. The unit of observation is a school. The dependent variable is the average test scores in language for natives students (i.e. fraction of correct answers). The instruments are a set of 15 dummies, one for each level of the theoretical number of natives in a class predicted by equation (2) in the paper according to the rules of class formation as a function of native enrolment at the school×grade level. The omitted category corresponds to a number of natives in a class equal to 25. There are no school-grades in which the number of natives in a class is less than 10. All regressions include a 2nd order polynomial of natives enrolment at the school×grade level as well as institution×grade fixed effects. Robust standard errors clustered at the institution-grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%. The table reports also: i) the p-value of the Hansen test; ii) the value of the F test of null hypothesis that the coefficients of the instruments are all zero in each first stage equation; and iii) the Sanderson-Windemeier first stage F statistic of each individual endogenous regressor (in the case of a model with one endogenous regressor this coincides with the F-test on excluded instruments) to test for weak identification ; and iv) the p-value of Sanderson-Windemeier χ^2 statistic of each individual endogenous regressor to test for under-identification.

Table B.8: First Stage for the number natives and immigrants (first and second generation); language sample.

	Pooled 2nd & 5th grades			2nd grade			5th grade		
	Two endogenous		One endogenous	Two endogenous		One endogenous	Two endogenous		One endogenous
	N	I	I	N	I	I	N	I	I
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$1(10 \leq C_{sg}^{N} < 11)$	-6.06***	-0.11	-1.31***	-5.66***	-0.12	-1.37***	-6.47***	-0.11	-1.23***
	(0.29)	(0.14)	(0.15)	(0.41)	(0.21)	(0.21)	(0.42)	(0.20)	(0.21)
$1(11 \leq C_{sg}^{N} < 12)$	-5.16***	-0.11	-1.13***	-4.76***	-0.26	-1.31***	-5.59***	0.05	-0.92***
	(0.29)	(0.14)	(0.14)	(0.41)	(0.20)	(0.19)	(0.41)	(0.18)	(0.19)
$1(12 \leq C_{sg}^{N} < 13)$	-3.97***	-0.12	-0.90***	-3.60***	-0.21	-1.01***	-4.35***	-0.01	-0.76***
	(0.29)	(0.14)	(0.14)	(0.40)	(0.22)	(0.21)	(0.41)	(0.18)	(0.18)
$1(13 \leq C_{sg}^{N} < 14)$	-2.77***	-0.01	-0.56***	-2.42***	-0.14	-0.68***	-3.14***	0.13	-0.42***
	(0.29)	(0.12)	(0.11)	(0.41)	(0.18)	(0.16)	(0.41)	(0.16)	(0.16)
$1(14 \leq C_{sg}^{N} < 15)$	-2.33***	-0.05	-0.51***	-2.06***	-0.16	-0.61***	-2.62***	0.06	-0.40**
	(0.28)	(0.12)	(0.11)	(0.39)	(0.18)	(0.16)	(0.40)	(0.15)	(0.15)
$1(15 \leq C_{sg}^{N} < 16)$	-1.58***	-0.09	-0.40***	-1.32***	-0.14	-0.43***	-1.83***	-0.06	-0.37**
	(0.28)	(0.12)	(0.11)	(0.39)	(0.18)	(0.16)	(0.40)	(0.15)	(0.15)
$1(16 \leq C_{sg}^{N} < 17)$	-0.72***	-0.07	-0.21*	-0.52	-0.22	-0.34**	-0.95**	0.10	-0.07
	(0.28)	(0.12)	(0.11)	(0.39)	(0.18)	(0.16)	(0.40)	(0.15)	(0.15)
$1(17 \leq C_{sg}^{N} < 18)$	0.29	0.01	0.07	0.58	-0.19	-0.07	-0.02	0.22	0.22
	(0.28)	(0.11)	(0.10)	(0.39)	(0.17)	(0.15)	(0.39)	(0.15)	(0.14)
$1(18 \leq C_{sg}^{N} < 19)$	0.97***	0.01	0.20*	1.19***	-0.15	0.11	0.72*	0.18	0.31**
	(0.28)	(0.11)	(0.10)	(0.39)	(0.17)	(0.15)	(0.40)	(0.15)	(0.14)
$1(19 \leq C_{sg}^{N} < 20)$	1.65***	-0.06	0.26**	1.93***	-0.21	0.21	1.36***	0.09	0.33**
	(0.28)	(0.11)	(0.10)	(0.39)	(0.17)	(0.15)	(0.39)	(0.15)	(0.14)
$1(20 \leq C_{sg}^{N} < 21)$	2.29***	-0.14	0.31***	2.64***	-0.31*	0.27*	1.93***	0.04	0.37**
	(0.28)	(0.11)	(0.11)	(0.40)	(0.17)	(0.15)	(0.40)	(0.15)	(0.15)
$1(21 \leq C_{sg}^{N} < 22)$	2.77***	-0.24**	0.31***	3.02***	-0.42**	0.25	2.51***	-0.05	0.38**
	(0.29)	(0.12)	(0.11)	(0.41)	(0.17)	(0.16)	(0.41)	(0.15)	(0.15)
$1(22 \leq C_{sg}^{N} < 23)$	2.58***	-0.30**	0.21*	3.04***	-0.46**	0.21	2.11***	-0.13	0.23
	(0.30)	(0.12)	(0.11)	(0.43)	(0.18)	(0.16)	(0.43)	(0.16)	(0.15)
$1(23 \leq C_{sg}^{N} < 24)$	2.50***	-0.33***	0.17	2.36***	-0.35*	0.17	2.63***	-0.30*	0.16
	(0.32)	(0.12)	(0.11)	(0.45)	(0.18)	(0.16)	(0.44)	(0.17)	(0.16)
$1(24 \leq C_{sg}^{N} < 25)$	1.74***	-0.39***	-0.04	2.01***	-0.57***	-0.12	1.45***	-0.19	0.06
	(0.33)	(0.13)	(0.12)	(0.46)	(0.19)	(0.16)	(0.47)	(0.17)	(0.17)
Institution×grade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Polynomial in natives enrolment	✓	✓	✓	✓	✓	✓	✓	✓	
School level controls	NO	NO	NO	NO	NO	NO	NO	NO	NO
Observations	16,100	16,100	16,100	8,014	8,014	8,014	8,086	8,086	8,086
F stat	396.27	3.06	18.04	190.49	1.64	9.68	213.32	2.24	8.93
SW F stat	16.85	2.82	18.04	3.88	1.23	9.68	47.22	2.33	8.93
SW χ^2 p-value	0.00	0.00	0.00	0.00	0.30	0.00	0.00	0.00	0.00

Notes: The table reports in each column a different first stage regression correspondent to the IV estimates of Table B.6. The unit of observation is a school. The dependent variable is the average number of natives (immigrants) per a class in a school. The instruments are a set of 15 dummies, one for each level of the theoretical number of natives in a class predicted by equation (2) according to the rules of class formation as a function of native enrolment at the school×grade level. The omitted category corresponds to a number of natives in a class equal to 25. There are no school-grades in which the number of natives in a class is less than 10. All regressions include a 2nd order polynomial of natives enrolment at the school×grade leve as well as institution×grade fixed effects. Robust standard errors clustered at the institution-grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%. The table reports also: i) the value of the F test of null hypothesis that the coefficients of the instruments are all zero in each first stage equation; and ii) the Sanderson-Windemeier first stage F statistic of each individual endogenous regressor (in the case of a model with one endogenous regressor this coincides with the F-test on excluded instruments) to test for weak identification ; and iii) the p-value of Sanderson-Windemeier χ^2 statistic of each individual endogenous regressor to test for under-identification.

Table B.9: First Stage for the number natives and immigrants (first and second generation); math sample.

	Pooled 2nd & 5th grades			2nd grade			5th grade		
	Two endogenous		One endogenous	Two endogenous		One endogenous	Two endogenous		One endogenous
	N	I	I	N	I	I	N	I	I
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$1(10 \leq C_{sg}^N < 11)$	-5.98***	-0.11	-1.29***	-5.60***	-0.12	-1.36***	-6.39***	-0.10	-1.20***
	(0.29)	(0.14)	(0.15)	(0.41)	(0.21)	(0.21)	(0.42)	(0.20)	(0.21)
$1(11 \leq C_{sg}^N < 12)$	-5.08***	-0.11	-1.12***	-4.68***	-0.27	-1.30***	-5.50***	0.05	-0.90***
	(0.29)	(0.14)	(0.14)	(0.41)	(0.20)	(0.19)	(0.42)	(0.18)	(0.19)
$1(12 \leq C_{sg}^N < 13)$	-3.88***	-0.12	-0.89***	-3.51***	-0.23	-1.00***	-4.26***	-0.00	-0.74***
	(0.29)	(0.14)	(0.14)	(0.40)	(0.22)	(0.20)	(0.42)	(0.18)	(0.18)
$1(13 \leq C_{sg}^N < 14)$	-2.69***	-0.02	-0.55***	-2.36***	-0.15	-0.67***	-3.04***	0.13	-0.40**
	(0.29)	(0.12)	(0.11)	(0.41)	(0.18)	(0.16)	(0.42)	(0.16)	(0.16)
$1(14 \leq C_{sg}^N < 15)$	-2.24***	-0.06	-0.50***	-1.97***	-0.17	-0.60***	-2.53***	0.06	-0.38**
	(0.28)	(0.12)	(0.11)	(0.39)	(0.18)	(0.16)	(0.41)	(0.15)	(0.15)
$1(15 \leq C_{sg}^N < 16)$	-1.48***	-0.10	-0.39***	-1.21***	-0.15	-0.42***	-1.74***	-0.05	-0.35**
	(0.28)	(0.12)	(0.11)	(0.39)	(0.18)	(0.16)	(0.41)	(0.15)	(0.15)
$1(16 \leq C_{sg}^N < 17)$	-0.63**	-0.07	-0.19*	-0.42	-0.23	-0.32**	-0.86**	0.10	-0.05
	(0.28)	(0.12)	(0.11)	(0.39)	(0.18)	(0.16)	(0.40)	(0.15)	(0.15)
$1(17 \leq C_{sg}^N < 18)$	0.38	0.00	0.08	0.68*	-0.21	-0.06	0.07	0.22	0.23
	(0.28)	(0.11)	(0.10)	(0.38)	(0.17)	(0.15)	(0.40)	(0.15)	(0.14)
$1(18 \leq C_{sg}^N < 19)$	1.07***	0.01	0.22**	1.28***	-0.16	0.12	0.83**	0.19	0.33**
	(0.28)	(0.11)	(0.10)	(0.39)	(0.17)	(0.15)	(0.40)	(0.15)	(0.15)
$1(19 \leq C_{sg}^N < 20)$	1.73***	-0.07	0.27***	2.00***	-0.23	0.22	1.45***	0.10	0.34**
	(0.28)	(0.11)	(0.10)	(0.39)	(0.17)	(0.15)	(0.40)	(0.15)	(0.14)
$1(20 \leq C_{sg}^N < 21)$	2.37***	-0.15	0.32***	2.72***	-0.33*	0.27*	2.02***	0.04	0.39***
	(0.28)	(0.11)	(0.11)	(0.40)	(0.17)	(0.15)	(0.40)	(0.15)	(0.15)
$1(21 \leq C_{sg}^N < 22)$	2.86***	-0.25**	0.32***	3.11***	-0.43**	0.25	2.59***	-0.06	0.39***
	(0.29)	(0.12)	(0.11)	(0.41)	(0.17)	(0.16)	(0.41)	(0.15)	(0.15)
$1(22 \leq C_{sg}^N < 23)$	2.67***	-0.30**	0.23**	3.12***	-0.46**	0.23	2.19***	-0.13	0.25
	(0.31)	(0.12)	(0.11)	(0.43)	(0.18)	(0.16)	(0.44)	(0.16)	(0.15)
$1(23 \leq C_{sg}^N < 24)$	2.60***	-0.33***	0.18	2.47***	-0.36**	0.18	2.72***	-0.29*	0.18
	(0.32)	(0.12)	(0.11)	(0.45)	(0.18)	(0.16)	(0.45)	(0.17)	(0.16)
$1(24 \leq C_{sg}^N < 25)$	1.83***	-0.39***	-0.03	2.09***	-0.57***	-0.11	1.54***	-0.18	0.08
	(0.33)	(0.13)	(0.12)	(0.46)	(0.19)	(0.16)	(0.47)	(0.17)	(0.17)
Institution×grade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Polynomial in natives enrolment	✓	✓	✓	✓	✓	✓	✓	✓	
School level controls	NO	NO	NO	NO	NO	NO	NO	NO	NO
Observations	16,091	16,091	16,091	8,006	8,006	8,006	8,085	8,085	8,085
F stat	400.74	3.04	17.82	194.38	1.65	9.62	213.33	2.21	8.79
SW F stat	15.93	2.77	17.82	3.78	1.22	9.62	44.88	2.30	8.79
SW χ^2 p-value	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.00

Notes: The table reports in each column a different first stage regression correspondent to the IV estimates of Table B.7. The unit of observation is a school. The dependent variable is the average number of natives (immigrants) per a class in a school. The instruments are a set of 15 dummies, one for each level of the theoretical number of natives in a class predicted by equation (2) according to the rules of class formation as a function of native enrolment at the school×grade level. The omitted category corresponds to a number of natives in a class equal to 25. There are no school-grades in which the number of natives in a class is less than 10. All regressions include a 2nd order polynomial of natives enrolment at the school×grade leve as well as institution×grade fixed effects. Robust standard errors clustered at the institution-grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%. The table reports also: i) the value of the F test of null hypothesis that the coefficients of the instruments are all zero in each first stage equation; and ii) the Sanderson-Windemeier first stage F statistic of each individual endogenous regressor (in the case of a model with one endogenous regressor this coincides with the F-test on excluded instruments) to test for weak identification ; and iii) the p-value of Sanderson-Windemeier χ^2 statistic of each individual endogenous regressor to test for under-identification.

Table B.10: Descriptive statistics for the language sample.

	2nd grade		5th grade	
	Mean	S.D.	Mean	S.D.
	Panel A. Class characteristics			
Fraction of correct answers:				
language (natives)	0.67	0.11	0.71	0.08
math (natives)	0.62	0.12	0.65	0.10
language (immigrants)	0.53	0.17	0.60	0.15
math (immigrants)	0.55	0.16	0.58	0.15
Number of natives in class	16.7	3.97	16.9	3.95
Number of immigrants in class	2.23	2.30	2.21	2.25
Class size	19	3.86	19.1	3.95
Share (0-1) of natives in class with				
low educated mother	0.34	0.20	0.38	0.21
low educated father	0.43	0.21	0.46	0.21
employed mother	0.64	0.23	0.62	0.23
employed father	0.96	0.08	0.96	0.07
Share (0-1) of natives in class who				
attended kindergarten	0.99	0.04	0.99	0.04
are male	0.51	0.13	0.51	0.13
Cheating propensity (Quintano et al., 2009)	0.04	0.14	0.04	0.15
Cheating indicator (Angrist et al., forthcoming)	0.04	0.18	0.05	0.20
Sample size (number of classes)	13,292		13,449	
	Panel b. School characteristics			
Enrolment (natives)	28.2	15.2	28.5	15.4
Enrolment (immigrants)	3.75	4.68	3.71	4.46
Average number of classes	1.68	0.77	1.68	0.78
Sample size (number of schools)	8,012		8,082	
	Panel C. Institution characteristics			
External monitored institutions	0.25	0.43	0.25	0.43
Average number of classes	4.76	1.71	4.79	1.7
Average number of schools	2.79	0.97	2.81	0.97
Sample size (number of institutions)	2,867		2,880	

Notes: The unit of observation is a class in Panel A a school in Panel B and an institution in Panel C. Institutions are groups of schools managed by the same principal. The family and individual background characteristics in Panel A are the class-average shares of natives in a class who have that specific characteristic over the total number of natives in the class. Missing values do not contribute to the computation of these shares. Regressions based on this data in the paper include the class-average shares of missing values for each characteristic as an additional control. All these variables come from the school administration through the data file that we received from INVALSI, except for the cheating indicator that was computed by Angrist et al. (forthcoming) and kindly given to us by these authors.

Table B.11: Descriptive statistics for the math sample.

	2nd grade		5th grade	
	Mean	S.D.	Mean	S.D.
	Panel A. Class characteristics			
Fraction of correct answers:				
language (natives)	0.67	0.11	0.71	0.09
math (natives)	0.62	0.12	0.65	0.11
language (immigrants)	0.53	0.19	0.60	0.15
math (immigrants)	0.54	0.16	0.57	0.15
Number of natives in class	16.7	3.96	16.9	3.95
Number of immigrants in class	2.24	2.30	2.21	2.25
Class size	19	3.86	19.1	3.95
Share (0-1) of natives in class with				
low educated mother	0.34	0.20	0.38	0.21
low educated father	0.43	0.21	0.46	0.21
employed mother	0.64	0.23	0.62	0.23
employed father	0.96	0.07	0.96	0.07
Share (0-1) of natives in class who				
attended kindergarten	0.99	0.04	0.99	0.04
are male	0.51	0.13	0.51	0.13
Cheating propensity (Quintano et al., 2009)	0.04	0.15	0.04	0.15
Cheating indicator (Angrist et al., forthcoming)	0.04	0.20	0.05	0.21
Sample size (number of classes)	13,284		13,449	
	Panel b. School characteristics			
Enrolment (natives)	28.1	15.2	28.5	15.4
Enrolment (immigrants)	3.76	4.69	3.71	4.46
Average number of classes	1.68	0.77	1.68	0.78
Sample size (number of schools)	8,004		8,082	
	Panel C. Institution characteristics			
External monitored institutions	0.25	0.43	0.25	0.43
Average number of classes	4.76	1.71	4.79	1.7
Average number of schools	2.79	0.97	2.81	0.97
Sample size (number of institutions)	2,865		2,881	

Notes: The unit of observation is a class in Panel A a school in Panel B and an institution in Panel C. Institutions are groups of schools managed by the same principal. The family and individual background characteristics in Panel A are the class-average shares of natives in a class who have that specific characteristic over the total number of natives in the class. Missing values do not contribute to the computation of these shares. Regressions based on this data in the paper include the class-average shares of missing values for each characteristic as an additional control. All these variables come from the school administration through the data file that we received from INVALSI, except for the cheating indicator that was computed by Angrist et al. (forthcoming) and kindly given to us by these authors.

Table B.12: School-fixed effect estimates of the effect of the number of natives and immigrants (first and second generation) on the test scores of natives; language and mathematics samples (classes).

	Language			Mathematics		
	Pooled	2nd grade	5th grade	Pooled	2nd grade	5th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Number of natives: $\hat{\beta}$	0.0014***	0.0011*	0.0017***	0.0018***	0.0008	0.0027***
	(0.0004)	(0.0006)	(0.0005)	(0.0004)	(0.0006)	(0.0005)
Number of immigrants: $\hat{\gamma}$	-0.0006	-0.0005	-0.0006	0.0012**	0.0011	0.0013*
	(0.0006)	(0.0009)	(0.0006)	(0.0006)	(0.0009)	(0.0007)
Confounded ethnic composition effect: $\hat{\delta}$	-0.0020***	-0.0017***	-0.0022***	-0.0006	0.0003	-0.0014***
	(0.0005)	(0.0008)	(0.0005)	(0.0005)	(0.0007)	(0.0006)
Observations	26,741	13,292	13,449	26,733	13,284	13,449
School×grade FE	✓	✓	✓	✓	✓	✓
Class level controls	✓	✓	✓	✓	✓	✓

Notes: The table reports in each column a different regression based on the language and mathematics samples. The unit of observation is a class. The dependent variable is the average test scores in language (mathematics) for natives students in a class, i.e. the fraction of correct answers. The controls are covariates aggregated at the class level: the share of natives with mothers and fathers who attended, at most, a lower secondary school, the shares of natives with employed mothers and fathers, the share of natives who attended kindergarten and the share of male natives in the class. All regressions include also the share of native students who report missing values in each of these variables as well as institution×grade fixed effects. Robust standard errors clustered at the institution-grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%.

Table C.1: Descriptive statistics for the language sample

	2nd grade		5th grade	
	Mean	S.D.	Mean	S.D.
	Panel A. School characteristics			
Fraction of correct answers:				
language (natives)	0.67	0.10	0.71	0.08
math (natives)	0.62	0.11	0.65	0.10
language (1st gen. immigrants)	0.51	0.21	0.58	0.16
math (1st gen. immigrants)	0.54	0.17	0.56	0.16
Number of quasi-natives	17.9	3.81	17.7	3.82
Number of 1st gen. immigrants in class	0.80	0.92	1.13	1.21
Class size	18.7	3.89	18.8	3.99
Share (0-1) of natives in class with				
low educated father	0.44	0.19	0.47	0.20
low educated mother	0.34	0.19	0.38	0.19
employed father	0.97	0.06	0.96	0.06
employed mother	0.65	0.21	0.63	0.21
Share (0-1) of natives in class who				
attended kindergarten	0.99	0.04	0.99	0.04
are male	0.51	0.12	0.51	0.12
Cheating propensity (Quintano et al., 2009)	0.04	0.13	0.04	0.14
Cheating indicator (Angrist et al., forthcoming)	0.04	0.16	0.04	0.18
Enrolment (natives)	28.2	15.2	28.5	15.3
Enrolment (1st gen. immigrants)	1.44	1.87	2.03	2.56
Average number of classes	1.69	0.78	1.69	0.78
Sample size (number of schools)	7,037		7,496	
	Panel C. Institution characteristics			
External monitored institutions	0.26	0.43	0.25	0.43
Average number of classes	4.89	1.72	4.86	1.71
Average number of schools	2.85	1.00	2.84	0.98
Sample size (number of institutions)	2,469		2,637	

Notes: The unit of observation is a school in Panel A and an institution in Panel B. Institutions are groups of schools managed by the same principal. The family and individual background characteristics in Panel A are the school-average shares of natives in a class who have that specific characteristic over the total number of natives in the class. Missing values do not contribute to the computation of these shares. All regressions in the following tables include the school-average shares of missing values for each characteristic as an additional control. All these variables come from the school administration through the data file that we received from INVALSI, except for the cheating indicator that was computed by Angrist et al. (forthcoming) and kindly given to us by these authors. Quasi-natives are natives (children born from at least one Italian parent) and 2nd generation immigrants (children born in Italy from non Italian parents).

Table C.2: Descriptive statistics for the math sample

	2nd grade		5th grade	
	Mean	S.D.	Mean	S.D.
	Panel A. School characteristics			
Fraction of correct answers:				
language (natives)	0.67	0.10	0.71	0.08
math (natives)	0.62	0.11	0.65	0.10
language (1st gen. immigrants)	0.51	0.21	0.58	0.16
math (1st gen. immigrants)	0.53	0.17	0.56	0.16
Number of quasi-natives	17.9	3.81	17.7	3.83
Number of 1st gen. immigrants in class	0.80	0.92	1.13	1.21
Class size	18.7	3.89	18.8	4.0
Share (0-1) of natives in class with				
low educated father	0.44	0.19	0.46	0.20
low educated mother	0.34	0.19	0.38	0.19
employed father	0.97	0.06	0.96	0.06
employed mother	0.65	0.21	0.63	0.21
Share (0-1) of natives in class who				
attended kindergarten	0.99	0.04	0.99	0.04
are male	0.51	0.12	0.51	0.12
Cheating propensity (Quintano et al., 2009)	0.04	0.13	0.04	0.13
Cheating indicator (Angrist et al., forthcoming)	0.04	0.17	0.04	0.18
Enrolment (natives)	28.2	15.2	28.5	15.3
Enrolment (1st gen. immigrants)	1.44	1.87	2.03	2.56
Average number of classes	1.69	0.78	1.69	0.78
Sample size (number of schools)	7,030		7,494	
	Panel C. Institution characteristics			
External monitored institutions	0.26	0.43	0.25	0.43
Average number of classes	4.89	1.72	4.86	1.71
Average number of schools	2.85	1.00	2.84	0.98
Sample size (number of institutions)	2,467		2,637	

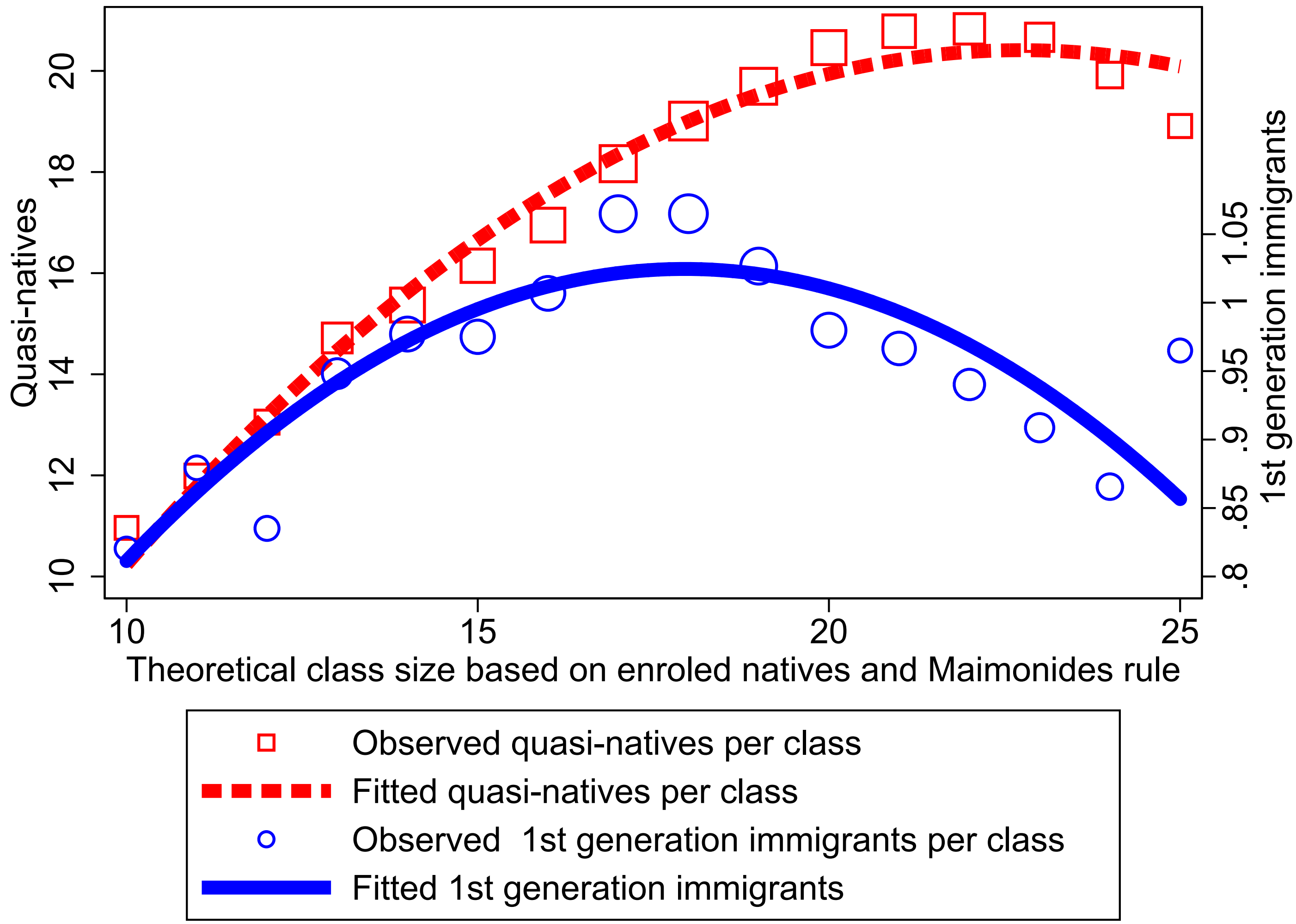
Notes: The unit of observation is a school in Panel A and an institution in Panel B. Institutions are groups of schools managed by the same principal. The family and individual background characteristics in Panel A are the school-average shares of natives in a class who have that specific characteristic over the total number of natives in the class. Missing values do not contribute to the computation of these shares. All regressions in the following tables include the school-average shares of missing values for each characteristic as an additional control. All these variables come from the school administration through the data file that we received from INVALSI, except for the cheating indicator that was computed by Angrist et al. (forthcoming) and kindly given to us by these authors. Quasi-natives are natives (children born from at least one Italian parent) and 2nd generation immigrants (children born in Italy from non Italian parents).

Table C.3: OLS-FE estimates of the effect of the number of quasi-natives and first generation immigrants on the test scores of natives; language and math samples.

	Language			math		
	Pooled	2nd grade	5th grade	Pooled	2nd grade	5th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Quasi-natives: $\hat{\beta}$	-0.0010***	-0.0014***	-0.0006**	-0.0010***	-0.0012***	-0.0009**
	(0.0002)	(0.0004)	(0.0003)	(0.0003)	(0.0004)	(0.0004)
Number of 1st generation immigrants: $\hat{\gamma}$	-0.0033***	-0.0041***	-0.0029***	-0.0040***	-0.0043***	-0.0039***
	(0.0008)	(0.0015)	(0.0009)	(0.0009)	(0.0016)	(0.0010)
Confounded ethnic composition effect: $\hat{\delta}$	-0.0024***	-0.0027***	-0.0023***	-0.0030***	-0.0030***	-0.0031***
	(0.0008)	(0.0015)	(0.0009)	(0.0009)	(0.0016)	(0.0011)
Observations	14,533	7,037	7,496	14,524	7,030	7,494
Institution×grade FE	✓	✓	✓	✓	✓	✓
School level controls	✓	✓	✓	✓	✓	✓

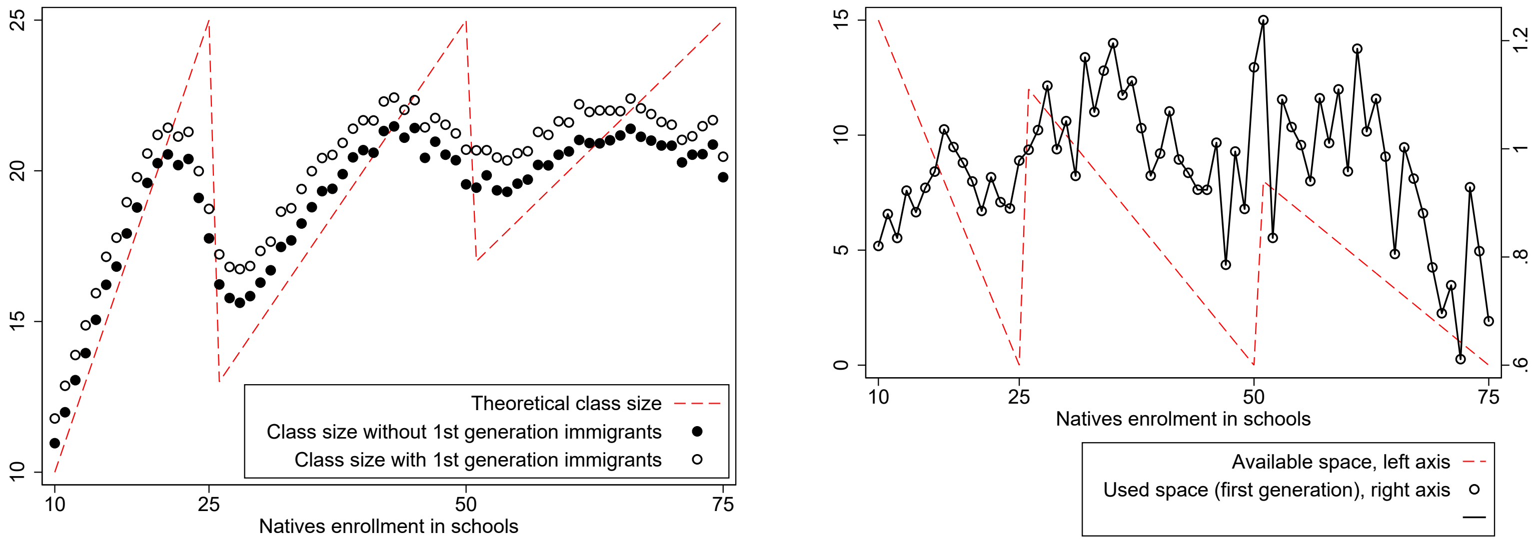
Notes: The table reports in each column a different regression based on the language and maths samples (described respectively in Tables C.1 and C.2). The unit of observation is a school. Quasi-natives are natives (children born from at least one Italian parent) and 2nd generation immigrants (children born in Italy from non Italian parents). The dependent variable is the the average test scores in language (math) for natives students, i.e. the fraction of correct answers. The controls are school-level averages of the following set class-level covariates covariates: the share of natives with mothers and fathers who attended, at most, a lower secondary school, the shares of natives with employed mothers and fathers, the share of natives who attended kindergarten and the share of male natives in the class. All regressions include also the share of native students who report missing values in each of these variables as well as institution×grade fixed effects. Robust standard errors clustered at the institution-grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%.

Figure C.1: Number of quasi-natives and first generation immigrants in a class as a function of theoretical class size based on native enrolment; language sample, 2nd and 5th grade.



Notes: Quasi-natives are natives (children born from at least one Italian parent) and 2nd generation immigrants (children born in Italy from non Italian parents). First generation immigrants are children not born in Italy from not Italian parents. In this figure, squares (left vertical axis) indicate the average number of quasi-natives per class in schools with the correspondent theoretical class size based on enrolled natives (horizontal axis). The dashed line is a quadratic fit of these averages. Circles (right vertical axis) indicate the average number of first generation immigrants per class in schools with the correspondent theoretical class size based on enrolled natives (horizontal axis). The continuous line is a quadratic fit of these averages. The size of squares and circles is proportional to the number of schools used to compute the averages that they represent. The quadratic fitted lines have been estimated with weights equal to the number of schools for each value of theoretical class size based on enrolled natives (horizontal axis).

Figure C.2: Number of quasi-natives and first generation immigrants in a class and class size as a function of native enrolment; language sample, 2nd and 5th grades.



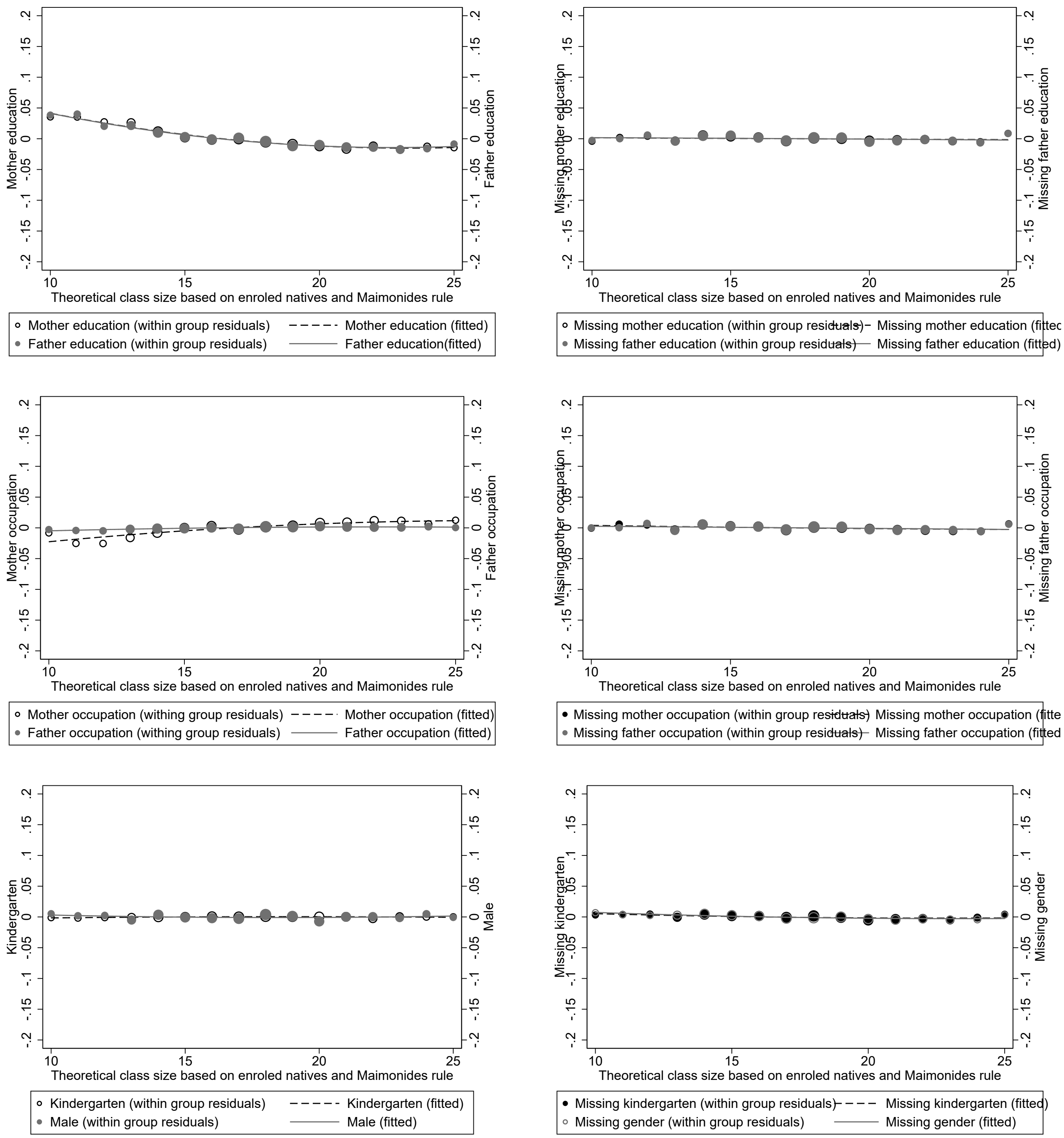
Notes: The left panels report the theoretical class size (dashed line), the class size without first generation immigrants (dark dots) and the class size with first generation immigrants (light dots) as a function of native enrollment in schools. In the right panels, the line connecting light dots represent the vertical distance between the light and dark dots of the left panels (the actual number of first generation immigrants per class) as a function of native enrolment. The right panels also plot the theoretically available space for immigrants (dashed line), defined as the maximum number of students in a class (25) minus the theoretical class size based on the number of natives C_{sg}^{N}. Quasi-natives are natives (children born from at least one Italian parent) and 2nd generation immigrants (children born in Italy from non Italian parents). First generation immigrants are children not born in Italy from not Italian parents.

Table C.4: First Stage for the number quasi-natives N and first generation immigrants I; language sample.

	Pooled 2nd & 5th grades			2nd grade			5th grade		
	Two endogenous		One endogenous	Two endogenous		One endogenous	Two endogenous		One endogenous
	N	I	I	N	I	I	N	I	I
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$1(10 \leq C^N_{sg} < 11)$	-6.01***	-0.06	-0.40***	-5.61***	-0.13	-0.38***	-6.42***	0.01	-0.41***
	(0.29)	(0.09)	(0.09)	(0.40)	(0.11)	(0.12)	(0.43)	(0.13)	(0.14)
$1(11 \leq C^N_{sg} < 12)$	-5.13***	-0.03	-0.31***	-4.88***	-0.18*	-0.40***	-5.40***	0.13	-0.22*
	(0.29)	(0.08)	(0.08)	(0.40)	(0.10)	(0.10)	(0.42)	(0.13)	(0.13)
$1(12 \leq C^N_{sg} < 13)$	-3.87***	-0.06	-0.28***	-3.53***	-0.19*	-0.34***	-4.19***	0.07	-0.21
	(0.29)	(0.08)	(0.08)	(0.40)	(0.10)	(0.10)	(0.42)	(0.13)	(0.13)
$1(13 \leq C^N_{sg} < 14)$	-2.71***	0.00	-0.15**	-2.43***	-0.16*	-0.27***	-2.99***	0.16	-0.03
	(0.28)	(0.07)	(0.07)	(0.39)	(0.08)	(0.08)	(0.42)	(0.11)	(0.11)
$1(14 \leq C^N_{sg} < 15)$	-2.27***	-0.04	-0.17**	-2.10***	-0.14*	-0.23***	-2.43***	0.06	-0.10
	(0.28)	(0.07)	(0.07)	(0.38)	(0.08)	(0.08)	(0.40)	(0.11)	(0.11)
$1(15 \leq C^N_{sg} < 16)$	-1.46***	-0.09	-0.17**	-1.30***	-0.16*	-0.22**	-1.61***	-0.02	-0.13
	(0.28)	(0.07)	(0.07)	(0.37)	(0.09)	(0.09)	(0.40)	(0.11)	(0.11)
$1(16 \leq C^N_{sg} < 17)$	-0.65**	-0.02	-0.05	-0.59	-0.13	-0.16*	-0.71*	0.10	0.06
	(0.27)	(0.07)	(0.07)	(0.37)	(0.09)	(0.08)	(0.40)	(0.11)	(0.11)
$1(17 \leq C^N_{sg} < 18)$	0.33	0.07	0.09	0.43	-0.07	-0.06	0.23	0.22**	0.23**
	(0.27)	(0.07)	(0.07)	(0.37)	(0.08)	(0.08)	(0.39)	(0.10)	(0.10)
$1(18 \leq C^N_{sg} < 19)$	1.07***	0.02	0.08	1.12***	-0.12	-0.07	1.01**	0.15	0.22**
	(0.27)	(0.07)	(0.07)	(0.37)	(0.08)	(0.08)	(0.40)	(0.11)	(0.11)
$1(19 \leq C^N_{sg} < 20)$	1.69***	-0.00	0.09	1.87***	-0.18**	-0.10	1.52***	0.18*	0.28***
	(0.27)	(0.07)	(0.07)	(0.38)	(0.08)	(0.08)	(0.40)	(0.11)	(0.11)
$1(20 \leq C^N_{sg} < 21)$	2.28***	-0.05	0.07	2.39***	-0.18**	-0.07	2.15***	0.07	0.21*
	(0.28)	(0.07)	(0.07)	(0.39)	(0.08)	(0.08)	(0.40)	(0.11)	(0.11)
$1(21 \leq C^N_{sg} < 22)$	2.68***	-0.11	0.04	2.71***	-0.21**	-0.09	2.64***	0.01	0.18*
	(0.28)	(0.07)	(0.07)	(0.39)	(0.08)	(0.08)	(0.41)	(0.11)	(0.11)
$1(22 \leq C^N_{sg} < 23)$	2.42***	-0.11	0.02	2.66***	-0.23***	-0.11	2.18***	0.01	0.15
	(0.30)	(0.07)	(0.07)	(0.41)	(0.09)	(0.09)	(0.43)	(0.11)	(0.11)
$1(23 \leq C^N_{sg} < 24)$	2.28***	-0.10	0.02	1.93***	-0.11	-0.02	2.61***	-0.09	0.08
	(0.31)	(0.07)	(0.07)	(0.44)	(0.08)	(0.08)	(0.45)	(0.11)	(0.11)
$1(24 \leq C^N_{sg} < 25)$	1.51***	-0.14*	-0.05	1.52***	-0.20**	-0.13	1.46***	-0.07	0.03
	(0.32)	(0.07)	(0.07)	(0.44)	(0.09)	(0.09)	(0.47)	(0.12)	(0.12)
Institution×grade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Polynomial in natives enrolment	✓	✓	✓	✓	✓	✓	✓	✓	
School level controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	14,533	14,533	14,533	7,037	7,037	7,037	7,496	7,496	7,496
F stat	305.13	2.23	5.26	141.60	1.08	2.12	166.50	2.17	4.09
SW F stat	61.57	2.29	5.26	16.68	1.08	2.12	75.11	2.27	4.09
SW χ^2 p-value	0.00	0.00	0.00	0.00	0.37	0.01	0.00	0.00	0.00

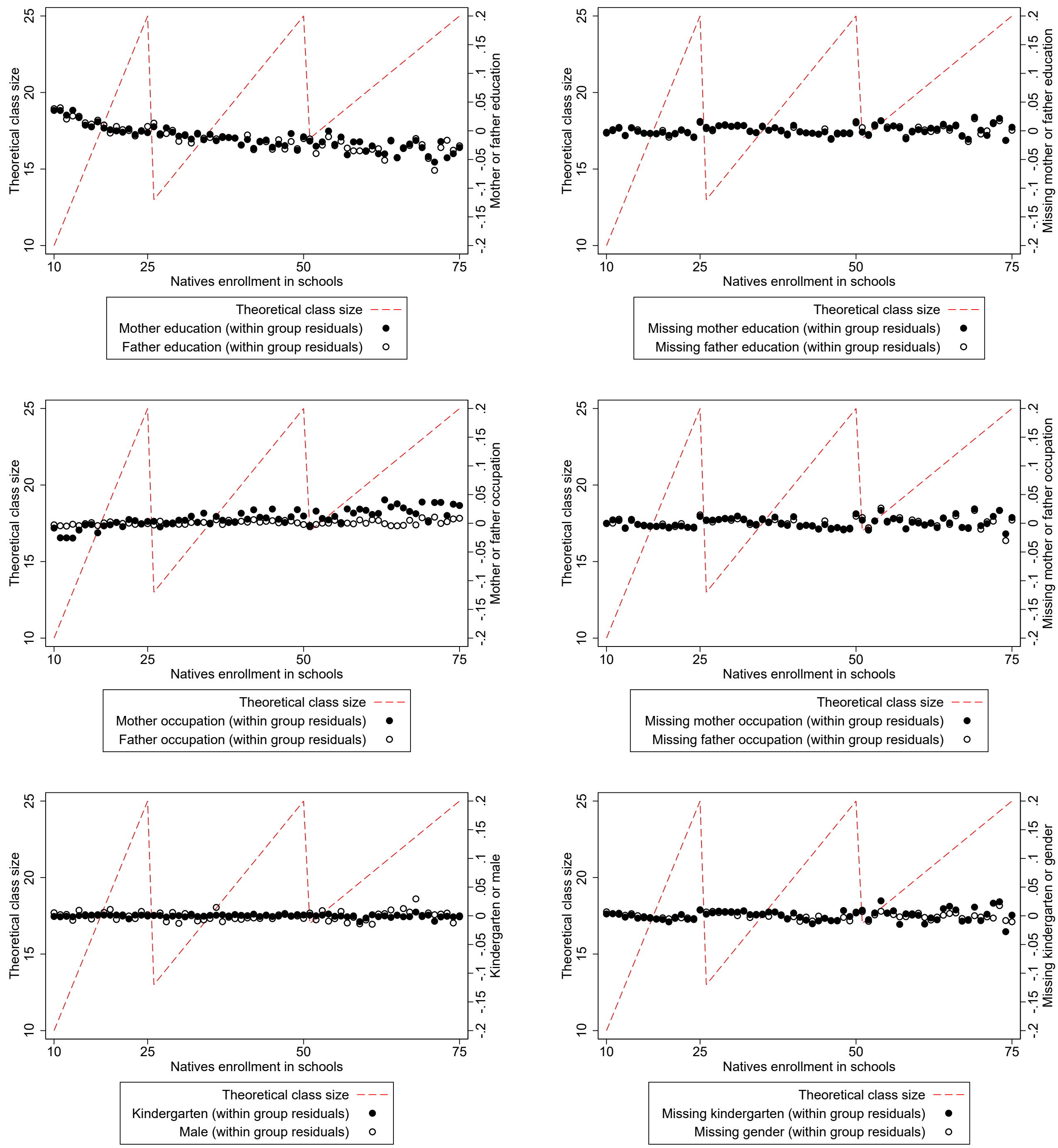
Notes: First generation immigrants are children not born in Italy from not Italian parents. Natives are children born from at least one Italian parent. Quasi-natives are natives and 2nd generation immigrants (children born in Italy from non Italian parents). The table reports in each column a different first stage regression correspondent to the IV estimates of Table 10. The unit of observation is a school. The dependent variable is the average number of natives and second generation immigrants N (1st generation immigrants I) per a class in a school. The instruments are a set of 15 dummies, one for each level of the theoretical number of natives in a class predicted by equation (2) in the paper according to the rules of class formation as a function of native enrolment at the school×grade level. The omitted category corresponds to a number of natives in a class equal to 25. There are no school-grades in which the number of natives in a class is less than 10. All regressions include a 2nd order polynomial of natives enrolment at the school×grade level. The controls are aggregated at the school level and include the following set of family and individual covariates: the share of natives with mothers and fathers who attended, at most, a lower secondary school, the shares of natives with employed mothers and fathers, the share of natives who attended kindergarten and the share of male natives in the school. All regressions include also the share of native students who report missing values in each of these variables as well as institution×grade fixed effects. Robust standard errors clustered at the school-grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%. The table reports also: i) the value of the F test of null hypothesis that the coefficients of the instruments are all zero in each first stage equation; and ii) the Sanderson-Windemeier first stage F statistic of each individual endogenous regressor (in the case of a model with one endogenous regressor this coincides with the F-test on excluded instruments) to test for weak identification ; and iii) the p-value of Sanderson-Windemeier χ^2 statistic of each individual endogenous regressor to test for under-identification.

Figure C.3: Included covariates as a function of theoretical class size based on native enrolment; language sample, 2nd and 5th grade.



Notes: First generation immigrants are children not born in Italy from not Italian parents. Natives are children born from at least one Italian parent. Quasi-natives are natives and 2nd generation immigrants (children born in Italy from non Italian parents). In these panels, rounds indicate the average of the within-group (institution-grade) residual of each covariate included in the OLS (Table C.3) and IV (Table 10 in the paper) regression that focus on first generation immigrans-indicated on the left or right vertical axes according to the legend- in schools at the correspondent theoretical class size based on enrolled natives (horizontal axis).The solid and dashed lines represent quadratic fits of these averages. The size of squares and circles is proportional to the number of schools used to compute the averages that they represent. The quadratic fitted lines have been estimated with weights equal to the number of schools for each value of theoretical class size based on enrolled natives (horizontal axis).

Figure C.4: Included covariates as a function of theoretical class size based on native enrolment; language sample, 2nd and 5th grade.



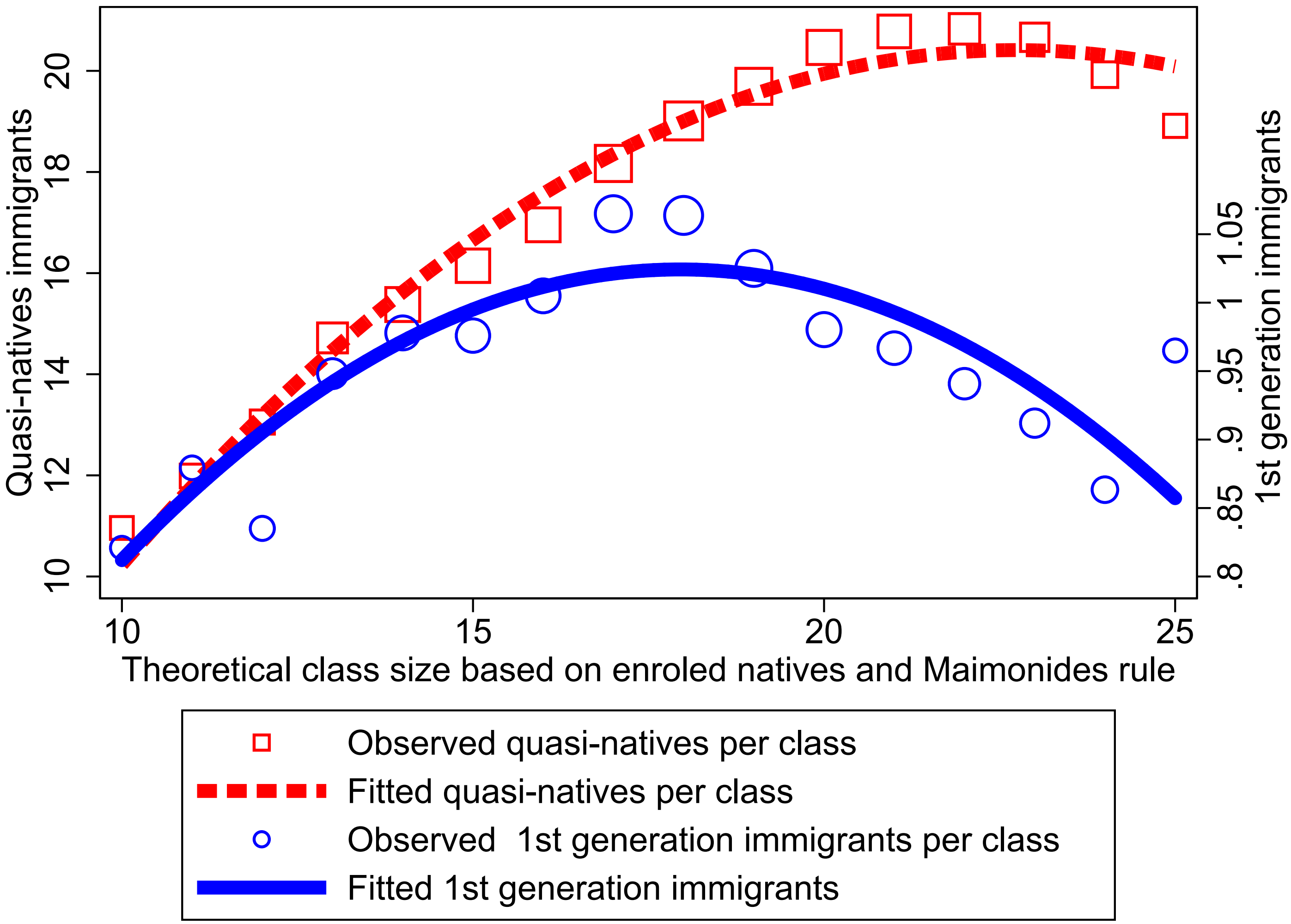
Notes: In these panels, we report theoretical class size (dashed line) and the average of the within-group (institution-grade) residual of each covariate included in the OLS (Table C.3) and IV (Table 10 in the paper) regression of the paper (dark or light dots) as a function of native enrollment in schools. First generation immigrants are children not born in Italy from not Italian parents. Natives are children born from at least one Italian parent. Quasi-natives are natives and 2nd generation immigrants (children born in Italy from non Italian parents).

Table C.5: Effect of the instruments on covariates; language sample; pooled 2nd and 5th grade.

Outcomes	Education		Employed		Kindergarten	Male	Missing					
							Education		Employed		Kindergarten	Male
Regressors	Mother	Father	Mother	Father			Mother	Father	Mother	Father		
$1(10 \leq C_{sg}^N < 11)$	0.02**	0.01	0.00	-0.00	-0.00	-0.00	-0.02**	-0.02*	-0.01	-0.01	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$1(11 \leq C_{sg}^N < 12)$	0.02**	0.02	-0.03***	-0.00	-0.00	-0.00	-0.01	-0.01	0.00	-0.01	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$1(12 \leq C_{sg}^N < 13)$	0.01	-0.01	-0.03***	-0.00	-0.00	-0.00	-0.01	-0.00	-0.00	-0.00	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$1(13 \leq C_{sg}^N < 14)$	0.02*	-0.00	-0.02**	0.00	-0.00	-0.01	-0.02***	-0.02***	-0.02**	-0.02***	-0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(14 \leq C_{sg}^N < 15)$	0.01	-0.00	-0.01	-0.00	-0.00	0.00	-0.01	-0.01	-0.01	-0.01	-0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(15 \leq C_{sg}^N < 16)$	-0.00	-0.02**	0.00	-0.00	-0.00	-0.00	-0.01*	-0.01*	-0.01	-0.01**	-0.01	-0.01*
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(16 \leq C_{sg}^N < 17)$	-0.01	-0.02*	0.00	0.00	0.00	-0.01	-0.02***	-0.02***	-0.02**	-0.02***	-0.01	-0.01**
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(17 \leq C_{sg}^N < 18)$	-0.00	-0.01	-0.01	-0.00	0.00	-0.00	-0.02***	-0.02***	-0.02***	-0.02***	-0.01*	-0.01***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(18 \leq C_{sg}^N < 19)$	-0.00	-0.01	-0.00	0.00	0.00	0.00	-0.02***	-0.02***	-0.02**	-0.02***	-0.01	-0.01***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(19 \leq C_{sg}^N < 20)$	-0.01	-0.02**	-0.00	0.01	0.00	0.00	-0.03***	-0.03***	-0.02***	-0.02***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(20 \leq C_{sg}^N < 21)$	-0.01	-0.02*	0.00	0.00	0.00	-0.01	-0.03***	-0.03***	-0.02***	-0.03***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(21 \leq C_{sg}^N < 22)$	-0.01*	-0.02***	0.00	0.00	0.00	-0.00	-0.03***	-0.03***	-0.03***	-0.03***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(22 \leq C_{sg}^N < 23)$	-0.01	-0.02**	0.01	0.00	-0.00	0.00	-0.03***	-0.02***	-0.03***	-0.02***	-0.02**	-0.02***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(23 \leq C_{sg}^N < 24)$	-0.02*	-0.03***	-0.00	0.00	0.00	-0.00	-0.03***	-0.03***	-0.03***	-0.03***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(24 \leq C_{sg}^N < 25)$	-0.00	-0.01	0.00	0.00	0.00	0.01	-0.03***	-0.04***	-0.03***	-0.03***	-0.02**	-0.02***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Observations							14,533					
Joint significance of instruments (p-value)	0.0003	0.0059	0.0069	0.2837	0.5087	0.4740	0.0006	0.0002	0.0000	0.0002	0.0001	0.0000

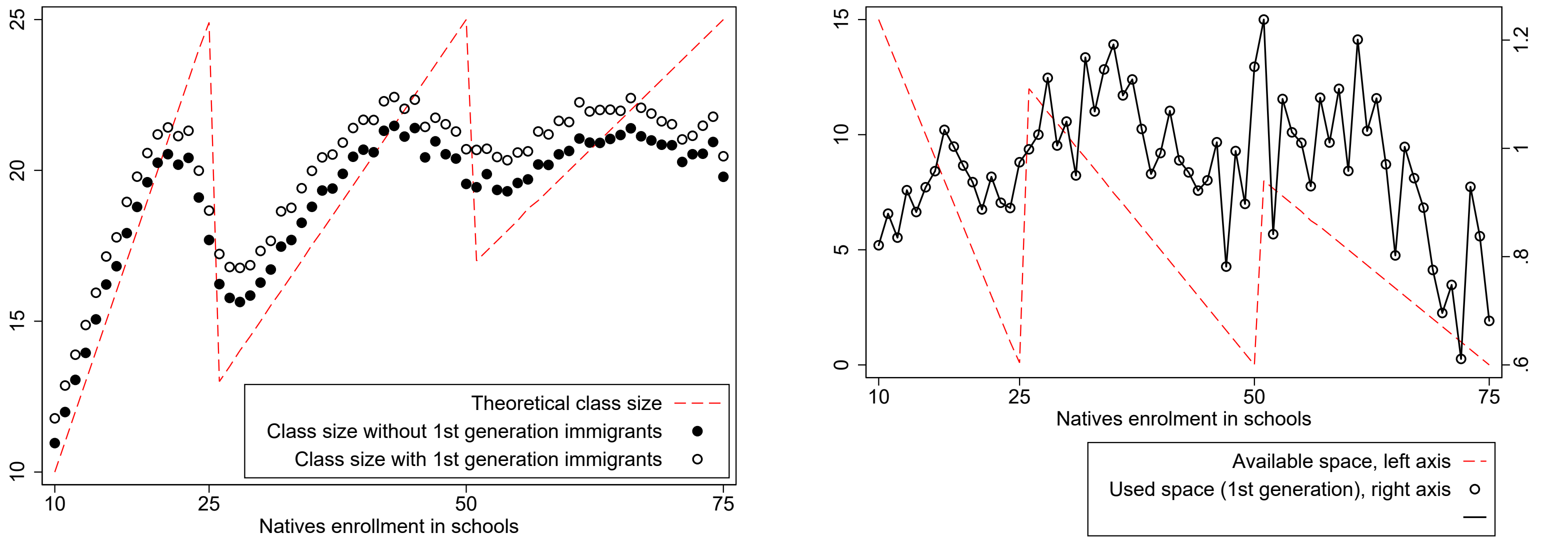
Notes: The table reports in each column the estimates of a system of equations (Seemingly Unrelated Regression Estimates) with one equation for each control variable included in included in the OLS (Table C.3) and IV (Table 10 in the paper) regression included in the paper that focus on first generation immigrants. The unit of observation is a school. The dependent variable in each column is the within-group (institution-grade) residual of the covariate indicated in the heading, i.e. the observed value of the variable in the school minus the institution-grade average of the same variable. The controls include the following set of within-group (institution-grade) residuals: a 2nd order polynomial of natives enrolment at at the school×grade level and the instruments. The instruments are a set of 15 dummies, one for each level of the theoretical number of natives in a class predicted by equation (2) in the paper according to the rules of class formation as a function of native enrolment at the school×grade level. The omitted category corresponds to a number of natives in a class equal to 25. There are no school-grades in which the number of natives in a class is less than 10. Robust standard errors clustered at the institution-grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%. First generation immigrants are children not born in Italy from not Italian parents. Natives are children born from at least one Italian parent. Quasi-natives are natives and 2nd generation immigrants (children born in Italy from non Italian parents).

Figure C.5: Number of quasi-natives and first generation immigrants in a class as a function of theoretical class size based on native enrolment; math sample, 2nd and 5th grade.



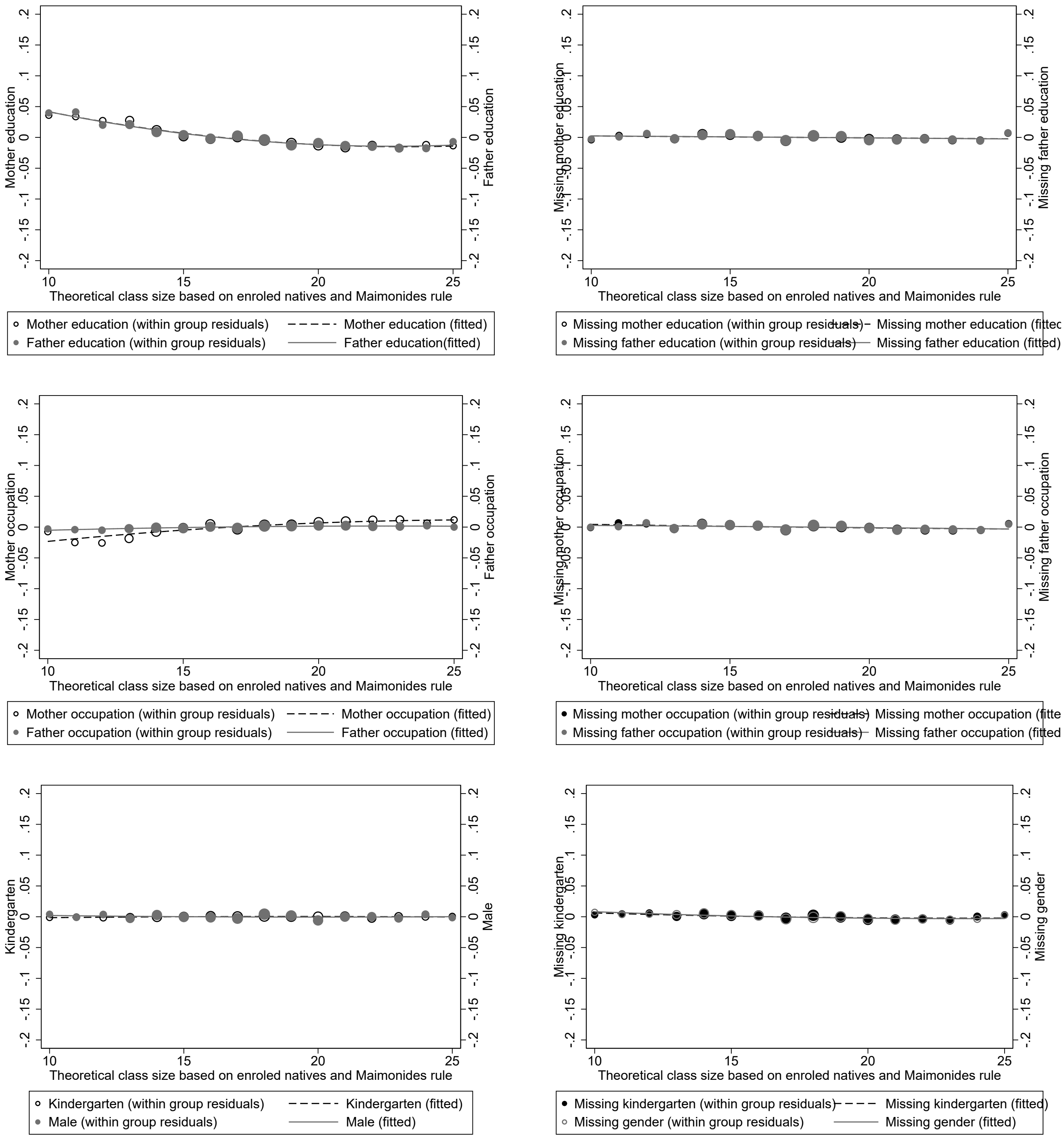
Notes: First generation immigrants are children not born in Italy from not Italian parents. Natives are children born from at least one Italian parent. Quasi-natives are natives and 2nd generation immigrants (children born in Italy from non Italian parents). In this figure, squares (left vertical axis) indicate the average number of quasi-natives per class in schools with the correspondent theoretical class size based on enrolled natives (horizontal axis). The dashed line is a quadratic fit of these averages. Circles (right vertical axis) indicate the average number of first generation immigrants per class in schools with the correspondent theoretical class size based on enrolled natives (horizontal axis). The continuous line is a quadratic fit of these averages. The size of squares and circles is proportional to the number of schools used to compute the averages that they represent. The quadratic fitted lines have been estimated with weights equal to the number of schools for each value of theoretical class size based on enrolled natives (horizontal axis).

Figure C.6: Number of quasi-natives and first generation immigrants in a class and class size as a function of native enrolment; math sample, 2nd and 5th grades.



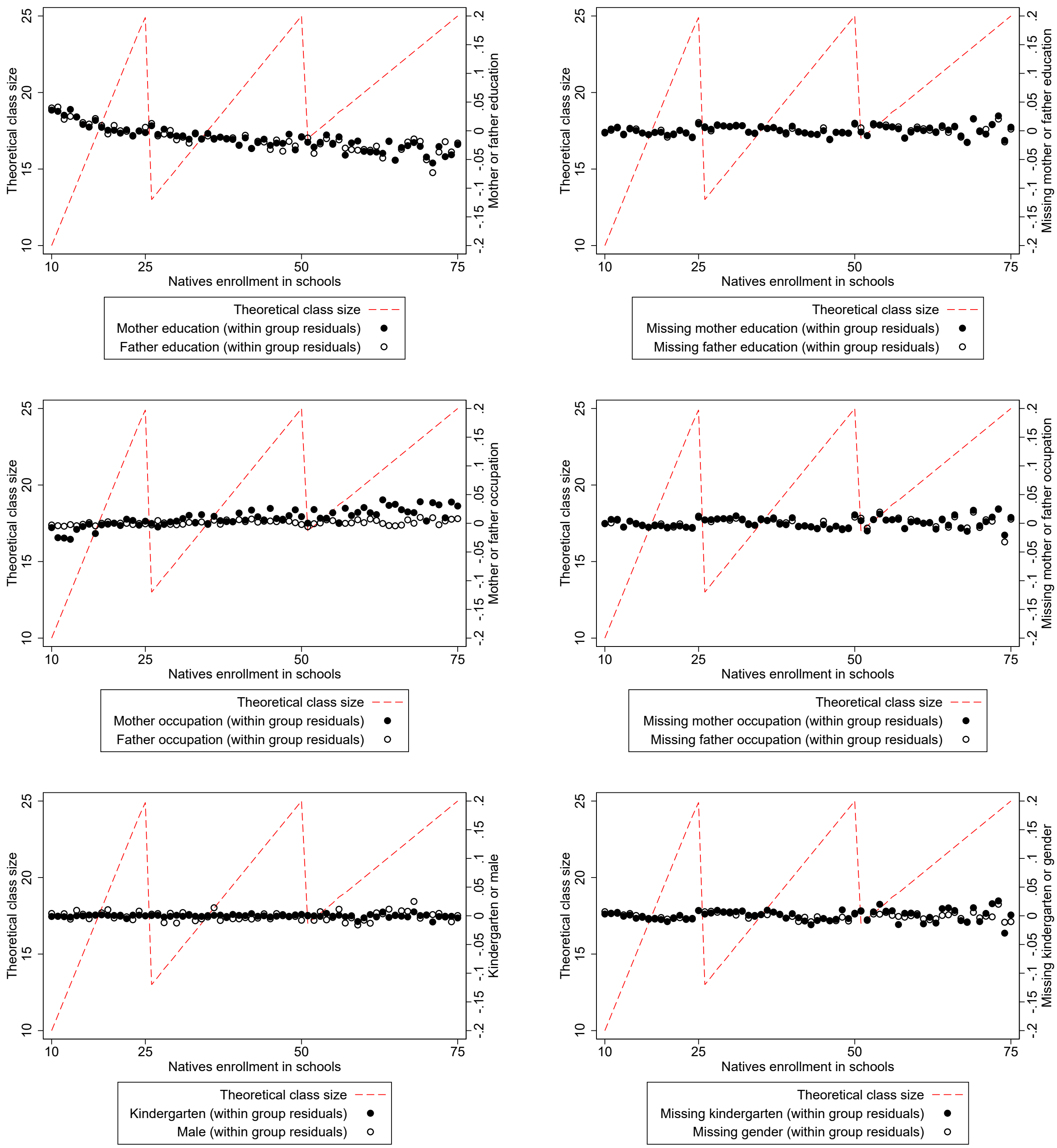
Notes: The left panels report the theoretical class size (dashed line), the class size without first generation immigrants (dark dots) and the class size with first generation immigrants (light dots) as a function of native enrollment in schools. In the right panels, the line connecing light dots represent the vertical distance between the light and dark dots of the left panels (the actual number of first generation immigrants per class) as a function of native enrolment. The right panels also plot the theoretically available space for immigrants (dashed line), defined as the maximum number of students in a class (25) minus the theoretical class size based on the number of natives C^N_{sg}. First generation immigrants are children not born in Italy from not Italian parents. Natives are children born from at least one Italian parent. Quasi-natives are natives and 2nd generation immigrants (children born in Italy from non Italian parents).

Figure C.7: Included covariates as a function of theoretical class size based on native enrolment; math sample, 2nd and 5th grade.



Notes: First generation immigrants are children not born in Italy from not Italian parents. Natives are children born from at least one Italian parent. Quasi-natives are natives and 2nd generation immigrants (children born in Italy from non Italian parents). In these panels, rounds indicate the within-group (institution-grade) residual of each covariate included in the OLS (Table C.3) and IV (Table 11 in the paper) regression included in the paper -names are indicated on the left or right vertical axes according to the legend- in schools at the correspondent theoretical class size based on enrolled natives (horizontal axis).The solid and dashed lines represent quadratic fits of these averages. The size of squares and circles is proportional to the number of schools used to compute the averages that they represent. The quadratic fitted lines have been estimated with weights equal to the number of schools for each value of theoretical class size based on enrolled natives (horizontal axis).

Figure C.8: Included covariates as a function of theoretical class size based on native enrolment; math sample, 2nd and 5th grade.



Notes: In these panels, we report theoretical class size (dashed line) and the average of the within-group (institution-grade) residual of each covariate included in the OLS (Table C.3) and IV (Table 11 in the paper) regression of the paper (dark or light dots) as a function of native enrollment in schools. First generation immigrants are children not born in Italy from not Italian parents. Natives are children born from at least one Italian parent. Quasi-natives are natives and 2nd generation immigrants (children born in Italy from non Italian parents).

Table C.6: First Stage for for the number quasi-natives N and first generation immigrants I; math sample.

	Pooled 2nd & 5th grades			2nd grade			5th grade		
	Two endogenous		One endogenous	Two endogenous		One endogenous	Two endogenous		One endogenous
	N	I	I	N	I	I	N	I	I
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$1(10 \leq C^N_{sg} < 11)$	-5.96***	-0.05	-0.38***	-5.60***	-0.11	-0.35***	-6.33***	0.01	-0.41***
	(0.30)	(0.09)	(0.09)	(0.40)	(0.11)	(0.12)	(0.44)	(0.13)	(0.14)
$1(11 \leq C^N_{sg} < 12)$	-5.07***	-0.03	-0.31***	-4.86***	-0.17*	-0.39***	-5.30***	0.13	-0.22*
	(0.29)	(0.08)	(0.08)	(0.40)	(0.10)	(0.10)	(0.43)	(0.13)	(0.13)
$1(12 \leq C^N_{sg} < 13)$	-3.82***	-0.07	-0.28***	-3.51***	-0.19*	-0.34***	-4.12***	0.06	-0.21
	(0.29)	(0.08)	(0.08)	(0.40)	(0.10)	(0.10)	(0.43)	(0.13)	(0.13)
$1(13 \leq C^N_{sg} < 14)$	-2.64***	-0.00	-0.15**	-2.40***	-0.16*	-0.27***	-2.89***	0.16	-0.03
	(0.29)	(0.07)	(0.07)	(0.38)	(0.08)	(0.08)	(0.43)	(0.11)	(0.11)
$1(14 \leq C^N_{sg} < 15)$	-2.21***	-0.04	-0.16**	-2.07***	-0.13	-0.22***	-2.36***	0.05	-0.10
	(0.28)	(0.07)	(0.07)	(0.38)	(0.08)	(0.08)	(0.41)	(0.11)	(0.11)
$1(15 \leq C^N_{sg} < 16)$	-1.40***	-0.09	-0.17**	-1.26***	-0.16*	-0.21**	-1.52***	-0.03	-0.13
	(0.28)	(0.07)	(0.07)	(0.37)	(0.09)	(0.08)	(0.41)	(0.11)	(0.11)
$1(16 \leq C^N_{sg} < 17)$	-0.58**	-0.02	-0.05	-0.57	-0.13	-0.16*	-0.62	0.10	0.06
	(0.28)	(0.07)	(0.07)	(0.37)	(0.09)	(0.08)	(0.41)	(0.11)	(0.11)
$1(17 \leq C^N_{sg} < 18)$	0.41	0.07	0.10	0.47	-0.07	-0.05	0.32	0.22**	0.24**
	(0.27)	(0.07)	(0.07)	(0.37)	(0.08)	(0.08)	(0.40)	(0.10)	(0.10)
$1(18 \leq C^N_{sg} < 19)$	1.14***	0.02	0.08	1.16***	-0.12	-0.07	1.11***	0.15	0.22**
	(0.27)	(0.07)	(0.07)	(0.37)	(0.08)	(0.08)	(0.41)	(0.11)	(0.11)
$1(19 \leq C^N_{sg} < 20)$	1.75***	-0.01	0.09	1.89***	-0.20**	-0.11	1.60***	0.18*	0.29***
	(0.28)	(0.07)	(0.07)	(0.38)	(0.08)	(0.08)	(0.40)	(0.11)	(0.11)
$1(20 \leq C^N_{sg} < 21)$	2.34***	-0.06	0.07	2.43***	-0.18**	-0.07	2.25***	0.06	0.21*
	(0.28)	(0.07)	(0.07)	(0.38)	(0.08)	(0.08)	(0.41)	(0.11)	(0.11)
$1(21 \leq C^N_{sg} < 22)$	2.73***	-0.10	0.05	2.73***	-0.21**	-0.09	2.71***	0.01	0.19*
	(0.29)	(0.07)	(0.07)	(0.39)	(0.08)	(0.08)	(0.42)	(0.11)	(0.11)
$1(22 \leq C^N_{sg} < 23)$	2.48***	-0.11	0.03	2.70***	-0.22***	-0.10	2.26***	0.00	0.15
	(0.30)	(0.07)	(0.07)	(0.41)	(0.09)	(0.08)	(0.44)	(0.11)	(0.11)
$1(23 \leq C^N_{sg} < 24)$	2.34***	-0.11	0.03	1.96***	-0.12	-0.03	2.70***	-0.08	0.09
	(0.32)	(0.07)	(0.07)	(0.44)	(0.08)	(0.08)	(0.45)	(0.11)	(0.11)
$1(24 \leq C^N_{sg} < 25)$	1.57***	-0.14*	-0.05	1.56***	-0.20**	-0.13	1.56***	-0.07	0.03
	(0.32)	(0.07)	(0.07)	(0.44)	(0.09)	(0.09)	(0.48)	(0.12)	(0.12)
Institution×grade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Polynomial in natives enrolment	✓	✓	✓	✓	✓	✓	✓	✓	
School level controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	14,524	14,524	14,524	7,030	7,030	7,030	7,494	7,494	7,494
F stat	307.07	2.19	5.20	144.07	1.07	2.03	165.50	2.20	4.18
SW F stat	55.60	2.24	5.20	12.23	1.06	2.03	84.96	2.31	4.18
SW χ^2 p-value	0.00	0.00	0.00	0.00	0.39	0.01	0.00	0.00	0.00

Notes: First generation immigrants are children not born in Italy from not Italian parents. Natives are children born from at least one Italian parent. Quasi-natives are natives and 2nd generation immigrants (children born in Italy from non Italian parents). The table reports in each column a different first stage regression correspondent to the IV estimates of Table 11. The unit of observation is a school. The dependent variable is the average number of natives and second generation immigrants N (1st generation immigrants I) per a class in a school. The instruments are a set of 15 dummies, one for each level of the theoretical number of natives in a class predicted by equation (2) in the paper according to the rules of class formation as a function of native enrolment at the school×grade level. The omitted category corresponds to a number of natives in a class equal to 25. There are no school-grades in which the number of natives in a class is less than 10. All regressions include a 2nd order polynomial of natives enrolment at the school×grade level. The controls are aggregated at the school level and include the following set of family and individual covariates: the share of natives with mothers and fathers who attended, at most, a lower secondary school, the shares of natives with employed mothers and fathers, the share of natives who attended kindergarten and the share of male natives in the school. All regressions include also the share of native students who report missing values in each of these variables as well as institution×grade fixed effects. Robust standard errors clustered at the school-grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%. The table reports also: i) the value of the F test of null hypothesis that the coefficients of the instruments are all zero in each first stage equation; and ii) the Sanderson-Windemeier first stage F statistic of each individual endogenous regressor (in the case of a model with one endogenous regressor this coincides with the F-test on excluded instruments) to test for weak identification ; and iii) the p-value of Sanderson-Windemeier χ^2 statistic of each individual endogenous regressor to test for under-identification.

Table C.7: Effect of the instruments on covariates; math sample; pooled 2nd and 5th grade.

Outcomes	Education		Employed		Kindergarten	Male	Missing					
							Education		Employed		Kindergarten	Male
Regressors	Mother	Father	Mother	Father			Mother	Father	Mother	Father		
$1(10 \leq C^N_{sg} < 11)$	0.02*	0.01	0.00	0.00	-0.00	0.00	-0.01	-0.01	-0.01	-0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$1(11 \leq C^N_{sg} < 12)$	0.02**	0.02*	-0.02**	-0.00	-0.00	-0.01	-0.00	-0.01	0.00	-0.01	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$1(12 \leq C^N_{sg} < 13)$	0.01	-0.01	-0.03***	-0.00	-0.00	0.00	-0.00	-0.00	0.00	0.00	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$1(13 \leq C^N_{sg} < 14)$	0.02**	-0.00	-0.02**	0.00	-0.00	-0.00	-0.01**	-0.02**	-0.01*	-0.01**	-0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(14 \leq C^N_{sg} < 15)$	0.01	-0.00	-0.01	-0.00	-0.00	0.00	-0.00	-0.01	-0.00	-0.01	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(15 \leq C^N_{sg} < 16)$	-0.01	-0.02**	0.00	-0.00	-0.00	0.00	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(16 \leq C^N_{sg} < 17)$	-0.01	-0.02*	0.00	0.00	0.00	-0.01	-0.02**	-0.02**	-0.02**	-0.02**	-0.01	-0.01*
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(17 \leq C^N_{sg} < 18)$	0.00	-0.01	-0.01	-0.00	-0.00	0.00	-0.02***	-0.02***	-0.02***	-0.02***	-0.01*	-0.01***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(18 \leq C^N_{sg} < 19)$	-0.00	-0.01	-0.00	0.00	0.00	0.01	-0.01**	-0.02**	-0.01**	-0.01**	-0.00	-0.01**
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(19 \leq C^N_{sg} < 20)$	-0.01	-0.02***	-0.00	0.01*	-0.00	0.00	-0.02***	-0.02***	-0.02***	-0.02***	-0.02***	-0.01***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(20 \leq C^N_{sg} < 21)$	-0.01	-0.01	0.00	0.01	-0.00	-0.00	-0.02***	-0.03***	-0.02***	-0.02***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(21 \leq C^N_{sg} < 22)$	-0.02*	-0.02***	0.01	0.00	0.00	0.00	-0.03***	-0.03***	-0.03***	-0.03***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(22 \leq C^N_{sg} < 23)$	-0.01	-0.02**	0.01	0.00	-0.00*	0.00	-0.02***	-0.02***	-0.03***	-0.02***	-0.02**	-0.02***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(23 \leq C^N_{sg} < 24)$	-0.02*	-0.03***	0.00	0.00	0.00	-0.00	-0.03***	-0.03***	-0.03***	-0.03***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
$1(24 \leq C^N_{sg} < 25)$	-0.00	-0.02	-0.00	0.00	0.00	0.01	-0.03***	-0.03***	-0.03***	-0.03***	-0.01*	-0.02***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Observations	14,524											
Joint significance of instruments (p-value)	0.0005	0.0016	0.0024	0.2560	0.4494	0.7387	0.0006	0.0005	0.0000	0.0002	0.0001	0.0000

Notes: First generation immigrants are children not born in Italy from not Italian parents. Natives are children born from at least one Italian parent. Quasi-natives are natives and 2nd generation immigrants (children born in Italy from non Italian parents). The table reports in each column the estimates of a system of equations (Seemingly Unrelated Regression Estimates) with one equation for each control variable included in included in the OLS (Table C.3) and IV (Table 11 in the paper) regression included in the paper. The unit of observation is a school. The dependent variable in each column is the within-group (institution-grade) residual of the covariate indicated in the heading, i.e. the observed value of the variable in the school minus the institution-grade average of the same variable. The controls include the following set of within-group (institution-grade) residuals: a 2nd order polynomial of natives enrolment at at the school×grade level and the instruments. The instruments are a set of 15 dummies, one for each level of the theoretical number of natives in a class predicted by equation (2) in the paper according to the rules of class formation as a function of native enrolment at the school×grade level. The omitted category corresponds to a number of natives in a class equal to 25. There are no school-grades in which the number of natives in a class is less than 10. Robust standard errors clustered at the institution-grade level are reported in parentheses. A * denotes significance at 10%; a ** denotes significance at 5%; a *** denotes significance at 1%.