The Tower of Babel in the Classroom?

Immigrants and Natives in Italian Schools

Rosario M. Ballatore [1] Margherita Fort [2] Andrea Ichino [3]

[1]Bank of Italy
[2]University of Bologna, CESifo and IZA
[3]EUI & University of Bologna, CEPR, CESifo and IZA

Workshop on Human Capital

Bank of Italy

Rome, November 29^{th}, 2016

Research question

The causal effect on school performance of a change in the ethnic composition of a class due to an immigrant inflow

Research question

The causal effect on school performance of a change in the ethnic composition of a class due to an immigrant inflow

net of:

Research question

The causal effect on school performance of a change in the ethnic composition of a class due to an immigrant inflow

net of:

- the endogenous adjustments of class size and students' quality generated by principals' reactions to this inflow

Research question

The causal effect on school performance of a change in the ethnic composition of a class due to an immigrant inflow

net of:

- the endogenous adjustments of class size and students' quality generated by principals' reactions to this inflow

We call it the **Pure Ethnic Composition (PEC) effect**

Research question: the PRF

$$V_j = \alpha + \beta N_j + \gamma I_j + \lambda Q_j^n + \mu Q_j^i + \epsilon_j \qquad (1)$$

The effects of changing class size by varying the number of natives (immigrants) keeping immigrants (natives) constant for given quality of the two ethnicities:

$$\beta = \left(\frac{dV_j}{dN}\right)_{I_j=\bar{I};Q_j^n=\bar{Q}^n;Q_j^i=\bar{Q}^i} \qquad \gamma = \left(\frac{dV_j}{dI}\right)_{N_j=\bar{N};Q_j^n=\bar{Q}^n;Q_j^i=\bar{Q}^i} \qquad (2)$$

The PEC effect is then given by

$$\delta = \left(\frac{dV_j}{dI}\right)_{C_j=\bar{C};Q_j^n=\bar{Q}^n;Q_j^i=\bar{Q}^i} = \gamma - \beta \qquad (3)$$

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

This paper in a nutshell

We explore features of the institutional setting:

- rules of class formation (with a cap of 25 students per class)
- Ministry of Education instructions to allocate immigrants where there is more space for them (e.g. where classes are smaller)
- differences in enrolment between February (pre-enrolment) and September (final enrolment)

the interaction between these features allows us to compare classes:

- that have different numbers of natives and immigrants
- for given students' quality

Main findings

The Pure Ethnic Composition (PEC) effect on native performance

- is negative and statistically significant at age 7
- $\approx$ -1.6% for both language and math
- it does not vanish when children grow up to age 10

When we use instead a more conventional identification strategy

- our estimates of the effects of immigrant inflows on native performance are smaller
- because they are confounded by the endogenous adjustments implemented by principals

The data

We use INVALSI data for Italian primary schools in 2009-10.

For each student in grades 2 and 5 the data set contains:

- test scores in language and mathematics
- educational institution, school, grade, class and student identifiers
- class size and class composition at the beginning of the year
- immigrant status based parents' nationality
- some individual and family background information

The unit of analysis is the class

We restrict the analysis to

- educational institutions with immigrants and more than one school (80% of the students)

Descriptive statistics for the language sample

	2nd grade		5th grade	
	Mean	S.D.	Mean	S.D.
Fraction of correct answers:				
- language (natives)	0.67	0.11	0.71	0.08
- mathematics (natives)	0.61	0.11	0.65	0.10
- language (immigrants)	0.53	0.19	0.60	0.15
- mathematics (immigrants)	0.54	0.16	0.58	0.16
Number of natives in class	16.37	4.09	16.65	4.02
Number of immigrants in class	3.08	2.21	3.02	2.17
Class size	19.44	3.84	19.67	3.89
Sample size (number of classes)	12,859		13,084	
Sample size (number of schools)	7,387		7,496	
Sample size (number of institutions)	2,734		2,776	

Identification: institutional framework

In February

Students should pre-enrol in a given school for the year that starts in the following September

The number of classes are tentatively formed by principals according to natives pre-enrolment, following a "Maimonides-type rule" with a cap at 25

The Ministry of Education instructs principals to put immigrants in the schools where, depending on natives enrolment, classes are smaller

In September

Additional splitting of classes occurs in September if late enrolment requires any further adjustment

No more room for endogenous reaction of principals

An example

Consider grade g of school s.

Predicted class size $\bar{C}^N_{sg}$, based on native pre-enrolment and rules of class formation, can take three equally likely values:

$$H > M > L = \frac{H}{2}$$

.

The principal knows that

- if $\bar{C}^N_{sg} = H$ in February, with probability π that class will be split in September
- splitting will originate two classes each one with (approximately) $L = \frac{H}{2}$ natives
- in the other two cases, instead, there is no risk of splitting

Principals' decisions in February

Each principal manages three classes (in different schools)

and has to allocate a total of I immigrants

With probability $1 - \pi$ the class expected to be large in February

- will remain large
- the principal will not put immigrants in it
- immigrants will end up in the other two classes

Therefore,

the February allocation of immigrants across the three classes, is:

$$I_{sg} = \begin{cases} 0 & \text{if} \quad \bar{C}^N_{sg} = H \\ \frac{I}{2} & \text{if} \quad \bar{C}^N_{sg} = M \\ \frac{I}{2} & \text{if} \quad \bar{C}^N_{sg} = L \end{cases} \tag{4}$$

Final allocation in September

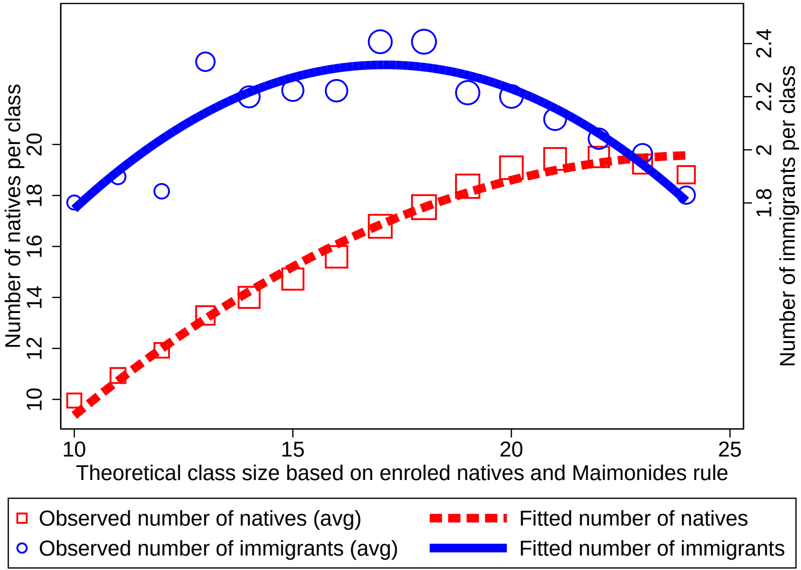
The allocation of immigrants based on the *final* number of natives per class C^N_{sg}, after late enrolment has occurred, is,

$$I_{sg} = \begin{cases} \approx 0 & \text{if} \quad C^N_{sg} \approx H \\ \frac{I}{2} & \text{if} \quad C^N_{sg} \approx M \\ \approx \frac{I}{2}\frac{1}{(1+2\pi)} & \text{if} \quad C^N_{sg} \approx L \end{cases} \tag{5}$$

The number of immigrants is a hump-shaped function of the final number of natives per class

- some small classes originate from final splitting of classes expected to be large in February and that were thus without immigrants
- classes that remain large have no immigrants
- the highest number of immigrants remains allocated to classes with an intermediate number of natives

Natives and immigrants in a class as a function of predicted class size based on native enrolment, pooling grade 2 and 5



Estimated equation

We apply our identification strategy to equation

$$V_{jskg} = \alpha + \beta N_{jskg} + \gamma I_{jskg} + \mu X_{jskg} + \eta_{kg} + f(N_{sg}) + u_{jskg}, \quad (14)$$

which includes

- fixed effects defined at the institution×grade level
- a polynomial in native enrolment at the school×grade level
  - to control for the systematic and continuous components of the relationship between native enrolment and native performance

Instruments

Instruments are constructed as Angrist and Lang (2004):

$$\Psi \in \{1(1 \leq C_{sg}^{N} < 2), 1(2 \leq C_{sg}^{N} < 3), ..., 1(24 \leq C_{sg}^{N} < 25\}, \quad (15)$$

They are indicators defined for each possibile level of the theoretical number of natives in a class, C_{sg}^{N}, predicted by

$$C_{sg}^{N} = \frac{N_{sg}}{\text{Int}\left(\frac{N_{sg}-1}{25}\right) + 1} \quad (13)$$

They capture in the most flexible way the

- non-linearities and
- discontinuities

generated by the rules of class formation, that relate native enrolment to the numbers of natives and immigrants in a class

IV-FE estimates: Language

	OLS-FE	IV-FE
Number of natives: $\hat{\beta}$	0.0001	-0.0019**
	(0.0003)	(0.0007)
Number of immigrants: $\hat{\gamma}$	-0.0049***	-0.0177***
	(0.0006)	(0.0054)
PEC: $\hat{\delta}$	-0.0050***	-0.0158***
	(0.0006)	(0.0052)
Observations	25,943	25,943
Institution×grade FE	✓	✓
Polynomial in natives enrolment		✓
Class level controls	✓	✓
Hansen (p-value)	n.a.	0.715
Natives (F-test)	n.a.	382.52
Natives (AP-test)	n.a.	187.98
Immigrants (F-test)	n.a.	299.19
Immigrants (AP-test)	n.a.	88.01

IV-FE estimates: Mathematics

	OLS-FE	IV-FE
Number of natives: $\hat{\beta}$	0.0002	-0.0010
	(0.0003)	(0.0009)
Number of immigrants: $\hat{\gamma}$	-0.0042***	-0.0168***
	(0.0007)	(0.0056)
PEC: $\hat{\delta}$	-0.0044***	-0.0158***
	(0.0006)	(0.0053)
Observations	25,943	25,936
Institution×grade FE	✓	✓
Polynomial in natives enrolment		✓
Class level controls	✓	✓
Hansen (p-value)	n.a.	0.972
Natives (F-test)	n.a.	221.76
Natives (AP-test)	n.a.	193.84
Immigrants (F-test)	n.a.	46.45
Immigrants (AP-test)	n.a.	40.90

Analysis by grade

- Qualitatively similar results
- Sometimes less precise
- $\hat{\gamma}$ and $\hat{\beta}$ typically larger in 5th than 2nd grade

Language: Instrumental Variable estimates by grade

Mathematics: Instrumental Variable estimates by grade

The threat of test scores manipulation by teachers

Angrist, Battistin and Vuri (2014) find evidence that test scores are

- manipulated by teachers in some southern regions of the country
- more as a result of shirking than because of self-interested cheating

Our results are essentially unchanged when we restrict the analysis to different sub-samples in which, according to ABV, score manipulation is likely to be minimal, if at all present

IV-FE estimates in the north and centre: Language

	Baseline specification	Classes with cheating indicator = 0	Externally monitored institutions
Number of natives: $\hat{\beta}$	-0.0026***	-0.0022***	-0.0033**
	(0.0008)	(0.0008)	(0.0015)
Number of immigrants: $\hat{\gamma}$	-0.0155***	-0.0172***	-0.0180**
	(0.0053)	(0.0055)	(0.0075)
PEC: $\hat{\delta}$	-0.0129**	-0.0149***	-0.0147**
	(0.0051)	(0.0052)	(0.0072)
Observations	19,001	18,636	4730
Institution×grade FE	✓	✓	✓
Polynomial in natives enrolment	✓	✓	✓
Class level controls	✓	✓	✓
Hansen (p-value)	0.373	0.197	0.475
Natives (F-test)	243.61	230.40	92.60
Natives (AP-test)	130.66	125.43	49.28
Immigrants (F-test)	159.90	156.49	13.17
Immigrants (AP-test)	58.39	57.49	8.91

IV-FE estimates in the north and centre: Math

	Baseline specification	Classes with cheating indicator = 0	Externally monitored institutions
Number of natives: $\hat{\beta}$	-0.0008	-0.0010	-0.0019
	(0.0009)	(0.0009)	(0.0017)
Number of immigrants: $\hat{\gamma}$	-0.0129**	-0.0155***	-0.0160*
	(0.0055)	(0.0055)	(0.0084)
PEC: $\hat{\delta}$	-0.0121**	-0.0145***	-0.0141*
	(0.0052)	(0.0052)	(0.0079)
Observations	19,005	18,697	4733
Institution×grade FE	✓	✓	✓
Polynomial in natives enrolment	✓	✓	✓
Class level controls	✓	✓	✓
Hansen (p-value)	0.878	0.832	0.694
Natives (F-test)	165.09	169.27	103.96
Natives (AP-test)	137.02	135.53	50.37
Immigrants (F-test)	39.65	34.42	14.35
Immigrants (AP-test)	33.28	28.70	9.52

Concluding remarks

Anecdotal evidence of class disruption involving immigrants often generates concerns in the public opinion and drives policy reactions

We clarify that a useful policy parameter is the PEC effect

- the effect of substituing one native with one immigrant in class
- net of endogenous principals' reactions (in numbers and quality)
- net of the mechanical class size effects that these inflows entail

The institutional setting in Italy allows us to identify the PEC

Adding one immigrant to a class while taking away one native,

- reduces native performance in both language and math by approximately 1.6%
- these estimates are larger than conventional ones because they are not confounded by principals' reactions

Thank You

Questions and comments
are welcome
You can contact us at

rosariomaria.ballatore@bancaditalia.it
margherita.fort@unibo.it
andrea.ichino@eui.eu

IV-FE estimates: Language

	Pooled (1)	2nd grade (2)	5th grade (3)
Number of natives: $\hat{\beta}$	-0.0019**	-0.0025**	-0.0011
	(0.001)	(0.001)	(0.001)
Number of immigrants: $\hat{\gamma}$	-0.0177***	-0.0150*	-0.0182***
	(0.005)	(0.008)	(0.006)
(Pure) composition effect: $\hat{\delta}$	-0.0158***	-0.0125	-0.0171**
	(0.005)	(0.008)	(0.006)
Observations	25,943	12,859	13,084
Institution×grade FE	✓	✓	✓
Polynomial in natives enrolment	✓	✓	✓
Class level controls	✓	✓	✓
Hansen (p-value)	0.716	0.717	0.718
F test (excluded instruments)			
Natives	382.52	185.09	231.93
Immigrants	299.19	90.97	86.91

Back

IV-FE estimates: Math

	Pooled (4)	2nd grade (5)	5th grade (6)
Number of natives: $\hat{\beta}$	-0.0010	-0.0013	-0.0005
	(0.001)	(0.001)	(0.001)
Number of immigrants: $\hat{\gamma}$	-0.0168***	-0.0134	-0.0174**
	(0.005)	(0.008)	(0.006)
(Pure) composition effect: $\hat{\delta}$	-0.0158***	-0.0121	-0.0169**
	(0.005)	(0.008)	(0.007)
Observations	25,936	12,854	13,082
Institution×grade FE	✓	✓	✓
Polynomial in natives enrolment	✓	✓	✓
Class level controls	✓	✓	✓
Hansen (p-value)	0.972	0.778	0.866
F test (excluded instruments)			
Natives	221.76	141.52	194.88
Immigrants	46.46	48.59	76.97

Back

Actual and predicted number of natives in a class based on native enrolment, pooling grade 2 and 5