

Second generation at school

Integration and citizenship

Daniela Cocchi Francesco Giovinazzi



Department of Statistical Sciences
University of Bologna

Pisa, 8-10 May 2018

Workshop “Small Area Methods and Living Conditions Indicators in European Poverty Studies in the Era of Data Deluge and Big Data”

Final Event of the Jean Monnet Chair SAMPLEU

Outline of the presentation

- The idea: different level of integration by nationality.
- ISG data: brief description.
- Citizenship in the sample: distribution and association.
- Proposed small area models: multinomial and latent class.
- Final remarks

The *Integration of the Second Generation* survey

2015

Istat carries out the survey on *Integration of the Second Generation* (ISG), co-financed by the Italian Ministry of Interior and the European Fund for the Integration of third-country nationals (EFI).

ISG involved about 1400 lower and upper secondary schools attended by at least 5 foreign students in 20 Italian regions.

2018

Istat releases public microdata for research purposes.

2019

Istat will carry out a second wave of the survey.

The idea

Based on the survey on *Integration of the Second Generation* (ISG) (Istat, 2017) we aim to model social and educational integration of foreign students in Italy, considering citizenship as a study domain.

Is the process of integration promoted by the Italian Education system equally effective on children belonging to different nationalities?

ISG microdata for research purposes

Data consists of 68127 observations on 255 items.

The questionnaire investigated many different dimensions of social inclusion, and can be divided in 6 broad sections:

- A. Administrative data and migration history.
- B. Use of native and local languages.
- C. Relationship with schoolmates and teachers.
- D. Relationship with friends, free time and social habits.
- E. Composition of the family and relationship with its members.
- F. Household conditions.

Citizenship distribution in the sample

Foreigners
31687
(46.5%)

Italians
36440
(53.5%)

Born in Italy
9002 (28.4%)

Born abroad
22685 (71.5%)

Albania	1811 (20.1%)
Morocco	1080 (12%)
China	1015 (11.3%)
Romania	750 (8.3%)
Philippines	581 (6.4%)
Others	3765 (41.9 %)

Romania	5879 (25.9%)
Albania	2852 (12.6%)
Morocco	1555 (6.8%)
Moldova	1361 (6%)
Ukraine	1056 (4.7%)
Others	9982 (44%)

Denoting integration by a proxy

Among the many items in the questionnaire we selected A11 as best proxy of integration for a foreign student.

Item A11 – Do you feel more...

Italian

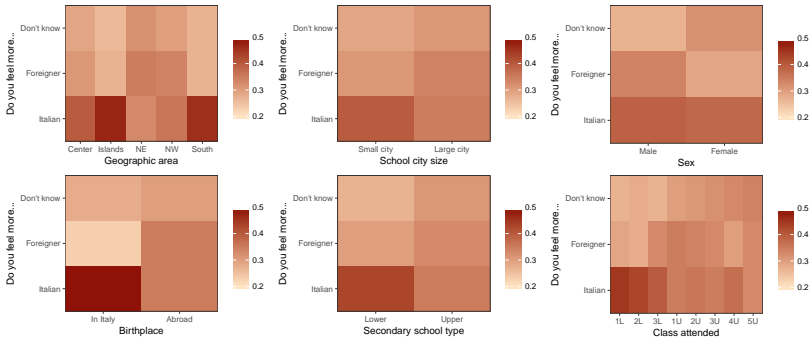
Foreigner

Don't know

Answers to A11 by citizenship

		A11 – Do you feel more...		
		Italian	Foreigner	Don't know
Total		12298 (38.8)	10039 (31.7)	9350 (29.5)
By birthplace	Born in Italy	4397 (48.8)	2105 (23.4)	2500 (27.8)
	Born abroad	7901 (37.8)	7934(35)	6850 (30.2)
By nationality	Albania	2031 (43.6)	1347 (28.9)	1285 (27.6)
	Romania	3002 (45.3)	1771 (26.7)	1856 (28)
	Ukraine	505 (44.2)	308 (26.9)	330 (28.9)
	Moldova	484 (33.9)	460 (32.2)	484 (33.9)
	China	390 (22.7)	696 (40.6)	630 (36.7)
	Philippines	330 (30.9)	387 (36.2)	351 (32.9)
	India	187 (31.6)	164 (27.7)	240 (40.6)
	Morocco	954 (36.2)	878 (33.3)	803 (30.5)
	Ecuador	262 (30.9)	332 (39.2)	254 (30)
	Peru	187 (26)	284 (39.5)	248 (34.5)
Others	3966 (38.7)	3412 (33.3)	2869 (28.0)	

A11 and some institutional items



A multinomial mixed-effects small area model

The area level model is (López-Vizcaíno *et al.*, 2013):

$$\log \left(\frac{p_{ijk}}{p_{ijK}} \right) = \alpha_{ik} + \beta_{1k} \mathbf{GA}_{ij} + \beta_{2k} \mathbf{SCS}_{ij} + \beta_{3k} \mathbf{S}_{ij} + \beta_{4k} \mathbf{BP}_{ij} + \beta_{5k} \mathbf{ST}_{ij},$$

where:

$i = 1, \dots, 10$ is the nationality of the student (the domain)

$j = 1, \dots, n_i$ is the size of the i -th domain

$k = 1, 2$ is the response level ($K = 3$ baseline of A11)

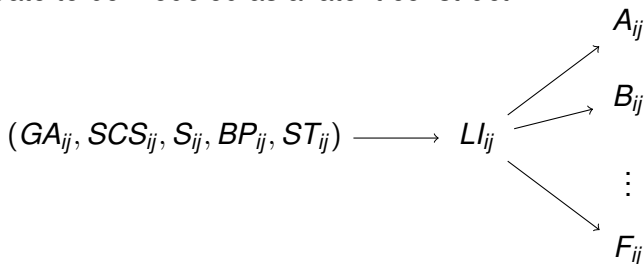
p_{ijk} is the probability to give answer k to item A11

α_{ik} is a normally distributed random effect

Choice of area level model motivated by the nature of the available population information.

A latent class small area model

The level of integration (LI_{ij}), as a multivariate unobservable phenomenon influencing many observable variables (items A_{ij}, B_{ij}, \dots following the questionnaire sectioning) is a very good candidate to be modeled as a latent construct:



A latent class small area model

The next step will be to replace A11 in the multinomial small area model with a categorical latent variable, implementing a Latent Class small area model, in the spirit of Fabrizi *et al.* (2016).

Final remarks

- We are currently waiting to receive the population official information that will allow to fit the proposed models.
- The work integrates with the European Cohort Development Project (ECDP), on the well-being of children and young adults in Europe, to which we participate.
- Difficulty of record linkage due to privacy constraint adopted by Istat (data are masked).

References

- FABRIZI, ENRICO, MONTANARI, GIORGIO E, & GIOVANNA RANALLI, MARIA. 2016. A hierarchical latent class model for predicting disability small area counts from survey data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **179**(1), 103–131.
- ISTAT. 2017. *Indagine sull'integrazione delle seconde generazioni: obiettivi, metodologia e organizzazione*. Letture statistiche - Metodi.
- LÓPEZ-VIZCAÍNO, ESTHER, LOMBARDÍA, MARÍA JOSÉ, & MORALES, DOMINGO. 2013. Multinomial-based small area estimation of labour force indicators. *Statistical Modelling*, **13**(2), 153–178.