

# HOW TO SAMPLE MIGRANTS IN GENERAL POPULATION SURVEYS WHEN POPULATION LISTS ARE NOT AVAILABLE?

**Ferruccio Biolcati, Riccardo Ladini, Francesco Molteni**

*University of Milan*

ESRA - European Survey Research Association - Conference 2023  
Survey research in times of crisis: Challenges, opportunities, and new directions

University of Milano Bicocca, Milan, 17-21 July 2023

## **BULLET POINTS LIST**

- **Sampling migrants in general population surveys**
- **Sure problems**
- **Possible solutions**
- **COValues survey**
- **Some results**
- **Preliminary evaluation and lessons learnt**



# SAMPLING MIGRANTS IN GENERAL POPULATION SURVEYS

- **Aim**

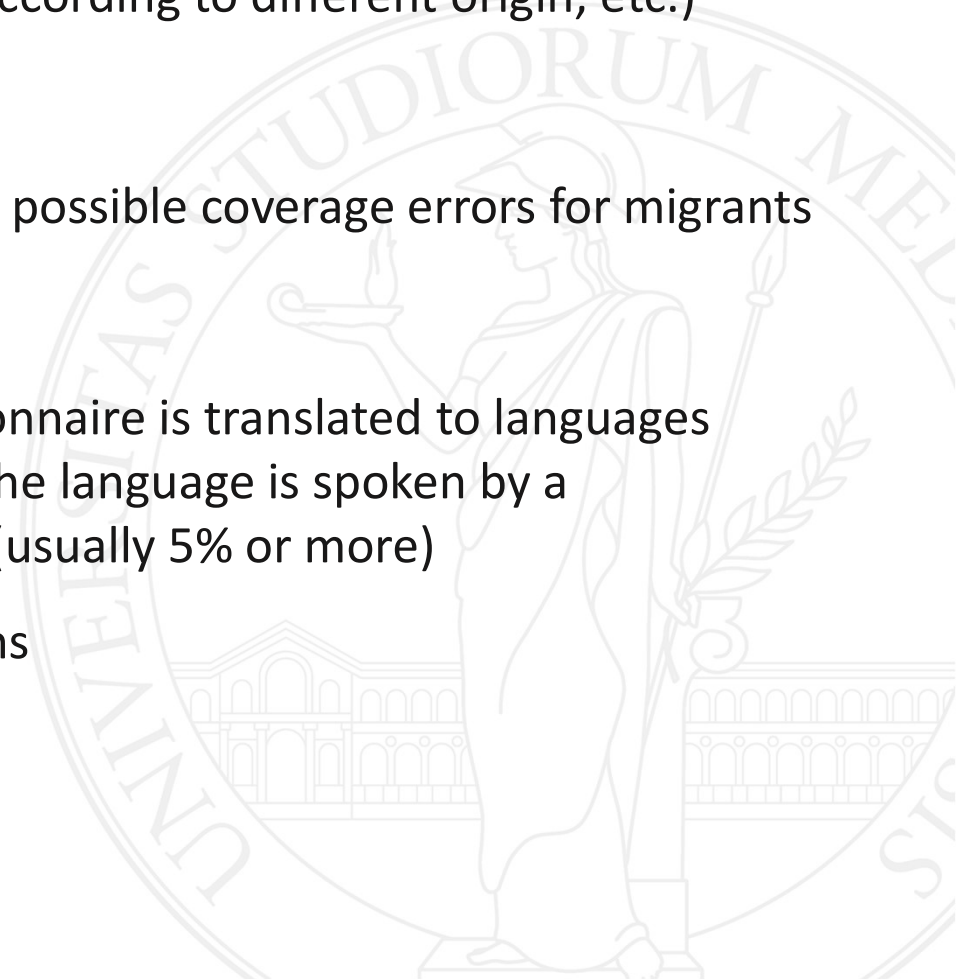
- Avoiding coverage error, comparing natives and migrants; not focusing on migrant population (according to different origin, etc.)

- **Sampling**

- Usually based on population lists with possible coverage errors for migrants

- **Questionnaire**

- Language comprehension: the questionnaire is translated to languages other to the national one only when the language is spoken by a considerable share of the population (usually 5% or more)
- Even attitudinal and complex questions

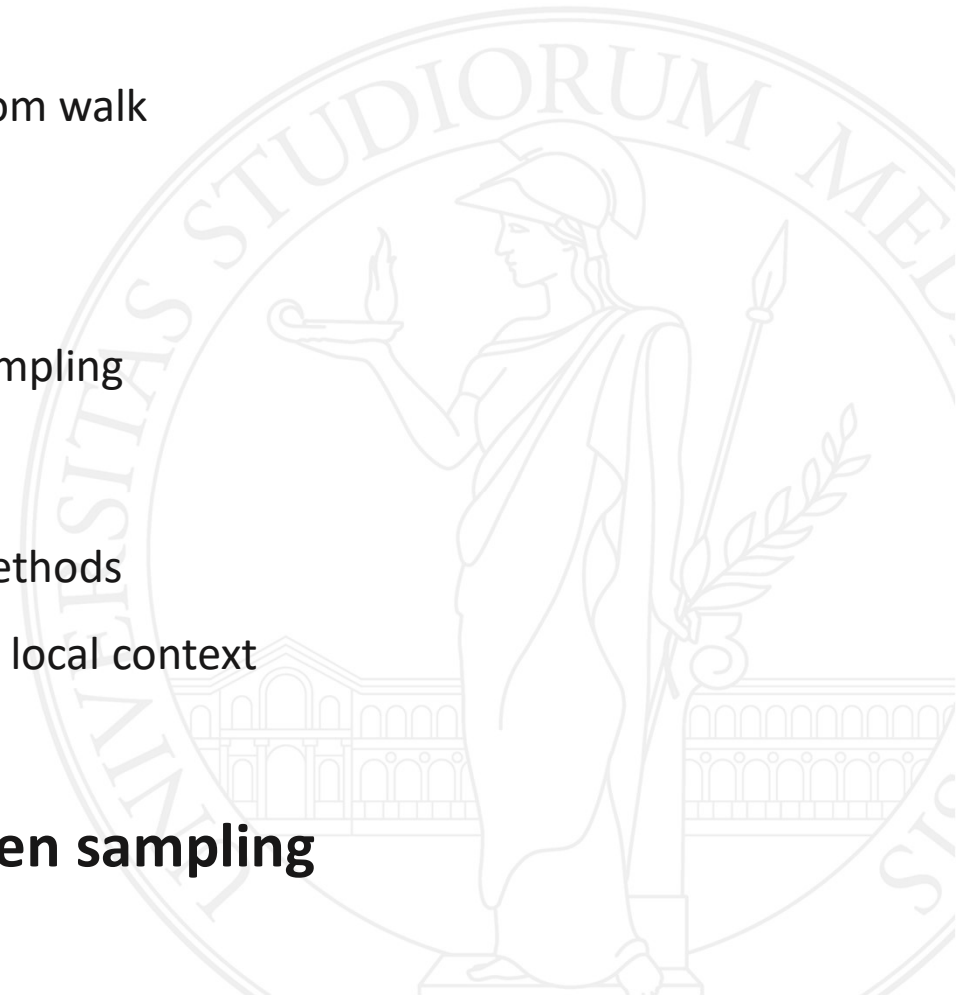


# SURE PROBLEMS: THE ITALIAN CASE

- **Target population**
  - People residing in a country
- **Register**
  - Population registers only available when the institution in charge of the survey is part of the National Statistical System (e.g. ESS), unfortunately academic institutions are not part of it
- **Electoral lists**
  - The current procedure foresees the use of electoral registers (e.g. EVS, ISSP, national surveys) where non-citizens are absent
- **Problem**
  - No representation of migrants (non-citizens) in general population surveys in some countries
- **Relevance**
  - Non-citizens are 8.5% of people over-18 residing in Italy (10.6% in Lombardy)
- **Further issues**
  - In comparative perspective, different target populations across countries; impossibility of comparisons with people with a migratory background residing in other countries (Italy excluded from the analysis); impossibility of comparing natives VS. migrants within the country;

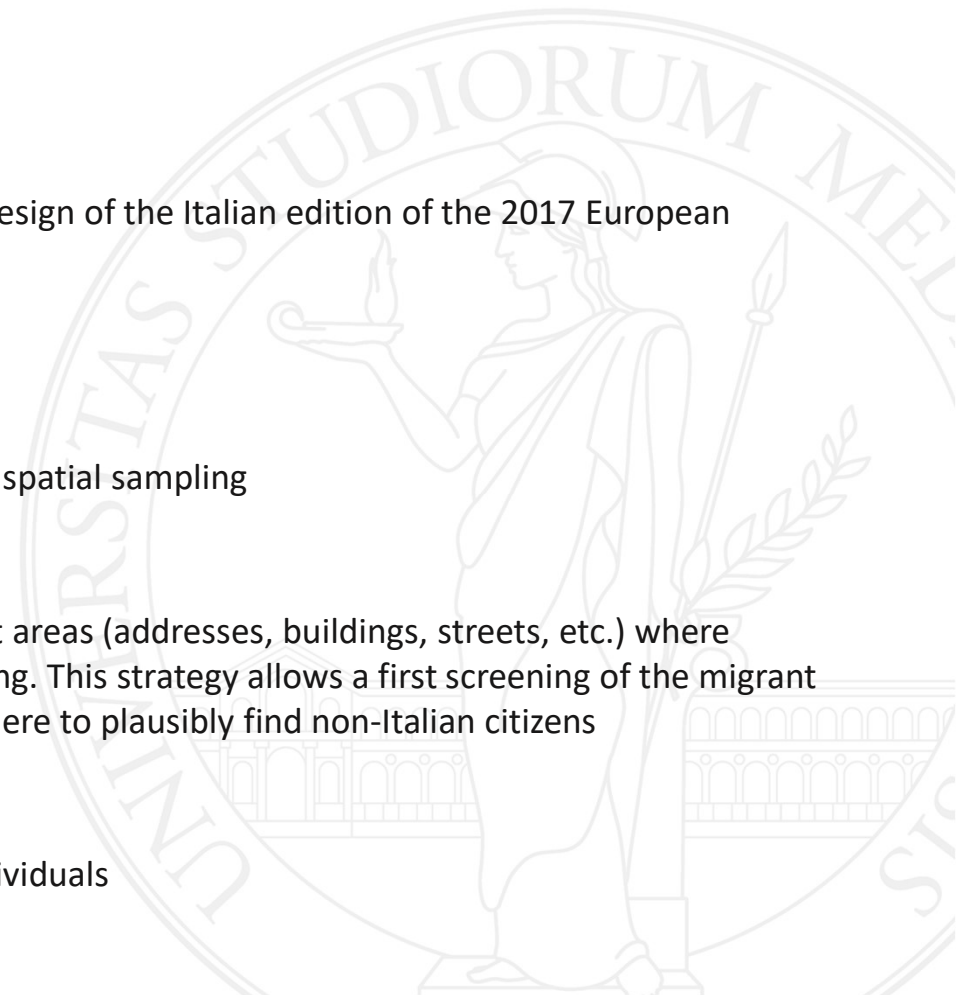
## POSSIBLE SOLUTIONS

- **Quota sampling or opt-in panel**
  - Loosing the probabilistic framework
- **Spatial sampling**
  - Random household, random routes, random walk
  - Different density of migrant population
  - Screening procedures
  - Focused enumeration, adaptive cluster sampling
- **Time-location sampling**
  - Centres of aggregation sampling, other methods
  - Necessary ethnographic knowledge of the local context
  - Interview context
- **Snowball sampling, Respondent-driven sampling**
  - Network bias



# COValues - VALUE CHANGE AFTER COVID-19 PANDEMICS IN LOMBARDY

- **Aim**
  - Studying value change after Covid-19 pandemics and native-migrant divide in values
- **Field**
  - Lombardy in late 2022
- **Italian population**
  - The survey employs the same probabilistic sampling design of the Italian edition of the 2017 European Values Study base on electoral lists
- **Foreign population**
  - Oversampling
  - To sample non-Italian citizens, it employs an informed spatial sampling
- **Informed spatial sampling**
  - Each respondent was asked to name the three nearest areas (addresses, buildings, streets, etc.) where he/she thoughts people with a foreign origin were living. This strategy allows a first screening of the migrant population, by providing to interviewers addresses where to plausibly find non-Italian citizens
- **Protocol**
  - A rather rigid protocol was employed for sampling individuals



# THE PROTOCOL

- **Question**

- “Please indicate three areas, the closest to your home, where you know that families with at least one member of foreign origin reside or assume that they reside. By foreign origin we mean the fact that the person was born abroad or from at least one parent born abroad, both in other European countries and in non-European countries.”

- **Indicating**

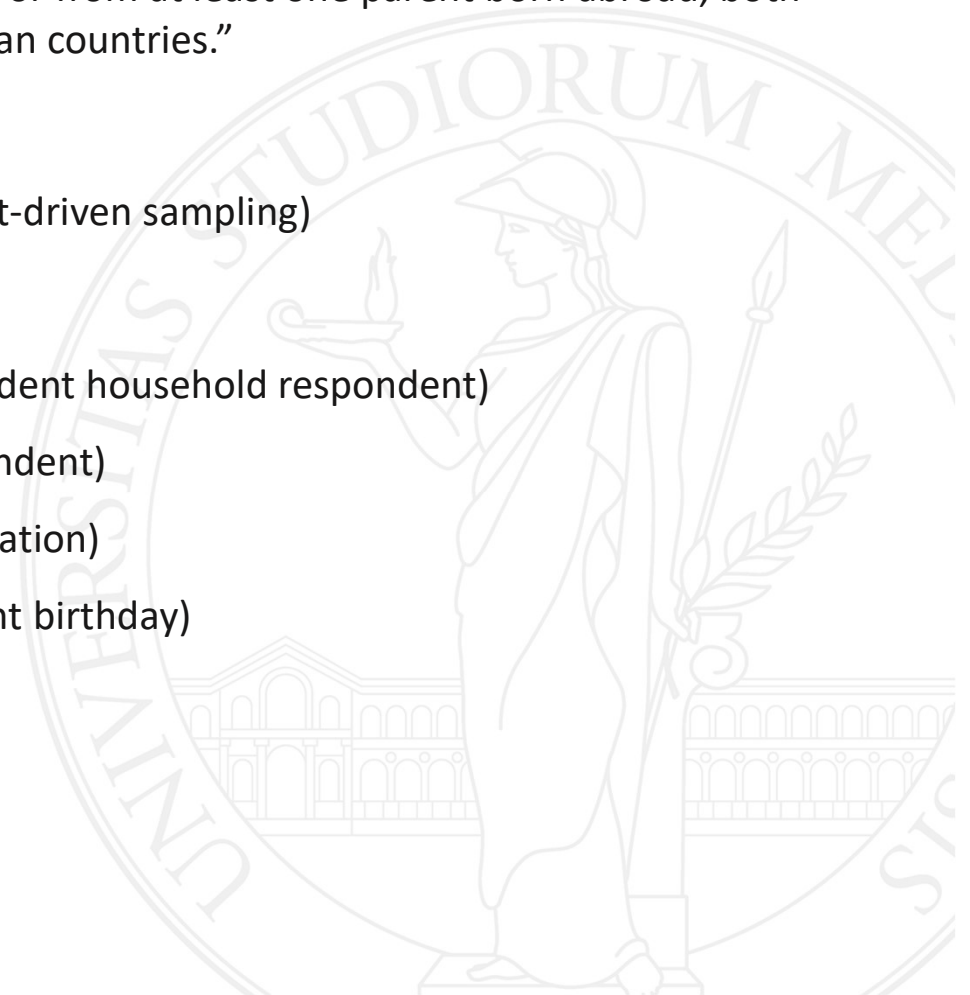
- Areas, not families or individuals (as respondent-driven sampling)

- **Selection criteria**

- Streets, squares, etc. (e.g. most close to respondent household respondent)
- Street numbers (e.g. odd number if male respondent)
- Households (issue of foreign surnames identification)
- Individual inside the household (e.g. most recent birthday)

- **Issues**

- Complexity
- Reproducibility but rigidity



## SOME RESULTS – Area of origin of foreign citizens coming from countries with strong migratory pressure, according to different sources

	ISTAT official statistics	ISTAT survey: CICS	ORIM	Labour Force Survey	ESS 9	COValues
Year	2022	2011	2021	2020	2018	2022
Region	Lombardy	Lombardy	Lombardy	Lombardy	Italy	Lombardy
Age	>= 0	>= 15	>= 18	>= 15	>= 16	>= 18
N	1,094,475	15,164	2,227	4,890	130	136
Sampling	Register	Register (HH)	Centres of aggregation	Register (HH)	Register	Informed spatial
Notes	At least 1,000 individuals				At least 3 respondents	
AREA OF ORIGIN						
Eastern Europe – EU						
Eastern Europe – extra-EU						
Asia						
North Africa						
Africa Others						
Latin America						
COMPARISON TO OFFICIAL STATISTICS						
Eastern Europe – EU						
Eastern Europe – extra-EU						
Asia						
North Africa						
Africa Others						
Latin America						
Sum of absolute values						



## SOME RESULTS – Area of origin of foreign citizens coming from countries with strong migratory pressure, according to different sources

	ISTAT official statistics	ISTAT survey: CICS	ORIM	Labour Force Survey	ESS 9	COValues
Year	2022	2011	2021	2020	2018	2022
Region	Lombardy	Lombardy	Lombardy	Lombardy	Italy	Lombardy
Age	>= 0	>= 15	>= 18	>= 15	>= 16	>= 18
N	1,094,475	15,164	2,227	4,890	130	136
Sampling	Register	Register (HH)	Centres of aggregation	Register (HH)	Register	Informed spatial
Notes	At least 1,000 individuals				At least 3 respondents	
AREA OF ORIGIN						
Eastern Europe – EU	17,3	28,4	18,8	20,0	28,5	12,7
Eastern Europe – extra-EU	17,0	28,5	21,0	19,6	24,6	15,9
Asia	27,0	15,2	19,5	23,8	19,2	13,5
North Africa	18,8	14,8	17,2	15,9	11,5	25,3
Africa Others	7,9	6,0	10,7	6,7	10,8	17,5
Latin America	12,0	7,1	12,8	14,0	5,4	15,1
COMPARISON TO OFFICIAL STATISTICS						
Eastern Europe – EU		11.1	1.5	2.7	11.2	-4.6
Eastern Europe – extra-EU		11.5	4.0	2.6	7.6	-1.1
Asia		-11.8	-7.5	-3.2	-7.7	-13.5
North Africa		-4.0	-1.6	-2.9	-7.3	6.5
Africa Others		-1.9	2.8	-1.2	2.9	9.6
Latin America		-4.9	0.8	2.0	-6.7	3.1
Sum of absolute values		45.3	18.2	14.6	43.4	38.2

## SOME RESULTS – Area of origin of foreign citizens coming from countries with strong migratory pressure, according to different sources

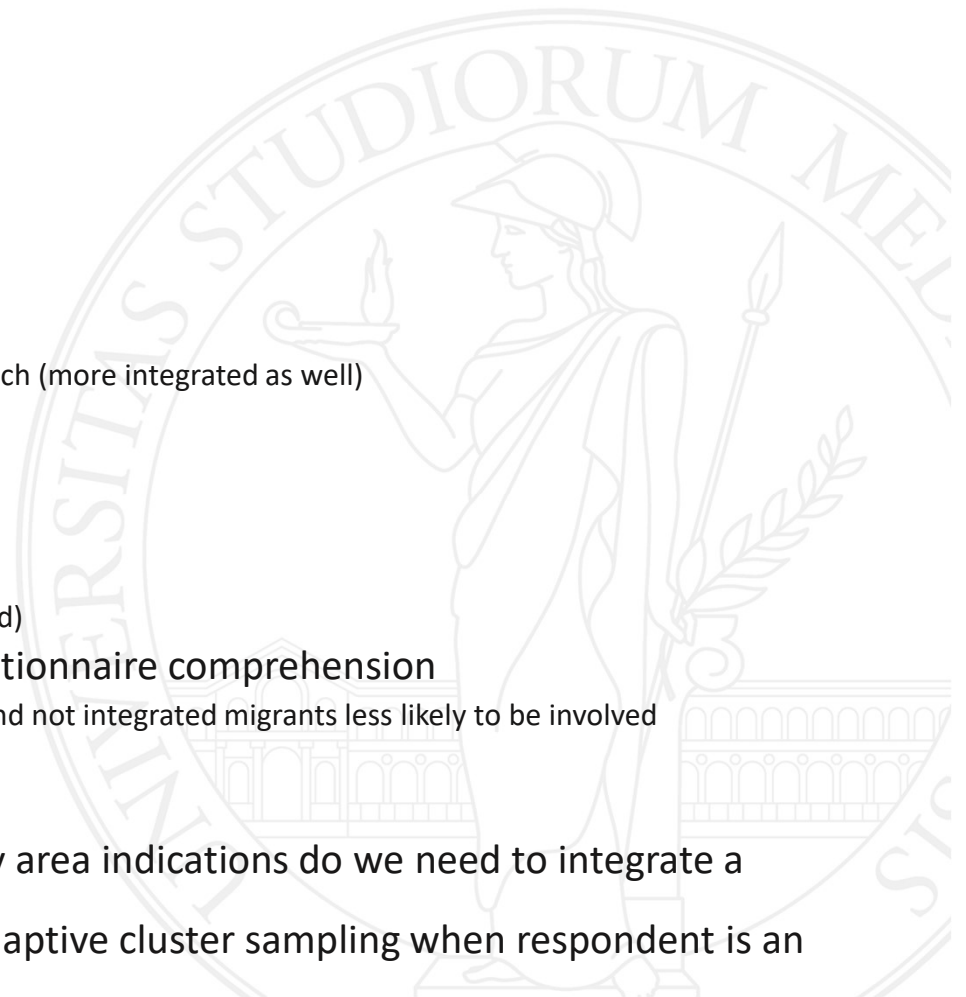
	ISTAT official statistics	ISTAT survey: CICS	ORIM	Labour Force Survey	ESS 9	COValues
Year	2022	2011	2021	2020	2018	2022
Region	Lombardy	Lombardy	Lombardy	Lombardy	Italy	Lombardy
Age	>= 0	>= 15	>= 18	>= 15	>= 16	>= 18
N	1,094,475	15,164	2,227	4,890	130	136
Sampling	Register	Register (HH)	Centres of aggregation	Register (HH)	Register	Informed spatial
Notes	At least 1,000 individuals				At least 3 respondents	
AREA OF ORIGIN						
Eastern Europe – EU	17,3	28,4	18,8	20,0	28,5	12,7
Eastern Europe – extra-EU	17,0	28,5	21,0	19,6	24,6	15,9
Asia	27,0	15,2	19,5	23,8	19,2	13,5
North Africa	18,8	14,8	17,2	15,9	11,5	25,3
Africa Others	7,9	6,0	10,7	6,7	10,8	17,5
Latin America	12,0	7,1	12,8	14,0	5,4	15,1
COMPARISON TO OFFICIAL STATISTICS						
Eastern Europe – EU		11.1	1.5	2.7	11.2	-4.6
Eastern Europe – extra-EU		11.5	4.0	2.6	7.6	-1.1
Asia		-11.8	-7.5	-3.2	-7.7	-13.5
North Africa		-4.0	-1.6	-2.9	-7.3	6.5
Africa Others		-1.9	2.8	-1.2	2.9	9.6
Latin America		-4.9	0.8	2.0	-6.7	3.1
Sum of absolute values		45.3	18.2	14.6	43.4	38.2

## SOME RESULTS – Questionnaire comprehension according to the interviewer by migration status (row %)

	< 50% questions	> 50% questions	Almost all questions	(n)
Italian citizens born in Italy	1.3	11.9	86.8	(395)
Italian citizens born abroad	5.6	19.4	75.0	(72)
Foreign citizens born abroad	5.2	45.6	49.2	(136)

# A PRELIMINARY EVALUATION

- **Suboptimal but a possibility to overcome the issue of non-representativeness of migrants in general population surveys**
  - Reducing coverage error
  - Building on traditional sampling design
  - Mix of techniques
- **Limitations**
  - Probability sampling
  - Biases
    - Possible underestimation of migrants not perceived as such (more integrated as well)
    - Possible overestimation of more visible migrants
  - Protocol difficulties
    - Complex protocol and difficult control over interviewers
  - Screening needed
    - Some of the respondents are Italian citizens (about a third)
  - Trade-off between representativeness and questionnaire comprehension
    - Questionnaire only in Italian (and interviewers): recent and not integrated migrants less likely to be involved
- **Next steps**
  - Implementation to a national survey: how many area indications do we need to integrate a general population survey in Italy?
  - Possible alternative: focused enumeration or adaptive cluster sampling when respondent is an Italian citizen born abroad



**THANK YOU!**

 [ferruccio.biolcati@unimi.it](mailto:ferruccio.biolcati@unimi.it)

 [@FerruBiolca](https://twitter.com/FerruBiolca)

**spSTREND**

 [www.spstrend.it](http://www.spstrend.it)

 [@SpsTrend](https://twitter.com/SpsTrend)



## SOME RESULTS - Next steps

- **Considering countries with weak migratory pressure**
- **Answers quality analysis**
  - Scales reliability
  - Non-response
- **Second generation migrants**

