



Using SAE methods to estimate the economic living conditions of immigrant communities in Italy

Francesco Schirripa^{1,2} Antonella D'Agostino^{1,3}
Nicola Salvati^{2,3}

¹DISAQ, Università Napoli "Parthenope"

²DEM, Università di Pisa

³Centro ASESD Camilo Dagum

Workshop Jean Monnet Chair SAMPLEU

Pisa, 9 May 2018



Overview

- 1 Introduction
- 2 Small area estimation method
- 3 Data, model specification and diagnostics
- 4 Results
- 5 Conclusions and future works



- 1 Introduction
- 2 Small area estimation method
- 3 Data, model specification and diagnostics
- 4 Results
- 5 Conclusions and future works



Foreign population in Italy

- The migrant population, particularly those born outside the EU, tends to have a **higher risk of poverty and is more likely to be socially excluded than the native population**
- Moreover, the living conditions of immigrants are even more difficult in Italy for several reasons:
 - ⇒ International migration flow is quite recent in Italy, but the share of foreign people in Italy has been increasing over time: in 2001 they were about 1 million, at the end of 2014 they were over 5 million (8,2% of the population)
 - ⇒ Imperfect immigrant-integration policies
 - ⇒ Economic crisis



Consequences

- Social problems
- Marginalization and racism
- Excessive income inequality in the overall population
- Other poverty related issues (the rising of criminality, the decreasing of educational opportunities of future generations...)

Study the living conditions of immigrants may be very interesting for the Italian political agenda in order to provide useful data and implement anti-poverty policies.



Economic Standard of Living

- Economic aspects of the standard of living
- Economic condition is relatively easy to quantify
- Disposable income is the target variable used in poverty analysis
- Economic disparities between migrants and natives play a fundamental role in the assimilation process of international migrants



Role of communities

- Foreign population is very heterogeneous
- Immigrants come from a wide range of countries with different cultures and their level of assimilation can vary: **immigrants tend to be integrated in the host Country as communities (ethnic groups) and not as individuals.**
- According to the last Population Census 2011 there are 124 different ethnic groups and the majority of them (about 80%) identify international migrants from developing and low-income countries:



Role of communities

⇒ Considering immigrant communities as the core of the analysis is an interesting way to find homogeneous subgroups for measuring the economic standard of living of immigrant population

⇒ Measuring the differences in economic standard of living of immigrant communities is an useful indicator in order to evaluate disparities between immigrants and the native population



Sample size problem

- Small sample size in the majority of communities (even zero!)
- High variability of direct estimators
- Policy-maker do not have a correct view of the phenomenon in order to establish their decisions

Solution: Small Area Estimation Methods (SAE)

We propose to treat communities as “small areas” and use SAE methods for computing an estimation of monetary conditions (represented by mean of household equivalised income) of every immigrant community.



- 1 Introduction
- 2 Small area estimation method
- 3 Data, model specification and diagnostics
- 4 Results
- 5 Conclusions and future works



EBLUP: Unit Level Approach

A vector of p auxiliary variables \mathbf{x}_{dj} is known for each population unit j in small area d and the values y_{dj} are available only for the sampled units.

Super population model (Battese et al, 1988):

$$y_{jd} = \mathbf{x}_{jd}\boldsymbol{\beta} + z_d u_d + e_{dj}$$

- u_d are area-specific random effects, $u_d \sim N(0, \sigma_u^2)$
- e_{dj} are the unit error, $e_{dj} \sim N(0, \sigma_{e_d}^2)$
- u_d and e_{dj} are independent
- z_d is a vector with the value 1 in d -th position and zeros otherwise.



M-quantile approach

A different approach has been proposed by Chambers and Tzavidis (2006) based on M-quantile regression:

- relaxes the assumptions of the mixed effect models
- assures outlier robust estimation

A linear M-quantile regression for the q^{th} quantile of $f(y|\mathbf{x})$ is defined by

$$MQ_y(\mathbf{x}; \psi) = \mathbf{x}^T \beta_\psi(q)$$

- estimates of $\beta_\psi(q)$ are obtained via an iterative weighted least squares algorithm
- ψ is an influence function that controls the effect of outliers and usually is assumed to be Huber Proposal 2 function, $\psi(r) = r I(|r| \leq c) + \text{sgn}(r) c I(|r| > c)$, with tuning constant $c = 1.345$



M-quantile models in Small Area Estimation (1)

- Given individual data of y and x , each sampled value will lie on **one and only one** regression M-quantile line.
- So each sampled value have a specific M-quantile coefficient (q_{dj})
- The M-quantile coefficient q_d for area d can be calculated by suitable averaging the M-quantile coefficients, q_{dj} , of the units in the area d and the M-Quantile small area model is:

$$y_{ij} = \mathbf{x}^T \beta_{\psi}(\hat{q}_d) + \varepsilon_{ij}$$

A mixed effects model uses random effects u_j to capture the differences between groups. M-quantile models try to capture these differences via the group-specific M-quantile coefficients \hat{q}_d .

M-quantile models in Small Area Estimation (2)

A bias-adjusted M-quantile predictor for the mean of y in the area d is (Tzavidis et al., 2010):

$$\hat{Y}_d^{MQ/CD} = N_d^{-1} \left[\sum_{j \in s_d} y_{dj} + \sum_{j \in r_d} \hat{y}_{dj}^{\psi} + \frac{N_d - n_d}{n_d} \sum_{j \in s_d} (y_{dj} - \mathbf{x}_{dj}^T \hat{\beta}_{\psi}(\hat{q}_d)) \right]$$

where $\hat{y}_{dj}^{\psi} = \mathbf{x}_{dj}^T \hat{\beta}_{\psi}(\hat{q}_d)$ is the prediction based on the M-quantile model for non-sampled values.

For the non-sampled areas a synthetic M-median regression estimator can be used (Chambers and Tzavidis 2006):

$$\hat{Y}_d^{MQ} = \bar{\mathbf{X}}_d^T \hat{\beta}_{\psi}(0.5)$$

where $\bar{\mathbf{X}}_d$ represents the vector of covariates for small area d obtained from census.



- 1 Introduction
- 2 Small area estimation method
- 3 Data, model specification and diagnostics
- 4 Results
- 5 Conclusions and future works



Data - EU SILC-FP

- Survey on Income and Living Conditions of Household with Foreign People (EUSILC-FP) carried out by Istat for the first time in 2009 on a sample of about 6,000 households with at least one foreign member
- Immigrants are individuals living in Italy without Italian citizenship: we analyse the household income of households with foreign household head
- Sample counts $n = 5031$ households that are grouped in 115 communities according to the ethnic identity of their household head
- A set of explanatory variables is available for each community from the Population Census 2011



Preliminary analysis

Table: Communities by CV of Direct Estimates

CV of Direct Estimates	Acceptable	Marginal	Unacceptable	Total
Number of communities	13	26	76	115

- Acceptable: CV is in the range of 0.0% to 16.5%.
- Marginal: CV is in the range of 16.6% to 33.3%.
- Unacceptable: CV greater than 33.3%.

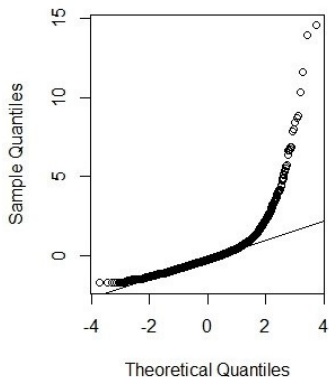


Supplementary information

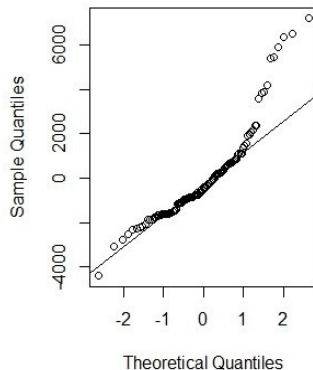
Supplementary information is fundamental in order to increase the efficiency of estimates for the small areas. This supplementary information is represented by common covariates at unit and area level. In this study we used as supplementary information the following set covariates:

- the region of residence (South Italy vs North-Central Italy)
- the educational level of the head of household (graduated – e.g. ISCED level 05 – vs not-graduated)
- the employment status of the head of household (working vs not working)
- the duration of stay in Italy of the head of household (less than three years vs more than three years)

Diagnostics for Mixed Effects Model (1)

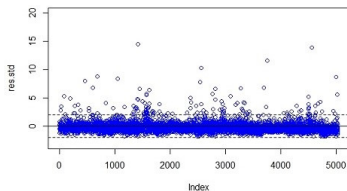


(a)

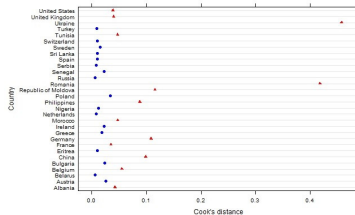


(b)

Diagnostics for Mixed Effects Model (2)



(c)



(d)



- 1 Introduction
- 2 Small area estimation method
- 3 Data, model specification and diagnostics
- 4 Results
- 5 Conclusions and future works

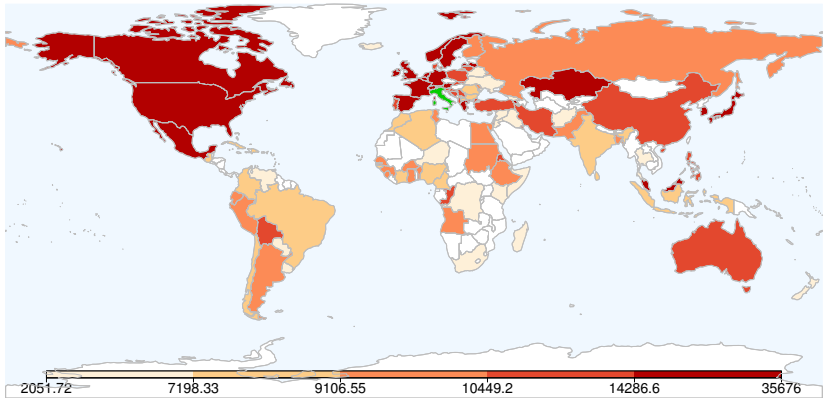


Estimation using M-quantile approach

Table: Communities by CV of MQ Estimates

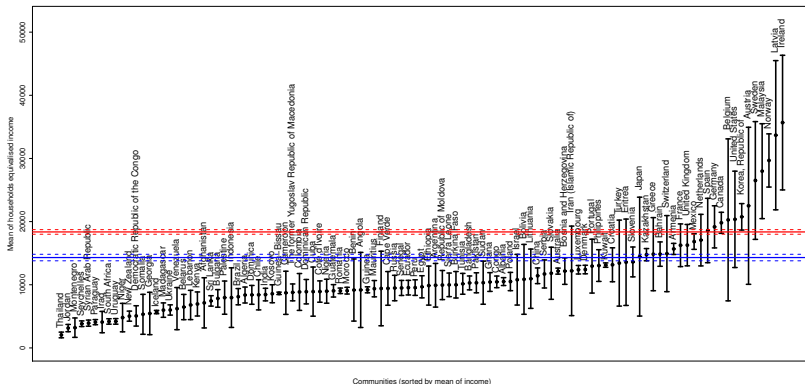
CV of MQ Estimates	Acceptable	Marginal	Unacceptable	Total
Number of communities	75	30	10	115

Mapping household income in each community





Comparing income of immigrant communities





- 1 Introduction
- 2 Small area estimation method
- 3 Data, model specification and diagnostics
- 4 Results
- 5 Conclusions and future works



Conclusions (1)

- SAE methods allow us to measure with a good level of precision the economic conditions of each immigrant community living in Italy
- Ensuring immigrants adequate economic resources can reduce integration problems and other poverty-related issues
- Having a level of income lower than that of the Italian population is a great obstacle for the full integration in the economic and social life and the risk of social exclusion for migrants of these communities is very high
- There are differences among citizenship groups in term of economic conditions and if policy makers disregard this evidence there could be a waste of resources.



Conclusions (2)

- The idea of one size fits all policy could not be the best solution for preventing the social exclusion of specific communities and findings could be used for implementing more adequate policies targeting on the most disadvantaged ethnic groups
- Study other aspects of the integration (poverty and social exclusion are multidimensional phenomena)
- Irregular immigration

Francesco Schirripa



francesco.schirripa@ec.unipi.it