Small Area Estimation of Latent Economic Wellbeing

Angelo Moretti, Natalie Shlomo and Joseph Sakshaug

Social Statistics, School of Social Sciences, University of Manchester, United Kingdom

Summary. Small area estimation (SAE) plays a crucial role in the social sciences due to the growing need for reliable and accurate estimates for small domains. In the study of wellbeing, for example, policy-makers need detailed information about the geographical distribution of a range of social indicators. We investigate data dimensionality reduction using factor analysis models and implement SAE on the factor scores under the empirical best linear unbiased prediction approach. We contrast this approach with the standard approach of providing a dashboard of indicators, or a weighted average of indicators at the local level. We demonstrate the approach in a simulation study and a real application based on the European Union Statistics for Income and Living Conditions (EU-SILC) for the municipalities of Tuscany.

Keywords: Composite estimation; Direct estimation; EBLUP; Factor analysis; Factor scores; Model-based estimation.

1. Introduction

Measuring poverty and wellbeing is a key issue for policy makers requiring a detailed understanding of the geographical distribution of social indicators. This understanding is essential for the formulation of targeted policies that address the needs of people in specific geographical locations. Most large-scale social surveys provide accurate estimates only at a national level. A relevant survey in the European Union (EU) for analyzing wellbeing is the European Union Statistics for Income and Living Conditions (EU-SILC). However, this data can be used to produce accurate direct estimates only at the NUTS (Nomenclature of Territorial Units for Statistics) 2 level. In Italy the NUTS 2 level refers to regions (e.g. Tuscany). Hence, if the goal is to measure poverty and well-being indicators at a sub-regional level, such as NUTS 3 or LAU (Local Administrative Units) 2, the indicators cannot be directly estimated from EU-SILC. In fact, the domains corresponding to the levels under NUTS 2 are so called *unplanned domains* where domain membership is not incorporated in the sampling design, and therefore the sample size in each domain is random and in many cases zero. In this case, indirect model-based estimation methods, in particular small area estimation approaches can be used to predict target parameters for the small domains.

Small area estimation (SAE) is defined as a set of statistical procedures with the goal of producing efficient and precise estimates for small areas, as well as for domains with zero sample size (Rao and Molina, 2015). According to Rao and Molina (2015), an area is *small*, if the area-specific sample size is not large enough to provide *precise* and efficient direct design-based estimates. Small areas can also be defined by the cross-classification of geographical areas by social, economic or demographic characteristics.

SAE methods can be classified into two approaches: the unit-level and the area-level approach. The unit-level approach is used when covariates are available for each unit of the population, for example from census or administrative data, while the area-level approach is used when covariate information is known only at the area level. The use of the error-components model by Battese Harter and Fuller (1988), also known as Battese, Harter and Fuller (BHF) model, is commonly used for the unit-level SAE approach. In the SAE literature, estimation methods include empirical best linear unbiased prediction (EBLUP), empirical Bayes (EB), and hierarchical Bayes (HB). The EBLUP method can be used under linear mixed models, while the EB and HB methods can be used under generalized linear mixed models. For a good review of these methodologies and their extensions we refer to Rao and Molina (2015).

A second important issue we consider in this paper is the multidimensionality of wellbeing indicators. An interesting work on the statistical properties of multidimensional indicators obtained by multivariate statistical methods is Krishnakumar and Nagar (2008). Although it is generally agreed that wellbeing are multidimensional phenomena (OECD, 2013), there is a continuing debate about the suitability of combining the social indicators using composite estimates based on averaging social indicators, or using a dashboard of single indicators. On the one hand, Ravallion (2011) argues that a single multidimensional index in the context of poverty leads to a loss of information, and on the other hand, Yalonetsky (2012) points out that composite estimates are necessary when the aim is measuring multiple deprivations within the same unit (individual or household). For a theoretical review of the multivariate statistical techniques and the related problems we refer to Bartholomew et al. (2008).

Taking this latter view, an approach to reducing data dimensionality is to consider the multidimensional phenomena as a latent variable construct measurable by a set of observed variables, and estimated using a factor analysis (FA) model. Factor scores are estimated from a FA model and are defined as a *composite variable* computed from more than one response variable. Indeed, factor scores provide details on each unit's placement on the factor. Although there is some literature in the exploratory FA modeling approach, the use of factor scores in the confirmatory FA modeling approach is not very diffused in the literature (DiStefano, Zhu, and Mindrila 2009). When we have a substantive framework where a set of variables explains a latent construct, the confirmatory FA modeling approach can be used. This is also true when a measurement framework is provided by official statistics or international organizations. In the context of wellbeing measurement the vector of unobserved variables represents a set of variables that jointly describe the underlying phenomenon (Sosa-Escudero, Caruso, and Svarc 2013). Other works on the use of factor analysis in latent wellbeing measurement are Ferro Luzzi, Fluckiger, and Weber (2008) and Gasparini et al (2011). These authors propose the use of factor analytic models to obtain a set of relevant factors resulting in factor scores in order to reduce the data dimensionality.

Once factor scores are estimated, they can be used to conduct other statistical analysis. For instance, they can be used as part of regression or predictive analyses to answer particular research questions. Kawashima and Shiomi (2007) use factor scores in order to conduct an ANOVA analysis on high school students' attitudes towards critical thinking and tested differences by grade level and gender. In addition, Bell, McCallum, and Cox (2003) investigated reading and writing skills where they extracted the factors and estimated factor scores in order to do a multiple regression analysis.

In the current literature on SAE of social indicators, there is a research gap on the estimation of multidimensional indicators. In particular, the use of factor scores and factor analysis in SAE models is an open area of research. This research area is important when we deal with having to reduce the data dimensionality in the estimation of social indicators at a local level.

In this paper, we consider economic wellbeing as a latent variable construct with the aim of reducing the dimensionality of wellbeing indicators. We then implement the SAE method of BHF on the factor scores in both a simulation study and on real data from EU-SILC for the region of Tuscany, Italy. This paper is organized as follows. In section 2, we describe the FA model for reducing data dimensionality on a dashboard of economic wellbeing indicators. In section 3 we review the unit-level SAE approach, and present the point estimation under the EBLUP for small area means. In section 4, we show results of a simulation study considering factor scores to deal with data dimensionality reduction and contrast them to the approach of weighting single univariate EBLUPs on the original variables using simple weights and weights defined by the factor loadings. In section 5, we discuss multidimensional economic wellbeing in Italy considering indicators from the Italian framework BES (Equitable and Sustainable Wellbeing) 2015 (ISTAT 2015). In section 6, using real data from EU-SILC 2009 for the area of Tuscany, we apply the proposed method and compute small area EBLUPs for factor score means and their mean squared error (MSE) for each Tuscany municipality (LAU 2). Finally, in section 7, we conclude the work with some final remarks and a general discussion.

2. Using Factor Scores for Data Dimensionality Reduction

Multivariate statistical methods aim to reduce the dimensionality of a multivariate random variable $\boldsymbol{Y}$. Formally, starting from a R^Kspace, where K denotes the number of variables, and we want to represent the observations in a reduced space $R^M$with $M \ll K$. Bartholomew et al. (2008) suggests several multivariate statistical techniques in order to deal with data dimensionality reduction in the social sciences (e.g. principal component analysis, factor analysis models, multiple correspondence analysis, etc.). In this work we consider the linear single-factor model, where the factor can be interpreted as a latent characteristic of the individuals revealed by the original variables.This model allows us to make inference on the population since the observable variables are linked to the unobservable factor by a probabilistic model (Bartholomew et al. 2008).

2.1 The Linear Single-factor Analysis Model

Let us consider K observed continuous variables, $y_1, y_2, \dots, y_K$, and we assume that these variables are dependent on a single factor f. Thus, we can write the following linking model:

$$y_k = \mu_k + \lambda_k f + \epsilon_k, k = 1, \dots, K, \tag{1}$$

where f denotes the common factor, common to all the observed variables, ϵ_k **the** residual term, and μ_k denotes the mean of the k observed variables. λ_k is the factor loading related to y_k and this tells us how strongly y_k depends on the factor.

We assume the following (Bartholomew et al. 2008):

1. $cov(f, \epsilon_k) = 0$**;**
2. $\epsilon_k \sim N(0, \sigma_k)$;
3. ϵ_k with $k = 1, \dots, K$ are independent.

2.2 Factor Scores

After the model parameters have been estimated we can estimate factor scores, i.e. the estimate of the unobserved latent variable for each unit i *(i=1,…n)*. Many estimation methods for the factor scores are available in the current literature and for a technical review we refer to Johnson and Wichern (1998). Clearly, the factor score estimation method depends on the type of the observed variable (binary, continuous etc.). For continuous variables, we use the *regression method* where an estimate of the individual unit *i* factor score is given by the following estimator:

$$\hat{f}_i = c_1(y_{1i} - \bar{y}_1) + \cdots . + c_k(y_{ki} - \bar{y}_k), i = 1, \dots, n \tag{2}$$

where c_k are the factor coefficients from the single factor model calculated as the loadings λ_k multiplied by the inverse of the covariance matrix (see Bartholomew, et al. 2008). The extension to factor models with multiple factors is straight-forward. The factor analysis and estimation of factor scores were carried out in Mplus, Version 7.4 (Muthén and Muthén, 2012).

In the application presented in section 5, we also have binary dependent variables. According to Muthén and Muthén (2012) *logistic regression* is employed for binary dependent variables where the following transformation is applied in a single-factor model:

$$logit\ [\pi_k(f)] = log \frac{\pi_k(f)}{1-\pi_k(f)} = \lambda_k f + \mu_k + \epsilon_k\ , k = 1, \dots, K. \tag{3}$$

where $\pi_k(f)$ denotes the probability that the dependent variable is equal to one, and $\frac{\pi_k(f)}{1-\pi_k(f)}$ the odds. We can then write the following expression:

$$\pi_k(f) = \frac{\exp(\lambda_k f + \mu_k + e_k\ )}{1+\exp(\lambda_k f + \mu_k + e_k\ )}, \tag{4}$$

monotonic in f and with domain in the interval [0,1].

In the presence of binary and continuous observed variables and under a maximum likelihood estimation approach, the factor scores may be estimated via the expected posterior method described in Muthén (2012) and applied in Mplus, Version 7.4.

3. Small Area Estimation using Empirical Best Linear Unbiased Prediction (EBLUP)

A class of models for SAE is the mixed effects models where we include random area-specific effects in the models and take into account the *between* area variation.

3.1. Notation

Let $d = 1, \dots, D$ denote small areas for which we want to compute estimates of the target population parameter for each d, in our case the population mean $\bar{Y}_d$ of a wellbeing variable or $\bar{F}_d$ of the factor score. For a sample $s \subset \Omega$ of size n drawn from the target population of size N, the non-sampled units, $N - n$ are denoted by r. Hence, $s_d = s \cap \Omega_d$ is the sub-sample from the small area d of size n_d, $n = \sum_{d=1}^{D} n_d$, and $s = \cup_d s_d$. r_d denotes the non-sampled units for the small area d of $N_d - n_d$ dimension.

3.2. Model based prediction using EBLUP

We consider the small area estimation problem for the mean under the EBLUP approach in the BHF model. Focusing on the population parameter of factor score means $\bar{F}_d, d = 1, \dots, D$, and as the population mean is a linear quantity, we can write the following decomposition:

$$\bar{F}_d = N_d^{-1}\left(\sum_{i \in s_d} f_{di} + \sum_{i \in r_d} f_{di}\right). \tag{5}$$

where f_{di} is the population factor score for unit *i* within small area *d* assuming that the factor model is implemented on the whole population. We note that we drop the '^' from the factor scores estimated in formula (2) as we assume that any error in the dependent variable estimated as a *composite* variable will be accounted for in the error terms of the regression model in the BHF approach.

When auxiliary variables are available at the unit level the BHF model can be used in order to predict the out-of-sample units. Considering the data for unit i in area d being $(f_{di}, \boldsymbol{x}_{di}^T)$ where $\boldsymbol{x}_{di}^T$ denotes a vector of p auxiliary variables, the nested error regression model is the following:

$$\begin{gathered} f_{di} = \boldsymbol{x}_{di}^T \boldsymbol{\beta} + u_d + e_{di}, i = 1, \dots, N_D, d = 1, \dots, D \\ u_d \sim \mathrm{iidN}(0, \sigma_u^2), e_{di} \sim \mathrm{iidN}(0, \sigma_e^2), \textit{independent}. \end{gathered} \tag{6}$$

In this model there are two error components, u_d and e_{di}, the random effect and the residual error term, respectively.

According to Royall (1970), we can write the best linear unbiased predictor (BLUP) for the mean as follows:

$$\hat{\bar{F}}_d^{BLUP} = N_d^{-1}\left(\sum_{i\in s_d} f_{di} + \sum_{i\in r_d} \tilde{f}_{di}\right). \tag{7}$$

Where $\tilde{f}_{di} = \boldsymbol{x}_{di}^T\widetilde{\boldsymbol{\beta}} + \tilde{u}_d$ is the BLUP of f_{di}, and $\tilde{u}_d = \gamma_d(\bar{f}_{ds} - \bar{\boldsymbol{x}}_{ds}^T\widetilde{\boldsymbol{\beta}})$ the BLUP of u_d. Here, $\bar{f}_{ds} = n_d^{-1}\sum_{i\in s_d} f_{di}$, $\bar{x}_{ds} = n_d^{-1}\sum_{i\in s_d} x_{di}$, and $\gamma_d = \frac{\sigma_u^2}{\sigma_u^2 + \frac{\sigma_e^2}{n_d}} \in (0,1)$. γ_d is the *shrinkage estimator* measuring the unexplained between-area variability on the total variability.

Since in practice the variance components are unknown, we replace these quantities by estimates, so we calculate the EBLUP of the mean:

$$\hat{\bar{F}}_d^{EBLUP} = N_d^{-1}\left(\sum_{i\in s_d} f_{di} + \sum_{i\in r_d} \boldsymbol{x}_{dj}^T\widehat{\boldsymbol{\beta}} + \hat{u}_d\right). \tag{8}$$

It is clear that if the sample size in a small area is zero, it holds that $\bar{F}_d^{EBLUP} = \overline{\boldsymbol{X}}_d\hat{\beta} = \bar{F}_d^{SYNTHETIC}$ where $\overline{\boldsymbol{X}}_d$ denotes the means of the covariates in the population.

The mean squared error can be estimated via parametric bootstrap according to Gonzalez-Manteiga et al. (2008). However, we also account for the factor analysis model variability, as the factor scores are estimated from a model. The steps are provided in appendix A.

4 Simulation Study

The simulation study was designed to assess the behaviour of the EBLUP of factor score means under a FA model. We compare this approach with the use of the factor loadings to calculate a weighted average from a dashboard of standardised univariate EBLUPs calculated from the original variables. In addition, we compare to the case where a simple average is taken from the standardized univariate EBLUPs.

In the case of outliers, the regression S-estimator by Rousseeuw and Yohai (1984) and Salibian-Barrera and Yohai (2006) can be applied in order to estimate the model parameters under a robust approach, denoted by REBLUP. The random effects can be robustly predicted by using the Copt and Victoria-Feser (2009) predictor. This is shown in Schoch (2011). We also applied the REBLUP method approach to evaluate the robustness of the EBLUP method in the presence of outliers .

4.1 Generating the population

The target population was generated from a multivariate mixed-effects model, the *natural extension* of the BHF model (Fuller and Harter, 1987) with $N = 20{,}000$, $D = 80$, and $130 \leq N_d \leq 420$. N_d was generated from the discrete uniform distribution, $N_d \sim \mathcal{U}(a = 130, b = 420)$, with $\sum_{d=1}^{D} N_d = 20{,}000$ where the parameters are obtained from the Italian EU-SILC 2009 dataset used in the application in section 5. Here the multivariate BHF model that we use to generate the population for the original variables is:

$$\boldsymbol{y}_{di} = \boldsymbol{x}_{di}^T\boldsymbol{\beta} + \boldsymbol{u}_d + \boldsymbol{e}_{di}, i = 1, \dots, N_D, d = 1, \dots, D \tag{9}$$

$$\boldsymbol{u}_d \sim \text{iid} MVN(\boldsymbol{0}, \boldsymbol{\Sigma_u}), \boldsymbol{e}_{di} \sim \text{iid} logMVN(\boldsymbol{0}, \boldsymbol{\Sigma_e}), \textit{independent}.$$

Two uncorrelated covariates are generated from the Normal distribution:

$$X_1 \sim N(9.93, 4.98^2), \qquad X_2 \sim N(57.13, 17.07^2).$$

These parameters reflect two real variables in the Italian EU-SILC 2009 dataset: the years of education and age. We selected three response variables from the Italian EU-SILC 2009 data: the log of the income, squared metres of the house, and the number of rooms, and fit regression models using the covariates X_1 and X_2. We generate the residuals term $\boldsymbol{e}_{di}$ from a log-multivariate Normal distribution in order to introduce outliers in the distributions to evaluate the robustness of the EBLUP.

From these models, we derive the beta coefficient matrix and standard errors to build the simulation population by the model in (9). The $\boldsymbol{\beta}(3 \times 3)$ matrix of coefficients is given by:

$$\boldsymbol{\beta} = \begin{bmatrix} 3.983 & 0.018 & 0.001 \\ 1.263 & 0.007 & 0.005 \\ 0.404 & 0.006 & 0.002 \end{bmatrix}$$

The response vector $\boldsymbol{Y}_{(3\times1)} = (\boldsymbol{Y}_1, \boldsymbol{Y}_2, \boldsymbol{Y}_3)^{\boldsymbol{T}}$ was generated according to the following variance components, where the correlation was set at 0.5 as derived from the Italian EU-SILC 2009 data:

$$\boldsymbol{\Sigma}_e = \begin{bmatrix} 0.063 & 0.028 & 0.021 \\ 0.028 & 0.049 & 0.018 \\ 0.021 & 0.018 & 0.027 \end{bmatrix}.$$

We control the intra-class correlation ρ defined as $\rho_{y_k} = \sigma^2_{uy_k}/(\sigma^2_{uy_k} + \sigma^2_{ey_k})$, and obtain the random effects matrices. We chose three levels of intra-class correlations: 0.1, 0.3 and 0.8, and obtain the following matrices:

$$\boldsymbol{\Sigma}_u^{0.1} = \begin{bmatrix} 0.00693 & 0.00306 & 0.00227 \\ 0.00306 & 0.00539 & 0.00200 \\ 0.00227 & 0.00200 & 0.00297 \end{bmatrix},$$
$$\boldsymbol{\Sigma}_u^{0.3} = \begin{bmatrix} 0.02709 & 0.01195 & 0.00887 \\ 0.01195 & 0.02107 & 0.00782 \\ 0.00887 & 0.00782 & 0.01161 \end{bmatrix},$$
$$\boldsymbol{\Sigma}_u^{0.8} = \begin{bmatrix} 0.25500 & 0.11112 & 0.08249 \\ 0.11112 & 0.19600 & 0.07275 \\ 0.08249 & 0.07275 & 0.10800 \end{bmatrix}.$$

In order to have *unplanned domains*, we select $S = 1, \dots, 500$ simple random samples without replacement (SRSWOR) from the population, and therefore we incur zero sample size domains as well as small domains.

We run the FA model on the population to derive the population factor scores $F_i, i = 1, \dots, N$. We note that although FA models have been developed for multilevel structures within domains, it is not possible to use these models for *unplanned domains* given small and zero sample size domains.

The factor analysis model on the population shows that the first factor explains a good amount of the total variability. Table 1 shows the eigenvalues under different levels of the intra-class correlations and Figure 1 the scree plots. Hence, we fit a one-factor confirmatory FA model on the population, which provides good fit statistics: $RMSEA = 0$ and $CFI = 1, TLI = 1$ (Hu and Bentler, 1999).

Factor	**Eigenvalues $\boldsymbol{\rho = 0.1}$**	**Eigenvalues $\boldsymbol{\rho = 0.3}$**	**Eigenvalues $\boldsymbol{\rho = 0.8}$**
1	2.059	2.059	2.142
2	0.476	0.486	0.451
3	0.465	0.455	0.407

Table 1 Eigenvalues from the FA of the simulation population

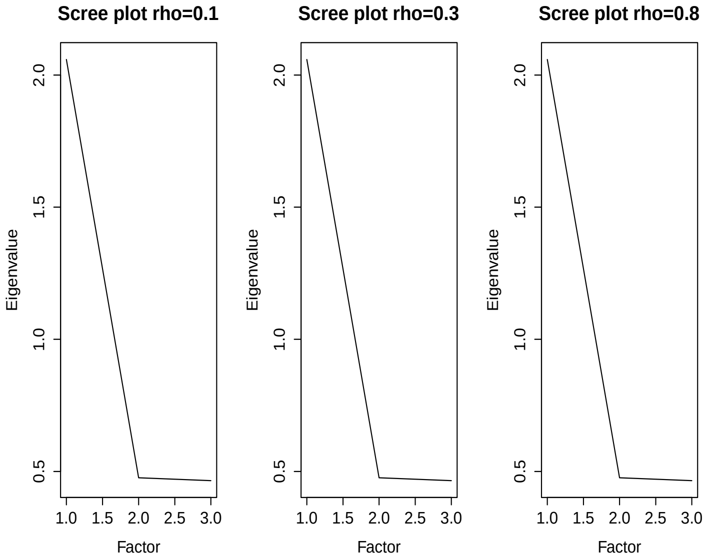

Figure 1 Scree plots from the FA of the simulation population

We now calculate the following true values for each of the small areas d from our simulated population: the factor score means denoted by $\bar{F}_d$, simple average of the standardized original variable means and weighted average using the FA loadings denoted by $\bar{Y}_d^{S\_Averages}$ and $\bar{Y}_d^{W\_Averages}$ respectively. We note that the factor scores are strongly linearly related to the observed variables and have the same economic interpretation as the observed variables.

The generated population may result in estimated factor scores with outliers and in this case we have introduced outliers by use of the log multivariate normal distribution when generating the population. Figure 2 show the histograms and box-plots of the factor scores with outliers.

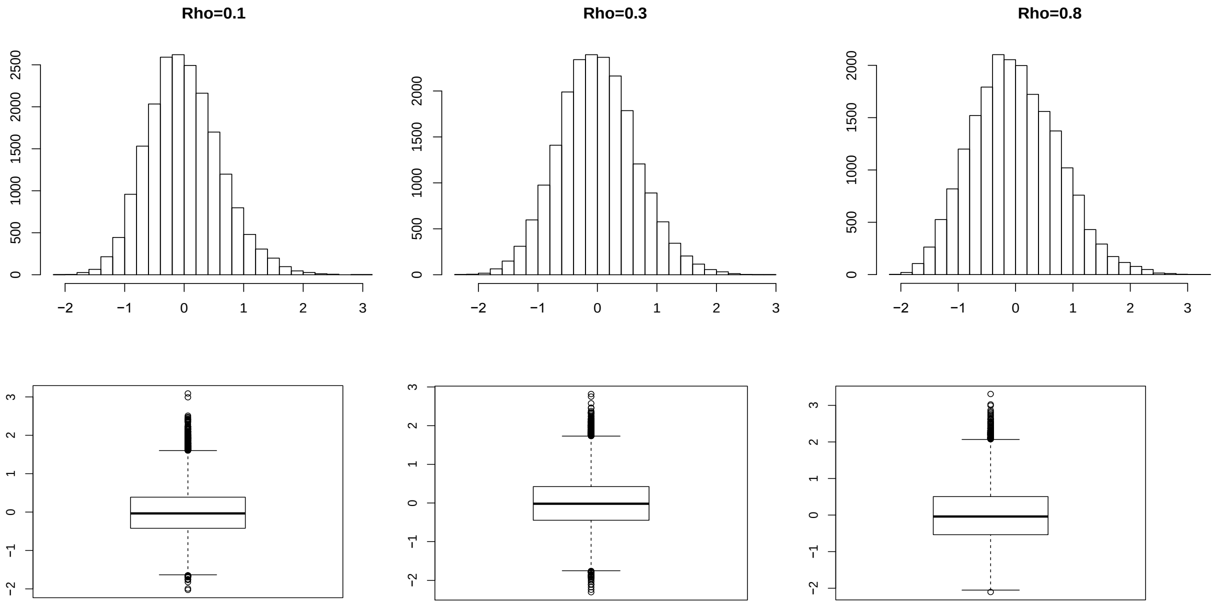

Figure 2 Histograms and box-plots of the factor scores.

4.2 Simulation steps

The simulation study consists of the following steps:

1. Draw $S = 500$ samples using simple random sampling without replacement;
2. Fit the one-factor confirmatory FA model and estimate the EBLUP and robust REBLUP of factor score means for each area *d* in each sample. We also calculate Horwitz-Thompson (HT) (Horvitz and Thompson, 1952) direct estimates of the factor score means for those areas with a non-zero sample size. In addition, the EBLUP and REBLUP for each of the original variables are also estimated in order to construct a simple average of the standardized small area EBLUPs and REBLUPs and a weighted average using the factor loadings;
3. As the true values are known from the simulation population, we are able to calculate the root mean squared error and the relative bias for each area *d* for the three types of estimates: EBLUPs and REBLUPs of factor score means, simple and weighted average of EBLUPs. For example, for the EBLUPs of factor score means:

$$RMSE\left(\hat{\bar{F}}_{d}^{EBLUP}\right) = \sqrt{S^{-1}\sum_{s=1}^{S}\left(\bar{F}_{ds}^{TRUE} - \hat{\bar{F}}_{ds}^{EBLUP}\right)^{2}} \tag{10}$$

where $\bar{F}_{ds}^{TRUE}$ denotes the true mean of the factor scores for the d_{th} area, and the relative bias:

$$RBIAS\left(\hat{\bar{F}}_d^{EBLUP}\right)_d = S^{-1}\sum_{s=1}^{S}\frac{\left(\hat{\bar{F}}_{ds}^{EBLUP} - \bar{F}_{ds}^{TRUE}\right)}{\bar{F}_{ds}^{TRUE}}, \tag{11}$$

4. For the overall comparison across all areas, we rank the small areas according to the estimates averaged across the 500 samples and compare each to the ranking in the population. We also examine the average of the RMSE and RBIAS across all areas.

The matrix of rotated (varimax rotation) factor loadings matrix can be used to weight the single indicators (OECD, 2008).. Hence, a simple way to reduce data dimensionality is to construct the following weighted mean after standardizing the EBLUP estimates:

$$\hat{\bar{Y}}_d^{EBLUP\_W\_Averages} = \frac{\sum_{k=1}^{K}\left(\hat{\bar{Y}}_{dk}^{standard\_EBLUP}\lambda_k\right)}{\sum_{k=1}^{K}\lambda_k}, d = 1, \dots, D, k = 1, \dots, K, \tag{12}$$

where k denotes the kth variable and λ_k the factor loading

We estimate the EBLUP and REBLUP for each original variable separately on each of 500 samples, and then standardized them and constructed weighted and simple averages. These are compared to the true values in the simulation population. The REBLUP's results are shown in Appendix D, since the EBLUP performs well in our study.

4.3 Results

In this section we show the main results of the simulation study. Table 2 can be compared to Table 1 and we can see that we are able to obtain good estimates for the eigenvalues across the samples (average over 500 samples) compared to the true eigenvalues. Table 3 presents the intra-class correlation coefficients estimated from the SAE model (average across 500 samples) showing that we approximate the known intra-class correlation coefficients as expected.

Factor	Eigenvalues $\boldsymbol{\rho = 0.1}$	Eigenvalues $\boldsymbol{\rho = 0.3}$	Eigenvalues $\boldsymbol{\rho = 0.8}$
1	2.081	2.074	2.179
2	0.469	0.479	0.426
3	0.450	0.447	0.395

Table 2 Average eigenvalues across 500 samples from confirmatory FA model

Scenario	$\boldsymbol{\rho = 0.1}$	$\boldsymbol{\rho = 0.3}$	$\boldsymbol{\rho = 0.8}$
$\widehat{ICC}$	0.110	0.322	0.799

Table 3 Average ICC estimates across 500 samples

Figure 3 show the QQ-plot of the level-1 residuals obtained from the BHF model. It can be seen that there are some outliers in the distributions as expected.

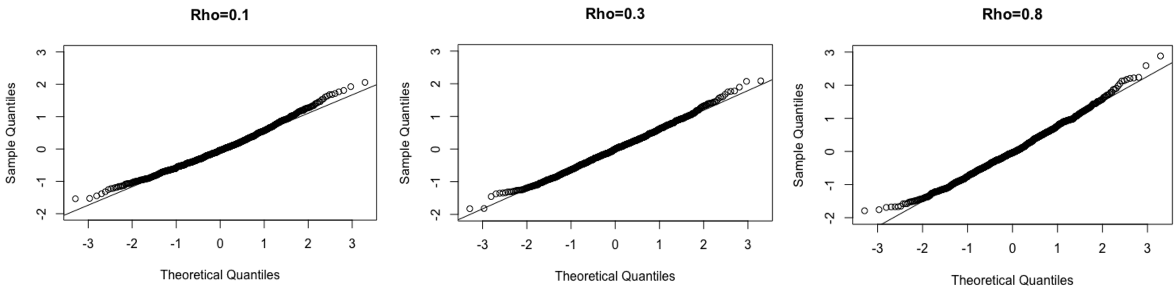

Figure 3 QQ-plot from the BHF model across the samples – Level-1 residuals.

In order to compare the ranking of the small area domains obtained from the samples to the true ranking obtained in the population, we calculate the Spearman's correlation coefficient for the estimates under the different approaches across the 500 samples and their true values. These are shown in Table 4. In spite of having outliers in the distribution of the variables and their error terms, all the approaches provide a similar ranking to the population with a small deviation for the low intra-class correlation. The factor scores approach shows an improvement over the averages of EBLUPs.

Spearman's correlation	$\boldsymbol{\rho = 0.1}$	$\boldsymbol{\rho = 0.3}$	$\boldsymbol{\rho = 0.8}$
Simple EBLUPs averages	0.997	0.999	1.000
Weighted EBLUPs averages	0.997	0.999	1.000
Factor scores means EBLUP	0.998	1.000	1.000

Table 4 Spearman's correlation coefficients for the three approaches

Figure 4 shows the individual RMSE of the small areas for those areas with non-zero sample size. In line with SAE literature the EBLUP approach produces estimates with a lower variability than direct HT estimates.

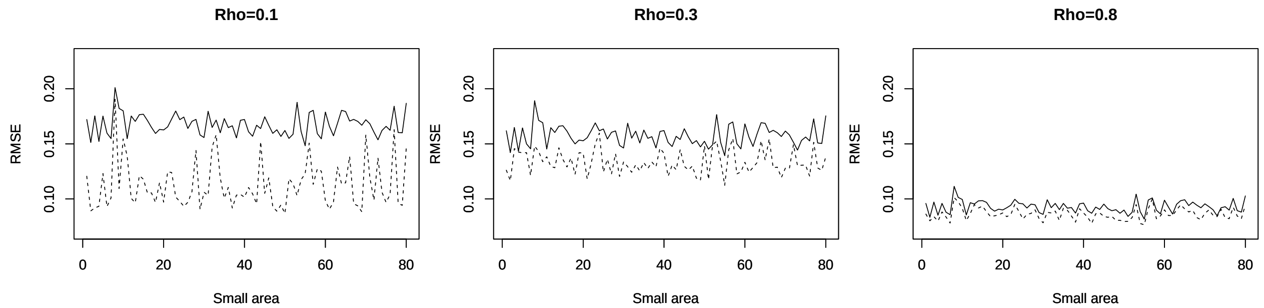

Figure 4 RMSE for Direct estimates and EBLUP of factor scores means for small areas with $n_d > 0$, *{(__) direct, (---) EBLUP }*

Table 5 shows the overall RMSE comparison defined in (10) across 500 samples for the EBLUPs of factor scores, simple and weighted standardized EBLUPs. We do not show the overall relative bias across the samples and areas since the estimates are all unbiased.

RMSE	$\boldsymbol{\rho = 0.1}$	$\boldsymbol{\rho = 0.3}$	$\boldsymbol{\rho = 0.8}$
Simple average of EBLUPs	0.542	0.340	0.119
Weighted average of EBLUPs	0.544	0.334	0.112
EBLUP factor scores means	0.113	0.133	0.080

Table 5 RMSE comparison across 500 samples for the three approaches

The overall RMSEs for the EBLUP factor score means are lower than in the case of the simple and weighted averages of the dashboard of single EBLUPs for all levels of intra-class correlations even taking into account the extra modeling step of estimating factor scores. Hence, applying the EBLUP method on the factor score means provides more precise estimates whilst reducing the data dimensionality of multiple observed variables.

In Appendix D, we compare the EBLUP approach to the REBLUP approach given that we have introduced outliers. In Table D1 we see that the REBLUP approach distorts the rankings of the sample estimates to the true population rankings and in addition, increases the overall RMSE as shown in Table D2. Although there are outliers in the data, it appears that the EBLUP estimates are robust to these outliers and outperform the REBLUP estimates.

Based on these results, we will use the EBLUP of the factor scores means approach to reduce the dimensionality of observed variables in a real application using the Italian 2009 EU-SILC data for the Tuscany region.

5 The Economic Wellbeing in Italy: a Multidimensional Approach

Income and economic resources can be seen as conditions by which an individual is able to have a sustainable standard of life. One of the dimensions in the Italian Equitable and Sustainable Wellbeing (BES) framework is dedicated to Economic Wellbeing (ISTAT 2015). It consists of ten single economic-related indicators (a dashboard of indicators). In this work, we focus on a subset of these highly correlated variables:

- Severe material deprivation according to Eurostat;
- Equivalized disposable income;
- Housing ownership;
- Housing density.

Appendix B contains the variables nomenclature for the 2009 Tuscany EU-SILC dataset used in our study and described in section 6.1 and descriptive statistics of these study variables which are explained in the next sections.

5.1 Severe material deprivation index

Material deprivation can be defined as the inability to afford some items considered to be desirable, or even necessary, to achieve an adequate standard of life. Indicators related to this are absolute measures useful to analyses and compare aspects of poverty in and across EU countries (Eurostat, 2012). According to Eurostat, material deprivation in the EU can be measured by the proportion of people whose living conditions are severely affected by a lack of basic resources. Technically, the severe material deprivation rate shows the proportion of people living in households that cannot afford *at least four* of the following nine items because of financial difficulty:

1. Mortgage or rent payments, utility bills, hire purchase installments or other loan payments;

2. One week holiday away from home;
3. A meal with meat, chicken, fish or vegetarian equivalent every second day;
4. Unexpected financial expenses;
5. A telephone (including mobile telephone);
6. A color TV;
7. A washing machine;
8. A car;
9. Heating to keep the home sufficiently warm.

It can be argued that some of these indicators (e.g. 5 and 6)are nowadays less relevant than in the past. Nevertheless, these indicators are still used to describe the difficulties that households face in achieving a standard of life considered to be sufficient by society. This index is described in Table B3 in Appendix B.

5.2 Disposable equivalized income

Disposable household income is the sum of all household members of gross personal income components plus gross income components at the household level minus employer's social insurance contributions, interest paid on mortgage, regular taxes on wealth, regular inter-household cash transfer paid and tax on income.

In order to take into account differences in household size and composition, we consider disposable equivalized income y^{DE} defined as follows:

$$y_i^{DE} = \frac{y_i^D}{n_i^E}, i = 1, \dots, N, \tag{13}$$

where $i = 1, \dots, N$ denotes households, y_i^D is the disposable household income, n_i^E is the equivalized household size calculated in the following way (Haagenars et al., 1994):

$$n_i^E = 1 + 0.5 \cdot (HM_{14+} - 1) + 0.3 \cdot HM_{13-}, \tag{14}$$

where HM_{14+} and HM_{13+} are the numbers of household members aged 14 and over and 13 or younger at the end of the income reference period, respectively. This so-called 'OECD modified scaling' procedure is crucial to taking into account the economy of scales in the household. Due to the skewness of the variable, we use the log transformation in the factor model and SAE. The histograms are in Figure B2 and descriptive statistics in Table B1 of Appendix B.

5.3 Housing ownership

Housing ownership is measured by a dichotomous variable (0,1) where 0 denotes that the property where the household lives is not owned. According to the 2009 Tuscany EU-SILC data, 73.96% of households own their property where they live.

5.4 Overcrowding and housing density

Overcrowding is one of the indicators that National Statistics Institutes include in their wellbeing measurement frameworks. A very simple indicator of housing density is given by the ratio between the number of rooms in the household (excluding kitchen, bathroom and rooms used for work purposes) and the household size:

$$\bar{r}_i = \frac{R_i}{M_i} \tag{15}$$

where i is the household, M_i denotes the number of people in the *ith* household, and R_i the number of rooms in the household.

The histogram of this variable is in Figure B3 and descriptive statistics in Table B2 of Appendix B.

6 Estimation of Economic Wellbeing for Tuscany Municipalities

The aim of this section is to estimate an economic wellbeing indicator following the BES guidelines for Tuscany municipalities. In our application, data from the EU-SILC 2009 was combined with the Italian Census 2001 as described below. We note that the 2009 EU-SILC data were collected seven years after the census (2009 EU-SILC data refers to data collected in 2007). However, the years 2001–2007 in Italy were a period of relatively slow growth and low inflation (Marchetti et al., 2015). Thus, we consider the Census covariates in this application.

6.1 The data and variables

EU-SILC is conducted yearly by ISTAT for Italy, and coordinated by EUROSTAT at the EU level. The survey is designed to produce accurate estimates at the national and regional levels (NUTS-2).

Hence, for the Italian geography the survey is not representative of provinces, municipalities (NUTS-3 and LAU-2 levels, respectively), and lower geographical levels. The regional samples are based on a stratified two-stage sample design. The Primary Sampling Units (PSUs) are the municipalities within the provinces, and households are the Secondary Sampling Units (SSUs). The PSUs are stratified according to their population size and SSUs are selected by systematic sampling in each selected PSU. The total number of households in the sample for Tuscany is 1448.

The 14th Population and Housing Census 2001 surveyed 1,388,252 households of persons living in Tuscany permanently or temporarily, including the homeless population and persons without a dwelling. The census data is linked to the EU-SILC 2009 data for Tuscany and we use the Census data covariates for the out-of-sample units based on the head of the household.

6.2 The construction of the factor scores

The first step in our analysis consists of identifying the underlying latent variable. Here we perform an exploratory FA model (EFA) on the 4 study variables presented in Section 5. The eigenvalues from the EFA on the Tuscany EU-SILC 2009 dataset are presented in Table 6 and the scree plot in Figure 5. From these results, we can see that a two-factor FA model would be desirable. However, according to the goodness of fit statistics estimated on the one-factor model solution, Root Mean Squared Error of Approximation (RMSEA=0.047) and the Comparative Fit Index criteria (CFI=0.966) the one-factor model provides good fitness (Hu and Bentler 1999). Thus, for the purpose of our study, we fit a confirmatory FA model with only one factor. This choice can be justified also substantively as our variables relate to economic wellbeing according to the BES framework, which is the phenomenon we want to measure.

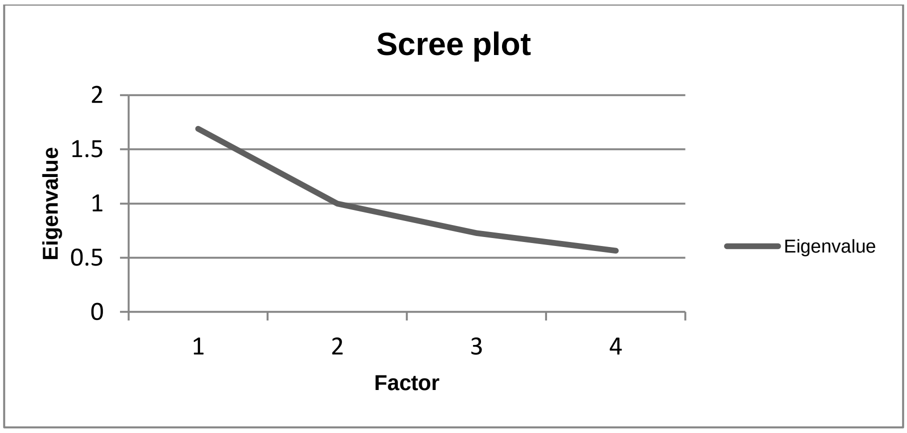

Figure 5 Scree plot exploratory FA model on Tuscany EU-SILC 2009

Factor	Eigenvalue
1	**1.791**
2	**1.000**
3	0.727
4	0.566

Table 6 Eigenvalues from exploratory FA model on Tuscany EU-SILC 2009

The histogram, Q-Q plot, and box-plot of the factor scores are shown in Figure 6 as well as descriptive statistics in Table 7. We see evidence of skewness in the factor scores likely due to discrete variables included in the FA model. One interesting thing to note based on Table B4 in Appendix B is that the estimated ICC for the factor scores is 0.1987 which is considerably higher than the estimated ICC's for the single study variables, thus as seen in the simulation study, we expect that the EBLUP of the factor scores will provide good rankings of the small areas.

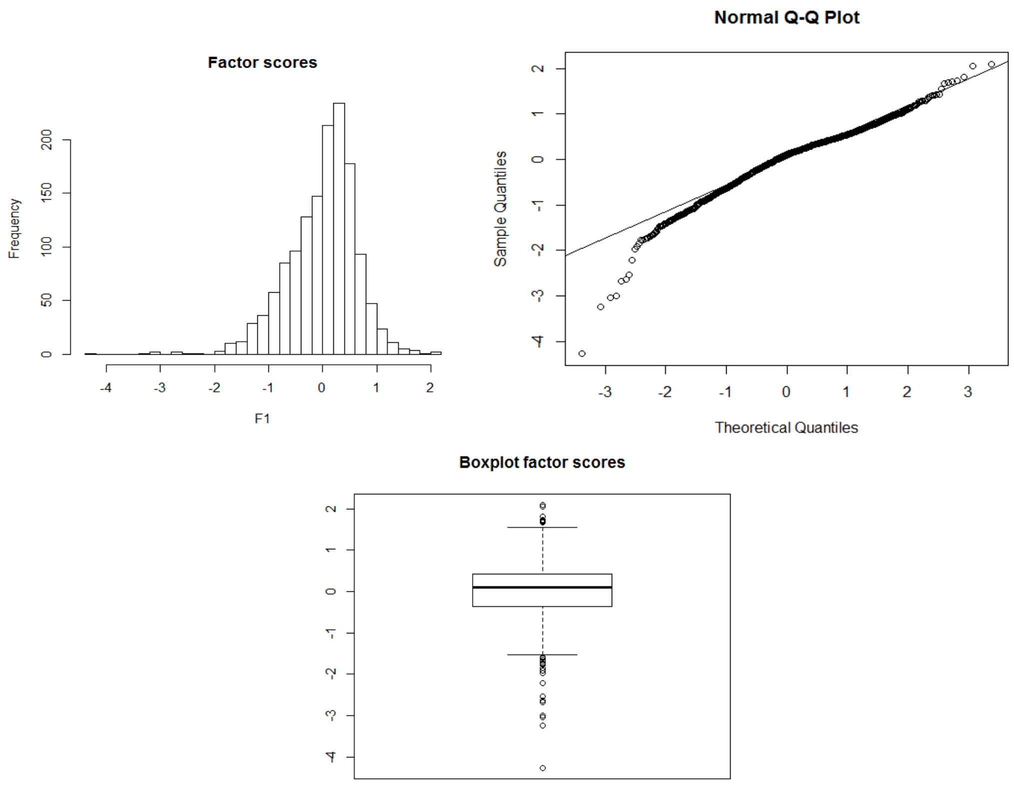

Figure 6 Factor scores distribution graphs.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max	S.d.	$\widehat{\mathrm{ICC}}$
-4.2630	-0.3712	0.1050	0.0034	0.4120	2.0940	0.6436	0.1987

Table 7 Descriptive statistics of the factor scores

6.3 Small area estimates

In this application we treat municipalities as our small areas of interest. The municipalities within Tuscany are unplanned domains in EU-SILC and only 59 out of 287 were sampled. Sample sizes in municipalities range from 0 to 135 households.

First, we provide direct estimates for the small areas with $n_d > 0$. After this, we build a SAE model under the BHF approach where the response variable is the factor score interpreted as the latent economic wellbeing construct. The exploratory variables in the model relate to the head of the household and are those common to both the survey and Census data. In particular, after a preliminary analysis of the available data we chose gender, age, year of education, household size, size of the flat (in squared meters), and employment status as the explanatory variables.

The single EBLUPs of the dashboard indicators have been estimated to construct the simple and weighted averages, as was done in the simulation study. In the case of binary variables the following linear logistic mixed effects model was fitted (MacGibbon and Tomber 1989):

$$logit(p_{id}) = log\left(\frac{p_{id}}{1 - p_{id}}\right) = \boldsymbol{x}_{id}^T\boldsymbol{\beta} + u_d, \tag{17}$$

where p_{id} is the probability that $y_{id} = 1$ and $u_d \sim \mathrm{iidN}(0, \sigma_u^2)$.

Here, we compare the MSE of the EBLUPs of factor scores means with the direct estimates. The estimates of the MSE for the predictions are obtained via the bootstrap with $B = 500$ bootstrap samples as described in Appendix A. We can see the gain in efficiency (in terms of reduction in the root MSE) obtained by the EBLUP compared to the direct estimates as shown in Figure 8.

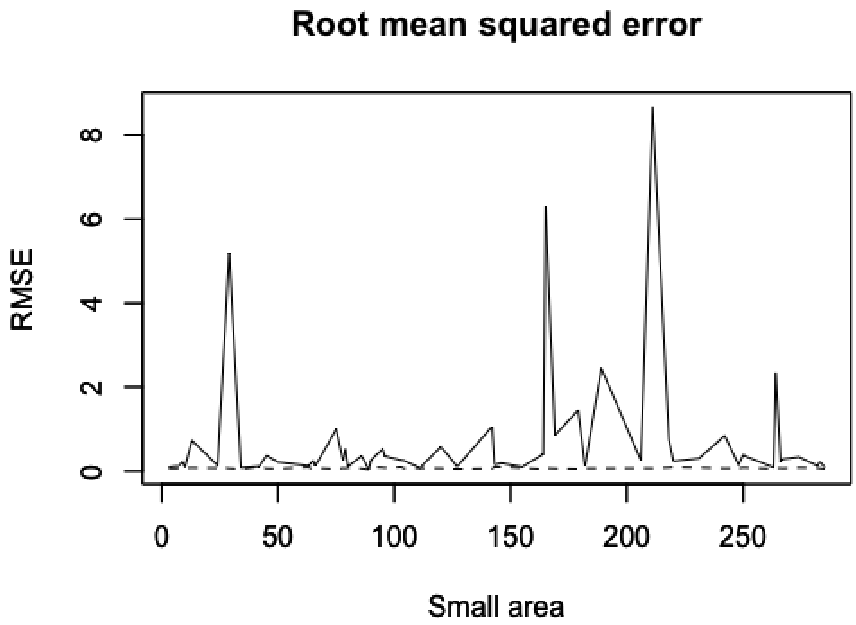

Figure 8 RMSE direct estimates (---) and EBLUPs (___) for small areas with $n_d > 0$

The overall RMSEs of the estimates and predictions across the small areas are 0.7336 and 0.0726 for the direct estimates and EBLUPs, respectively.

To facilitate the interpretation and provide a comparison between the different economic wellbeing indicators obtained from the EBLUP factor score means and the simple and weighted averages of the dashboard of EBLUPs, we have normalized the EBLUPs using the ‘Min-Max’ method (OECD, 2008), with range [0,1]. For the factor score EBLUPs, the normalisation (denoted with a ‘*’) is as follows

$$\hat{\bar{F}}_d^{*EBLUP} = \frac{\hat{\bar{F}}_d^{EBLUP} - \min(\hat{\bar{F}}_d^{EBLUP})}{\max\left(\hat{\bar{F}}_d^{EBLUP}\right) - \min(\hat{\bar{F}}_d^{EBLUP})}, d = 1, \ldots, D, \tag{16}$$

and similarly for the simple and weighted averages of the dashboard of EBLUPs.

Table 8 shows the percentiles for the latent economic wellbeing indicator based on the normalized EBLUP factor scores and the normalized averages of the dashboard of EBLUPs. Figure 9 and Figure 10 depict the maps of the quartiles of the EBLUPs under the different approaches for the Tuscany region.

Percentile	0%	25%	50%	75%	100%
EBLUP	0.0000	0.5110	0.5468	0.5819	1.0000
Simple	0.0000	0.4297	0.5297	0.6061	1.0000
Weighted	0.0000	0.4796	0.6006	0.7184	1.0000

Table 8 Percentiles for transformed latent economic wellbeing indicator based on EBLUP of factor score means and simple and weighted averages

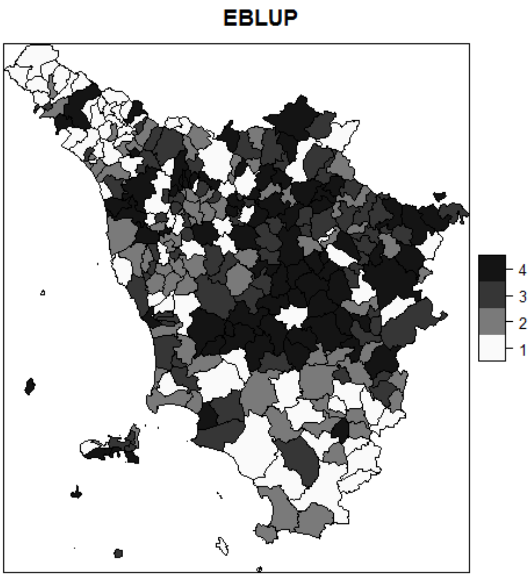

Figure 9 Latent economic wellbeing indicator based on transformed EBLUP of factor scores means {1=1st quartile; 2=2nd quartile; 3=3rd quartile; 4=4th quartile}

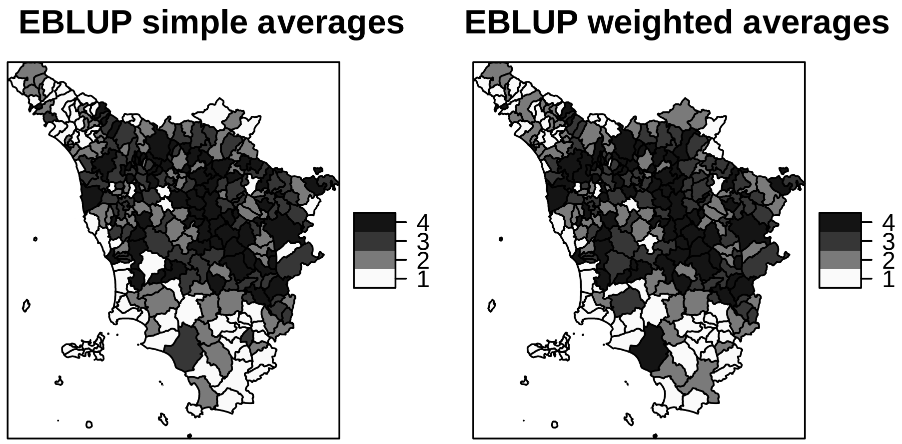

Figure 10 Latent economic wellbeing indicator based on simple and weighted averages of single EBLUPs {1=1st quartile; 2=2nd quartile; 3=3rd quartile; 4=4th quartile}.

In the maps of Figure 9 and Figure 10 a darker color denotes a better wellbeing phenomena. Looking at these figures we can draw some interesting conclusions on economic wellbeing in the Tuscany region.

The municipalities located in the Massa-Carrara province, which is based in the North of Tuscany (i.e. Pontremoli and Zeri municipalities), and municipalities based in Grosseto province (south of Tuscany), are the poorest ones. The small areas based in Florence province are wealthy municipalities, as well as the ones located in the center of the region (Siena province).The lowest point estimates of the latent economic wellbeing indicator are estimated for Carrara and Seravezza municipalities, and the highest values for Firenze and Arezzo municipalities. Our results based on the EBLUPs of the factor scores in Figure 9 are more comparable with other SAE studies on welfare and poverty in Tuscany (Marchetti, Tzavidis, and Pratesi 2012; Giusti et al. 2015) compared to the averages of a dashboard of EBLUPs in Figure 10, though previous SAE studies consider only income variables rather than a composite indicator used here. This is not surprising given the low ICCs for each of the individual EBLUPs that form the dashboard which may result in more distortions on the rankings, particularly since some of the individual EBLUPs are based on discrete variables.

6.4 Model diagnostics

We assess the fit of the model by analyzing the *level-1* standardized residuals. In particular, the histogram and Q-Q plot of the residuals are shown in Figure 11 and Figure 12 shows the leverage measures versus standardized scaled residuals from the linear model are shown. Both these figures show a presence of outliers in the right tail, although the factor scores distribution is approximately symmetric. Figure 12 also shows the contour of the Cook's distance (red smooth line) which does not deviate much from zero and hence we can conclude that the outliers are not influential.

Attempts to use the REBLUP as shown in the simulation study provided distorted and inconsistent rankings. However, based on the results in the simulation study, the presence of outliers does not appear to significantly impact the EBLUP estimates.

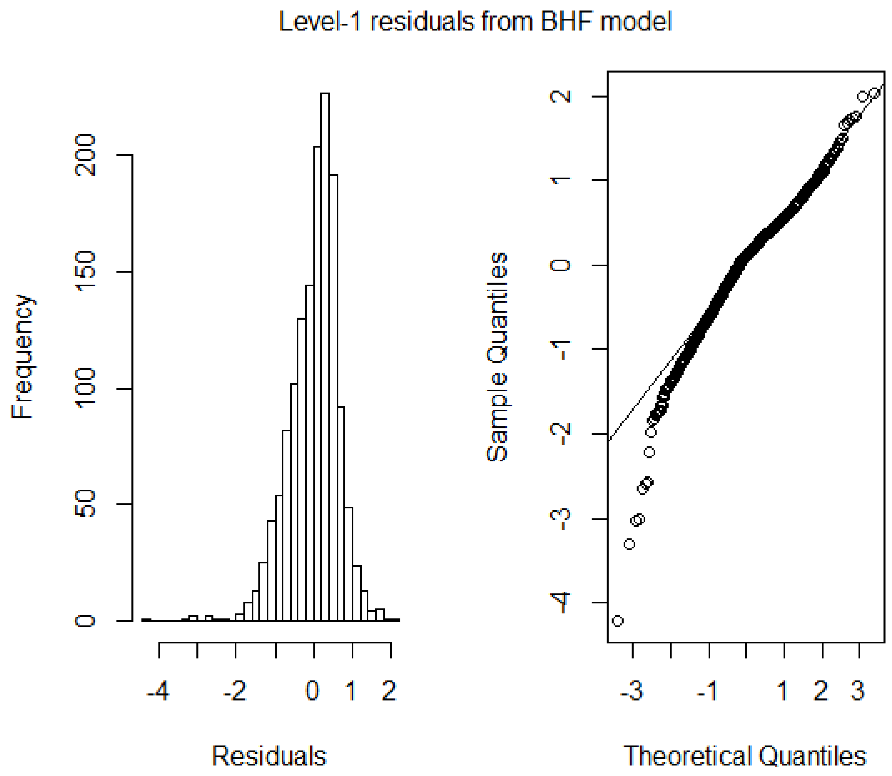

Figure 11 Level-1 residuals analysis plots for the BHF model fitting

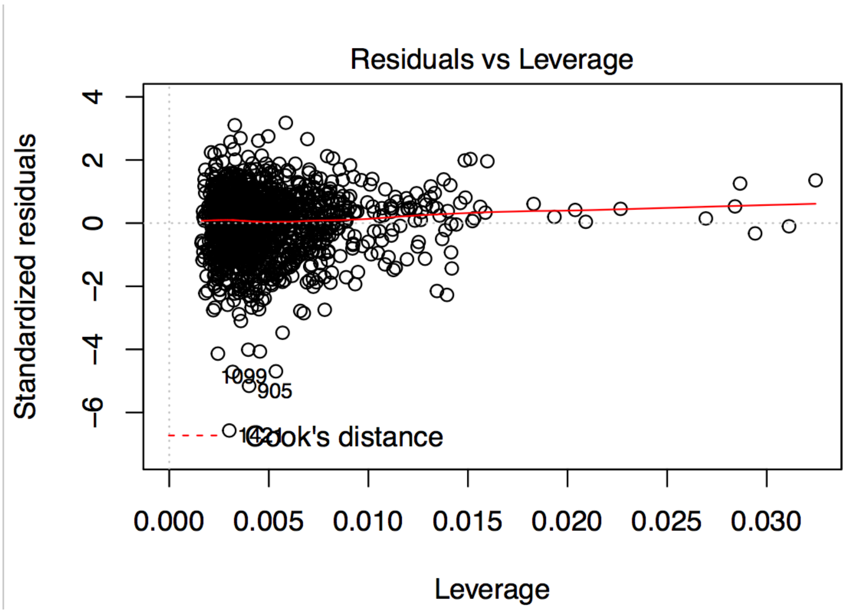

Figure 12 Standardized residuals versus leverage measure.

7 Conclusion and Discussion

In this paper we evaluated a method to estimate the mean of a latent economic wellbeing indicator at a local level for Tuscany using factor scores to reduce data dimensionality. We focused on the factor scores because they can be seen as a composite variable - in our context, a *latent* economic wellbeing composite variable. The simulation study demonstrated that although simple or weighted averages of univariate standardized EBLUPs provide good rankings, with the use of factor scores we can achieve more precise estimates in terms of MSE when we measure a multidimensional phenomenon. Moreover, the simulation study showed that the EBLUP method performs reasonably well for factor scores mean predictions even in the presence of outliers.

There are several areas where this work could be extended. Future work might consider other geographical levels, such as SLL (Sistemi Locali del Lavoro – Labor Local System), by looking at the flow of daily travel home/work (commuting) detected during the General Census of Population and Housing. Further interesting applications would involve comparisons between Italian regions in the North, Centre, and South.

Another worthwhile extension is accounting for more than one factor. When the goal is to reduce the dimensionality of the original data by identifying latent factors, one might face the issue of identifying multiple factors. This means that more than one factor explains the total variability of the original data. In our study, we only consider the case of reducing multiple dimensions into a *single* latent factor. The univariate EBLUP and REBLUP were then applied in order to estimate the mean factor score for economic wellbeing for Tuscany municipalities. However, multiple latent factors can arise, particularly when we have many indicators referring to the same phenomenon which can be grouped substantively into subdomains. For example, if the goal is to study housing quality we may want to consider the following dimensions: type of dwelling and tenure status, housing affordability, and housing quality (e.g. overcrowding, housing deprivation, problems in the residential area).

The orthogonality assumption of the factors is restrictive in welfare measurement as wellbeing dimensions can be related (Thorbecke 2008). This means that the factor scores may be correlated, thus, the use of the multivariate mixed effect model (Fuller and Harter, 1987) might be more appropriate than the BHF model for a univariate EBLUP. In fact, by using multivariate models one can take into account the correlation between responses. Interestingly, Datta, Day, and Basawa (1999) showed that the use of the multivariate mixed-effects model might lead to a gain in efficiency in terms of MSE for the EBLUP compared to the BHF model. Thus, the multivariate small area estimation method might provide better dashboard estimates and averages if the correlation between the single variables is taken into account. These extensions will be taken into account in future work.

Acknowledgements

This analysis was carried out on confidential data released by ISTAT. Data were analyzed by respecting all of the Italian confidential restriction regulations (D.Lgs. 196/03 – Codice Privacy).
The authors thank Dr. Luca Faustini and Dr. Linda Porciani from the ISTAT regional office of Florence for their kind help and suggestions during the data request process. Also, we would like to thank Caterina Santi, researcher at Scuola Superiore Sant'anna – Institute of Economics in Pisa for her kind suggestions on the software.
This work was financially supported by the following grants: ESRC DTC Award and Advanced Quantitative Methods (also known as AQM).

Appendix A: Parametric bootstrap procedure for the EBLUP's MSE.

Here we show the bootstrap steps for the EBLUP's MSE. The bootstrap procedure is the one proposed by Gonzalez-Manteiga et al. (2008). We particularize this bootstrap procedure by taking into account for the FA model variability here.

1. Draw $b = 1, \dots, 500$ simple random samples with replacement from the EUSILC sample and estimate FA models to obtain factor scores. After this the usual parametric bootstrap proposed by Gonzalez-Manteiga et al. (2008) is run for the $b = 1 \dots, B$ bootstraps.
2. Fit the Battese, Harter and Fuller model to the sampled units $\boldsymbol{f}_s = (\boldsymbol{f}'_{1s}, \dots, \boldsymbol{f}'_{Ds})'$, and estimate the model parameters $\hat{\beta}, \hat{\sigma}_u^2$ and $\hat{\sigma}_e^2$.
3. Generate $u_d^{*(b)} \sim \mathrm{iidN}(0, \hat{\sigma}_u^2), d = 1, \dots, D$, which are the bootstrap area effects.
4. Generate the bootstrap errors for the sample units $e_{di}^{*(b)} \sim \mathrm{iidN}(0, \hat{\sigma}_e^2), i \in s_d$ independently of the $u_d^{*(b)}$ and the error domain means $\bar{E}_d^{*(b)} \sim \mathrm{iidN}\left(0, \frac{\hat{\sigma}_e^2}{N_d}\right), d = 1, \dots, D$.
5. Calculate the true means for each small area of the bootstrap population as follows:
$$\bar{F}_d^{*(b)} = \overline{\boldsymbol{X}}'_d \hat{\beta} + u_d^{*(b)} + \bar{E}_d^{*(b)}, d = 1, \dots, D,$$
$\overline{\boldsymbol{X}}'_d$ denotes the covariates from the population (auxiliary variables).

6. Generate the responses for the sample units by using the sample covariates vectors $\boldsymbol{x}_{di}, i \in s_d$:
$$\bar{F}_d^{*(b)} = \bar{\boldsymbol{x}}'_d \hat{\beta} + u_d^{*(b)} + \bar{e}_d^{*(b)}, i \in s_d, d = 1, \dots, D.$$
7. Fit the nested errors model to the bootstrap sample data $\boldsymbol{f}_s^{*(\boldsymbol{b})}$ and obtain the bootstrap EBLUPs $\hat{\bar{F}}_d^{*(b)}, d = 1, \dots, D$.
8. The final MSEs bootstraps estimates of the EBLUPs are given by the following:

$$mse\left(\hat{\bar{F}}_d^{EBLUP}\right) = \frac{1}{B} \sum_{b=1}^{B} \left(\hat{\bar{F}}_d^{*(b)} - \bar{F}_d^{*(b)}\right)^2, d = 1, \dots, D.$$

$\bar{F}_d^{*(b)}$ denotes the true mean and $\hat{\bar{F}}_d^{*(b)}$ the EBLUP for the area d for replicate b.

Appendix B: EU-SILC data study variables

Here we describe the Italian EU-SILC 2009 data nomenclature and show some descriptive statistics on the study variables.

Variable name	*Description*
FCOM	Area code: comune (municipality)
HOUSEHOLD CROSS-SECTIONAL WEIGHT	Cross-sectional survey weight
TOTAL DISPOSABLE HOUSEHOLD INCOME	Total disposable household income
STANZE	Rooms in the flat (except: kitchen, toilet and bathroom, hallway, corridor, rooms used for work purposes).
GODAB_B	House ownership variable indicator
Material deprivation variables	
IMPREV	Ability to deal with unexpected expenses of €1000
FERIE	Affordability of one week per year away from home
PASTO	Affordability of a meat or chicken, or fish (or equivalent vegetarian) every two days
RISADE	Capacity of heating the house properly
LAVATR	Washing machine ownership
TV	TV ownership
AUTO	Car ownership
CELL	Telephone ownership
PAGAFF	Difficulties in paying the rent
PAGBOL	Difficulties in paying bills
PAGALDEB	Difficulties in paying loans or something similar
PAGMUT	Difficulties in paying the mortgage

Figure B1. Italian EU-SILC variables nomenclature

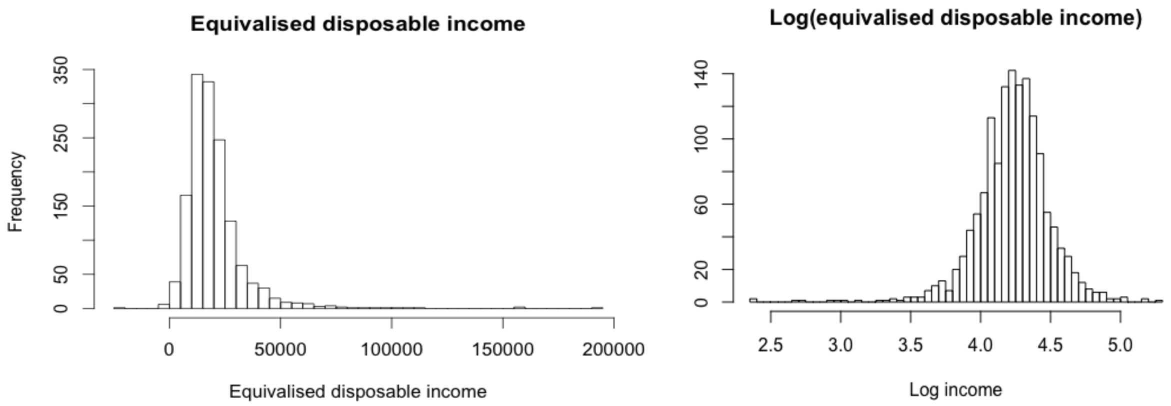

Figure B2. Disposable equivalized income histogram

Min.	1st Qu.	Median	Mean	3rd Qu.	Max	S.d.
-24,670	12,200	17,410	20,090	23,740	190,800	13,990.88
2.398	4.087	4.243	4.231	4.377	5.280	0.264

Table B1. Equivalized disposable income and log equivalized disposable income descriptive statistics

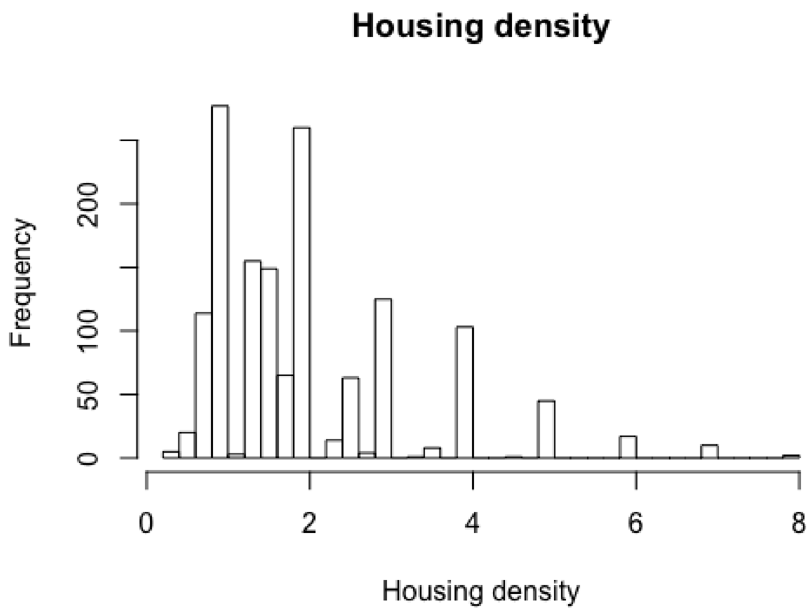

Figure B3. Housing density

Min.	1^{st} Qu.	Median	Mean	3^{rd} Qu.	Max	S.d.
0.250	1.000	1.600	1.989	2.500	8.000	1.239

Table B2. Descriptive statistics of the housing density

Variable	Frequency %
Material deprivation	3.94%
House ownership	73.96%

Table B3. Frequencies of the binary variables

Variable	Estimated ICC
Factor scores	0.1987
Disposable equivalized income	0.0019
Room average	0.0680
Material deprivation	0.0189
House ownership	0.0410

Table B4. Estimation of the ICCs of the study variables and factor scores

Appendix C: Specification of the main R functions

Here we describe the main R packages we used for the small area estimates. All the other analyses were programmed manually.

C.1 Estimation of small area means and MSE under EBLUP approach with the "sae" package (Molina and Marhuenda 2015)

- Required packages: nlme, MASS
- Functions: eblupBHF( ) and pbmseBHF( ).

C.2 Robust estimation and prediction of small area means and MSE under REBLUP approach with the "rsae" package (Schoch 2014).

- Functions: fitsaemodel() and robpredict().

C.3 Running Mplus models in the R environment via MplusAutomation (Muthén and Muthén, (2012), Hallquist and Wiley (2014))

- Functions: mplusObject( ), mplusModeler( ).

C.4 Mapping using spdep, maptools, sp, Hmisc

- Functions: readShapePoly( ), spplot( )

Appendix D: More on the simulation study: REBLUP and EBLUP comparisons

Spearman's correlation	$\rho = 0.1$	$\rho = 0.3$	$\rho = 0.8$
Simple EBLUPs averages	0.997	0.999	1.000
Simple REBLUPs averages	0.981	0.989	0.994
Weighted EBLUPs averages	0.997	0.999	1.000
Weighted REBLUPs averages	0.982	0.989	0.995
Factor scores means EBLUP	0.998	1.000	1.000
Factor scores means REBLUP	0.984	0.991	0.994

Table D1. Spearman's correlation coefficient.

RMSE	$\rho = 0.1$	$\rho = 0.3$	$\rho = 0.8$
Simple EBLUPs averages	0.542	0.340	0.119
Simple REBLUPs averages	0.631	0.496	0.382
Weighted EBLUPs averages	0.544	0.334	0.112
Weighted REBLUPs averages	0.689	0.487	0.350
EBLUP factor scores means	0.113	0.133	0.080
REBLUP factor scores means	0.516	0.398	0.280

Table D2. RMSE of the different methods.

References

Bartholomew, David J., Fiona Steele, Irini Moustaki, and Jane I. Galbraith, I. 2008. Analysis of multivariate social science data. Chapman & Hall/CRC statistics in the social and behavioural sciences series. New York.

Battese, George. E., Rachel M. Harter, and Wayne A. Fuller. 1988. "An Error-Components model for Prediction of County crop areas using Survey and Satellite data." Journal of the American Statistical Association 83 (401): 28-36.

Bell, Sherry M., Steve R. McCallum, Elizabeth Cox. 2003, "Toward a research-based assessment of dyslexia: Using cognitive measures to identify reading disabilities." Journal of Learning Disabilities 36(6): 505-516.

Copt, Samuel. and Maria-Pia Victoria-Feser. 2009. Robust Predictions in Mixed Linear Models. Research Report, University of Geneva.

Datta, Gauri. S., Bannmo, Day, and Ishwar Basawa. 1999. "Empirical best linear unbiased and empirical Bayes prediction in multivariate small area estimation." Journal of Statistical Planning and Inference 75; 269-279.

DiStefano, Christine., Min Zhu, Diana Mindrila. 2009. “Understanding and using factor scores: considerations for the applied researcher.” Practical assessment, Research and Evaluation, 14(20).

Eurostat. 2012. “Measuring material deprivation in the EU”. Methodologies and Working papers, 2012. Retrieved February 13, 2017 (http://ec.europa.eu/eurostat/documents/3888793/5853037/KS-RA-12-018-EN.PDF/390c5677-90a6-4dc9-b972-82d589df77c2).

Ferro Luzzi, Giovanni, Yves Fluckiger, and Sylvain Weber. 2008. A cluster analysis of multidimensional poverty in Switzerland in Kakwani, Nanak. and Jacques Silber. 2008. Quantitative Approaches to Multidimensional Poverty Measurement. Palgrave Macmillan, Basingstoke.

Fuller, Wayne A., and Rachel M. Harter. 1987. The Multivariate Components of Variance Model for Small Area Estimation, in Richard, Platek, J. N. K. Rao, Carl-Erik. Sarndal, and M. P. Singh (Eds.), Small Area Statistics, New York: Wiley, 103-123.

Gasparini, Leonardo, Walter Sosa Escudero, Mariana Marchionni, and Sergio Olivieri. 2011 “Multidimensional social poverty in Latin America and the Caribbean: new evidence from the Gallup World Poll.” Journal of Economic Inequality 11(2): 1-20.

Giusti, Caterina, Lucio Masserini, and Monica Pratesi. 2015. “Local Comparisons of Small Area Estimates of Poverty: An Application within the Tuscany Region of Italy.” Social Indicators Research 1-20.

Gonzalez-Manteiga, Wenceslao., Maria J. Lombardia, Isabel Molina, Domingo Morales and Laureano Santamaria. 2008. “Bootstrap mean squared error of a small-area EBLUP. ” Journal of Statistical Computation and Simulation (78): 443–462.

Haagenars, Aldi J. M., Klaas de Vos, K. and Ashar Zaidi. 1994. Poverty Statistics in the Late 1980s: Research Based on Micro-data. Office for Official Publications of the European Communities. Luxembourg.

Hallquist, Michael and Joshua Wiley. 2014. MplusAutomation: Automating Mplus Model Estimation and Interpretation. R package version 0.6-3. http://CRAN.R-project.org/package=MplusAutomation.

Horvitz Daniel G. and Donovan J. Thompson. 1952. “A Generalization of Sampling without Replacement from Finite Universe”. Journal of the American Statistical Association 47: 663-685.

Hu, Li-tze., and Peter M. Bentler. 1999. “Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives.” Structural Equation Modeling. 6(1): 1-55.

ISTAT. 2015. BES 2015. Il benessere equo e sostenibile in Italia. *Report* by ISTAT and CNEL in 2015. Available from http://www.istat.it/it/files/2015/12/Rapporto_BES_2015.pdf In Italian.

Johnson, Richard. A. and Dean W. Wichern. 1998. Applied Multivariate Analysis, 4th edition, Prentice.

Kawashima, Noriaki and Kunio Shiomi. 2007. "Factors of the thinking disposition of Japanese high school students." Social Behaviour and Personality, 35(2): 187-194.

Krishnakumar, Jaya and Anirudh Lal Nagar. 2008. "On Exact Statistical Properties of Multidimensional Indices Based on Principal Components, Factor Analysis, MIMIC and Structural Equation Models." Social Indicators Research 86: 481-496.

MacGibbon, Brenda and Tomberlin, Thomas. J. 1989. "Small Area Estimation of Proportions Via Empirical Bayes Techniques." Survey Methodology 15: 237-252.

Marchetti, Stefano, Caterina Giusti, Monica Pratesi, Nicola Salvati, Fosca Giannotti, Dino Pedreschi, Salvatore Rinzivillo, Luca Pappalardo, and Lorenzo Gabrielli. 2015. "Small Area Model-Based Estimators Using Big Data Sources. Journal of Official Statistics." 31(2): 263-281.

Marchetti, Stefano, Nikos Tzavidis, and Monica Pratesi. 2012. "Non-parametric bootstrap mean squared error estimation for M-quantile estimators of small area averages, quantiles and poverty indicators." Computational Statistics and Data Analysis, 56: 2889–2902.

Molina, Isabel. And Yolanda Marhuenda. 2015. sae: "An R Package for Small Area Estimation." The R Journal, **7**(1): 81-98.

Muthén, Bengt O. 2012. Mplus Technical Appendices. Los Angeles, CA: Muthén and Muthén.

Muthén, Linda .K. and Bengt O Muthén. 2012. Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén and Muthén.

OECD. 2008. Handbook on constructing composite indicators: methodology and user guide. OECD Statistics report. (Available from http://www.oecd.org/std/42495745.pdf).

OECD. 2013. OECD framework for statistics on the distribution of household income, consumption and wealth OECD Statistics report.

Rao, J. N. K. and Isabel Molina. 2015. Small area estimation. New York: Wiley.

Ravallion, Martin. 2011. "On multidimensional indices of poverty." Journal of Economic Inequality 9(2): 235-248.

Rousseeuw, Peter J., and Victor J. Yohai. 1984. Robust regression by means of S-estimators, in Robust and Nonlinear Time Series Analysis, Franke, J., W. Hardle, and R.D. Martin. 1984. New York: Springer.

Royall, Richard M. 1970. "On finite population sampling theory under certain linear regression." Biometrika 57: 377-387.

Salibian-Barrera, Matias. and Victor J. Yohai. 2006. "A fast algorithm for S-regression estimates." Journal of Computational and Graphical Statistics 15: 414-427.

Sosa-Escudero, Walter, German Caruso, and Marcela Svarc. 2013. Poverty and the dimensionality of welfare in Betti, Giulio and Achille Lemmi. 2013. Poverty and Social Exclusion. London: Routledge.

Schoch, Thomas. 2014. rsae: Robust Small Area Estimation. R package version 0.1-5.

Schoch, Thomas. 2011. "The robust basic unit-level small area model. A simple and fast algorithm for large datasets." Proceedings of the New Technologies and Techniques Conference (NTTS), Eurostat, Brussels.

Thorbecke, Erik. 2008. Multidimensional Poverty: Conceptual and Measurement Issues in Kakwani, Nanak, Jacques Silber. 2008. The Many Dimensions of Poverty. Palgrave Macmillan. Basingstoke.

Yalonetzky, Gaston. 2012. Conditions for the most robust multidimensional poverty comparisons using counting measures and ordinal variables. ECINEQ working paper, 2012-2257.