SSOAR

Open Access Repository
www.ssoar.info

Estimating the Uncertainty of a Small Area Estimator Based on a Microsimulation Approach

Moretti, Angelo; Whitworth, Adam

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:
Moretti, A., & Whitworth, A. (2021). Estimating the Uncertainty of a Small Area Estimator Based on a Microsimulation Approach. *Sociological Methods & Research*, OnlineFirst, 1-31. https://doi.org/10.1177/0049124120986199

Nutzungsbedingungen:
Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
https://creativecommons.org/licenses/by/4.0/deed.de

Terms of use:
This document is made available under a CC BY Licence (Attribution). For more Information see:
https://creativecommons.org/licenses/by/4.0

gesis
Leibniz-Institut
für Sozialwissenschaften

Mitglied der
Leibniz-Gemeinschaft

Diese Version ist zitierbar unter / This version is citable under:
https://nbn-resolving.org/urn:nbn:de:0168-ssoar-84589-7

Check for updates

Original Article

Sociological Methods & Research
1-31
© The Author(s) 2021

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0049124120986199
journals.sagepub.com/home/smr


Estimating the Uncertainty of a Small Area Estimator Based on a Microsimulation Approach

Angelo Moretti[1] **and Adam Whitworth**[2]

Abstract

Spatial microsimulation encompasses a range of alternative methodological approaches for the small area estimation (SAE) of target population parameters from sample survey data down to target small areas in contexts where such data are desired but not otherwise available. Although widely used, an enduring limitation of spatial microsimulation SAE approaches is their current inability to deliver reliable measures of uncertainty—and hence confidence intervals—around the small area estimates produced. In this article, we overcome this key limitation via the development of a measure of uncertainty that takes into account both variance and bias, that is, the mean squared error. This new approach is evaluated via a simulation study and demonstrated in a practical application using European Union Statistics on Income and Living Conditions data to explore income levels across Italian municipalities. Evaluations show that the approach proposed delivers accurate estimates of uncertainty and is robust to nonnormal distributions. The

[1] Department of Computing and Mathematics, Manchester Metropolitan University, United Kingdom
[2] Department of Geography, University of Sheffield, United Kingdom

Corresponding Author:
Angelo Moretti, Department of Computing and Mathematics, Manchester Metropolitan University, John Dalton Building, Chester Street, Manchester M15 6BH, United Kingdom.
Email: a.moretti@mmu.ac.uk

approach provides a significant development to widely used spatial microsimulation SAE techniques.

Keywords

calibration, weighting, synthetic, indirect estimator, raking, resampling

Large-scale surveys are designed to obtain reliable estimates at national level or, in some instances, for large subnational scales such as regions. These can be considered to be the planned domains of the survey sampling design (Benavent and Morales 2016). However, there is a growing demand from both research and policy communities for various local estimates at more detailed spatial resolutions such as municipalities or neighborhoods due to the absence of data at such small area scales from existing census or administrative sources. However, this small area desire frequently encounters a problem of unplanned domains, given that for cost reasons such small areas typically have small or zero sample sizes in the survey sampling design. In these circumstances, commonly used direct estimators such as the Horvitz–Thompson estimator (Horvitz and Thompson 1952) that only use sample survey information either cannot be used (in the case of zero sample size domains) or provide unacceptably large variability in the estimates to be practically useful (in the case of small sample size domains).

In such scenarios, indirect small area estimation (SAE) of target population parameters has become a relatively widely used and increasingly demanded methodological technique via a range of SAE approaches. We refer to Rao and Molina (2015), Whitworth (2013), Rahman and Hardin (2017), and Marshall (2010) for useful methodological reviews on both regression-based and microsimulation-based SAE methods. Spatial microsimulation approaches, sometimes referred to as survey calibration approaches (Espuny-Pujol, Morrissey, and Williamson 2018), represent a family of reweighting approaches to SAE in which the challenge is to reweight the survey units such that they optimally fit the demographic and socioeconomic profile of each small area according to a selected set of benchmark constraints. Part of the appeal of spatial microsimulation approaches to SAE for both researchers and policy users is their intuitive and accessible appeal without much of the complex statistical expertise required within many regression-based SAE methods, particularly as assumptions fail or more complex outcomes are desired. In those

circumstances, one advantage of spatial microsimulation approaches over regression-based SAE estimators is that they tend to be more robust to failures in model assumptions, given that as model-assisted estimators it is only necessary that the population be reasonably well described by an assumed model for that model to be valid (Särndal, Swensson, and Wretman 1992).

Spatial microsimulation SAE approaches have been used to produce small area estimates across a range of policy areas including child malnutrition (Johnson et al. 2012), obesity (Edwards et al. 2010), fuel poverty (Office for National Statistics 2019), income and poverty (Bell, Basel, and Maples 2016; Pratesi 2016; World Bank 2018), regional planning (Clarke and Holm 1987), participation in sport (Ipsos MORI 2018), and transport (Lovelace, 2016; Ravulaparthy and Goulias 2011; Tribby and Zandbergen 2012). Spatial microsimulation SAE approaches have been well validated against known external data and against alternative regression-based SAE techniques (Moretti and Whitworth 2019; Tanton, Williamson, and Harding 2014; Whitworth and Carter 2015). Spatial microsimulation SAE has also been used effectively to assess the spatial impacts of differing "what if" policy scenarios (Chin and Harding 2006; Cullinan, Hynes, and O'Donoghue 2006; Tanton and Edwards 2013; Williamson, Birkin, and Rees 1998). For example, Campbell and Ballas (2013) introduce the SimAlba spatial microsimulation model Scotland in order to estimate the simulated impact of various policy scenarios on individual's health outcomes. In an Australian context, Tanton and Edwards (2013) use spatial microsimulation SAE to assess the geographical impacts on expected elderly poverty levels from changes to the state pension (Tanton et al. 2009).

However, despite these widespread applications and contributions, an important long-standing limitation of spatial microsimulation SAE approaches is their continued inability to deliver estimates of uncertainty around central point estimates at small area level. More specifically, one faces something of a trade-off with any SAE approach in seeking via the indirect SAE estimates to reduce variance compared to the direct estimates while acknowledging that those direct estimates are unbiased compared to the indirect SAE estimates that are inherently biased. As such, it is essential when calculating the uncertainty of any SAE estimator that the mean squared error (MSE) is used, given that this takes into account both variance and bias. However, calculating MSE is often challenging in an SAE content and particularly so in spatial microsimulation approaches. In the case of design-based estimators, the design weights are known before the sample selection and are therefore nonrandom meaning that analytical approximations of MSE are available (Särndal et al. 1992). However, this is no longer the case in spatial microsimulation approaches to SAE when reweighing algorithms are used as the weights

become random variables themselves. In these scenarios, analytical approximation of MSE becomes highly challenging, given that bias and variance cannot be computed in closed form such that empirically based resampling techniques are instead required to estimate their uncertainty (Chen and Shen 2015). Some analytical approximations have been suggested in the literature (D'Arrigo and Skinner 2010; Deville and Särndal 1992), but Chen and Shen (2015) highlight important practical challenges around the requirement for joint selection probabilities that are rarely computed or known in practice.

Empirical attempts to estimate the uncertainty around spatial microsimulation SAE estimates have been proposed in recent years (Chen and Shen 2015; Nagle et al. 2014; Whitworth et al. 2017), though none entirely successful. In response, this article develops a novel modified parametric bootstrap technique in order to estimate the uncertainty of spatial microsimulation small area estimates based on the MSE such that it captures both bias and variance components in the uncertainty estimate, unlike previous attempts. Our approach benefits from clear statistical properties under a linear model and can be used flexibly across alternative spatial microsimulation SAE techniques.

The remainder of this article is structured as follows. In the second section, the general problem of SAE of the population mean using iterative proportional fitting (IPF) is outlined. In the third section, the challenge of uncertainty estimation in SAE contexts is further described, and our approach to MSE estimation via bootstrap is detailed. In the fourth section, the bootstrap approach is evaluated via a simulation study, and in the fifth section, a practical data application focusing on Italian Statistics on Income and Living Conditions (SILC) data is presented to illustrate the approach. The sixth section provides a concluding discussion focusing on wider implications for spatial microsimulation SAE and potential next steps for research.

SAE With IPF

This section sets out the SAE problem of the population mean and formally introducing the IPF spatial microsimulation approach used in the later development of our proposed uncertainty estimator.

The General SAE Problem for a Small Area Mean

Let us consider a sample $s \subset \Omega$ of size n drawn from the target finite population Ω of size N. Let $d = 1, \ldots, D$ denote the small areas for which we want to compute the small area estimates. $N - n$ are the nonsampled units and these are denoted by r, hence $s_d = s \cap \Omega_d$ is the subsample 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 in small area d with $N_d - n_d$ dimension. Here, the target parameter is the population mean $\bar{Y}_d = N_d^{-1} \sum_{i=1}^{N_d} y_{di}$ of the variable Y for area d, with y_{di} denoting the value of variable Y for ith unit from dth area.

Due to the unplanned domain problem, n_d may be too small (even zero) for many small areas in the survey data to compute reliable direct estimates of $\bar{Y}_d$ from the survey data based on $\hat{\bar{Y}}_d^{Direct} = \left(\sum_{i \in s_d} w_{di} \right)^{-1} \sum_{i \in s_d} y_{di} w_{di}$, where w_{di} denotes the design weight for ith unit from dth area in s_d. The direct survey estimate comes from a standard direct estimator and is based on sample survey information only. They are weighted averages where the weights are the design weight based on the complex survey design. As a consequence, there is a need in such circumstances to consider indirect SAE estimation techniques using auxiliary information if one wishes small area estimates either at all (in the case of zero survey sample sizes) or with reduced uncertainty (in the case of low small area survey sample sizes). There is a bias-variance trade-off in operation when doing so such that any reductions in variability must naturally be balanced with acknowledgment of increased bias in the indirect SAE estimates compared to the unbiased direct estimates (Rahman and Harding 2017; Rao and Molina 2015).

Reweighing Using the IPF Algorithm

IPF is one of three main spatial microsimulation approaches to SAE—IPF, generalized regression reweighting and combinatorial optimization. To demonstrate our approach to the estimation of uncertainty in spatial microsimulation SAE, we focus on the IPF algorithm, given that this is both widely used and a constructively challenging test of our MSE estimator, given that its iterative nature renders the survey weights random variables themselves (Chen and Shen 2015; Rahman et al. 2013; Simpson and Tranmer 2005).

Like all spatial microsimulation SAE approaches, IPF can be understood as a reweighting optimization problem where the aim is to reweight the survey units (e.g., individuals or households) such that they optimally fit the demographic and socioeconomic profile of each small area according to a selected set of benchmark constraints (e.g., age-sex, employment status, ethnicity, health, education). Deville and Särndal (1992) provide a statistical theory of these reweighting techniques and alternative approaches. For each local area, the result is a tailored set of reweighted survey cases that fit to the benchmark characteristics of that small area in terms both of total

population and the profile of that population across the benchmark constraints. A key data set created during IPF is a weights matrix giving new weights for each survey unit in each separate small area. For each small area, those final IPF weights show how representative each survey unit is of each area given their respective characteristics across the benchmarks. Across all survey individuals, these reweighted units sum to the small area population total and map onto its population profile across the benchmarks. As such, these reweighted data provide a valuable synthetic micropopulation for each small area that can be employed in further local analyses (e.g., "what if" policy simulations) as desired (Anderson 2013; Lovelace and Dumont 2016).

Formally, IPF can be understood as follows. Let w_i be the initial weight (usually the survey design weight) for $i \in s$. The calibration problem is area-specific and therefore generates new weights denoted by w_i^* for $i \in s_d$ for area d that satisfy the calibration equation given by $\sum_{i \in s_d} w_i^* \mathbf{x}_i = \sum_{i \in \Omega_d} \mathbf{x}_i = \mathbf{X}_d$, where x_i is a vector of auxiliary variables. Here, w_i^* minimizes a given distance function between $\left\{w_i^*; i \in s_d\right\}$ and $\{w_i; i \in s_d\}$. IPF is the exponential case within a wider family of synthetic reweighting algorithms (Deville and Särndal 1992). Its constrained optimization problem is given as follows, where a_i denotes the initial weights, usually the design weights (Chen and Shen 2015; Deville and Särndal 1992):

$$\begin{aligned} &\min : \sum_{i \in s_d} \left[w_i \ln\left(\frac{w_i}{a_i}\right) - w_i + a_i \right], \\ &\text{such that} \sum_{i \in s_d} w_i \mathbf{x}_i = \sum_{i \in \Omega_d} \mathbf{x}_i = \mathbf{X}_d. \end{aligned} \tag{1}$$

As noted above, (1) unfortunately does not have a closed-form solution such that solution via analytical approximation is required. This is however highly challenging. The IPF algorithm is therefore employed iteratively across the benchmark constraints in order to estimate the final weights for each survey unit in order to derive a solution empirically. To describe the IPF method more fully, the terminology and notation used by Kolenikov (2014) is followed:

1. Initialize the iteration counter $t \leftarrow 0$ and the weights as $w_i^{0,p} \leftarrow w_i$.
2. Increment the iteration counter $t \leftarrow t + 1$, thus updating the weights as $w_i^{t,0} \leftarrow w_i^{t-1,p}$.

3. Update the weights through each of the benchmark constraint variables in turn, $v = 1, \ldots, p$:

$$w_i^{t,v} = \begin{cases} w_i^{t,v-1} \dfrac{T(\mathbf{X}_v)}{\sum_{l \in s} w_l^{t,v-1} x_{vl}}, & x_{vi} \neq 0 \\ w_i^{t,v-1}, & x_{vi} = 0 \end{cases}.$$

4. If the discrepancies between $\sum_{i \in s} w_i^{t,p} x_v$ (i.e., the sample totals with the new weights) and $T(\mathbf{X}_v)$ are within a priori defined tolerance for all $v = 1, \ldots, p$, then declare convergence and go to step 5, otherwise return to step 2.
5. The weights $w_i^{t,p}$ are the final calibrated weights and are denoted by $w_i^{t,p} = w_i^*$.

The benchmark constraints used for the survey reweighting are usually categorical variables in real applications. Therefore,

$$\mathbf{x}_i' = \left(\delta_{1i}^{(1)}, \ldots, \delta_{F_1 i}^{(1)}, \delta_{1i}^{(2)}, \ldots \delta_{1i}^{(p)}, \ldots, \delta_{F_p i}^{(p)}\right),$$

where $l = 1, \ldots, p$ denotes the lth benchmark constraint and $\delta_{ki}^{(l)} = 1$ if I is in the category k of lth control variable. F_l is the number of categories of the lth benchmark constraint. Anderson (2007) suggests that $R = 20$ is sufficient as a conservative guide to the number of loops through the benchmark constraints in order to optimize the calibration to the set of benchmarks. The IPF algorithm is area-specific, and the IPF reweighting therefore needs to be iterated for each small area $d =, 1 \ldots, D$.

The IPF estimator can therefore be defined as follows:

$$\hat{\bar{Y}}_d^{IPF} = \frac{\sum_{i=1}^{n} w_{di}^* y_i}{\sum_{i=1}^{n} w_{di}^*}, \; d = 1, \ldots, D, \; i =, 1 \ldots, n, \tag{2}$$

where w_{di}^* denotes the IPF-calibrated survey weight for unit ith from area dth. It can be noted that y_i appears for $i = 1, \ldots, n$, which means that $\hat{\bar{Y}}_d^{IPF}$ belongs to the class of small area synthetic estimators (Rao and Molina 2015). Of course, in order for $\hat{\bar{Y}}_d^{IPF}$ to be more efficient than $\hat{\bar{Y}}_d^{Direct}$ in terms of its combination of variance and bias as measured by the MSE, the auxiliary variables used in the calibration problem need to be related sufficiently to the target variable Y, as with all such model-based or model-assisted small area estimators (Fuller 2002).

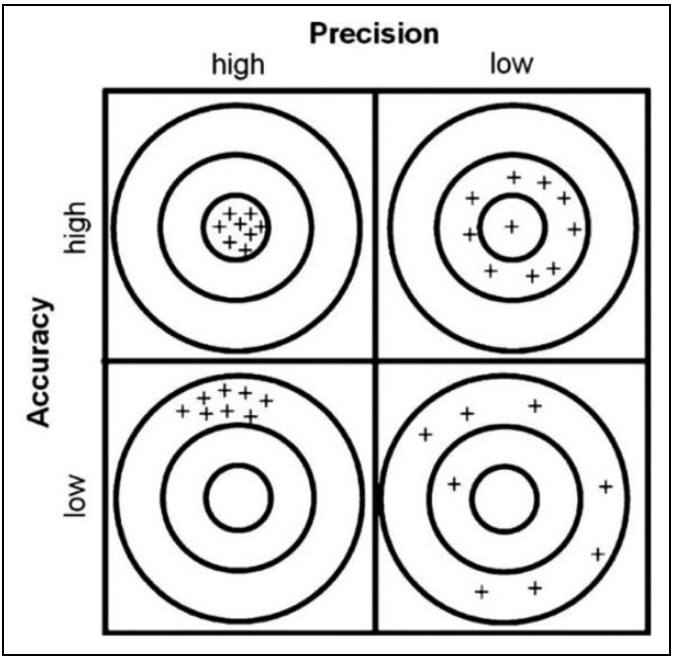

Figure 1. Precision and accuracy in estimates.

Measuring the Uncertainty: the MSE Estimator of $\hat{\bar{Y}}_d^{IPF}$

The quality of an estimate is assessed with reference both to its accuracy (bias) and to its precision (variability). It is therefore important to capture both aspects in any measure of uncertainty (Dodge and Commenges 2006; Statistics Canada 2009). The bias of an estimate can be defined as its degree to which it describes the measured phenomena correctly, in other words its difference from the true (though often unobserved) population value. In contrast, the variability of an estimate relates to how closely repeated observations confirm themselves (e.g., under random sampling). Figure 1 reproduces a visual summary of these two considerations from Ferrante and Cameriere (2009).

The MSE is the second moment about the origin of the error and thus takes into account both bias and variance. It is therefore an appropriate measure for our proposed estimation of uncertainty around spatial microsimulation SAE estimator. For the design unbiased direct survey estimates, the MSE is equal to the variance, while for the indirect SAE estimator, the MSE is equal to the

bias squared plus the variance. As such, the attractiveness of any indirect SAE estimator compared to the unbiased direct estimator is dependent upon the reductions in variance in any SAE approach exceeding its increases to bias such that the MSE of the indirect SAE estimator is smaller than the MSE of the direct estimator. This is possible in SAE contexts where small areas have low or no survey sample sizes such that direct small area estimates are either nonviable or come with large variance.

Capturing the MSE via resampling techniques such as the bootstrap is common within regression-based SAE approaches (González-Manteiga et al. 2008b; Marchetti et al. 2018; Moretti, Shlomo, and Sakshaug 2018). Indeed, González-Manteiga et al. (2008b) point out that even when analytical approximations are available, bootstrap resampling might provide more accurate estimates due to its second-order accuracy, a property discussed further in Efron and Tibshirani (1993). However, no similar MSE measures have yet been considered in the spatial microsimulation SAE context where MSE expressions are not available in closed form and where empirical approaches are therefore necessary to explore (Chen and Shen 2015; D'Arrigo and Skinner 2010). This is particularly relevant since the estimation of bias in particular has proven elusive in previous attempts (Chen and Shen 2015; Nagle et al. 2014; Whitworth et al. 2017).

This article responds to this gap through its modification of bootstrap ideas for regression-based small area estimators set out in González-Manteiga et al. (2008b), so that they become suited to the differing technical processes and requirements of spatial microsimulation approaches. The use of models in the estimation of MSE of model-assisted estimators can be found in the regression estimator context where a model unbiased conditional MSE estimator is proposed (Kott 2009). This provides initial motivation for this article to develop and adapt such an approach in the context of spatial microsimulation in order to estimate the MSE of $\hat{\bar{Y}}_d^{IPF}$.

In order to provide an estimator of the MSE of $\hat{\bar{Y}}_d^{IPF}$, denoted by $MSE\left(\hat{\bar{Y}}_d^{IPF}\right)$, we assume that the observations y_{di} for unit i in area d are related to $\boldsymbol{x_{di}} = \left(x_{di1}, \ldots, x_{dip}\right)^T$ denoting a vector of p auxiliary variables, via the Battese, Harter, and Fuller (1988) linear nested-error (i.e., multilevel) regression model:

$$\begin{gathered} y_{di} = \boldsymbol{x}_{di}^T\boldsymbol{\beta} + u_d + e_{di},\ i = 1, \ldots, N_D,\ d = 1, \ldots, D \\ u_d \overset{iid}{\sim} N\left(0, \sigma_u^2\right),\ e_{di} \overset{iid}{\sim} N\left(0, \sigma_e^2\right),\ \textit{independent}, \end{gathered} \tag{3}$$

where u_d and e_{di} are the area random effect and the residual error term, respectively. As with all SAE methods, these are assumed to be independent (Rao and Molina 2015). The model assumes that the population has a two-level structure where units are nested in areas. This is reasonable in the SAE context given the aim to estimate target parameters of small domains in the population. By doing so, the approach recognizes explicitly that the MSE will be based on a multilevel (two-level) structure and that the intraclass correlation (ICC) will therefore have relevance. The ICC describes the extent to which units (e.g., individuals) within the same higher level unit (e.g. areas) are similar to one another (R. Koch 2008). To sensitivity test this issue, the simulation study below explicitly tests the effect of differing ICCs on the MSE estimator.

Under model (3), our proposed estimator of $MSE\left(\hat{\bar{Y}}_d^{IPF}\right)$ can be derived via a parametric bootstrap by adapting the principles in González-Manteiga et al. (2008b) for the different challenges of the spatial microsimulation context. The algorithm steps for the bootstrap MSE for IPF are listed below for $b = 1, \ldots, B$ bootstrap replications where the symbol * is used to denote the bootstrap quantities and for $d = 1, \ldots, D$ small areas:

1. Fit model (3) to the observed sample data, denoted by s, and estimate the model parameters. The estimates are denoted by $\hat{\boldsymbol{\beta}}$, $\hat{\sigma}_u^2$, $\hat{\sigma}_e^2$;
2. Generate the bootstrap area effects $u_d^{*(b)} \overset{iid}{\sim} N\left(0, \hat{\sigma}_u^2\right)$;
3. Generate the bootstrap residual error term $e_{di}^{*(b)} \overset{iid}{\sim} N\left(0, \hat{\sigma}_e^2\right)$, independently of $u_d^{*(b)}$, for every unit i in the sample in area d, for the sample units, $i \in s_d$;
4. Calculate the true population means for each small area of the bootstrap population as follows:

$$\bar{Y}_d^{*(b)} = \bar{\mathbf{x}}_{d,\,pop}^T \hat{\boldsymbol{\beta}} + u_d^{*(b)} \tag{4}$$

where $\bar{x}_{d,\,pop}$ denotes the means of the known population auxiliary variables for each area d. These may be taken, for instance, from the census or administrative data.

5. Generate the bootstrap data as follows , $i \in s_d$:

$$y_{di}^{*(b)} = \mathbf{x}_{di}^T \hat{\boldsymbol{\beta}} + u_d^{*(b)} + \mathbf{e}_{di}^{*(b)}, \tag{5}$$

noting that (5) follows model (3).

6. Compute the IPF estimator defined in (2) on $y_{di}^{*(b)}$ and obtain the IPF estimates on the bootstrap data $\hat{\bar{Y}}_d^{IPF*(b)}$;
7. Repeat steps (2) through (6) for $b = 1, \ldots, B$ for each area $d = 1, \ldots, D$.

An estimator of $MSE\left(\hat{\bar{Y}}_d^{IPF}\right)$ is given by the following Monte Carlo approximation:

$$\widehat{MSE}_{boot}\left(\hat{\bar{Y}}_d^{IPF}\right) = B^{-1}\sum_{b=1}^{B}\left(\hat{\bar{Y}}_d^{IPF*(b)} - \bar{Y}_d^{*(b)}\right)^2. \tag{6}$$

Simulation Study

This section presents the findings from a simulation study to examine the performance of our proposed MSE bootstrap estimator for spatial microsimulation SAE. For a classificatory work on simulation studies in SAE and further theoretical details, we refer to Münnich (2014). In this model-based simulation, $S = 1,000$ populations are generated from model (3), given that estimators such as the MSE estimator depend on model assumptions and hence that it is important to evaluate the statistical properties of our proposed approach under the model. There are no significant computational barriers to the approach, and this is an important practical consideration. Using a standard modern machine, the simulation study took around 10 hours to perform and the application of municipality income in Tuscany in the fifth section took around 20 minutes to perform.

The parameters for the simulation are selected from the LANDSAT data that are widely used in SAE simulation settings. These are survey and satellite data for corn and soybeans in 12 Iowa counties obtained from the 1978 June survey of the U.S. Department of Agriculture and from land observatory satellites (see Battese et al. 1988; Datta, Day, and Basawa 1999; Moretti et al. 2018). The small area problem arises in these data since small area sample sizes are small. The simulations are computationally intensive in large population dimensions and are therefore controlled for the purposes of this simulation. All analyses are conducted in R, and details on the code and functions used for the bootstrap are provided in the Online Appendix (which can be found at http://smr.sagepub.com/supplemental/).

Generating the Population

The population is generated using the following parameters: $N = 20,000$, $D = 80$, and $130 \leq N_d \leq 420$. $N_d,\ d = 1, \ldots, D$ is generated from the discrete uniform distribution, $N_d \sim dUnif(130,\ 420)$, with $\sum_{d=1}^{D} N_d = 20,000$. y_{di} observations are generated according to the following model:

$$\begin{gathered} y_{di} = \mathbf{x}_{di}^T \boldsymbol{\beta} + u_d + e_{di},\ i = 1, \ldots, N_D,\ d = 1, \ldots, D, \\ u_d \overset{iid}{\sim} N\left(0, \sigma_u^2\right), e_{di} \overset{iid}{\sim} F\left(0, \sigma_e^2\right),\ \textit{independent}, \end{gathered} \tag{7}$$

where $F \in \{Normal,\ Gumbel, Logistic\}$. The rationale for sensitivity testing the performance of our bootstrap estimator across these three distribution types is that the MSE is based on a normality assumption. In line with good practice in previous SAE literature (González-Manteiga et al. 2008b), distributions are chosen deliberately in order to sensitivity test how the estimators perform when the error term e_{di} is not normal but is instead skewed (Gumbel) or symmetric with heavy tails (Logistic). All three distribution types are common in real data applications.

The auxiliary variables are defined as follows:

$$\mathbf{x}_{di} = \left(1\ x_{di1}\ x_{di2}\right)^T,\ \text{with } x_{di1} \sim dUnif(145,\ 459) \text{ and } x_{di2} \sim dUnif(55,\ 345.)$$

and the regression coefficients are given by the following vector:

$$\boldsymbol{\beta} = \left(17.97\ 0.36\ -0.03\right)^T.$$

As noted above, since the data are assumed to take a multilevel structure (units inside areas), the ICC will have relevance and requires sensitivity testing. The ICC varies across applications dependent upon the extent to which the variability in the data observes a hierarchal structure with, for example, values ranging across 0.005, 0.05, and 0.2 with respect to mortality (Ambugo and Hegn, 2015), fear of crime (Whitworth 2012), and well-being (Moretti et al. 2019), respectively. Given that the ICC plays a role in the performance of the model-based MSE estimator, it is important that sensitivity tests are performed around its value within the simulation (Molina, Nandram, and Rao 2014; Moretti et al. 2018). We use the following relationships to explore the role of the ICC in the case of the Normal distribution:

$$\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2},\ \sigma_u^2 = -\frac{\rho}{\rho - 1}\sigma_e^2 \text{ with } \sigma_e^2 = 297.71 \text{ and}$$
$$\rho \in \{0.01, 0.03,\ 0.05, 0.08,\ 0.10,\ 0.15,\ 0.20,\ 0.50\}.$$

Due to space constraints, in the cases of the Gumbel and Logistic distributions, the ICC is set at a realistic value of 0.05 only (see, e.g., Moretti et al. 2019).

In order to produce the small area IPF estimates, we create the following classes identifying the benchmark constraints related to the covariates x_1 and x_2:

$$145 \leq x_{1i} \leq 224.20,\ 224.20 < x_{1i} \leq 380.70,\ 380.70 < x_{1i} \leq 459,$$

$$55 \leq x_{2i} \leq 126.30,\ 126.30 < x_{2i} \leq 272.10,\ 272.10 < x_{2i} \leq 345.$$

Simulation Steps

The simulation consists of the following steps:

1. Population generation: Generate the responses y_{dis} according to model (7) for $s = 1, \ldots, S$, $(S = 1,000)$ with parameters presented above;
2. Draw a stratified random sample with simple random sample without replacement selection in each area d from each simulated population, $n_d \sim dUnif(7,\ 21)$, with $n = \sum_d n_d = 1,129$. The overall sampling fraction is given by $f = \frac{n}{N} = 5.6\%$. Across the areas, these takes values between 1.91 percent and 15.33 percent;
3. Estimate $\bar{Y}_d$ via the IPF estimator given in (2) in each sample and obtain $\hat{\bar{Y}}_{ds}^{IPF}$;
4. Estimate $\bar{Y}_d$ via the direct estimator given by $\hat{\bar{Y}}_d^{Direct}$ in each sample and obtain $\hat{\bar{Y}}_{ds}^{Direct}$;
5. Estimate the MSE of $\hat{\bar{Y}}_{ds}^{IPF}$ via parametric bootstrap described in the fourth section (with $B = 500$), the estimate is denoted by $\hat{MSE}_{boot}\left(\hat{\bar{Y}}_{ds}^{IPF}\right)$.

In order to evaluate the performance of the proposed bootstrap MSE estimator, the following quality measures are calculated:

Empirical MSE of $\hat{\bar{Y}}_d^{IPF}$ (the true MSE):

$$EMSE\left(\hat{\bar{Y}}_d^{IPF}\right) = S^{-1}\sum\nolimits_s \left(\hat{\bar{Y}}_{ds}^{IPF} - \bar{Y}_{ds}\right)^2. \tag{8}$$

Empirical MSE of $\hat{\bar{Y}}_d^{Direct}$:

$$EMSE\left(\hat{\bar{Y}}_d^{Direct}\right) = S^{-1}\sum_s \left(\hat{\bar{Y}}_{ds}^{Direct} - \bar{Y}_{ds}\right)^2. \tag{9}$$

Relative bias of $\widehat{\mathrm{MSE}}_{boot}\left(\hat{\bar{Y}}_d^{IPF}\right)$:

$$RB\left(\widehat{\mathrm{MSE}}_{boot}\left(\hat{\bar{Y}}_d^{IPF}\right)\right) = S^{-1}\sum\nolimits_s \frac{mse_{boot}\left(\hat{\bar{Y}}_{ds}^{IPF}\right) - EMSE\left(\hat{\bar{Y}}_d^{IPF}\right)}{EMSE\left(\hat{\bar{Y}}_d^{IPF}\right)}. \tag{10}$$

Relative bias of $\hat{\bar{Y}}_d^{IPF}$:

$$\frac{RB\left(\hat{\bar{Y}}_d^{IPF}\right) = S^{-1}\sum_s \hat{\bar{Y}}_{ds}^{IPF} - \bar{Y}_{ds}}{\bar{Y}_{ds}}. \tag{11}$$

Relative bias of $\hat{\bar{Y}}_d^{Direct}$:

$$\frac{RB\left(\hat{\bar{Y}}_d^{Direct}\right) = S^{-1}\sum_s \hat{\bar{Y}}_{ds}^{Direct} - \bar{Y}_{ds}}{\bar{Y}_{ds}}. \tag{12}$$

where $\bar{Y}_{ds} = \sum_{i=1}^{N_d} y_{dis}/N_d$. The true small area means are denoted by $\bar{Y}_d = \sum_{s=1}^{S}\sum_{i=1}^{N_d} y_{dis}/N_d$

In the following section, $ERMSE\left(\hat{\bar{Y}}_d^{IPF}\right) = \sqrt{EMSE\left(\hat{\bar{Y}}_d^{IPF}\right)}$ is used to denote the empirical root MSE (and $ERMSE\left(\hat{\bar{Y}}_d^{Direct}\right)$ for $\hat{\bar{Y}}_d^{Direct}$). These quality measures are evaluated and compared across the areas using the median as a robust central tendency measure (Chambers, Chandra, and Tzavidis 2011; Giusti et al. 2013).

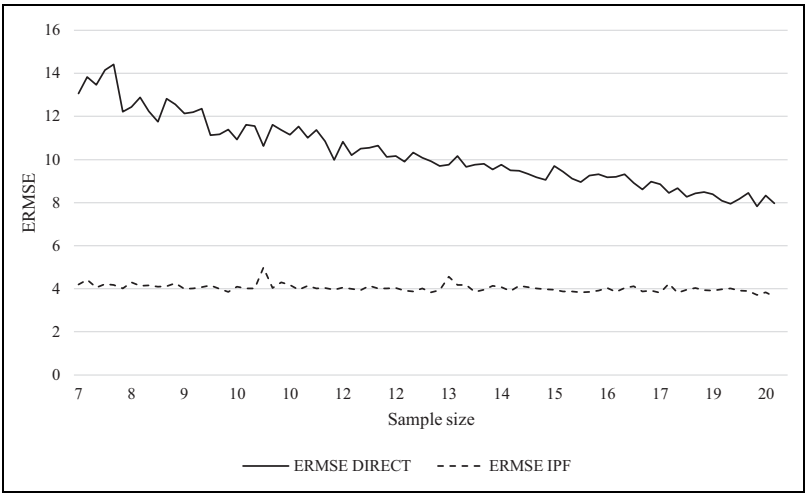

Figure 2. Empirical root mean squared error comparisons: Direct versus iterative proportional fitting estimator for the Normal case.

Results

Performance of the bootstrap MSE estimator under different distributional assumptions of e_{di} *(*$\rho = 0.05$*).* Figure 2 presents the empirical root MSE (ERMSE) of the IPF estimator and the direct estimator. Since the performance is very similar across all three distributions, Figure 2 presents the findings for the Normal case only, ordered by increasing small area sample size.

It can be seen that the IPF synthetic estimator provides estimates with lower MSE than the direct estimator due to the use of related auxiliary variables in reducing variance. This is particularly true for small areas with smaller survey sample sizes where the performance gains of IPF are large relative to direct estimator. This occurs because the ERMSE of the direct estimator naturally depends on the survey sample size: When the survey sample size in the small area d is smaller, the ERMSE tends to increase due to the larger variance around such estimates. As the small area survey sample size increases, the performance gains of the synthetic IPF estimator decline relative to the direct estimator until a point where its performance converges with that of the direct estimator. For reference, Figure 2 looks identical when ordered by sampling fraction rather than sample size.

For each distribution, Figure 3 shows the IPF point estimates across the small areas plotted against the true values observed in the population.

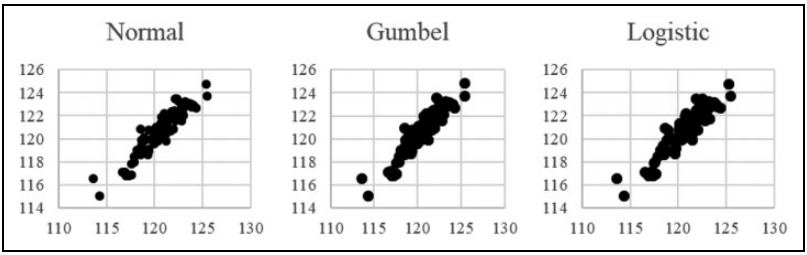

Figure 3. Comparisons of iterative proportional fitting estimates versus true means.

Table 1. Point Estimates Comparisons and Relative Biases Across the Small Area, Median Values Where Intraclass Correlation = 0.05.

	Scenario		
Performance Measure	Normal	Gumbel	Logistic
$\bar{Y}_d$	120.696	120.595	120.714
$\hat{\bar{Y}}_d^{Direct}$	120.720	120.756	120.964
$\hat{\bar{Y}}_d^{IPF}$	120.662	120.677	120.687
$RB\left(\hat{\bar{Y}}_d^{Direct}\right)$	0.000	0.000	0.000
$RB\left(\hat{\bar{Y}}_d^{IPF}\right)$	0.006	0.004	0.004

Table 1 shows the median estimate comparisons obtained using $\hat{\bar{Y}}_d^{Direct}$ and $\hat{\bar{Y}}_d^{IPF}$ under the different distributional scenarios. The true value $\bar{Y}_d$ is shown in the first row, while the direct and IPF central estimates are shown in rows 2 and 3, respectively. The relative bias of those direct and IPF estimates across all the small areas are then shown in the penultimate two rows. It can be seen that the IPF small area estimator returns only negligible biases across the small area even in cases of Gumbel and Logistic distributions of the error term.

Figure 4 moves on to focus on the performance of the bootstrap MSE estimator to calculate the uncertainty around those central small area point estimates in the three Normal, Gumbel, and Logistic distributions, respectively. It can be seen that our proposed bootstrap approach provides nearly unbiased estimates of the true MSE with relative bias centered on and close to zero across the small areas. No association is found between estimate bias and sampling fraction across Figure 4.

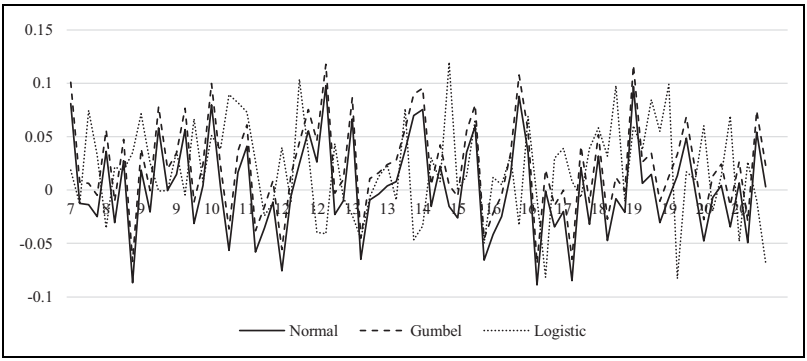

Figure 4. $RB\left(\hat{MSE}_{boot}\left(\hat{\bar{\mathbf{Y}}}_{\mathbf{d}}^{IPF}\right)\right)$, Normal, Gumbel, and Logistic cases.

Table 2. Performance Measures of the Bootstrap MSE Estimates.

	Scenario		
Performance Measure	Normal	Gumbel	Logistic
$EMSE\left(\hat{\bar{Y}}_d^{IPF}\right)$	16.114	17.159	15.990
$\widehat{MSE}_{boot}\left(\hat{\bar{Y}}_d^{IPF}\right)$	14.592	15.865	14.001
$RB\left(\widehat{MSE}_{boot}\left(\hat{\bar{Y}}_d^{IPF}\right)\right)$	−0.076	−0.080	−0.079
Coverage rates	0.918	0.915	0.914

Note: MSE = mean squared error; EMSE = empirical mean squared error.

Table 2 provides further details of the performance of the bootstrap MSE estimator across the three distributions. In particular, the true MSE (i.e., empirical MSE) in each distribution is compared to our bootstrap MSE estimate and coverage rates (of 95 percent confidence intervals) are also presented. It can be seen that the relative bias values are close to zero and that the MSE bootstrap estimator is nearly unbiased across the small areas in each of the three distributions.

On the role of the ICC in the case of the Normal distribution. On the basis of these analyses, the performance of our proposed MSE estimator is encouraging in terms both of the small relative bias of the MSE estimates and the quality of

the coverage. However, as noted above, it may be that the magnitude of the ICC affects the performance of the MSE estimator.

Table 3 shows the findings of ICC sensitivity analyses with a focus on the Normal distribution where the ICC is varied at several points from 0.01 to 0.50. The results for $\rho = 0.05$ shown above are repeated to aid comparison. It can be seen that the relative bias of $\hat{\bar{Y}}_d^{IPF}$ increases slightly when the ICC increases beyond around 0.15, ranging from -0.001 when $\rho = 0.01$–0.006 when $\rho = 0.2$ and 0.021 when $\rho = 0.50$. In terms of the bootstrap MSE estimator, the penultimate row shows that this estimator delivers consistently small relative bias in the MSE though with somewhat weaker performance in coverage at very low levels of ICC as displayed in the final row. When the ICC is small, the MSE is slightly underestimated as seen by looking the relative bias and by comparing the empirical MSE (line two) with our bootstrap MSE (line 3). However, these relative bias estimates remain acceptable.

Application to Small Area Income Estimation in Italian Municipalities

This section provides a real-world application of an IPF small area estimator and, more centrally for this article, of our proposed bootstrap estimator of its MSE. The application used is the estimation of mean equivalized annual household disposable income (in Euros) for the municipalities of Tuscany region ($D = 287$). The survey data used are provided from the 2009 European Union SILC (EU-SILC). These EU-SILC data contain a sample of 1,448 households for Tuscany. EU-SILC is designed to deliver estimates at the national and also regional (NUTS-2) level (Giusti, Masserini, and Pratesi 2015). Therefore, this situation is typical of most survey situations in that EU-SILC cannot be used to derive usable income estimates at smaller sub-regional geographies such as municipalities due to low or zero survey sample sizes. The household income variable of interest is given in the EU-SILC data and equivalized using Eurostat's official modified Organization for Economic Cooperation and Development equivalence scale (Haagenars, de Vos, and Zaidi 1994; Marchetti et al. 2018). The auxiliary variables for the Tuscan municipalities come from the Population Census of Italy.

Model Fitting and Internal Validation

The explanatory variables used in this application are working status, years of education, gender, and age of the survey identified head of household. These have been informed by preliminary model investigations and findings

Table 3. Performance Measures of Our Bootstrap MSE Estimator at Varying Levels of Intraclass Correlation in the Normal Distribution.

	Scenario							
Performance Measure	$\rho = 0.01$	$\rho = 0.03$	$\rho = 0.05$	$\rho = 0.08$	$\rho = 0.10$	$\rho = 0.15$	$\rho = 0.20$	$\rho = 0.50$
$RB\left(\hat{\bar{Y}}_d^{IPF}\right)$	−0.001	0.002	−0.001	0.002	0.001	0.002	0.006	0.021
$EMSE\left(\hat{\bar{Y}}_d^{IPF}\right)$	5.300	11.001	16.114	16.115	16.116	53.426	64.551	293.397
$\widehat{\mathrm{MSE}}_{\mathrm{boot}}\left(\hat{\bar{Y}}_d^{IPF}\right)$	4.734	10.197	14.592	14.650	14.600	52.247	63.220	293.835
$RB\left(\widehat{MSE}_{boot}\left(\hat{\bar{Y}}_d^{IPF}\right)\right)$	−0.080	−0.079	−0.076	−0.075	−0.075	−0.019	−0.018	0.006
Coverage rates	0.917	0.917	0.918	0.924	0.939	0.942	0.950	0.950

Note: MSE = mean squared error; EMSE = empirical mean squared error.

Table 4. Model Results.

Coefficient	Estimates	exp(β)	Standard Error	*p* Value
Intercept	8.460	4,754.748	.105	.000
Gender	0.215	1.236	.031	.000
Working status	0.352	1.422	.041	.000
Age	0.010	1.012	.001	.000
Years of education	0.034	1.035	.003	.000

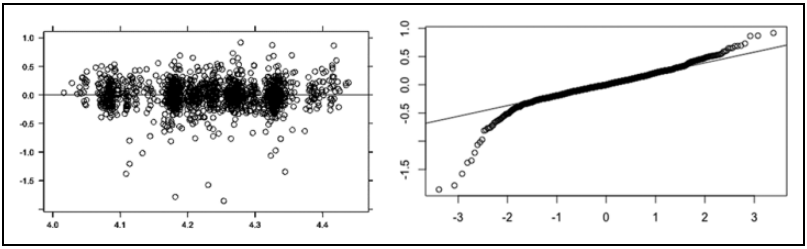

Figure 5. Fitted values versus residuals (left) and *Q*–*Q* plot of the residuals from the model used to produce the mean squared error of the iterative proportional fitting estimates.

from previous studies (Giusti et al. 2013). Model diagnostics identified some skewness and outliers in the distribution of the income outcome variable, as is common with such distributions, and this variable was therefore log transformed. No evidence of leverage was found. Table 4 presents the results from the log-linear linear model in EU-SILC based on (3).

Validation is an important step in any SAE study. SAE models can be validated internally in terms of the underlying model and externally against some known other external data of the target outcome variable. In terms of the internal validation, Figure 5 shows the fitted values versus the residuals as well as the *Q*–*Q* plots of the residuals from the log-linear model used to produce the MSE of the IPF estimates. These show good behavior with respect to the normality assumption. External validation is discussed below.

Estimating Municipality Income in Tuscany

This section discusses the results of the IPF SAE of the mean equivalized annual household disposable income across Tuscan municipalities along

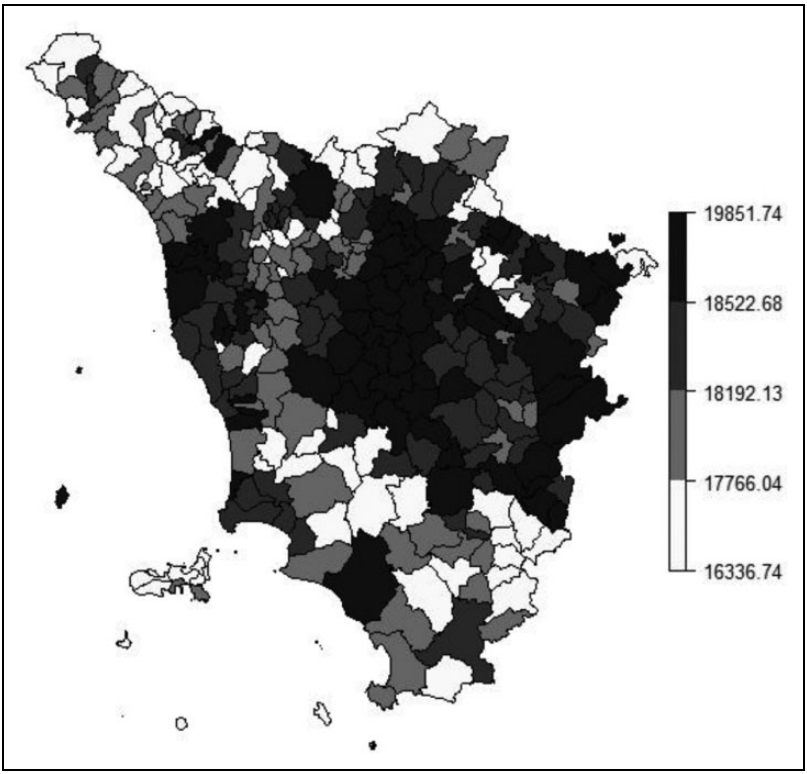

Figure 6. Iterative proportional fitting income estimates for Tuscan municipalities.

with their uncertainty estimates. Figure 6 maps the mean IPF estimates across the 287 Tuscan municipalities. The map displays municipalities in four quartiles and shows a range in 2009 municipal income estimates from a low of just over 16,000 Euros per annum to a high of just under 20,000 Euros per annum. Municipalities located in the provinces of Massa Carrara (North West), Grosseto (South), and Prato and Pistoia (North) show the lowest estimated municipal income levels. On the contrary, municipalities around Florence, Arezzo, Pisa, and Livorno show the highest estimated municipal income levels.

In terms of external validation of these IPF estimates, a frequent inherent challenge, as here, is the typical lack of any such existing small area data against which to validate (hence the motivation for the SAE). External validation of these IPF estimates is provided in two ways. Firstly, the spatial

Table 5. Summary Statistics of the Performance Gains from the Synthetic IPF Estimator Compared to the Direct Estimator for Small Area Income Estimates Across Tuscan Municipalities.

Areas with	Summary Statistic	RMSE IPF	*SD* Direct	RRMSE IPF %	CV Direct %	Gains %
$n_d > 0$	Min.	2,197.50	2,444.02	10.63	11.21	26.46
	Mean	3,289.50	6,663.51	17.50	32.10	32.71
	Median	3,356.00	4,802.52	18.48	26.20	32.80
	Max.	4,431.00	51,750.48	23.51	99.05	94.19
$n_d = 0$	Min.	2,196.80	—	13.11	—	—
	Mean	3,308.03	—-	19.27	—	—
	Median	3,346.46	—-	18.98	—	—
	Max.	4,756.51	—-	24.51	—	—

Note: MSE = mean squared error; CV = coefficient of variation; RRMSE = relative root mean squared error; IPF = iterative proportional fitting.

patterns in Figure 6 are in line with known geographical patterns of similar indicators across Tuscany seen in previously published research (Giusti et al. 2015; Moretti et al. 2019). Secondly, no identical income indicators exist at this municipality scale, and no direct survey estimates to municipality level are viable from the EU-SILC survey data. However, it is viable to produce direct survey estimates from the EU-SILC survey data to Tuscany's 10 larger provinces and to compare these with indirect IPF estimates also to province level. The Spearman's rank correlation between these two sets of estimates is.93, and this is statistically significant at below the 1 percent level, although acknowledging the limited sample size involved.

Table 5 presents summary statistics of the uncertainty of the direct survey estimates compared with the IPF estimates. In particular, it shows the root mean squared error (column 1) and, expressed as a percentage of the estimates, the relative root mean squared error (RRMSE%; column three) of the small area estimates. Given that the direct estimates are unbiased, the standard deviation (*SD*; column 2) and coefficient of variation (CV%; column 4) of the direct estimates enable a comparison of uncertainty of the direct estimates with the bootstrap MSE estimates of the IPF estimator. The coefficient of variation (CV) is a standardized measure of the dispersion of a distribution and is calculated as a ratio of its *SD* to its mean. In the present analyses, it is obtained as the ratio between the *SD* of the direct survey estimate and the direct survey estimate for every area. Since direct estimates are unbiased, their CVs represent measures of uncertainty (Rao and Molina

2015). RRMSE and CV are standard measures of uncertainty that are required in many official statistics institutes (see Schirripa-Spagnolo, D'Agostino, and Salvati 2018; Statistics Canada 2009).

Table 5 shows that the IPF estimates are more reliable than the direct estimates across all points of the income distribution, as depicted by the lower values of the RRMSE IPF (column 1) and RRMSE IPF (column 3) compared to *SD* direct (column 2) and CV direct (column 4), respectively. The IPF small area estimates can also be considered reliable in absolute terms. Values of RRMSE below a threshold of 20 percent are often taken by statistical agencies as acceptable (Commonwealth Department of Social Services 2015), and almost all of this municipality distribution of small area estimates is below this level. The final column of Table 5 summarizes the gains in efficiency of the IPF estimates over the direct survey estimates by comparing the MSE for the IPF estimator with the variance of the unbiased direct estimator. Technically, these are calculated by

$$Gain\left(\hat{\bar{Y}}_d^{IPF}\right) = \frac{MSE\left(\hat{\bar{Y}}_d^{IPF}\right) - Var\left(\hat{\bar{Y}}_d^{Direct}\right)}{Var\left(\hat{\bar{Y}}_d^{Direct}\right)} \times 100,\ d = 1, \ldots, D, \tag{13}$$

where $Var\left(\hat{\bar{Y}}_d^{Direct}\right)$ denotes the variance of $\hat{\bar{Y}}_d^{Direct}$. Equation (13) denotes a measure of gain in efficiency of using an estimator with higher precision compared to the direct estimator. We refer to Särndal et al (1992) for measures of efficiency in survey statistics and to Moretti et al. (2018) and González-Manteiga et al. (2008a) for some examples of their use. Results for $n_d = 0$ (municipalities with zero sample size) and $n_d > 0$ (municipalities with some households in the survey) have been separated because it is not possible to compute direct estimates (and as consequently the gains) for areas with zero sample size. Table 5 shows that the small area estimates from the indirect IPF estimator provide significant performance gains compared to the direct estimator at all points of the municipality income distribution.

Figure 7 drills down to focus on the extent to which these performance gains vary according to the size of the municipality sample size in the EU-SILC survey, a key driver of the variance of the direct estimator and key limiter of the viability of using the direct estimator to produce reliable survey estimates for small areas. To aid comparison, Figure 7 is ordered from left to right by municipality survey sample size in the EU-SILC. RRMSE is shown for the IPF estimates, while CV is shown for the direct estimates. Figure 7

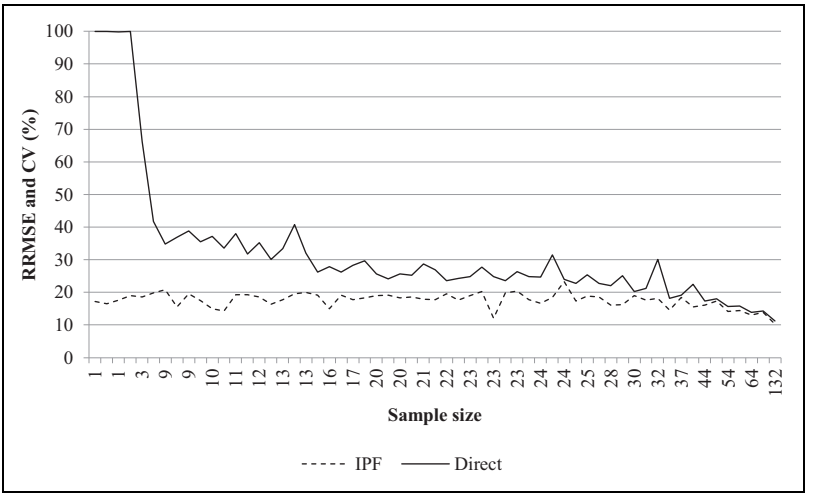

Figure 7. Performance comparison sorted by small area sample size.

illustrates that the RRMSE of the indirect IPF estimates does not depend on the sample size in contrast to the direct estimator. As such, Figure 7 highlights that while performance gains from the IPF estimator are seen across the whole distribution, they increase as the municipality sample size in the survey decreases. Among those municipalities with the smallest survey sample sizes, there is a marked increase in the performance gains available from the IPF estimator compared to the direct survey estimator. Figure 7 is naturally only able to display comparative results for municipalities with nonzero sample sizes, given that direct estimates cannot be produced for small areas with zero sample size. Estimates for these municipalities do of course become viable with synthetic IPF estimator. For reference, Figure 7 looks identical when ordered by sampling fraction rather than sample size.

Discussion

The combination of high costs of survey data collection and increasing demands for ever more spatially detailed data from policy makers and scholars alike mean growing demands for SAE techniques. Spatial microsimulation approaches to SAE continue to be widely used across diverse domains including transport, health, physical activity, and income. However, their continuing inability to produce reliable estimates of uncertainty alongside their central point estimates remains a pressing limitation to their utility for

practitioners and scholars alike. This is understandable in part given that the estimation of MSE is difficult in a spatial microsimulation context since it cannot be estimated in a closed form and analytical approximations are highly challenging.

Widely discussed in the SAE literature is the importance of providing measures of uncertainty such as MSE or confidence intervals alongside the central point estimates in order to assess the reliability of the small area estimates (Pratesi 2016). This is particularly important where policy decisions are taken on the basis of the small area estimates since the consequences of real-world decision making without a clear sense of the uncertainty around the point estimates can be misleading and potentially harmful (Goedemé et al. 2013). Klevmarken (2002:264) argues that "[T]he credibility of [microsimulation models] with the research community as well as with users will in the long run depend on the application of sound principles of inference in the estimation, testing and validation of these models."

This article provides a significant development in this context by presenting a novel parametric bootstrap approach for the estimation of uncertainty in spatial microsimulation SAE techniques. Importantly, the measure of uncertainty estimated is the MSE that contains both the variance and bias of the estimate. Simulation results demonstrate that under model assumptions, our proposed MSE estimator is relatively unbiased and displays good coverage properties against known true population values. The approach delivers substantial performance gains compared to the direct estimator across all portions of the distribution. In doing so, our approach enables researchers and policy makers alike to quantify both the performance gains potentially available through the use of spatial microsimulation approaches to SAE compared to direct survey estimates and to quantify the extent of uncertainty around those small area estimates. The simulation results show that those performance gains exist irrespective of the target small area sample size but are especially large at low sample sizes (below 10 in this simulation) and, naturally, when small areas have zero sample size such that direct estimates are nonviable but synthetic small area estimates are possible. Sensitivity tests confirm that these performance gains are maintained both across nonnormal Gumbel and Logistic distributions as well as across differing values of ICC, though coverage performance falls slightly at very low levels of the ICC in our simulation. A practical data application using EU-SILC data to municipality income across Tuscany is presented and validated in order to demonstrate the applicability and similar performance of the MSE bootstrap estimator in a real-world setting.

While the performance and sensitivity analyses confirm that our proposed approach evaluates well and marks a significant contribution to the field, it serves also to open up opportunities for further more advanced enquiry as a result. First, we focus here only on a linear model. Future work will need to explore the performance of other nonlinear models in this context. However, the bootstrap approach in these scenarios will follow the same steps as those proposed in our approach. Second, a clear next step is to extend the framework to different types of outcome variables beyond the scalar target variable assessed here. Third, in this study, the common case of nonnormal error terms is assessed, and our simulation study shows good performance of the MSE estimator in distributions with mild skew and heavy tails as well as in the normal case. However, future work could take into account more fully the implications of and potential responses to failures in model assumptions and of model failure. Our hope therefore is that this article's contributions will not only make a significant advance to research and policy practice in spatial microsimulation SAE, but that it will also stimulate further scholarly attention to these and other areas.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the UK Economic and Social Research Council National Centre for Research Methods (grant number ES/N011619/1).

ORCID iD


Angelo Moretti https://orcid.org/0000-0001-6543-9418

Supplemental Material

Supplemental material for this article is available online.

References

Anderson, B. 2007. *Creating Small Area Income Estimates for England: Spatial Microsimulation Modelling*. London, England: Department of Communities and Local Government.

Anderson, B. 2013. "Estimating Small-area Income Deprivation: An Iterative Proportional Fitting Approach" Pp.19 in *Spatial Microsimulation: A Reference Guide for Users*, edited by R. Tanton and K. L. Edwards. the Netherlands: Springer.

Ambugo, E. A. and T. P. Hagen. 2015. A multilevel analysis of mortality following acute myocardial infarction in Norway: do municipal health services make a difference? *BMJ Open* 5:e008764.

Battese, G. E., R. M. Harter, and W. 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, W. R., W. W. Basel, and J. J. Maples. 2016. "An Overview of the U.S. Census Bureau's Small Area Income and Poverty Estimates Program" Pp. 349 in *Analysis of Poverty Data by Small Area Estimation*, edited by M. Pratesi. London, England: Wiley.

Benavent, R. and D. Morales. 2016. "Multivariate Fay–Herriot Models for Small Area Estimation." *Computational Statistics & Data Analysis* 94:372-90.

Campbell, M. and D. Ballas. 2013. "A Spatial Microsimulation Approach to Economic Policy Analysis in Scotland." *Regional Science Policy and Practice* 5(3): 263-89.

Chambers, R., H. Chandra, and N. Tzavidis. 2011. "On Bias-robust Mean Squared Error Estimation for Pseudo-linear Small Area Estimators." *Survey Methodology* 37:153-70.

Chen, H. and R Shen. 2015. "Variance Estimation for Survey-weighted Data Using Bootstrap Resampling Methods: 2013 Methods-of-payment Survey Questionnaire." *Technical Reports 104*. Bank of Canada. Retrieved from (https://ideas.repec.org/s/bca/bocatr.html).

Chin, S. F. and A. Harding. 2006. *Regional Dimensions: Creating Synthetic Small-area Microdata and Spatial Microsimulation Models*. Online Technical Paper—TP33, NATSEM, University of Canberra.

Clarke, M. and E. Holm. 1987. "Microsimulation Methods in Spatial Analysis and Planning." *Geografiska Annaler, Series B, Human Geography* 69(2):145-64.

Commonwealth Department of Social Services. 2015. *Survey of Disability, Ageing and Carers, 2012: Modelled Estimates for Small Areas*, Projected 2015. Prepared by the Regional Statistics National Centre, ABS (Release 1: February 2015). Retrieved August 10, 2018 (https://www.health.gov.au/internet/main/publishing.nsf/Content/98DCE47FC10BDD51CA257F15000413F5/$File/SDAC%202012%20Modelled%20Estimates%20for%20Small%20Areas%20projected%202015_Explanatory%20Notes%20-%20Release%201.pdf).

Cullinan, J., S. Hynes, and C. O'Donoghue. 2006, "The Use of Spatial Microsimulation and Geographic Information Systems (GIS) in Benefit Function Transfer—An Application to Modelling the Demand for Recreational Activities in Ireland."

Paper presented at the 8th Nordic Seminar on Microsimulation models, Oslo, Norway, June 7–9.

D'Arrigo, J. and C. Skinner. 2010. "Linearization Variance Estimation for Generalized Raking Estimators in the Presence of Nonresponse," *Survey Methodology* 36(2):181-92.

Datta, G. S., B. Day, and I. Basawa. 1999. "Empirical Best Linear Unbiased and Empirical Bayes Prediction in Multivariate Small Area Estimation." *Journal of Statistical Planning and Inference* 75:269-79

Deville, J. C. and C. E. Särndal. 1992. "Calibration Estimators in Survey Sampling." *Journal of the American Statistical Association* 87(418):376-82.

Dodge, Y. and D. Commenges. eds. 2006. *The Oxford Dictionary of Statistical Terms*. Oxford, England: Oxford University Press on Demand.

Edwards, K. L., G. P. Clarke, J. K. Ransley, and J. E. Cade. 2010. "The Neighbourhood Matters: Studying Exposures Relevant to Childhood Obesity and the Policy Implications in Leeds, UK." *Journal of Epidemiology and Community Health* 64(3):194-201.

Efron, B. and R. Tibshirani. 1993. *An Introduction to the Bootstrap*. London, England: Chapman and Hall.

Espuny-Pujol, F., K. Morrissey, and P. Williamson. 2018. "A Global Optimization approach to Range-restricted Survey Calibration." *Statistics and Computing* 28: 427-39.

Ferrante, L. and R. Cameriere. 2009. "Statistical Methods to Assess the Reliability of Measurements in the Procedures for Forensic Age Estimation." *International Journal of Legal Medicine* 123(4):277-83.

Fuller, W.A.2002. "Regression Estimation for Survey Samples." *Survey Methodology* 28(1):5-23

Giusti, C., L. Masserini, and M. Pratesi. 2015. "Local Comparisons of Small Area Estimates of Poverty: An Application within the Tuscany Region of Italy." *Social Indicators Research* 131(1):235-54.

Giusti, C., N. Tzavidis, M. Pratesi, and N. Salvati. 2013. "Resistance to Outliers of m-Quantile and Robust Random Effects Small Area Models." *Communications in Statistics: Simulation and Computation* 43:549-68.

Goedemé, T., K. Van den Bosch, L. Salanauskaite, and G. Verbist. 2013. "Testing the Statistical Significance of Microsimulation Results: A Plea." *International Journal of Microsimulation* 6(3):50-77.

González-Manteiga, W., M. J. Lombardía, I. Molina, D. Morales, and L. Santamaría. 2008a. "Analytic and Bootstrap Approximations of Prediction Errors under a Multivariate Fay-Herriot Model." *Computational Statistics and Data Analysis* 52:5242-52.

González-Manteiga, W., M. J. Lombardía, I. Molina, D. Morales, and L. Santamaría. 2008b. "Bootstrap Mean Squared Error of a Small-area EBLUP." *Journal of Statistical Computation and Simulation* 78:443-62.

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

Horvitz, D. G. and D. J. Thompson. 1952. "A Generalization of Sampling without Replacement from Finite Universe." *Journal of the American Statistical Association* 47:663-85.

Ipsos MORI. 2018. *Small Area Estimation of Sport Participation and Activity Using Data from the Active Lives Survey*. London, England: Ipsos MORI.

Johnson, A., F. H. Chandra, J. Brown, and S. P. Sabu. 2012. "Small Area Estimation for Policy Development: A Case Study of Child Undernutrition in Ghana." *Journal of the Indian Society of Agricultural Statistics* 66(1):171-86.

Klevmarken, A. N. 2002. "Statistical Inference in Micro-simulation Models: Incorporating External Information." *Mathematics and Computers in Simulation* 59(1-3):255-65.

Koch, R. 2008. *The 80/20 Principle: The Secret of Achieving More with Less*. New York: Doubleday.

Kolenikov, S. 2014. "Calibrating Survey Data Using Iterative Proportional Fitting (Raking)." *The Stata Journal* 14(1):22-59.

Kott, P. S. 2009. "Calibration Weighting: Combining Probability Samples and Linear Prediction Models" Pp. 55 in *Handbook of Statistics 29B Sample Surveys: Inference and Analysis*, edited by D. Pfeffermann and C. R. Rao. North Holland: Elsevier.

Lovelace, R. and M. Dumont. 2016. *Spatial Microsimulation With R*. Boca Raton, FL: Chapman and Hall/CRC.

Marchetti, S., M. Beręsewicz, N. Salvati, M. Szymkowiak, and Ł. Wawrowski. 2018. "The Use of a Three-level M-quantile Model to Map Poverty at Local Administrative Unit 1 in Poland." *Journal of Royal Statistical Society—Series A* 181(4):1077-1104.

Marshall, A. 2010. "Small Area Estimation Using ESDS Government Surveys—An Introductory Guide." *Economic and Social Data Service*: P. 15.

Molina, I., B. Nandram, and J. N. K. Rao. 2014. "Small Area Estimation of General Parameters with Application to Poverty Indicators: A Hierarchical Bayes Approach." *The Annals of Applied Statistics* 8(2):852-85.

Moretti, A., N. Shlomo, and J. W. Sakshaug. 2018. "Parametric Bootstrap Mean Squared Error of a Small Area Multivariate EBLUP." *Communications in Statistics—Simulation and Computation* 49(6):1474-86

Moretti, A. and A. Whitworth. 2019. "Development and Evaluation of an Optimal Composite Estimator in Spatial Microsimulation Small Area Estimation," *Geographical Analysis* 52(3):351-70.

Moretti, A., N. Shlomo, and J. W. Sakshaug 2019. Small Area Estimation of Latent Economic Well-being. *Sociological Methods & Research*.

Münnich, R. 2014. "Small Area Applications: Some Results from a Design-based View." *International Small Area Estimation Conference, SAE 2014, Poznan, Poland*. Retrieved January 13, 2021 (http://www.sae2014.ue.poznan.pl/presentations/SAE2014_Ralf_Munnich_c330a31c0a.pdf).

Nagle, N., B. Buttenfield, S. Leyk, and S. Spielman. 2014. "Dasymetric Modeling and Uncertainty." *Annals of the Association of American Geographers* 104(1):80-95.

Office for National Statistics (ONS). 2019. *Research Outputs: Small Area Estimation of Fuel Poverty in England, 2013 to 2017*. London, England: Office for National Statistics.

Pratesi, M. 2016. *Analysis of Poverty Data by Small Area Estimation*. London, England: Wiley.

Rahman, A. and A. Harding. 2017. *Small Area Estimation and Microsimulation Modeling*. Boca Raton, FL: CRC Press Taylor & Francis Group.

Rahman, A., A. Harding, R. Tanton, and S. Liu. 2013. "Simulating the Characteristics of Populations at the Small Area Level: New Validation Techniques for a Spatial Microsimulation Model in Australia." *Computational Statistics & Data Analysis* 57(1):149-65.

Rao, J. N. K. and I. Molina. 2015. *Small Area Estimation*. New York: Wiley.

Ravulaparthy, S. and K. Goulias. 2011. *Forecasting with Dynamic Microsimulation: Design, Implementation, and Demonstration: Final Report on Review, Model Guidelines, and a Pilot Test for a Santa Barbara County Application*. Technical Report, May, California: University of California Transportation Center (UCTC).

Särndal, C. E., B. Swensson, and J. Wretman. 1992. *Model Assisted Survey Sampling*. the Netherlands: Springer.

Schirripa-Spagnolo, F., A. D'Agostino, and N. Salvati. 2018. "Measuring Differences in Economic Standard of Living between Immigrant Communities in Italy." *Quality and Quantity* 52:1643-67.

Simpson, L and M. Tranmer. 2005. "Combining Sample and Census Data in Small Area Estimates: Iterative Proportional Fitting with Standard Software." *The Professional Geographer* 57(2):222-34.

Statistics Canada. 2009. *Statistics Canada Quality Guidelines*. Catalogue no. 12-539-X. https://www150.statcan.gc.ca/n1/en/catalogue/12-539-X

Tanton, R. and K. L. Edwards. 2013. *Spatial Microsimulation: A Reference Guide for Users*. the Netherlands: Springer.

Tanton, R., J. McNamara, A. Harding, and T. Morrison. 2009. "Rich Suburbs, Poor Suburbs? Small Area Poverty Estimates for Australia's Eastern Seaboard in 2006" Pp. 79 in *New Frontiers in Microsimulation Modelling*, edited by A. Zaidi, Harding, and P. Williamson. London, England: Ashgate.

Tanton, R., P. Williamson, and A. Harding. 2014. "Comparing Two Methods of Reweighting a Survey File to Small Area Data." *International Journal of Microsimulation* 7(1):76-99.

Tribby, C. P. and P. A. Zandbergen. 2012. "High-resolution Spatio-temporal Modeling of Public Transit Accessibility." *Applied Geography* 34:345-55.

Whitworth, A. (2012) Sustaining evidence-based policing in an era of cuts: Estimating fear of crime at small area level in England. *Crime Prevention & Community Safety* 14(1): 48-68.

Whitworth, A. (Ed.). 2013. "*Evaluations and Improvements in Small Area Estimation Methodologies."* National Centre for Research Methods Methodological Review Paper, Economic and Social Research Council.

Whitworth, A. and E. Carter. 2015. *Understanding Wales at the Small Area Level: Maximising the Performance of Small Area Estimation*. Cardiff, Wales: Welsh Government.

Whitworth, A., E. Carter, D. Ballas, and G. Moon. 2017. "Estimating Uncertainty in Spatial Microsimulation Approaches to Small Area Estimation: A New Approach to Solving an Old Problem." *Computers, Environment and Urban Systems* 63: 50-57.

Williamson, P., M. Birkin, and P. Rees. 1998. "The Estimation of Population Microdata by Using Data from Small Area Statistics and Samples of Anonymised Records." *Environment and Planning A* 30(5):785-816.

World Bank. 2018. *Small Area Estimation of Poverty Under Structural Change*. Washington, DC: World Bank Group: Poverty and Equity Global Practice.

Author Biographies

Angelo Moretti is a lecturer (assistant professor) in the Department of Computing and Mathematics, Manchester Metropolitan University, UK. His research interests are in multivariate small area estimation, survey methodology, well-being and poverty measurement, multivariate statistics, and mixed-effects modeling. He has a PhD in social statistics from the University of Manchester, UK, and an MSc in marketing and market research and BSc in economics from the University of Pisa, Italy.

Adam Whitworth is a senior lecturer in the Department of Geography, University of Sheffield, UK. He is a spatial social policy scholar with ongoing programs of work focusing on the design and analysis of employment activation policies, local integration, and spatial statistical methodologies. He has a PhD in social policy, MSc in comparative social policy, and BA (Hons) degree in politics, philosophy, and economics all from the University of Oxford, UK.