# ECONSTOR 

# Working Paper <br> Area-level small area estimation with missing values 

Research Papers in Economics, No. 14/19

## Provided in Cooperation with:

University of Trier, Department of Economics

Suggested Citation: Burgard, Jan Pablo; Morales, Domingo; Wölwer, Anna-Lena (2019) : Arealevel small area estimation with missing values, Research Papers in Economics, No. 14/19, Universität Trier, Fachbereich IV - Volkswirtschaftslehre, Trier

This Version is available at: https://hdl.handle.net/10419/207046

## Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

## Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

## TUniversität Trier

# Area-Level Small Area Estimation <br> with Missing Values 

Jan Pablo Burgard
Domingo Morales
Anna-Lena Wölwer

Research Papers in Economics
No. 14/19

# Area-Level Small Area Estimation with Missing Values 

Jan Pablo Burgard*<br>Research Institute for Offcial and Survey Statistics (RIFOSS), Trier University, Germany<br>Domingo Morales<br>Operations Research Center, University Miguel Hernández de Elche, Spain<br>Anna-Lena Wölwer<br>Research Institute for Offcial and Survey Statistics (RIFOSS), Trier University, Germany

September 20, 2019


#### Abstract

Model-based small area predictors are derived under the assumption that data files are complete. In application to real data, files may contain missing values. We introduce a variant of the bivariate Fay-Herriot model that takes into account for missing values in one component of the target variable and give fitting algorithms to estimate the model parameters. Based on the new model, we introduce empirical best predictors of domain means and derive an approximation to the mean squared error.


Keywords: Multivariate models, Fay-Herriot model, small area estimation, missing values.

[^0]
## 1 Introduction

Fay and Herriot (1979) considered an area-level linear mixed model and derived an empirical best linear unbiased predictor (EBLUP) of a domain quantity that it is also called FayHerriot (FH) predictor, or FH-EBLUP. It is widely applied in the context of small area estimation (SAE). The idea of the FH predictor is to improve the precision of a direct estimator on a domain of interest by borrowing strength from other domains. This method, however, requires that for every statistic of interest there is a direct estimate available in all the domains of the study. In the common case of unplanned sample sizes within domain, this cannot be guaranteed. Therefore, it appears that for some domains direct estimates may be missing by chance. Furthermore, estimates in domains or table cells can be suppressed for confidentiality reasons, e.g. due to small cell sizes or high sampling errors as it was done in Zayatz (2007).

Since the publication of the FH predictor, many extensions were made to allow for different practical problems. Inter alia, Prasad and Rao (1990) and Datta and Lahiri (2000) proposed mean squared error (MSE) estimators for the FH predictor, Li and Lahiri (2010) and Yoshimori and Lahiri (2014) introduced new adjusted maximum likelihood fitting methods, Ybarra and Lohr (2008), Arima et al. (2017), Burgard et al. (2019a), and Burgard et al. (2019b) studied the effect of measurement errors in the covariates, Pratesi and Salvati (2008), González-Manteiga et al. (2010), Articus and Burgard (2014) and Morales et al. (2015) allow for a heterogeneous dependency structure in the FH model, Datta et al. (1996), González-Manteiga et al. (2008), Porter et al. (2015) and Benavent and Morales (2016) investigated and applied multivariate FH models, Esteban et al. (2012) and Marhuenda et al. (2013) estimated small area poverty proportions under temporal and spatio-temporal Fay-Herriot models respectively. Many other authors have studied further
variants of the Fay-Herriot model adapted to different setups. However, for the problem of missing values in the dependent variable, to our knowledge, there is no approach giving an empirical best predictor (EBP) also for the domains with missings.

This manuscript introduces an EBP for settings where direct estimates are partially missing. Furthermore, it derives an approximation to the MSE of the EBP and the corresponding MSE estimator and it illustrates the potential benefits of the proposed approach under different simulated data scenarios. In addition, the manuscript presents an application of the developed methodology to a SAE problem with publicly available county-level data from the U.S. American Community Survey (ACS).

The manuscript is structured as follows: Section 2 introduces the bivariate Fay-Herriot model, which is the basis for the development of the EBP theory when part of the values of the target variables are missing. Section 3 divides the set of domains in three groups depending on the existence or not of missing values in each of the dependent variables and gives the corresponding EBPs. Section 4 gives algorithms for calculating the maximum likelihood and the residual maximum likelihood estimators of the model parameters. Section 5 derives an approximation to the mean squared error of the best predictor and proposes an explicit-formula estimator. Section 6 presents a parametric bootstrap procedure for estimating the mean squared error.

## 2 The bivariate Fay-Herriot model

Let $U$ be a finite population partitioned into $D$ domains $U_{1}, \ldots, U_{D}$. Let $\mu_{d}=\left(\mu_{d 1}, \mu_{d 2}\right)^{\prime}$ be a vector of characteristics of interest in the domain $d$ and let $y_{d}=\left(y_{d 1}, y_{d 2}\right)^{\prime}$ be a vector of direct estimates of $\mu_{d}$ calculated by using the data of the target survey sample.

The bivariate Fay-Herriot model is defined in two stages. The first stage indicates that direct estimators $\left\{y_{d}\right\}$ are unbiased and follow the sampling model

$$
\begin{equation*}
y_{d}=\mu_{d}+e_{d}, \quad \forall d \in\{1, \ldots, D\} \tag{1}
\end{equation*}
$$

where the vectors $e_{d}=\left(e_{d 1}, e_{d 2}\right)^{\prime} \sim N_{2}\left(0, V_{e d}\right)$ are independent and the $2 \times 2$ covariance matrices $V_{e d}$ are known. In most cases, $V_{e d}$ is taken to be the design-based covariance matrix of direct estimators $y_{d}, \forall d \in\{1, \ldots, D\}$. The covariance matrices $V_{e d}$ are

$$
V_{e d}=\left(\begin{array}{cc}
\sigma_{e d 1}^{2} & \sigma_{e d 12} \\
\sigma_{e d 12} & \sigma_{e d 2}^{2}
\end{array}\right), \quad \sigma_{e d 12}=\rho_{e d 12} \sigma_{e d 1} \sigma_{e d 2}, \quad \forall d \in\{1, \ldots, D\}
$$

In the second stage the true area characteristic $\mu_{d k}$ is assumed to be linearly related to $p_{k}$ explanatory variables, $k=1,2, d \in\{1, \ldots, D\}$. Let $x_{d k}^{\prime}=\left(x_{d k 1}, \ldots, x_{d k p_{k}}\right)$ be a row vector containing the true aggregated (population) values of $p_{k}$ explanatory variables for $\mu_{d k}$ and let $X_{d}=\operatorname{diag}\left(x_{d 1}^{\prime}, x_{d 2}^{\prime}\right)$ be a $2 \times p$ block-diagonal matrix with $p=p_{1}+p_{2}$. Let $\beta_{k}=\left(\beta_{k 1}, \ldots, \beta_{k p_{k}}\right)^{\prime}$ be a column vector of size $p_{k}$ containing the regression parameters $\beta_{k j}$ for $\mu_{d k}$ and let $\beta=\left(\beta_{1}^{\prime}, \beta_{2}^{\prime}\right)_{p \times 1}^{\prime}$. The linking model is

$$
\begin{equation*}
\mu_{d}=X_{d} \beta+u_{d}, \quad u_{d}=\left(u_{d 1}, u_{d 2}\right)^{\prime} \sim N_{2}\left(0, V_{u d}\right), \quad \forall d \in\{1, \ldots, D\} \tag{2}
\end{equation*}
$$

where the vectors $u_{d}$ 's are independent of the vectors $e_{d}$ 's. The $2 \times 2$ covariance matrix $V_{u d}$ depends on three unknown parameters, $\theta_{1}=\sigma_{u 1}^{2}, \theta_{2}=\sigma_{u 2}^{2}$ and $\theta_{3}=\rho$, i.e.

$$
V_{u d}=\left(\begin{array}{cc}
\sigma_{u 1}^{2} & \rho \sigma_{u 1} \sigma_{u 2} \\
\rho \sigma_{u 1} \sigma_{u 2} & \sigma_{u 2}^{2}
\end{array}\right), \forall d \in\{1, \ldots, D\}
$$

The bivariate Fay-Herriot (BFH) model can be expressed as a single model in the form

$$
\begin{equation*}
y_{d}=X_{d} \beta+u_{d}+e_{d}, \quad \forall d \in\{1, \ldots, D\}, \tag{3}
\end{equation*}
$$

or in the matrix form

$$
y=X \beta+u+e
$$

where

$$
y=\operatorname{col}_{1 \leq d \leq D}\left(y_{d}\right), u=\operatorname{col}_{1 \leq d \leq D}\left(u_{d}\right), e=\operatorname{col}_{1 \leq d \leq D}\left(e_{d}\right), X=\operatorname{col}_{1 \leq d \leq D}\left(X_{d}\right)
$$

We finally assume that $u_{d}, e_{d}, d \in\{1, \ldots, D\}$, are independent. The BFH model (3) is a reparametrization of Model 3 introduced by Benavent and Morales (2016).

Let us define $V_{d}=V_{u d}+V_{e d}, \forall d \in\{1, \ldots, D\}$. Under model (3), it holds that

$$
E(y)=X \beta \quad \text { and } \quad V=\operatorname{var}(y)=Z^{\prime} V_{u} Z+V_{e}=V_{u}+V_{e}=\operatorname{diag}_{1 \leq d \leq D}\left(V_{d}\right)
$$

## 3 Prediction with missing target values

Let us assume that some of the $y_{d k}$ are missing. We define $y_{\bar{d} 1}=\left(y_{d 1}, 0\right)^{\prime}$ and $y_{\bar{d} 2}=\left(0, y_{d 2}\right)^{\prime}$, and partition the domains into three groups:
$\mathbb{D}_{1}=\left\{d \in \mathbb{N}: 1 \leq d \leq D_{1}\right\}$ contains the $D_{1}$ domains where only $y_{d 1}$ is observed.
$\mathbb{D}_{2}=\left\{d \in \mathbb{N}: D_{1}+1 \leq d \leq D_{1}+D_{2}\right\}$ contains the $D_{2}$ domains where only $y_{d 2}$ is observed.
$\mathbb{D}_{3}=\left\{d \in \mathbb{N}: D_{1}+D_{2}+1 \leq d \leq D\right\}$ contains the remaining domains where $y_{d}=\left(y_{d 1}, y_{d 2}\right)^{\prime}$ is fully observed.

If the BFH model (3) holds for $d \in\{1, \ldots, D\}$ and the missing data obey scheme $\{1, \ldots, D\}=$ $\mathbb{D}_{1} \cup \mathbb{D}_{2} \cup \mathbb{D}_{3}$, we say that target vectors $y_{d}$ obey a missing data $\mathrm{BFH}(\mathrm{MBFH})$ model. If the MBFH model holds, then

1. $y_{d 1} \sim N_{1}\left(x_{d 1}^{\prime} \beta_{1}, \sigma_{u 1}^{2}+\sigma_{e d 1}^{2}\right)$ and $\left.y_{d 1}\right|_{u_{d}} \sim N_{1}\left(x_{d 1}^{\prime} \beta_{1}+u_{d 1}, \sigma_{e d 1}^{2}\right)$ if $d \in \mathbb{D}_{1}$,
2. $y_{d 2} \sim N_{1}\left(x_{d 2}^{\prime} \beta, \sigma_{u 2}^{2}+\sigma_{e d 2}^{2}\right)$ and $\left.y_{d 2}\right|_{u_{d}} \sim N_{1}\left(x_{d 2}^{\prime} \beta_{2}+u_{d 2}, \sigma_{e d 2}^{2}\right)$ if $d \in \mathbb{D}_{2}$, and
3. $y_{d} \sim N_{2}\left(X_{d} \beta, V_{u d}+V_{e d}\right)$ and $\left.y_{d}\right|_{u_{d}} \sim N_{2}\left(X_{d} \beta+u_{d}, V_{e d}\right)$ if $d \in \mathbb{D}_{3}$.

In a real situation where the target data follows a MBFH model, the BFH model is strictly applicable to $\mathbb{D}_{3}$, but not to $\mathbb{D}_{1}$ or $\mathbb{D}_{2}$. For example, under the BFH model we can only calculate EBLUPs of $\mu_{d}$ or $u_{d}$ for $d \in \mathbb{D}_{3}$. However, in what follows we show that it is possible calculate EBPs for $d \in \mathbb{D}_{1} \cup \mathbb{D}_{2}$ under the MBFH model.

As the kernel of the $n$-variate normal distribution is

$$
f(y \mid \mu, \Sigma)=\frac{1}{(2 \pi)^{n / 2}|\Sigma|^{1 / 2}} \exp \left[-\frac{1}{2}(y-\mu)^{\prime} \Sigma^{-1}(y-\mu)\right] \propto \exp \left[-\frac{1}{2} y^{\prime} \Sigma^{-1} y+\mu^{\prime} \Sigma^{-1} y\right]
$$

we have the following three propositions.
Proposition 3.1. If $d \in \mathbb{D}_{1}$, then the best predictor ( BP ) of $u_{d}$ under the MBFH model is

$$
\hat{u}_{d}^{b p}=E\left[u_{d} \mid y_{d 1}\right]=\Phi_{d 1}\left(\begin{array}{cc}
\sigma_{e d 1}^{-2} & 0 \\
0 & 0
\end{array}\right)\left(y_{\bar{d} 1}-X_{d} \beta\right), \quad \Phi_{d 1}=\left[\left(\begin{array}{cc}
\sigma_{e d 1}^{-2} & 0 \\
0 & 0
\end{array}\right)+V_{u d}^{-1}\right]^{-1} .
$$

Proof. The conditional distribution of $u_{d}$, given $y_{d 1}$, is

$$
\begin{aligned}
f\left(u_{d} \mid y_{d 1}\right) & \propto f\left(y_{d 1} \mid u_{d}\right) f\left(u_{d}\right) \\
& \propto \exp \left\{-\frac{1}{2} u_{d}^{\prime}\left[\left(\begin{array}{cc}
\sigma_{e d 1}^{-2} & 0 \\
0 & 0
\end{array}\right)+V_{u d}^{-1}\right] u_{d}+u_{d}^{\prime} \Phi_{d 1}^{-1}\left[\Phi_{d 1}\left(\begin{array}{cc}
\sigma_{e d 1}^{-2} & 0 \\
0 & 0
\end{array}\right)\left(y_{\bar{d} 1}-X_{d} \beta\right)\right]\right\} .
\end{aligned}
$$

Therefore, $f\left(u_{d} \mid y_{d_{1}}\right)$ is a bivariate normal distribution with parameters
$\operatorname{var}\left(u_{d} \mid y_{d_{1}}\right)=\left[\left(\begin{array}{cc}\sigma_{e d 1}^{-2} & 0 \\ 0 & 0\end{array}\right)+V_{u d}^{-1}\right]^{-1}=\Phi_{d 1}, \quad E\left[u_{d} \mid y_{d_{1}}\right]=\Phi_{d 1}\left(\begin{array}{cc}\sigma_{e d 1}^{-2} & 0 \\ 0 & 0\end{array}\right)\left(y_{\bar{d} 1}-X_{d} \beta\right)$.

Proposition 3.2. If $d \in \mathbb{D}_{2}$, then the BP of $u_{d}$ under the MBFH model is

$$
\hat{u}_{d}^{b p}=E\left[u_{d} \mid y_{d 2}\right]=\Phi_{d 2}\left(\begin{array}{cc}
0 & 0 \\
0 & \sigma_{e d 2}^{-2}
\end{array}\right)\left(y_{\bar{d} 2}-X_{d} \beta\right), \quad \Phi_{d 2}=\left(\left(\begin{array}{cc}
0 & 0 \\
0 & \sigma_{e d 2}^{-2}
\end{array}\right)+V_{u d}^{-1}\right)^{-1}
$$

Proof. The proof is analogous as the one of Proposition 3.1.
Proposition 3.3. If $d \in \mathbb{D}_{3}$, then BP of $u_{d}$ under the MBFH model is

$$
\hat{u}_{d}^{b p}=E\left[u_{d} \mid y_{d}\right]=\Phi_{d} V_{e d}^{-1}\left(y_{d}-X_{d} \beta\right), \quad \Phi_{d}=\left(V_{e d}^{-1}+V_{u d}^{-1}\right)^{-1}
$$

Proof. The conditional distribution of $u_{d}$ given $y_{d}$, is

$$
\begin{aligned}
f\left(u_{d} \mid y_{d}\right) & \propto f\left(y_{d} \mid u_{d}\right) f\left(u_{d}\right) \\
& \propto \exp \left\{-\frac{1}{2} u_{d}^{\prime}\left(V_{e d}^{-1}+V_{u d}^{-1}\right) u_{d}+u_{d}^{\prime} \Phi_{d}^{-1}\left[\Phi_{d} V_{e d}^{-1}\left(y_{d}-X_{d} \beta\right)\right]\right\} .
\end{aligned}
$$

Therefore, $f\left(u_{d} \mid y_{d}\right)$ is a bivariate normal distribution with parameters

$$
\operatorname{var}\left(u_{d} \mid y_{d}\right)=\left(V_{e d}^{-1}+V_{u d}^{-1}\right)^{-1}=\Phi_{d}, \quad E\left[u_{d} \mid y_{d}\right]=\Phi_{d} V_{e d}^{-1}\left(y_{d}-X_{d} \beta\right) .
$$

Corollary 3.3. The BP of $\mu_{d}, d=1, \ldots, D$, under the MBFH model is

$$
\begin{equation*}
\hat{\mu}_{d}^{b p}=X_{d} \beta+\hat{u}_{d}^{b p} . \tag{4}
\end{equation*}
$$

Definition 3.1. The EBP of $\mu_{d}, d=1, \ldots, D$, under the MBFH model (MBFH-EBP) is obtained from formula (4) by plugging estimators $\hat{\boldsymbol{\beta}}, \hat{\sigma}_{u 1}^{2}, \hat{\sigma}_{u 2}^{2}$ and $\hat{\rho}$ in the places of $\boldsymbol{\beta}, \sigma_{u 1}^{2}$, $\sigma_{u 2}^{2}$ and $\rho$ respectively, i.e.

$$
\begin{equation*}
\hat{\mu}_{d}^{e b p}=X_{d} \hat{\beta}+\hat{u}_{d}^{e b p} \tag{5}
\end{equation*}
$$

## 4 Estimation of model parameters

This section presents the maximum likelihood (ML) and the residual maximum likelihood (REML) methods for estimating the model parameters.

### 4.1 Maximum likelihood method

The vector of model parameters is $\psi=\left(\beta^{\prime}, \theta^{\prime}\right)^{\prime}$, where $\theta=\left(\theta_{1}, \theta_{2}, \theta_{3}\right)^{\prime}=\left(\sigma_{u 1}^{2}, \sigma_{u 2}^{2}, \rho\right)^{\prime}$. The $\log$-likelihood is $l=\sum_{d=1}^{D} l_{d}$, where

$$
\begin{array}{ll}
l_{d}=-\frac{1}{2} \log 2 \pi-\frac{1}{2} \log \left(\sigma_{u 1}^{2}+\sigma_{e d 1}^{2}\right)-\frac{1}{2\left(\sigma_{u 1}^{2}+\sigma_{e d 1}^{2}\right)}\left(y_{d 1}-x_{d 1}^{\prime} \beta_{1}\right)^{2}, \quad \text { if } d \in \mathbb{D}_{1}, \\
l_{d}=-\frac{1}{2} \log 2 \pi-\frac{1}{2} \log \left(\sigma_{u 2}^{2}+\sigma_{e d 2}^{2}\right)-\frac{1}{2\left(\sigma_{u 2}^{2}+\sigma_{e d 2}^{2}\right)}\left(y_{d 2}-x_{d 2}^{\prime} \beta_{2}\right)^{2}, \quad \text { if } d \in \mathbb{D}_{2}, \\
l_{d}=-\log 2 \pi-\frac{1}{2} \log \left|V_{d}\right|-\frac{1}{2}\left(y_{d}-X_{d} \beta\right)^{\prime} V_{d}^{-1}\left(y_{d}-X_{d} \beta\right), & \text { if } d \in \mathbb{D}_{3} .
\end{array}
$$

The ML Fisher-scoring algorithm, with Fisher-information matrices $F_{\beta \beta}\left(\theta^{(r)}\right), F_{\theta \theta}\left(\theta^{(r)}\right)$ and score vectors $U_{\beta}\left(\psi^{(r)}\right), U_{\theta}\left(\psi^{(r)}\right)$, is

1. Set the initial values $\psi^{(0)}=\left(\beta_{1}^{(0)}, \beta_{2}^{(0)}, \theta^{(0)}\right)$ and $\varepsilon>0$.
2. Repeat the following steps until the tolerance or the boundary conditions are fulfilled.
(a) Updating equations: Do

$$
\beta^{(r+1)}=\beta^{(r)}+F_{\beta \beta}^{-1}\left(\theta^{(r)}\right) U_{\beta}\left(\psi^{(r)}\right), \quad \theta^{(r+1)}=\theta^{(r)}+F_{\theta \theta}^{-1}\left(\theta^{(r)}\right) U_{\theta}\left(\psi^{(r)}\right) .
$$

(b) Boundary condition: If $\theta_{1}^{(r+1)}>0, \theta_{2}^{(r+1)}>0$ and $\left|\theta_{3}^{(r+1)}\right|<1$, continue. Otherwise, do $\hat{\psi}=\psi^{(r)}$ and stop.
(c) Tolerance condition: If $\left|\psi_{j}^{(r+1)}-\psi_{j}^{(r)}\right|<\varepsilon, j=1, \ldots, p+3$, do $\hat{\psi}=\psi^{(r+1)}$ and stop. Otherwise, continue.

## 3. Output: $\hat{\psi}$.

Remark. As starting values, we take $\beta_{1}^{(0)}=\hat{\beta}_{1}^{(0)}, \beta_{2}^{(0)}=\hat{\beta}_{2}^{(0)}, \hat{\theta}_{3,0}=0, \hat{\theta}_{k, 0}=\hat{\sigma}_{u k, 0}^{2}, k=1,2$, where $\hat{\beta}_{1}^{(0)}, \hat{\beta}_{2}^{(0)}$ and $\hat{\sigma}_{u k, 0}^{2}$ are the REML or the ML estimators of the corresponding bivariate Fay-Herriot that uses only the data of group $\mathbb{D}_{3}$.

### 4.2 Residual maximum likelihood method

In the case that there are no missing values, the residual maximum likelihood (REML) log-likelihood is

$$
\begin{equation*}
l_{\text {reml }}(\theta)=-\frac{2 D-p}{2} \log 2 \pi+\frac{1}{2} \log \left|X^{\prime} X\right|-\frac{1}{2} \log |V|-\frac{1}{2} \log \left|X^{\prime} V^{-1} X\right|-\frac{1}{2} y^{\prime} P y \tag{6}
\end{equation*}
$$

where $\theta=\left(\theta_{1}, \theta_{2}, \theta_{3}\right), \theta_{1}=\sigma_{u 1}^{2}, \theta_{2}=\sigma_{u 2}^{2}, \theta_{3}=\rho, P=V^{-1}-V^{-1} X\left(X^{\prime} V^{-1} X\right)^{-1} X^{\prime} V^{-1}$, $P V P=P$ with $P X=0, y$ is the $2 D$-vector $y=\left(y_{1}^{\prime}, \ldots, y_{D}^{\prime}\right)^{\prime}$, and $X$ the corresponding $2 D \times p$-matrix $X=\left(X_{1}, \ldots, X_{D}\right)^{\prime} . P$ is defined as before on the reduced $X$ and $V$.

In the case that there are missing values, the residual maximum likelihood (REML) log-likelihood is

$$
\begin{array}{r}
l_{\text {reml }}(\theta)=-\frac{D_{1}+D_{2}+2\left(D-D_{1}-D_{2}\right)-p}{2} \log 2 \pi+\frac{1}{2} \log \left|X^{\prime} X\right|  \tag{7}\\
-\frac{1}{2} \log |V|-\frac{1}{2} \log \left|X^{\prime} V^{-1} X\right|-\frac{1}{2} y^{\prime} P y
\end{array}
$$

where now, the vector $y$ and the matrices $X$ and $V$ are reduced to those rows and columns that correspond to an observed value of $y_{d k}, \forall d \in \mathbb{D}$ and $k=1,2$. Let $\tilde{D}:=D_{1}+D_{2}+$ $2\left(D-D_{1}-D_{2}\right)$, then the length of $y$ is $\tilde{D}$ and the dimensions of $X$ and $V$ are $\tilde{D} \times p$ and $\tilde{D} \times \tilde{D}$ respectively.

The REML Fisher-scoring algorithm, with Fisher-information matrix $F\left(\theta^{(k)}\right)$ and score vector $S\left(\theta^{(k)}\right)$, is

1. Set the initial values $\beta^{(0)}, \theta^{(0)}$, and $\varepsilon_{k}>0, r=1, \ldots, p+3$.
2. Repeat the following steps until the tolerance or the boundary conditions are fulfilled.
(a) Updating equation for $\theta$ : Do $\theta^{(k+1)}=\theta^{(k)}+F^{-1}\left(\theta^{(k)}\right) S\left(\theta^{(k)}\right)$.
(b) Boundary condition: If $\theta_{1}^{(k+1)}>0, \theta_{2}^{(k+1)}>0$ and $\left|\theta_{3}^{(k+1)}\right|<1$, continue. Otherwise, do $\hat{\theta}=\theta^{(k)}$ and stop.
(c) Updating equation for $\beta$ : Do $\beta^{(k+1)}=\left(X^{\prime} V^{-1}\left(\theta^{(k+1)}\right) X\right)^{-1} X^{\prime} V^{-1}\left(\theta^{(k+1)}\right) y$.
(d) Tolerance condition: If $\left|\theta_{\ell}^{(k+1)}-\theta_{\ell}^{(k)}\right|<\varepsilon_{p+\ell},\left|\beta_{k}^{(k+1)}-\beta_{k}^{(k)}\right|<\varepsilon_{k}, r=1, \ldots, p$, $\ell=1,2,3$, do $\hat{\theta}_{\ell}=\theta_{\ell}^{(k+1)}, \hat{\beta}=\beta^{(k+1)}$ and stop. Otherwise, continue.
3. Output: $\hat{\theta}, \hat{\beta}, F^{-1}(\hat{\theta}),\left(X^{\prime} V^{-1}(\hat{\theta}) X\right)^{-1}$.

The asymptotic distributions of the REML estimators $\hat{\theta}$ and $\hat{\beta}$,

$$
\hat{\theta} \sim N_{3}\left(\theta, F^{-1}(\theta)\right), \quad \hat{\beta} \sim N_{k}\left(\beta,\left(X^{\prime} V^{-1}(\theta) X\right)^{-1}\right)
$$

can be used to construct $(1-\alpha)$-level confidence intervals for the components $\theta_{\ell}$ of $\theta$ and $\beta_{i}$ of $\beta$. Let $\beta_{i}$ denote the $i$-th component of the vector $\beta$, not the vector of regression parameters of the $i$-th component. The confidence intervals are given by

$$
\begin{equation*}
\hat{\theta}_{\ell} \pm z_{\alpha / 2} \nu_{\ell \ell}^{1 / 2}, \quad \ell=1,2,3, \quad \hat{\beta}_{i} \pm z_{\alpha / 2} q_{i i}^{1 / 2}, i=1, \ldots, r \tag{8}
\end{equation*}
$$

where $F^{-1}(\hat{\theta})=\left(\nu_{a b}\right)_{a, b=1,2,3},\left(X^{\prime} V^{-1}(\hat{\theta}) X\right)^{-1}=\left(q_{i j}\right)_{i, j=1, \ldots, k}$ and $z_{\alpha}$ is the $\alpha$-quantile of the $N(0,1)$ distribution. For $\hat{\beta}_{i}=\beta_{0}$, the $p$-value for testing the hypothesis $H_{0}: \beta_{i}=0$ is

$$
\begin{equation*}
p \text {-value }=2 P_{H_{0}}\left(\hat{\beta}_{i}>\left|\beta_{0}\right|\right)=2 P\left(N(0,1)>\left|\beta_{0}\right| / \sqrt{q_{i i}}\right) . \tag{9}
\end{equation*}
$$

## 5 Analytic approximation of the mean squared errors

### 5.1 Best predictors

Let us first consider the group $\mathbb{D}_{1}$. We use the notation

$$
\Phi_{d 1}=\left(A_{d 1}+V_{u d}^{-1}\right)^{-1} \triangleq\left(\begin{array}{ll}
\phi_{d 1,11} & \phi_{d 1,12} \\
\phi_{d 1,12} & \phi_{d 1,22}
\end{array}\right), \quad A_{d 1}=\left(\begin{array}{cc}
\sigma_{e d 1}^{-2} & 0 \\
0 & 0
\end{array}\right) .
$$

As $\hat{\mu}_{d}^{b p}-\mu_{d}=\hat{u}_{d}^{b p}-u_{d}$, the MSE matrix of $\hat{\mu}_{d}^{b p}$ is

$$
\begin{aligned}
\operatorname{MSE}\left(\hat{\mu}_{d}^{b p}\right) & =\operatorname{MSE}\left(\hat{u}_{d}^{b p}\right)=E\left[\left(\hat{u}_{d}^{b p}-u_{d}\right)\left(\hat{u}_{d}^{b p}-u_{d}\right)^{\prime}\right] \\
& =\Phi_{d 1} A_{d 1}\left(V_{u d}+V_{e d}\right) A_{d 1} \Phi_{d 1}+V_{u d}-2 \Phi_{d 1} A_{d 1} V_{u d} .
\end{aligned}
$$

For the group $\mathbb{D}_{2}$, we have similar mathematical derivations. For the sake of brevity, we omit them.

Let us consider the group $\mathbb{D}_{3}$. As $\hat{\mu}_{d}^{b p}-\mu_{d}=\hat{u}_{d}^{b p}-u_{d}$, the MSE matrix of $\hat{\mu}_{d}^{b p}$ is

$$
\begin{aligned}
\operatorname{MSE}\left(\hat{\mu}_{d}^{b p}\right) & =\operatorname{MSE}\left(\hat{u}_{d}^{b p}\right)=E\left[\left(\hat{u}_{d}^{b p}-u_{d}\right)\left(\hat{u}_{d}^{b p}-u_{d}\right)^{\prime}\right] \\
& =\Phi_{d} V_{e d}^{-1}\left(V_{u d}+V_{e d}\right) V_{e d}^{-1} \Phi_{d}+V_{u d}-2 \Phi_{d} V_{e d}^{-1} V_{u d} .
\end{aligned}
$$

### 5.2 Empirical best predictors

We sketch the derivation for the empirical best predictors for group $\mathbb{D}_{1}$. The mathematical derivations for group $\mathbb{D}_{2}$ and $\mathbb{D}_{2}$ are mostly analogous.

For $d \in \mathbb{D}_{1}$, we have the following approximation to $\operatorname{MSE}\left(\hat{\mu}_{d}^{e b p}\right)$.

$$
\operatorname{MSE}\left(\hat{\mu}_{d}^{e b p}\right)=G_{d 1}(\theta)+G_{d 2}(\theta)+G_{d 3}(\theta)+O_{2 \times 2}\left(D^{-1}\right),
$$

where $G_{d 2}(\theta)=G_{d 2,11}(\theta)+G_{d 2,22}(\theta)+G_{d 2,12}(\theta)+G_{d 2,12}^{\prime}(\theta)$ and

$$
\begin{aligned}
G_{d 1}(\theta) & =\Phi_{d 1}(\theta) A_{d 1}(\theta)\left(V_{u d}(\theta)+V_{e d}(\theta)\right) A_{d 1}(\theta) \Phi_{d 1}(\theta)+V_{u d}(\theta)-2 \Phi_{d 1}(\theta) A_{d 1}(\theta) V_{u d}(\theta), \\
G_{d 2, a b}(\theta) & =\left(\begin{array}{ll}
\operatorname{tr}\left\{H_{d \beta_{b} \beta_{a}, 11}(\theta) \operatorname{cov}\left(\hat{\beta}_{a}, \hat{\beta}_{b}\right)\right\} & \operatorname{tr}\left\{H_{d \beta_{b} \beta_{a}, 21}(\theta) \operatorname{cov}\left(\hat{\beta}_{a}, \hat{\beta}_{b}\right)\right\} \\
\operatorname{tr}\left\{H_{d \beta_{b} \beta_{a}, 12}(\theta) \operatorname{cov}\left(\hat{\beta}_{a}, \hat{\beta}_{b}\right)\right\} & \operatorname{tr}\left\{H_{d \beta_{b} \beta_{a}, 22}(\theta) \operatorname{cov}\left(\hat{\beta}_{a}, \hat{\beta}_{b}\right)\right\}
\end{array}\right), a, b=1,2, \\
G_{d 3}(\theta) & =\frac{\sigma_{u d 1}^{2}+\sigma_{e d 1}^{2}}{\sigma_{e d 1}^{4}}\left(\begin{array}{ll}
\operatorname{tr}\left\{G_{d \theta \theta, 11}(\theta) \operatorname{var}(\hat{\theta})\right\} & \operatorname{tr}\left\{G_{d \theta \theta, 21}(\theta) \operatorname{var}(\hat{\theta})\right\} \\
\operatorname{tr}\left\{G_{d \theta \theta, 12}(\theta) \operatorname{var}(\hat{\theta})\right\} & \operatorname{tr}\left\{G_{d \theta \theta, 22}(\theta) \operatorname{var}(\hat{\theta})\right\}
\end{array}\right) .
\end{aligned}
$$

The remaining vectors and matrices are derived as follows. The EBP is a function of the estimators $(\hat{\beta}, \hat{\theta})$ and of the target variable $y_{d 1}$. For the sake of brevity, we write

$$
h_{d}(\hat{\beta}, \hat{\theta}) \triangleq \hat{\mu}_{d}^{e b p}=X_{d} \hat{\beta}+\hat{\Phi}_{d 1} A_{d 1}\left(y_{\bar{d} 1}-X_{d} \hat{\beta}\right)
$$

where $\quad \hat{\Phi}_{d 1}=\Phi_{d 1}(\hat{\theta})=\left(A_{d 1}+\hat{V}_{u d}^{-1}\right)^{-1}, \quad \hat{V}_{u d}=V_{u d}(\hat{\theta})=\left(\begin{array}{cc}\hat{\sigma}_{u 1}^{2} & \hat{\rho} \hat{\sigma}_{u 1} \hat{\sigma}_{u 2} \\ \hat{\rho} \hat{\sigma}_{u 1} \hat{\sigma}_{u 2} & \hat{\sigma}_{u 2}^{2}\end{array}\right)$.
The derivatives of matrix $\Phi_{d 1}(\theta)$ with respect to $\theta_{\ell}, \ell=1,2,3$, are

$$
\frac{\partial \Phi_{d 1}}{\partial \theta_{\ell}}=\left(A_{d 1}+V_{u d}^{-1}\right)^{-1} V_{u d}^{-1} V_{u d \ell} V_{u d}^{-1}\left(A_{d 1}+V_{u d}^{-1}\right)^{-1}=\left(\begin{array}{ll}
\phi_{d 1 \ell, 11} & \phi_{d 1 \ell, 12} \\
\phi_{d 1 \ell, 12} & \phi_{d 1 \ell, 22}
\end{array}\right)
$$

The derivatives of $h_{d}(\beta, \theta)$ with respect to $\beta_{k j}$ and $\theta_{\ell}, k=1,2, j=1, \ldots, p_{k}, \ell=1,2,3$, are

$$
\begin{aligned}
\frac{\partial h_{d}}{\partial \beta_{1 j}} & =\binom{x_{d 1 j}}{0}-\frac{x_{d 1 j}}{\sigma_{e d 1}^{2}}\binom{\phi_{d 1,11}}{\phi_{d 1,12}} \triangleq\binom{h_{d \beta_{1 j}, 1}}{h_{d \beta_{1 j}, 2}}, \quad \frac{\partial h_{d}}{\partial \beta_{2 j}}=\binom{0}{x_{d 2 j}} \triangleq\binom{h_{d \beta_{2 j}, 1}}{h_{d \beta_{2 j}, 2}} \\
\frac{\partial h_{d}}{\partial \theta_{\ell}} & =\frac{\partial \Phi_{d 1}}{\partial \theta_{\ell}} A_{d 1}\left(y_{\overline{d 1}}-X_{d} \beta\right)=\frac{y_{d 1}-x_{d 1}^{\prime} \beta_{1}}{\sigma_{e d 1}^{2}}\binom{\phi_{d 1 \ell, 11}}{\phi_{d 1 \ell, 12}} \triangleq \frac{y_{d 1}-x_{d 1}^{\prime} \beta_{1}}{\sigma_{e d 1}^{2}}\binom{g_{d \theta_{\ell}, 1}}{g_{d \theta_{\ell, 2}}}
\end{aligned}
$$

The $3 \times 1$ vectors containing the derivatives with respect to $\theta$ are $g_{d \theta, 1}=\operatorname{col}_{1 \leq \ell \leq 3}\left(g_{d \theta_{\ell, 1}}\right), g_{d \theta, 2}=$ $\underset{1 \leq \ell \leq 3}{\operatorname{col}}\left(g_{d \theta_{\ell}, 2}\right)$ and the corresponding $3 \times 3$ matrices are $G_{d \theta \theta, a b}=g_{d \theta, a} g_{d \theta, b}^{\prime}, a, b=1,2$.

The $p_{k} \times 1$ vectors containing the derivatives with respect to $\beta_{k}, k=1,2$, are

$$
h_{d \beta_{k}, 1}=\operatorname{col}_{1 \leq j \leq p_{k}}\left(h_{d \beta_{k j}, 1}\right), \quad h_{d \beta_{k}, 2}=\operatorname{col}_{1 \leq j \leq p_{k}}\left(h_{d \beta_{k j}, 2}\right),
$$

and the corresponding $p_{k_{1}} \times p_{k_{2}}$ matrices are $H_{d \beta_{k_{1}} \beta_{k_{2}}, a b}=h_{d \beta_{k_{1}}, a} h_{d \beta_{k_{2}}, b}^{\prime}, k_{1}, k_{2}, a, b=1,2$. An estimator of $\operatorname{MSE}\left(\hat{\mu}_{d}^{e b p}\right)$ is

$$
m s e\left(\hat{\mu}_{d}^{e b p}\right)=G_{d 1}(\hat{\theta})+G_{d 2}(\hat{\theta})+2 G_{d 3}(\hat{\theta}) .
$$

## 6 Bootstrap approximations of the mean squared errors

This section introduces a parametric bootstrap procedure for approximating $\operatorname{MSE}\left(\hat{\mu}_{d}^{e b p}\right)$. Steps B1-B5 below describe the basic procedure for computing an approximation of $M S E\left(\hat{\mu}_{d}^{e b p}\right)$ called direct parametric bootstrap estimator.

## Parametric bootstrap procedure:

B1. Calculate the REML (or ML) estimates $\hat{\theta}$ and $\hat{\beta}$ of $\theta$ and $\beta$ respectively by using the observable data, i.e. $\left(y_{d 1}, X_{d}\right) \forall d \in \mathbb{D}_{1},\left(y_{d 2}, X_{d}\right) \forall d \in \mathbb{D}_{2}$, and $\left(y_{d}, X_{d}\right) \forall d \in \mathbb{D}_{3}$.

B2. $\forall d \in\{1, \ldots, D\}$, generate independent and identically distributed vectors $u_{d}^{*} \sim$ $N_{2}\left(0, V_{u d}(\hat{\theta})\right)$.

B3. $\forall d \in\{1, \ldots, D\}$, generate independent vectors $e_{d}^{*} \sim N_{2}\left(0, V_{e d}\right)$.
B4. Construct the bootstrap model

$$
y_{d}^{*}=X_{d} \hat{\beta}+u_{d}^{*}+e_{d}^{*}, \quad \forall d \in\{1, \ldots, D\} .
$$

For step B5 we introduce further notation. Let $E_{*}$ and $M S E_{*}$ denote the expectation and MSE under the probability distribution induced by bootstrap model B4, given the initial target vector $y$. The bootstrap mean vectors are

$$
\mu_{d}^{*}=X_{d} \hat{\beta}+u_{d}^{*}, \quad \forall d \in\{1, \ldots, D\} .
$$

Let $\hat{\beta}^{*}$ and $\hat{\theta}^{*}$ be the REML (or ML) estimators of the parameters $\hat{\beta}$ and $\hat{\theta}$ of bootstrap model B4. These estimators are calculated by using only the observable bootstrap data $\left(y_{d 1}^{*}, X_{d}\right)$ if $d \in \mathbb{D}_{1},\left(y_{d 2}^{*}, X_{d}\right)$ if $d \in \mathbb{D}_{2}$, and $\left(y_{d}^{*}, X_{d}\right)$ if $d \in \mathbb{D}_{3}$.

Let $\hat{\mu}_{d}^{* b p}$ and $\hat{\mu}_{d}^{* e b p}$ be the BP and EBP of $\mu_{d}^{*}$ under model B4 $\forall d \in\{1, \ldots, D\}$. In the same way, the bootstrap MSE of $\hat{\mu}_{d}^{* e b p}$ is

$$
\operatorname{MSE}_{*}^{1}\left(\hat{\mu}_{d}^{* e b p}\right)=E_{*}\left[\left(\hat{\mu}_{d}^{* e b p}-\mu_{d}^{*}\right)\left(\hat{\mu}_{d}^{* e b p}-\mu_{d}^{*}\right)^{\prime}\right], \quad \forall d \in\{1, \ldots, D\}
$$

These $2 \times 2$ matrices are called parametric bootstrap estimators. In practice, these estimators can be approximated via Monte Carlo as described in B5.

B5. Generate $B$ bootstrap vectors $y^{*(b)}=\left(y_{d}^{*(b)}: d \in\{1, \ldots, D\}\right), b=1, \ldots, B$, from model B4. From each vector $y^{*(b)}$, calculate the true means $\mu_{d}^{*(b)}$ and their EBPs $\hat{\mu}_{d}^{* e b p(b)}$ by using only the observable bootstrap data. Then compute the direct bootstrap estimators

$$
\begin{equation*}
\operatorname{mse}^{1}\left(\hat{\mu}_{d}^{e b p}\right)=B^{-1} \sum_{b=1}^{B}\left(\hat{\mu}_{d}^{* e b p(b)}-\mu_{d}^{*(b)}\right)\left(\hat{\mu}_{d}^{* e b p(b)}-\mu_{d}^{*(b)}\right)^{\prime}, \quad \forall d \in\{1, \ldots, D\} \tag{10}
\end{equation*}
$$

Observe that $\operatorname{mse}^{1}\left(\hat{\mu}_{d}^{* e b p}\right)$ is consistent for $\operatorname{MSE}_{*}^{1}\left(\hat{\mu}_{d}^{* e b p}\right)$ as $B \rightarrow \infty$.
We can also apply the bootstrap technique to approximate the terms $G_{2}$ and $G_{3}$ of $\operatorname{MSE}\left(\hat{\mu}_{d}^{e b p}\right)$. Following this idea, we define the term-to-term bootstrap estimator as

$$
\operatorname{MSE}_{*}^{2}\left(\hat{\mu}_{d}^{* e b p}\right)=G_{d 1}(\hat{\theta})+E_{*}\left[\left(\hat{\mu}_{d}^{* e b p}-\hat{\mu}_{d}^{* b p}\right)\left(\hat{\mu}_{d}^{* e b p}-\hat{\mu}_{d}^{* b p}\right)^{\prime}\right]
$$

As the plug-in estimator $G_{d 1}(\hat{\theta})$ of $G_{d 1}(\theta)$ is biased, we introduce the bias-corrected bootstrap estimator

$$
\operatorname{MSE}_{*}^{3}\left(\hat{\mu}_{d}^{* e b p}\right)=2 G_{d 1}(\hat{\theta})-E_{*}\left[G_{d 1}\left(\hat{\theta}^{*}\right)\right]+E_{*}\left[\left(\hat{\mu}_{d}^{* e b p}-\hat{\mu}_{d}^{* b p}\right)\left(\hat{\mu}_{d}^{* e b p}-\hat{\mu}_{d}^{* b p}\right)^{\prime}\right] .
$$

A Monte Carlo approximation $\operatorname{mse}^{a}\left(\hat{\mu}_{d}^{* e b p}\right)$ of the bootstrap matrix $\operatorname{MSE}_{*}^{a}\left(\hat{\mu}_{d}^{* e b p}\right)$, for $a=$ 2,3 , can be obtained similarly as (10).

## References

Arima, S., Bell, W. R., Datta, G. S., Franco, C., and Liseo, B. (2017). Multivariate FayHerriot Bayesian estimation of small area means under functional measurement error. Journal of the Royal Statistical Society: Series A (Statistics in Society), 180(4):11911209.

Articus, C. and Burgard, J. P. (2014). A finite mixture Fay Herriot-type model for estimating regional rental prices in Germany. Technical report, University of Trier, Department of Economics. Research Papers in Economics. No. 14/14.

Benavent, R. and Morales, D. (2016). Multivariate Fay-Herriot models for small area estimation. Computational Statistics $8 \mathcal{B}$ Data Analysis, 94:372-390.

Burgard, J. P., Esteban, M. D., Morales, D., and Pérez, A. (2019a). A Fay-Herriot model when auxiliary variables are measured with error. TEST. https://doi.org/10.1007/s11749-019-00649-3.

Burgard, J. P., Krause, J., Kreber, D., et al. (2019b). Regularized area-level modelling for robust small area estimation in the presence of unknown covariate measurement errors. Technical report, University of Trier, Department of Economics. Research Papers in Economics. No. 4/19.

Datta, G. S., Ghosh, M., Nangia, N., and Natarajan, K. (1996). Estimation of median income of four-person families: a Bayesian approach. In Berry, D. A., Chaloner, K. M., and Geweke, J. K., editors, Bayesian Analysis in Statistics and Econometrics, chapter 11, pages 129-140. Wiley, New York.

Datta, G. S. and Lahiri, P. (2000). A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems. Statistica Sinica, 110:613-627.

Esteban, M. D., Morales, D., Pérez, A., and Santamaría, L. (2012). Small area estimation of poverty proportions under area-level time models. Computational Statistics $8 \mathcal{B}$ Data Analysis, 56(10):2840-2855.

Fay, R. E. and Herriot, R. A. (1979). Estimates of income for small places: an application of James-Stein procedures to census data. Journal of the American Statistical Association, 74(366):269-277.

González-Manteiga, W., Lombarda, M. J., Molina, I., Morales, D., and Santamaría, L. (2010). Small area estimation under Fay-Herriot models with non-parametric estimation of heteroscedasticity. Statistical Modelling, 10(2):215-239.

González-Manteiga, W., Lombardía, M. J., Molina, I., Morales, D., and Santamaría, L. (2008). Analytic and bootstrap approximations of prediction errors under a multivariate Fay-Herriot model. Computational Statistics \& Data Analysis, 52(12):5242-5252.

Li, H. and Lahiri, P. (2010). An adjusted maximum likelihood method for solving small area estimation problems. Journal of Multivariate Analysis, 101(4):882-892.

Marhuenda, Y., Molina, I., and Morales, D. (2013). Small area estimation with spatiotemporal Fay-Herriot models. Computational Statistics \& Data Analysis, 58:308-325.

Morales, D., Pagliarella, M. C., and Salvatore, R. (2015). Small area estimation of poverty indicators under partitioned area-level time models. SORT: Statistics and Operations Research Transactions, 39(1):19-34.

Porter, A. T., Wikle, C. K., and Holan, S. H. (2015). Small area estimation via multivariate Fay-Herriot models with latent spatial dependence. Australian $\mathcal{E}$ New Zealand Journal of Statistics, 57(1):15-29.

Prasad, N. G. N. and Rao, J. N. K. (1990). The estimation of the mean squared error of small-area estimators. Journal of the American statistical association, 85(409):163-171.

Pratesi, M. and Salvati, N. (2008). Small area estimation: the EBLUP estimator based on spatially correlated random area effects. Statistical methods and applications, 17(1):113141.

Ybarra, L. M. R. and Lohr, S. L. (2008). Small area estimation when auxiliary information is measured with error. Biometrika, 95(4):919-931.

Yoshimori, M. and Lahiri, P. (2014). A new adjusted maximum likelihood method for the Fay-Herriot small area model. Journal of Multivariate Analysis, 124:281-294.

Zayatz, L. (2007). Disclosure avoidance practices and research at the US Census Bureau: An update. Journal of Official Statistics, 23(2):253-265.


[^0]:    *The authors gratefully acknowledge the Spanish grant PGC2018-096840-B-I00 and by the grant "Algorithmic Optimization (ALOP) - graduate school 2126" funded by the German Research Foundation

