Small Area Estimation Diagnostics: The Case of the Fay–Herriot Model

Maria Chiara Pagliarella

Abstract Leverage and Cook's distance are some of the most important tools in influence analysis, where the main target is to identify observations that might determine the character of model estimates and predictors. In the small area estimation setup, applied statisticians are interested in tools to identify observations that might influence the variance component and the regression parameter estimates, the empirical best linear unbiased predictor and its mean squared error estimate. For this reason, this paper discusses the leverage matrix, the influence on the mean squared error of the empirical predictor, and a Cook's Distance of the empirical predictor for the Fay–Herriot model, when the area-random effect variance is estimated by the restricted maximum likelihood method. Further, the validity of this approach is illustrated by means of an application to poverty data.

Keywords Influence analysis · Leverage · Cook's distance · Poverty

1 Introduction

In the model-based approach to small area estimation, data is assumed to be generated according to a specific model and the whole inferential process depends on this assumption. Therefore, it is quite important to check if some data points or groups of cases are particularly influential on the analysis. For this reason, diagnostics tools are needed to ensure that model parameters are properly estimated.

In classical linear models, this examination has been traditionally carried out by residual analysis and detection of influential cases. Many articles and books deal with influential observations and outliers. Some of them are [3, 5, 21], and important papers have been written by Chatterjee and Hadi [8], and Cook [6, 7]. Two main types of influence analysis for linear models have been developed. Within the first, the calculation of leverage and standardized residuals plays a key role (the leverage

M. C. Pagliarella (✉)
Istituto Nazionale per l'Analisi delle Politiche Pubbliche (INAPP), Rome, Italy
e-mail: mc.pagliarella@inapp.org

S. Balzano et al. (eds.), *Statistical Learning and Modeling in Data Analysis*,
Studies in Classification, Data Analysis, and Knowledge Organization,
https://doi.org/10.1007/978-3-030-69944-4_16

is the diagonal element of the hat matrix). The second one is based on measuring the effect on the estimates of deleting observations from the whole dataset, and it is called case deletion diagnostics. A third approach, less considered in applications, is based on the maximum curvature of log-likelihood displacement and it is called local influence (see [2, 7]).

In the context of mixed models, many contributions are available as well. Without this list being exhaustive, the following may be mentioned: Lesaffre and Verbeke [17] applied the local influence approach to linear mixed-effects models; Fung et al. [14] considered both case and subject deletion influence diagnostics for semi-parametric mixed models; Demidenko and Stukel [11] generalized common measures of influence for the fixed effects parameters of the linear mixed-effects models; Zewotir and Galpin [29] extended the ordinary linear regression influence diagnostics approach to linear mixed models; Nobre and Singer [22] covered a decomposition of the generalized leverage matrix for the linear mixed models; Pan et al. [23] proposed a case deletion approach to identify influential subjects and influential observations in linear mixed models.

A specific application of mixed models is small area estimation. Small area estimation refers to estimates over domains for which direct estimates are produced with unacceptably large standard errors due to the sample sizes available. Standard survey designs are typically carried out in order to achieve reliable estimates on planned domains (subpopulations) of the reference population. Direct estimates are those based only on the domain-specific sampling data. On the other hand, small area estimation produces indirect estimates for topic of interest on unplanned domains with too small or even zero sample sizes. Indirect estimators based on explicit linking models are called model-based estimators. They "borrow strength" by using values of the variables of interest from related small areas through supplementary information (auxiliary variables), such as data from other related areas or covariates from other sources.

Within this setting, case diagnostics requires special attention. Therefore, diagnostics for mixed models are an incomplete answer to diagnostics in small area estimation, because of the different population parameters of interest. This motivates our interest in diagnostics methods for area level linear mixed models appearing in small area estimation problems. In other words, the goal of small area estimation methods is to determine Empirical Best Linear Unbiased Predictor (EBLUP) for the mean or the total of the variable of interest and to minimize the Mean Squared Error (MSE) of the empirical predictor. Furthermore, case deletion diagnostics cannot be applied whenever there are few units for certain domain of interest.

However, while Battese, Harter, and Fuller [1] applied diagnostics methods for validating the small area estimation model, checking the normality of the error terms and the transformed residuals of the EBLUP, we found only a short note [20] on specific diagnostic measures for the Fay–Herriot model.

For these reasons, this paper has two main aims. On the one hand, it revises and makes it available to a larger audience the results in [20]. On the other, it shows the potential of such an approach by presenting an application of case diagnostics for the

Fay–Herriot model when the goal was to estimate poverty levels across small areas in Spain.

Fay–Herriot model is an area level linear mixed model, with random-area effects. It was first proposed by Fay and Herriot in 1979 [12] to estimate average per capita income in small places of the United States. Since then, the Fay–Herriot model has been widely used because of its flexibility in combining different sources of information with different error structures. It has been largely studied in small area estimation (e.g. [4, 9, 15, 18, 24]), and used to study poverty ([19, 25]) and other related socio-demographic variables [16].

The rest of the paper is organized as follows. Section 2 recalls the fundamentals of the area level Fay–Herriot model when we deal with Restricted Maximum Likelihood (REML) of the random-area effect variance estimator. Section 3 presents diagnostics for the Fay–Herriot model. More specifically Sect. 3.1 gives the leverage matrix on the fixed effects and the leverage matrix on the random-area effects; Sect. 3.2 shows the influence analysis on the first two terms of the estimated mean squared error of the EBLUP; Sect. 3.3 considers some case deletion diagnostics. Section 4 provides the application where diagnostics tools are tested on a model aiming at estimating small area poverty proportions in Spain, while Sect. 5 draws the conclusions. Lastly, an Appendix is provided with detailed formulas.

2 The Fay–Herriot Model

The Fay–Herriot model is a special case of a linear mixed model. We have

$$\widehat{y_i} = \mathbf{x}_i'\beta + b_i v_i + e_i, \quad v_i \overset{iid}{\sim} (0, \sigma_v^2), \quad e_i \overset{ind}{\sim} (0, \psi_i), \quad i = 1, ..., m \tag{1}$$

where the $\widehat{y_i}$'s are the direct estimates of the indicator of interest y for the i-th area, $\mathbf{x}_i$ is a vector containing the aggregated (population) values of p auxiliary variables with β regression coefficients, the random effects v_i and the sampling errors e_i are assumed to be independent with zero mean and known sampling variances ψ_i and unknown σ_v^2, respectively.

For our purposes, we rewrite the model in the general matrix form

$$\widehat{\mathbf{y}} = \mathbf{X}\beta + \mathbf{B}^{1/2}\mathbf{v} + \mathbf{e}, \tag{2}$$

where now $\mathbf{B} = \text{diag}(b_i^2)$ and the covariance matrix has a diagonal structure $\text{var}(\mathbf{y}) = \mathbf{V} = \text{diag}(V_i) = \text{diag}(\psi_i + \sigma_v^2 b_i^2)$. The vector of the Best Linear Unbiased Predictors (BLUPs) is given by

$$\widehat{\mathbf{y}}^H = \mathbf{X}\widehat{\beta} + \mathbf{B}^{1/2}\widehat{\mathbf{v}} = \mathbf{X}\widehat{\beta} + \mathbf{B}^{1/2}\sigma_v^2\mathbf{B}^{1/2}\mathbf{V}^{-1}(\widehat{\mathbf{y}} - \mathbf{X}\widehat{\beta}) = \mathbf{X}\widehat{\beta} + \Gamma(\widehat{\mathbf{y}} - \mathbf{X}\widehat{\beta}),$$

with $\Gamma = \text{diag}(\gamma_i) = \text{diag}(\sigma_v^2 b_i^2/(\psi_i + \sigma_v^2 b_i^2))$. The generalized least squares estimator of β is

$$\widehat{\beta} = \widehat{\beta}_{GLS} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{V}^{-1}\widehat{\mathbf{y}}).$$

By using the relation

$$\mathbf{V}^{-1} = (\Psi + \sigma_v^2\mathbf{B})^{-1} = \Psi^{-1} - \Psi^{-1}(\Psi^{-1} + (\sigma_v^2)^{-1}\mathbf{B}^{-1})^{-1}\Psi^{-1} = \Psi^{-1}(\mathbf{I} - \Gamma)$$

where $\Psi = \text{diag}(\psi_i)$ and $\mathbf{I}$ is the identity matrix. Denoting with $\widehat{\mathbf{y}}^* = \Psi^{-1/2}\widehat{\mathbf{y}}$ and $\mathbf{X}^* = \Psi^{-1/2}\mathbf{X}$, for this estimator the result is

$$\begin{aligned}\widehat{\beta}_{GLS} &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{V}^{-1}\widehat{\mathbf{y}})\\ &= (\mathbf{X}'\Psi^{-1}\mathbf{X} - \mathbf{X}'\Psi^{-1}\Gamma\mathbf{X})^{-1}(\mathbf{X}'\Psi^{-1}\widehat{\mathbf{y}} - \mathbf{X}'\Psi^{-1}\Gamma\widehat{\mathbf{y}})\\ &= (\mathbf{X}^{*\prime}\mathbf{X}^* - \mathbf{X}^{*\prime}\Gamma\mathbf{X}^*)^{-1}(\mathbf{X}^{*\prime}\widehat{\mathbf{y}}^* - \mathbf{X}^{*\prime}\Gamma\widehat{\mathbf{y}}^*).\end{aligned} \tag{3}$$

An Empirical Best Linear Unbiased Predictor (EBLUP) estimator is obtained from the BLUP by substituting suitable estimators of the variance and covariance parameters. Finally, the Restricted Maximum Likelihood (REML) estimator of σ_v^2 is (see [27] for more details)

$$\widehat{\sigma}^2_{v,REML} = \frac{a}{c^*}\frac{\sum(\widehat{y}_i^* - \overline{\widehat{y}}^*)^2 - (m-1)}{\sum(\widehat{y}_i^* - \overline{\widehat{y}}^*)^2}. \tag{4}$$

With reference to the error in the EBLUP estimator, Prasad and Rao in 1990 [24] gave an approximation to the mean squared error of the EBLUP under the Fay–Herriot model, which estimator includes three terms

$$mse(\hat{y}_i^H) = g_1(\hat{\sigma}_v^2) + g_2(\hat{\sigma}_v^2) + 2g_3(\hat{\sigma}_v^2). \tag{5}$$

It is worth noting that the terms g_2 and g_3, due to estimating β and σ_v, are of lower order than the leading term g_1.

The expressions (3), (4), and (5) will be used in next Section to derive the leverage matrix of the fixed and random effects, the influence on the MSE and a case deletion diagnostics for the empirical predictor.

3 Diagnostics for the Fay–Herriot Model

The main aim of a case diagnostics analysis is to identify observations or groups of observations that might determine the character of model estimates and predictors. In small area estimation, this means to identify the areas among the many that mostly affect the results of the estimates. In order to pursue that aim, after [20], we discuss three diagnostics measures: the leverage, the influential areas that affect the mean squared error estimates, and a Cook-type distance for the empirical predictor.

3.1 The Leverage Matrix

The aim is to investigate the influence of the domains (small areas) on the outcome of the analysis. We are therefore interested in the assessment of the effects of small perturbations in the data on the resulting BLUP estimates $\widehat{\mathbf{y}}^H$. For this reason, the leverage, that is the partial derivative of the predicted value with respect to the corresponding dependent variable, is considered here. In the framework of small area estimation under area level models, leverage is thus the partial derivative of the BLUP with reference to the corresponding direct estimator.

In order to obtain the leverage matrix of fixed and random effects, some useful results are provided below (more details are available within the Appendix). The leverage matrix for the traditional mixed model is given by definition as

$$L(\widehat{\beta}, \widehat{\mathbf{v}}) = \frac{\partial \widehat{\mathbf{y}}}{\partial \mathbf{y}}.$$

Assuming fixed $\mathbf{V}$, the leverage matrix $L(\widehat{\beta}, \widehat{\mathbf{v}})$ can be seen as sum of two components:

$$\begin{aligned} L(\widehat{\beta}, \widehat{\mathbf{v}}) &= L(\widehat{\beta}) + L(\widehat{\mathbf{v}}) \\ &= \mathbf{H}_1 + \mathbf{H}_2 \\ &= \mathbf{X}(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\widehat{\mathbf{V}}^{-1} + \widehat{\sigma}^2_{v,REML}\mathbf{B}\widehat{\mathbf{P}}, \end{aligned}$$

where the first component is the hat matrix $\mathbf{H}_1 = \mathbf{X}(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\widehat{\mathbf{V}}^{-1}$, also denominated generalized marginal leverage matrix, while the second component is given by $\mathbf{H}_2 = \widehat{\sigma}^2_{v,REML}\mathbf{B}\widehat{\mathbf{V}}^{-1}(\mathbf{I}_m - \mathbf{H}_1)$, the leverage matrix for the random component, with

$$\widehat{\mathbf{P}} = \widehat{\mathbf{V}}^{-1} - \widehat{\mathbf{V}}^{-1}\mathbf{X}(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\widehat{\mathbf{V}}^{-1}.$$

Under model (2), it is appropriate to evaluate the effect of each area level direct estimate on the final predictor $\widehat{\mathbf{y}}^H$. The explicit form of the joint leverage matrix that we denote as L^* can be thus decomposed in terms of sampled observations as follows

$$L^*(\widehat{\beta}, \widehat{\mathbf{v}}) = \frac{\partial \widehat{\mathbf{y}}^H}{\partial \widehat{\mathbf{y}}} = \frac{\partial (\mathbf{X}\widehat{\beta})}{\partial \widehat{\mathbf{y}}} + \frac{\partial (\mathbf{B}^{1/2}\widehat{\mathbf{v}})}{\partial \widehat{\mathbf{y}}} = L^*(\widehat{\beta}) + L^*(\widehat{\mathbf{v}}).$$

Based on the derivative

$$\frac{\partial \mathbf{H}_1}{\partial \widehat{\mathbf{y}}} = \left[\frac{\partial(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-1}}{\partial \widehat{\mathbf{y}}}(\mathbf{X}'\otimes\mathbf{X}')\right](\widehat{\mathbf{V}}^{-1}\otimes\mathbf{I}_m) + \frac{\partial\widehat{\mathbf{V}}^{-1}}{\partial\widehat{\mathbf{y}}}(\mathbf{I}_m\otimes[\mathbf{X}(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}']),$$

the leverage matrix for the fixed effects is given by [20]

$$\begin{aligned} L^*(\widehat{\beta}) &= \frac{\partial(\mathbf{X}\widehat{\beta})}{\partial\widehat{\mathbf{y}}} = \frac{\partial}{\partial\widehat{\mathbf{y}}}[\mathbf{X}(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\widehat{\mathbf{V}}^{-1}\widehat{\mathbf{y}}] \\ &= \left(\frac{\partial\mathbf{H}_1}{\partial\widehat{\mathbf{y}}}\right)(\widehat{\mathbf{y}}\otimes\mathbf{I}_m) + \mathbf{H}_1' = \mathbf{H}_1^* + \mathbf{H}_1'. \end{aligned} \tag{6}$$

While the leverage associated with the estimated random effects is

$$\begin{aligned} L^*(\widehat{\mathbf{v}}) &= \frac{\partial(\mathbf{B}^{1/2}\widehat{\mathbf{v}})}{\partial\widehat{\mathbf{y}}} = \frac{\partial}{\partial\widehat{\mathbf{y}}}[\mathbf{B}^{1/2}\widehat{\sigma}^2_{v,REML}\mathbf{B}^{1/2}\widehat{\mathbf{V}}^{-1}(\widehat{\mathbf{y}}-\mathbf{X}\widehat{\beta})] \\ &= [(\mathbf{B}\otimes\sigma^2_{\partial})(\widehat{\mathbf{V}}^{-1}\otimes\mathbf{I}_m) + (\frac{\partial\widehat{\mathbf{V}}^{-1}}{\partial\widehat{\mathbf{y}}})(\mathbf{I}_m\otimes\widehat{\sigma}^2_{v,REML}\mathbf{B})][(\widehat{\mathbf{y}}-\mathbf{X}\widehat{\beta})\otimes\mathbf{I}_m] \\ &\quad +[\mathbf{I}_m - L^*(\widehat{\beta})](\widehat{\sigma}^2_{v,REML}\mathbf{B}\widehat{\mathbf{V}}^{-1}). \end{aligned} \tag{7}$$

For the marginal leverage $\mathbf{H}_1$, as threshold value, it is suggested to use $2p/m$ (see [10]). Using $h^*_{1,ii}$ to indicate the diagonal elements of the matrix $\mathbf{H}_1^*$ (6) for the i-th area, and considering that $\mathrm{tr}(\mathbf{H}_1') = p$, by analogy with [22], in our case influential observations that affect the fixed effects estimates can be verified comparing the quantity $(h^*_{1,ii} - \frac{1}{m}\mathrm{tr}(\mathbf{H}_1^*))$ with $[L^*(\widehat{\beta})_{ii} - \frac{1}{m}\mathrm{tr}(L^*(\widehat{\beta}))]$, more directly through the estimation of the model variance.

In practice, high-leverage observations are also identified by visual examination of the plot of the diagonal values of the leverage matrix. When we assess the potentially influential values, this is very useful in analyzing the contribution to the leverage of the single observation (small area) in estimating the model variance, with reference to the marginal leverage $\mathbf{H}_1$.

3.2 *Influence on the MSE of the EBLUP*

The final purpose in small area estimation is to determine EBLUP estimates for the mean or the total of the variable of interest and to minimize the mean squared error of the empirical predictor. Because of that, the influence of some small areas on the estimation of the MSE plays a central role in the analysis. Once an influential observation has been identified, it could therefore be removed by the researcher in order to improve the precision of the estimates.

With regard to the first component of the mean squared error estimate of the EBLUP, we have the following term as matrix form

$$\mathbf{G}_1 = \text{diag}(g_{1i}) = \Gamma\Psi \tag{8}$$

so that influential area estimates can be detected by the following derivative

$$\begin{aligned}\frac{\partial \mathbf{G}_1}{\partial \widehat{\mathbf{y}}} &= \frac{\partial}{\partial \widehat{\mathbf{y}}}(\Gamma\Psi) = \frac{\partial}{\partial \widehat{\mathbf{y}}}(\widehat{\sigma}^2_{v,REML}\mathbf{B}\widehat{\mathbf{V}}^{-1}\Psi) \\ &= (\mathbf{B}\otimes\sigma^2_{\partial})(\widehat{\mathbf{V}}^{-1}\Psi\otimes\mathbf{I}_m) + [\frac{\partial\widehat{\mathbf{V}}^{-1}}{\partial\widehat{\mathbf{y}}}(\Psi\otimes\mathbf{I}_m)](\mathbf{I}_m\otimes\widehat{\sigma}^2_{v,REML}\mathbf{B}).\end{aligned} \tag{9}$$

With reference to the second component, related to the variation of the fixed effects, that is

$$\mathbf{G}_2 = (\mathbf{I}_m - \Gamma)\mathbf{X}(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'(\mathbf{I}_m - \Gamma)(\mathbf{I}_m - \Gamma)\mathbf{U}(\mathbf{I}_m - \Gamma)],$$

after [20] we have the following influence measure

$$\frac{\partial \mathbf{G}_2}{\partial \widehat{\mathbf{y}}} = \frac{\partial(\mathbf{I}_m - \Gamma)}{\partial\widehat{\mathbf{y}}}([\mathbf{U}(\mathbf{I}_m - \Gamma)]\otimes\mathbf{I}_m) + \frac{\partial[\mathbf{U}(\mathbf{I}_m - \Gamma)]}{\partial\widehat{\mathbf{y}}}[\mathbf{I}_m\otimes(\mathbf{I}_m - \Gamma)], \tag{10}$$

where

$$\begin{aligned}\frac{\partial(\mathbf{I}_m - \Gamma)}{\partial\widehat{\mathbf{y}}} &= -(\mathbf{B}\otimes\sigma^2_{\partial})(\widehat{\mathbf{V}}^{-1}\otimes\mathbf{I}_m) + (\frac{\partial\widehat{\mathbf{V}}^{-1}}{\partial\widehat{\mathbf{y}}}\mathbf{I}_{m^2})(\mathbf{I}_m\otimes\widehat{\sigma}^2_{v,REML}\mathbf{B}), \\ \frac{\partial[\mathbf{U}(\mathbf{I}_m - \Gamma)]}{\partial\widehat{\mathbf{y}}} &= [\frac{\partial(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-1}}{\partial\widehat{\mathbf{y}}}(\mathbf{X}'\otimes\mathbf{X}')][(\mathbf{I}_m - \Gamma)\otimes\mathbf{I}_m] + \frac{\partial(\mathbf{I}_m - \Gamma)}{\partial\widehat{\mathbf{y}}}(\mathbf{I}_m\otimes\mathbf{U}).\end{aligned}$$

These derivatives are important as they measure the increase (positive value) or the decrease (negative value) of the MSE of a small area, with reference to the direct estimate of another small area. For such an influential measure, no threshold values are available. Consequently, our suggestion is to first visualize the results for each area of interest (by column vector) from the $m \times m$ resulting matrix of (9) and (10) and then, for each column vector, investigate if there is any area showing particularly higher values.

3.3 Case Deletion Diagnostics and Cook's Distance

Here, we define Cook's distances for the REML estimate of $\hat{\sigma}^2_v$ and for the EBLUP $\hat{y}^H_i$.

After [20], Cook's distance for $\hat{\sigma}^2_v$ is given by

$$d^v_\ell = \frac{(\hat{\sigma}^2_v - \hat{\sigma}^2_{v(\ell)})^2}{\widehat{\text{var}}(\hat{\sigma}^2_v)},$$

where $\widehat{\text{var}}(\hat{\sigma}_v^2)$ is the asymptotic variance of $\hat{\sigma}_v^2$ that is obtained from the inverse of the REML Fisher information matrix, and the subscript $_{(\ell)}$ is used for those estimators that are calculated after deleting case ℓ.

The proposed Cook-type distance for the EBLUP $\hat{y}_i^H$ is

$$d_\ell^{eblup} = \frac{(\hat{y}_i^H - \hat{y}_{i(\ell)}^H)^2}{mse(\hat{y}_i^H)}$$

where $\hat{y}_{i(\ell)}^H$ is the EBLUP with case ℓ deleted and $mse(\hat{y}_i^H)$ is the Prasad-Rao [24] MSE estimator.

Cook's distance assesses the effects of a global change by removing an entire data point. It follows that large values of d_ℓ^{eblup} will point out that the corresponding area may affect the EBLUP estimate of the related deleted area.

4 An Application to Poverty Data

In order to design and implement poverty reduction policies and funding programs, there has been an increasing demand for poverty and living condition estimates at aggregate and local levels. Within such a framework, case diagnostics may have a very important role.

For this reason, this Section illustrates how the diagnostic tools introduced before can be exploited within a real data analysis performed on the official Spanish Living Condition Survey of the European Statistics on Income and Living Conditions (EU-SILC). The latter is a cross-sectional and longitudinal sample survey, coordinated by Eurostat, based on data from the European Union member states. It provides data on income, poverty, social exclusion, and living conditions in the European Union.

The analysis aims at estimating poverty levels in small area domains by the use of a Fay–Herriot area level model (1).

The dataset refers to the years 2004–2006 and contains 104 observations (areas in our context) obtained by crossing 52 Spanish provinces with 2 sex (men and women). The target variable is the direct estimate of the poverty indicator proposed by Foster et al. [13] (poverty incidence or proportion) at domain level (province $\times$ sex). Estimates of the domain means are used as responses in the area level model. The considered auxiliary variables are the known domain means of the category indicators of the following variables: age, education, citizenship, and labor. Finally, only 3 statistically significant variables that have a relevant meaning in a socio-economic sense are selected. They are age group 50–65, secondary education completed, and unemployment condition. The analysis was conducted with the open-source software R.

As discussed in Sect. 3.1, we start our influential analysis by computing the leverage values. We estimated the REML random-area effect variance, and we calculated

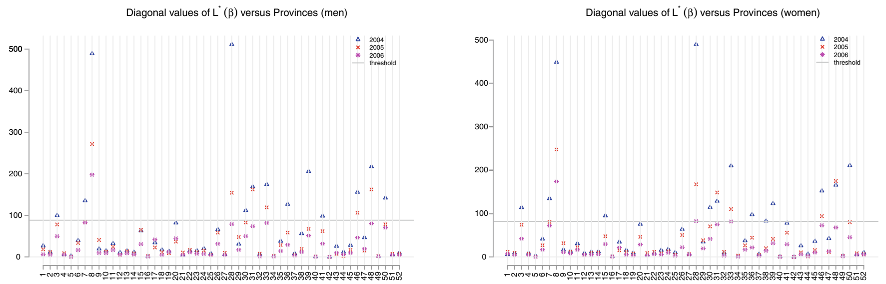

Fig. 1 The diagonal values of the leverage matrix for the fixed effects $L^*(\widehat{\beta})$ plotted for all the Spanish provinces, separately for men (left) and women (right)

the derivatives for the construction of the leverage matrix of the fixed effects as described in Eq. (6).

Results referring to the diagonal values of the leverage matrix for the fixed effects $L^*(\widehat{\beta})$ are presented in the scatter plot appearing in Fig. 1 for men and women, compared with a critical value of $2\left(\sum_{i=1}^{m} L^*_{ii}(\widehat{\beta})/m\right)$. By looking at the magnitude of the leverage values we conclude that the highest influential values are the provinces of Barcelona (8) and Madrid (28), for both sex categories, where the number in brackets indicates the corresponding numerical label on the abscissa axis of each plot. Differences also appear when focusing on to the three years taken into account. The leverage values tend to decrease from 2004 to 2006.

The relation between $L^*(\widehat{\beta})$ and the direct estimates of the poverty proportions along the three years 2004–2006 is shown in Fig. 2. The plot shows that the direct estimates with lower values generally correspond to higher level of leverage. Therefore, lower direct estimates of poverty proportion can be considered to be more influential than the higher ones. The values which stand out as the most influential are again the provinces of Barcelona (8) and Madrid (28).

Results of the influence analysis on the MSE of the EBLUP estimates, in particular that of the calculations of the derivative of $\mathbf{G}_1$ induced by the direct estimates (Eq. (9)), are illustrated in Fig. 3. By way of an example, according to their sample sizes (respectively small, large, and medium), the results for three provinces are presented: Alicante, Barcelona and Granada, for men and women, respectively, and for the year 2004. Similar results were also obtained for the years 2005 and 2006: for the sake of brevity they are not reported here. In this case, we observe a difference between men and women among the cases that are more influent on the MSE of the poverty level estimates. Among the three selected provinces, the most influential province for men is Granada, while for women it is Alicante. They are highlighted in the plots with dotted and dashed lines, respectively, which consistently dominate the others.

The derivative of $\mathbf{G}_1$ captures the influence that each small area can have on each other in terms of the power for increasing or decreasing the MSE of the EBLUP. The province of Granada suffers thus the increase of its MSE by the provinces of Badajoz (6), Barcelona (8), Madrid (28), Murcia (30), Sevilla (41), and Valencia

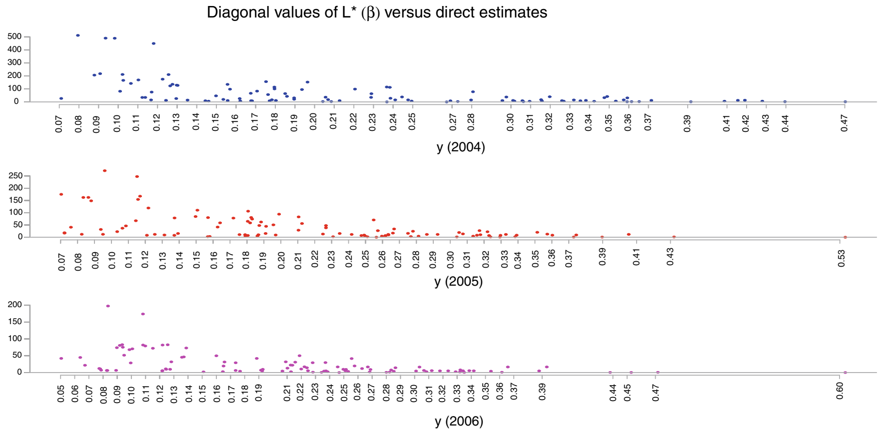

Fig. 2 The diagonal values of the leverage matrix for the fixed effects $L^*(\widehat{\beta})$ plotted against all the Spanish provinces direct estimates

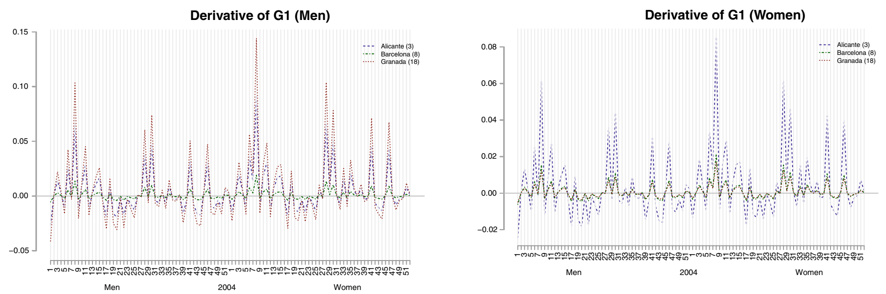

Fig. 3 Derivative of $\mathbf{G}_1$ for the provinces of Alicante (dashed line), Barcelona (dotdashed line), and Granada (dotted line) plotted against all the Spanish provinces in the year 2004, separately for men (left) and women (right)

(46). On the other side, the decrease of its MSE is caused by the provinces of Alava (1) and Gerona (17) (Fig. 3, left). As for the women (Fig. 3, right), the first part of the MSE of Alicante is affected by a positive influence by the same provinces that affect the plot of the men: Badajoz (6), Barcelona (8), Madrid (28), Murcia (30), Sevilla (41), and Valencia (46); the decrease instead is due to Alava (1), Gerona (17), Guipuzcoa (20), and Teruel (44).

Cook's Distance for the EBLUP as calculated in Sect. 3.3 are presented in Fig. 4. As done before, only the results of three provinces are presented: Alicante, Barcelona and Granada, for men and women, respectively, and for the year 2004. In the graph, three lines of Cook's distance are reported, which correspond to the deleted provinces of Alicante, Barcelona, and Granada. The peaks of Cook's distance represent the most influential values on the EBLUP estimates of the related deleted provinces. In particular, when looking at the results for men (Fig. 4, left), the lines of Alicante and

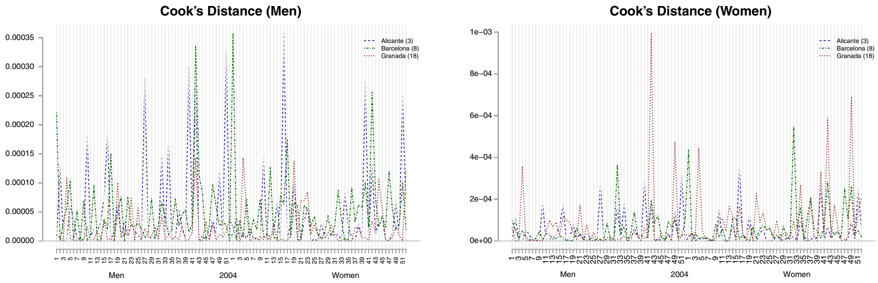

Fig. 4 Cook's Distance for the EBLUP $\hat{y}_i^H$ performed deleting the provinces of Alicante (dashed line), Barcelona (dotdashed line), and Granada (dotted line) plotted against all the Spanish provinces in the year 2004, separately for men (left) and women (right)

Barcelona show more peaks, and the deletion of these provinces shows that Alava (1), Cuenca (16), and Soria (42) are the provinces more influential for them. For women (Fig. 4, right), it shows that removing the province of Granada produces a high value of Cook's distance in correspondence with the province of Soria (42).

5 Concluding Remarks

A review of recent developments on diagnostic tools for the Fay–Herriot small area model when dealing with the restricted maximum likelihood estimate is proposed. Detailed formulae on fixed and random effects leverage matrices are reached in case of fixed **V**. Tools for an influence analysis on the Mean Squared Error (MSE) of the Empirical Best Linear Unbiased Predictor (EBLUP) and a Cook's distance for the empirical predictor are considered.

The problem of the leverage of observed values on predicted values by the EBLUP was observed when we consider that, even though we make use of convenient estimates, the latter depends on the same influential values. Therefore, the leverage matrix of the model can be affected by influential observations through the estimates of the model variance. On the other hand, influence analysis on MSE estimates is based on $m \times m^2$-order matrices, which can be very useful in assessing the contribution of single observations (the small area direct estimates) in the evaluation of the MSE of all areas.

An application to real data is offered to the reader. The case of the estimation of poverty proportions for the Spanish provinces is exploited to illustrate the benefits of using specific diagnostic tools in the context of small area estimation. This methodology is useful because once the influential areas have been identified through visual examination, the researcher can eliminate them to improve the accuracy of the estimates.

Results in this paper are intended to be extended by the author to some other small area models, in particular to models that borrow strength from time or spatial correlations. It is thought that this research line might be of great interest to applied statisticians.

Acknowledgements The author thanks the Spanish National Statistics Institute for supplying the data.

Appendix

Details on how Eqs. (6), (7), (9) and (10) are derived are provided below.

For Eq. (6), let us first define the matrix A as

$$A = \sum (\widehat{y}_i^* - \widehat{\bar{y}})^2 = \widehat{\mathbf{y}}^{*\prime}\widehat{\mathbf{y}}^* - \frac{1}{m}(\widehat{\mathbf{y}}^{*\prime}\mathbf{1}_m)^2$$

where $\mathbf{1}_m$ denote the unitary vector all of whose components are unity. Following [20], we have then

$$\begin{aligned}
\frac{\partial A}{\partial \widehat{\mathbf{y}}} &= \frac{\partial}{\partial \widehat{\mathbf{y}}} \sum (\widehat{y}_i^* - \widehat{\bar{y}})^2 = \frac{\partial}{\partial \widehat{\mathbf{y}}}\left[\widehat{\mathbf{y}}^{*\prime}\widehat{\mathbf{y}}^* - \frac{1}{m}(\widehat{\mathbf{y}}^{*\prime}\mathbf{1}_m)^2\right] \\
&= 2\widehat{\mathbf{y}}'\Psi^{-1} - \frac{2}{m}(\widehat{\mathbf{y}}'\mathbf{1}_m^{\Psi})(\mathbf{1}_m^{\Psi})', \text{ where } \mathbf{1}_m^{\Psi} = \Psi^{-1/2}\mathbf{1}_m.
\end{aligned}$$

For the Eqs. (7), (9), and (10), the derivative of the REML variance estimate is defined as follows:

$$\begin{aligned}
\frac{\partial \widehat{\sigma}^2_{v,REML}}{\partial \widehat{\mathbf{y}}} &= \frac{\partial}{\partial \widehat{\mathbf{y}}}\left[\frac{a}{c^*}\frac{A-(m-1)}{A}\right] \\
&= \frac{a}{c^*}A^{-2}\left[2\widehat{\mathbf{y}}'\Psi^{-1} - \frac{2}{m}(\widehat{\mathbf{y}}'\mathbf{1}_m^{\Psi})(\mathbf{1}_m^{\Psi})'\right] = \sigma^2_{\partial}\ .
\end{aligned}$$

For the derivative of $\widehat{\mathbf{V}}$ and its inverse, appearing in Eqs. (7) and (9), we have the following:

$$\begin{aligned}
\frac{\partial \widehat{\mathbf{V}}}{\partial \widehat{\mathbf{y}}} &= \frac{\partial}{\partial \widehat{\mathbf{y}}}\text{diag}(\psi_i + \widehat{\sigma}^2_{v,REML}b_i^2) \\
&= \frac{\partial}{\partial \widehat{\mathbf{y}}}(\Psi + \mathbf{B}^{1/2}\widehat{\sigma}^2_{v,REML}\mathbf{B}^{1/2}) = \mathbf{B} \otimes \sigma^2_{\partial},
\end{aligned}$$

$$\frac{\partial \widehat{\mathbf{V}}^{-1}}{\partial \widehat{\mathbf{y}}} = -(\widehat{\mathbf{V}}^{-1}\mathbf{B}\widehat{\mathbf{V}}^{-1}) \otimes \sigma^2_{\partial}.$$

Finally, the derivative of $(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-1}$ which refers to Eq. (10) is

$$\frac{\partial(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-1}}{\partial\widehat{\mathbf{y}}} = -\frac{\partial(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})}{\partial\widehat{\mathbf{y}}}[(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X}) \otimes (\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})]$$

$$= [(\widehat{\mathbf{V}}^{-1}\mathbf{B}\widehat{\mathbf{V}}^{-1}) \otimes \sigma_\partial^2](\mathbf{X} \otimes \mathbf{X})[(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X}) \otimes (\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})].$$

References

1. Battese, G.E., Harter, R.M., Fuller, W.A.: An error component model for prediction of county crop areas using survey and satelite data. J. Am. Stat. Assoc. **83**, 28–36 (1988)
2. Beckman, R.J., Nachtsheim, C.J., Cook, R.D.: Diagnostics for mixed-model analysis of variance. Technometrics **29**, 413–426 (1987)
3. Belsley, D.A., Kuh, E., Welsh, R.E.: Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. Wiley, Hoboken (1980)
4. Benavent, R., Morales, D.: Multivariate Fay? Herriot models for small area estimation. Comput. Stat. Data Anal. **94**, 372–390 (2016)
5. Cook, R.D., Weisberg, S.: Residuals and Influence in Regression. Chapman and Hall, New York (1982)
6. Cook, R.D.: Detection of influential observations in linear regression. Technometrics **19**, 15–18 (1977)
7. Cook, R.D.: Assessment of local influence. J. R. Stat. Soc. B **48**, 133–169 (1986)
8. Chatterjee, S., Hadi, A.S.: Influential observations, high leverage points, and outliers in linear regression. Stat. Sci. **1**, 379–416 (1986)
9. Datta, S., Lahiri, P.: A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems. Stat. Sinica **10**, 613–627 (2000)
10. Demidenko, E.: Mixed Models: Theory and Applications. Wiley, Hoboken (2004)
11. Demidenko, E., Stukel, T.A.: Influence analysis for linear mixed-effects models. Stat. Med. **24**, 893–909 (2005)
12. Fay, R.E., Herriot, R.A.: Estimates of income for small places: an application of James-Sein procedures to census data. J. Am. Stat. Assoc. **74**, 269–277 (1979)
13. Foster, J., Greer, J., Thorbecke, E.: A class of decomposable poverty measures. Econometrica **52**, 761–766 (1984)
14. Fung, W.K., Zhu, Z.Y., Wei, B.C., He, X.: Influence diagnostics and outliers tests for semiparametric mixed models. J. R. Stat. Soc. B **64**, 565–579 (2002)
15. Gonzalez-Manteiga, W., Lombardia, M.J., Molina, I., Morales, D., Santamaria, L.: Small area estimation under Fay?Herriot models with nonparametric estimation of heteroscedasticity. Stat. Model. **10**(2), 215–239 (2010)
16. Herrador, M., Esteban, M.D., Hobza, T., Morales, D.: A Fay?Herriot model with different random effect variances. Commun. Stat. Theory Methods **40**(5), 785–797 (2011)
17. Lesaffre, E., Verbeke, G.: Local Influence in Linear Mixed Models. Biometrics **54**(2), 570–582 (1998)
18. Marhuenda, Y., Molina, I., Morales, D.: Small area estimation with spatio-temporal Fay? Herriot models. Comput. Stat. Data Anal. **58**, 308–325 (2013)
19. Molina, I., Rao, J.N.K.: Small area estimation of poverty indicators. Can. J. Stat. **38**, 369–385 (2010)
20. Morales D., Pagliarella M.C., Salvatore R.: Influence analysis in small area estimation. In: Conference Proceedings of the 45th Meeting of the Italian Statistical Society. Cleup Editore, Padova (2010)

21. Neter, J., Kutner, M.H., Nachtsheim, C.J., Wasserman, W.: Applied Linear Statistical Models. IRWIN, Chicago (1990)
22. Nobre, J.S., Singer, M.: Leverage analysis for linear mixed models. J. Appl. Stat. **38**(5), 1063–1072 (2011)
23. Pan, J., Fei, Y., Foster, P.: Case-Deletion Diagnostics for Linear Mixed Models. Technometrics **56**(3), 269–281 (2014)
24. Prasad, N.G.N., Rao, J.N.K.: The estimation of the mean squared error of small-area estimators. J. Am. Stat. Assoc. **85**, 163–171 (1990)
25. Pratesi, M.: Analysis of Poverty Data by Small Area Estimation. Wiley, New York (2016)
26. Pregibon, D.: Logistic Regression Diagnostics. Ann. Stat. **4**, 705–724 (1981)
27. Rao, J.N.K.: Small Area Estimation. Wiley, Hoboken NJ (2003)
28. Singh, B., Shukla, G., Kundu, D.: Spatio-temporal models in small area estimation. Surv. Methodol. **31**, 183–195 (2005)
29. Zewotir, T., Galpin, J.S.: Influence diagnostics for linear mixed models. J. Data Sci. **3**, 153–177 (2005)