Outliers in the Annual Survey of Public Employment & Payroll

Small Area Estimation Approach

Giang Trinh
Giang.B.Trinh@census.gov
Bac Tran
Bac.Tran@census.gov
U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233

Abstract

The presence of outliers in the Small Area Estimation (SAE) raises serious concerns in the design-based part of population parameters prediction due to the violation of model-based assumptions. Various techniques have been introduced to mitigate the effect of outliers in the unit-level and area-level models in the SAE in the literature. In this paper, we introduced the square root transformation into the mixture models from [9, 10] in order to deal with those outliers and estimate the total number of employees in the Annual Survey of Public Employment & Payroll (ASPEP) data. We then compared our research method to the existing methods being used in the estimation of the ASPEP at the Census Bureau. The two Public Employment census data of 2007 and 2012 were used for the evaluation of this research.

Key Words: robust estimation, small area, mixture model

1. Introduction

There has been growing interest among researchers and practitioners in applying SAE to survey sampling. Small area (or domain) refers to populations for which sampling rates are small due to observational limitations. SAE provides reliable estimates by using both the survey data and auxiliary sources. There are many small area estimators available in the literature. These estimators typically use either implicit or explicit models to combine survey data with various administrative and Census records. The properties of such model-based estimators rely on assumptions applied to simulated data. There may be observations that carry important information in real data that cannot be ignored. However these observations would not be described appropriately by modeling assumptions. Recent advanced methods in SAE litterature (Fellner [7], Chambers et al. [4], Sinha and Rao [13], Gershunskaya and Lahiri [9, 10]) provide estimates that are robust to the existence of outliers. An estimator is robust if it provides good estimates even if some of the assumptions used to justify the estimation method are not applicable.

– – – – – – – – – – – – – –

The ASPEP survey is designed to produce reliable estimates of statistics on the number of federal, state, and local government civilian employees and their gross payroll each month at the national level and for large domains. However, it is also required to estimate the parameters for individual function codes within each state. This requirement leads us to explore SAE methodology that borrows strengths from previous Census data instead of collecting expensive additional data for small cells.

The purpose of this study is to apply SAE methodology to improve the estimates when outliers exist. We will explore the method of using models that are based on scale mixtures of two normal distributions [9, 10] to produce robust estimation for the ASPEP parameters in different areas. The models used the ASPEP data and auxiliary information from the preceding Census of Government Employment. This report is organized as follows: In Section 2, we summarize the algorithm for obtaining N2 and N2+OBC estimations proposed by Gershunskaya and Lahiri [9, 10]. Section 3 describes the application of this method where we introduced the square root transformation to the ASPEP data in order to provide the estimators of the total full-time employment for different government functions in California. Then we compare the relative root mean square errors (RRMSE) of the N2 and N2+OBC estimations and those of Horvitz-Thompson (HT) and of the empirical best linear unbiased predictor (EBLUP) used in ASPEP's production.

2. Estimation method

In this section, we briefly describe the N2 and N2+OBC robust estimators. (Readers interested in more details are referred to Gershunskaya and Lahiri [9, 10])

N2 and N2+OBC estimators

Let y_{mj} denote the population of the j^{th} unit within the m^{th} area. We are interested in estimating the total population $Y_m = \sum_{j=1}^{N_m} y_{mj}$; $m = 1, ..., M$ (N_m: number of units of the m^{th} area; M: number of areas). An estimator of Y_m is given by:

$$\widehat{Y}_m \quad = \quad N_m[f_m \overline{y}_m + (1 - f_m)\widehat{\overline{\boldsymbol{Y}}}_{\boldsymbol{mr}}], \tag{0.1}$$

where $\overline{y}_m = \frac{y_m}{n_m}$: the sampled population mean; $y_m = \sum_{j=1}^{n_m} y_{mj}$: the sampled population; $f_m = \frac{n_m}{N_m}$: the sampling rate; n_m: the sample size; and $\widehat{\overline{Y}}_{mr}$: a predictor of the mean of the non-sampled part of the m^{th} area.

Simplifying (0.1) we have:

$$\widehat{Y}_m \quad = \quad y_m + \widehat{Y}_{mr}, \tag{0.2}$$

where $\widehat{Y}_{mr}$ is a predictor of the total of the non-sampled part of the m^{th} area.

N2 Estimator

The predictor $\widehat{Y}_{mr}$ can be derived from a linear mixed model (denoted **N2**) that is based on a scale mixture of two normal distributions with common mean and different variances. The model is given by (0.3)-(0.5):

$$y_{mj} = \mathbf{x}_{\boldsymbol{mj}}^{\boldsymbol{T}}\boldsymbol{\beta} + u_m + \varepsilon_{mj},\ j = 1, ..., n_m,\ m = 1, ..., M, \tag{0.3}$$

$$u_m \overset{\text{iid}}{\sim} \mathcal{N}(0, \tau^2),\ \varepsilon_{mj}|\, z \overset{\text{iid}}{\sim} (1-z)\mathcal{N}(0, \sigma_1^2) + z\mathcal{N}(0, \sigma_2^2), \tag{0.4}$$

$$z|\, \boldsymbol{\pi} \sim \text{Bin}(1; \pi),\ \sigma_2 > \sigma_1. \tag{0.5}$$

where $\mathbf{x}_{\boldsymbol{mj}}$ is a vector of auxiliary variables for an observation mj; $\boldsymbol{\beta}$ is the vector of regression parameters associated with the vector $\mathbf{x}_{\boldsymbol{mj}}$ of auxiliary variables; u_m are random effects; ε_{mj} are errors in individual observations. The random variables u_m and ε_{mj} are assumed to be mutually independent. The mixture part indicator is a random binomial variable $z|\, \boldsymbol{\pi} \sim \text{Bin}(1; \pi)$ where $\boldsymbol{\pi}$ is the probability of the observation belonging to mixture part 2.

Step 1: Estimating the model parameters using the EM algorithm:

Let $\boldsymbol{\theta}^{(p)} = (\sigma_1^{(p)}, \sigma_2^{(p)}, \tau^{(p)}, \pi^{(p)}, \boldsymbol{\beta}^{\boldsymbol{(p)}})$ be a set of parameter values after the p^{th} iteration. At the $(p+1)^{\text{th}}$ iteration, compute:

E-step:

$$z_{mj}^{(p+1)} = \frac{\frac{1-\pi^{(p)}}{\sqrt{\sigma_2^{(p)2}+\tau^{(p)2}}}\exp\left[-\frac{(y_{mj}-\mathbf{x}_{\boldsymbol{mj}}^{\boldsymbol{T}}\boldsymbol{\beta}^{\boldsymbol{(p)}})^2}{2(\sigma_2^{(p)2}+\tau^{(p)2})}\right]}{\frac{\pi^{(p)}}{\sqrt{\sigma_1^{(p)2}+\tau^{(p)2}}}\exp\left[-\frac{(y_{mj}-\mathbf{x}_{\boldsymbol{mj}}^{\boldsymbol{T}}\boldsymbol{\beta}^{\boldsymbol{(p)}})^2}{2(\sigma_1^{(p)2}+\tau^{(p)2})}\right] + \frac{1-\pi^{(p)}}{\sqrt{\sigma_2^{(p)2}+\tau^{(p)2}}}\exp\left[-\frac{(y_{mj}-\mathbf{x}_{\boldsymbol{mj}}^{\boldsymbol{T}}\boldsymbol{\beta}^{\boldsymbol{(p)}})^2}{2(\sigma_2^{(p)2}+\tau^{(p)2})}\right]}$$

$$w_{mj}^{(p+1)} = \frac{1-z_{mj}^{(p+1)}}{\sigma_1^{(p)2}} + \frac{z_{mj}^{(p+1)}}{\sigma_2^{(p)2}} \tag{0.6}$$

$$\overline{y}_m^{(p+1)} = (\sum_{j=1}^{n_m} w_{mj}^{(p+1)} y_{mj}) / \sum_{j=1}^{n_m} w_{mj}^{(p+1)} \tag{0.7}$$

$$\overline{\mathbf{x}}_{\boldsymbol{m}}^{\boldsymbol{(p+1)}} = (\sum_{j=1}^{n_m} w_{mj}^{(p+1)} \mathbf{x}_{\boldsymbol{mj}}) / \sum_{j=1}^{n_m} w_{mj}^{(p+1)}$$

$$V_m^{(p+1)} = 1/(\sum_{j=1}^{n_m} w_{mj}^{(p+1)} + \frac{1}{\tau^{(p)2}})$$

$$u_m^{(p+1)} = V_m^{(p+1)}(\overline{y}_m^{(p+1)} - \overline{\mathbf{x}}_{\boldsymbol{m}}^{\boldsymbol{(p+1)T}}\boldsymbol{\beta}^{\boldsymbol{(p)}}) \sum_{j=1}^{n_m} w_{mj}^{(p+1)} \tag{0.8}$$

M-step

$$
\begin{aligned}
\pi^{(p+1)} &= (\sum_{m=1}^{M}\sum_{j=1}^{n_m} z_{mj}^{(p+1)})/n \\
\sigma_1^{(p+1)2} &= \frac{\sum_{m=1}^{M}\sum_{j=1}^{n_m}(1-z_{mj}^{(p+1)})\left[(y_{mj}-\mathbf{x}_{\boldsymbol{mj}}^{\boldsymbol{T}}\boldsymbol{\beta}^{(\boldsymbol{p})}-u_m^{(p+1)})^2+V_m^{(p+1)}\right]}{\sum_{m=1}^{M}\sum_{j=1}^{n_m}(1-z_{mj}^{(p+1)})} \\
\sigma_2^{(p+1)2} &= \frac{\sum_{m=1}^{M}\sum_{j=1}^{n_m} z_{mj}^{(p+1)}\left[(y_{mj}-\mathbf{x}_{\boldsymbol{mj}}^{\boldsymbol{T}}\boldsymbol{\beta}^{(\boldsymbol{p})}-u_m^{(p+1)})^2+V_m^{(p+1)}\right]}{\sum_{m=1}^{M}\sum_{j=1}^{n_m} z_{mj}^{(p+1)}} \\
\tau^{(p+1)2} &= (\sum_{m=1}^{M} u_m^{(p+1)2}+V_m^{(p+1)})/M \\
\boldsymbol{\beta}^{(p+1)} &= \frac{\sum_{m=1}^{M}\sum_{j=1}^{n_m} w_{mj}^{(p+1)}\mathbf{x}_{\boldsymbol{mj}}(y_{mj}-u_m^{(p+1)})}{\sum_{m=1}^{M}\sum_{j=1}^{n_m} w_{mj}^{(p+1)}\mathbf{x}_{\boldsymbol{mj}}^{\boldsymbol{T}}\mathbf{x}_{\boldsymbol{mj}}}.
\end{aligned} \tag{0.9}
$$

Step 2: The predictor of $\boldsymbol{Y_{mr}}$ is given by

$$
\widehat{\mathbf{Y}}_{\boldsymbol{mr}}^{\boldsymbol{N2}} = \mathbf{x}_{mr}^{T}\widehat{\beta}^{N2} + (N_m - n_m)\widehat{u}_m^{N2} \tag{0.10}
$$

where $\widehat{\boldsymbol{\beta}}^{\boldsymbol{N2}}$, $\widehat{u}_m^{N2}$ are computed from the EM algorithm; and $\mathbf{x}_{mr}^{T} = \sum_{j=n_m+1}^{N_m}\mathbf{x}_{\boldsymbol{mj}}^{\boldsymbol{T}}$.

Step 3: The estimate of $\boldsymbol{Y_m}$ is $\widehat{\mathbf{Y}}_m^{N2} = y_m + \widehat{\mathbf{Y}}_{mr}^{N2}$.

N2+OBC estimator (N2 with Overall Bias Correction)

The assumptions about the distribution of the random effects u_m and the error terms ε_{mj} may be over-specified. The influence of the outlying areas or outlying units on the estimates of model parameters can be reduced by using a bounded Huber's function, $\phi_b(r) = \min(b, \max(-b, r))$, for the corresponding residual terms (Fellner [7], Huber and Ronchetti [11]).

Let $\mathbf{e}_{\mathbf{mj}}^{\mathbf{N2}} = \mathbf{y_{mj}} - \mathbf{x}_{\boldsymbol{mj}}^{\boldsymbol{T}}\widehat{\boldsymbol{\beta}}^{\boldsymbol{N2}} - \widehat{\mathbf{u}}_{\mathbf{m}}^{\mathbf{N2}}$. The overall bias-corrected estimate (of $\mathbf{Y_{mr}}$) is given by

$$
\widehat{\mathbf{Y}}_{\mathbf{mr}}^{\mathbf{N2+OBC}} = \widehat{\mathbf{Y}}_{\mathbf{mr}}^{\mathbf{N2}} + (\boldsymbol{N_m} - \boldsymbol{n_m})\frac{\boldsymbol{s^R}\sum_{\boldsymbol{m=1}}^{\boldsymbol{M}}\sum_{\boldsymbol{j=1}}^{\boldsymbol{n_m}}\boldsymbol{\phi_b}(\boldsymbol{e_{mj}^{N2}}/\boldsymbol{s^R})}{\boldsymbol{n}}, \tag{0.11}
$$

where $m = 1, ...M$; $j = 1, ..., n_m$; ϕ_b is the bounded Huber's function with tuning parameter $b = 5$; and $\mathbf{s^R} = \frac{\mathrm{med}\left|\mathbf{e_{mj}^{N2}} - \mathrm{med}(\mathbf{e_{mj}^{N2}})\right|}{\mathbf{0.6745}}$ is a robust measure of scale for the set of residuals $\boldsymbol{e_{mj}^{N2}}$.

RRMSE

To assess the quality of the estimators, we used the relative root mean squared error, $\mathrm{RRMSE} = 100\sqrt{\frac{1}{\mathrm{rep}}\sum_{k=1}^{\mathrm{rep}}(\frac{\widehat{Y}_{m.k}-Y_m}{Y_m})^2}$, where $\widehat{Y}_{m.k}$ is the estimate of Y_m (the population total of area m from the k^{th} replicate).

3. Application to ASPEP data

The U.S. Census Bureau conducts Censuses of about 90,000 state and local government units every five years in order to collect data on the number of full-time and part-time state and local government employees and payroll. Between two consecutive Censuses (years ending with 2 and 7, e.g., 2002, 2007, and 2012), the Census Bureau also conducts the Annual Survey of Public Employment & Payroll, a nationwide sample survey covering all state and local governments in the United States, which include five types of governments: counties, cities, townships, special districts, and school districts. The first three types of government are referred to as general-purpose government, because they generally provide multiple government activities. Data on employment include the number of full-time and part-time employees and gross pay as well as hours paid for part-time employees. All data are reported for the government's pay period covering March 12. Data collection begins in March and continues for about seven months. For more information on the survey, we refer to http://www.Census.gov/govs/apes.

Missing data, small areas and outliers

In order to produce reliable estimation, a direct estimate requires a large enough sampling rate. Unfortunately, this is not the case, and missing data always exist in survey sampling. Table 1 shows the actual sample rates in 29 areas are less than 33%. Especially, the sample rates are small (less than 10%) in areas corresponding to function codes 016, 018, 059, 087, 092, and 093 where the data available may not be large enough to produce direct estimates that are reliable. Instead, it may be preferable to use a model-based small area approach. Small area estimation uses indirect estimates that utilize information from various sources such as samples, non-sampled parts, and also from areas with similar characteristics to the areas of interest. It borrows strength from data in different ways. The N2 (or N2+OBC) estimate combines information from sample in y_m (step), the non-sampled parts from all areas in the data, not just area m. See equations (0.9), (0.8), and (0.10). This estimate is robust because $\widehat{\mathbf{Y}}_m^{N2}$ (or $\widehat{\mathbf{Y}}_m^{N2+OBC}$) may not be affected by outliers. When z_{mj} increases, the equation (0.6) shows that outliers would be down-weighted. As a result the N2 (or N2+OBC) estimate is insensitive to these outliers.

Model-based estimation relies on some assumptions. SAE applies linear models such as the Fay-Herriot model [6] at the area level. These models assume normality of the error terms. In practice, this assumption is rarely met. Relaxing model assumptions may help in fitting more observations to the model. The following nested-error regression model at the unit level gives us more flexibilities in expressing the error terms (see [1]) :

$$\begin{aligned} y_{mj} &= \mathrm{x}_{mj}^T \beta + u_m + \varepsilon_{mj},\ m = 1, ..., M;\ j = 1, ..., n_m \\ & u_m \sim \mathcal{N}(0, \tau^2),\ \varepsilon_{mj} \sim \mathcal{N}(0, \sigma^2) \end{aligned}$$

Failure to meet model assumptions may lead to large errors in estimation when deleting the unfit outliers that are influential. The assumption of normality may be resolved in parts by using transformation. Relaxing normal assumptions may be made by using t-Distribution to describe either the random effect u_m (model error) or ε_m (survey errors), see Bell et al. [3], Staudenmayer et al. [14]. Survey errors may come from more than one normal distribution. Then the assumption that ε_m comes from one normal distribution may be too strong. This can be relaxed by using a scale mixture of two normal distributions (Fig. 1 and/or De Veaux and Krieger [15]) to describe the error term ε_m (Gershunskaya and Lahiri [9, 10]). ASPEP data is skewed (Fig. 2) and contains outliers. The box-plots show data after applying transformations, such as log and power functions, approximately meet normal assumptions (Fig. 3 and 4). In the following, we will describe the selection of transformations and their convergence rates before applying the method in section 2 to produce the N2 and N2+OBC estimates for the ASPEP data in California.

Square-root transformation

Let y_{mj} be the number of full-time employees for the j^{th} unit within the m^{th} area. We apply transformations, such as log and powers, to the data to approximately meet the assumption of normality before fitting the model (0.3)-(0.14). The convergence is significantly faster when using the square root transformation (or $\alpha-$power with $1/3 \leq \alpha < 1$) instead of the log transformation (extremely slow). To obtain parameter estimates with a precision of $\varepsilon = 10^{-5}$, the square root transformation requires no more than 100 iterations whereas the log transformation fails to converge at 10^5 interations. The simple linear regression model parameters are used as the initial values of parameters to be estimated.

The unit-level model is given by:

$$\sqrt{y_{mj}} = \beta_0 + \boldsymbol{\beta}_1 \sqrt{x_{mj}} + u_m + \varepsilon_{mj},\ j = 1, ..., n_m,\ m = 1, ..., M, \quad (0.12)$$

$$u_m \overset{\text{iid}}{\sim} \mathcal{N}(0, \tau^2),\ \varepsilon_{mj} |\, z \overset{\text{iid}}{\sim} (1 - z)\mathcal{N}(0, \sigma_1^2) + z\mathcal{N}(0, \sigma_2^2), \quad (0.13)$$

$$z |\, \boldsymbol{\pi} \sim \text{Bin}(1; \pi),\ \sigma_2 > \sigma_1. \quad (0.14)$$

The estimated parameters β_0, β_1 and the random effect u_m are used to predict y_{mj} $(m = 1, ..., M;\ j = n_m + 1, ..., N_m)$ using the inverse trasnformation

$$\widehat{y}_{mj}^{N2} = (\widehat{\beta}_0^{N2} + \widehat{\boldsymbol{\beta}}_{\mathbf{1}}^{\boldsymbol{N2}} \sqrt{x_{mj}} + \widehat{u}_m^{N2})^2 \quad (0.15)$$

We considered the ASPEP data for California and obtained estimates for 2012 ASPEP using the 2007 ASPEP data as auxiliary information x_{mj}. The resulting estimates were compared to the corresponding true employment from each small area.

Conclusion

From our research results, the performance of N2 (or N2+OBC) produces estimates with better RRMSE. In California, the N2 (or N2+OBC) estimator is more efficient than the Horvitz Thompson (HT) estimator. In small areas, where sampling rates are less than 10%, the N2 (or N2+OBC) is slightly better than EBLUP.

Figure 1: Standard Normal $\mathcal{N}(0,1)$ and Mixture $0.75\mathcal{N}(0,1) + 0.25\mathcal{N}(0,10)$

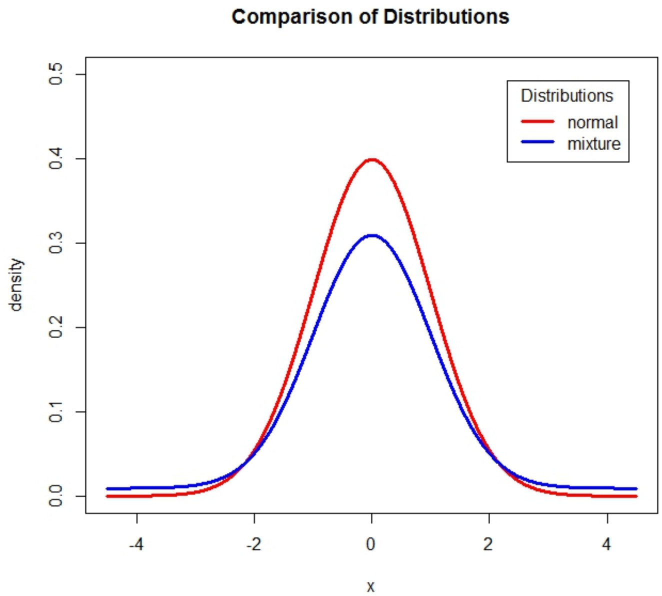

Figure 2: Skewed data (California)

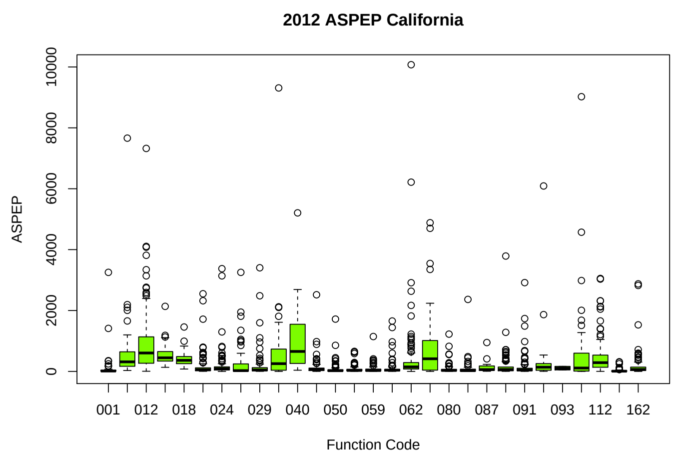

Figure 3: ASPEP after Square-root Transformed

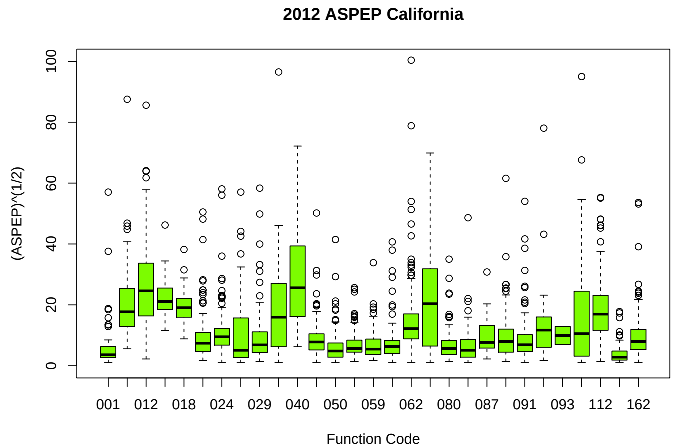

Figure 4: Normality of the Residuals of Transformed ASPEP

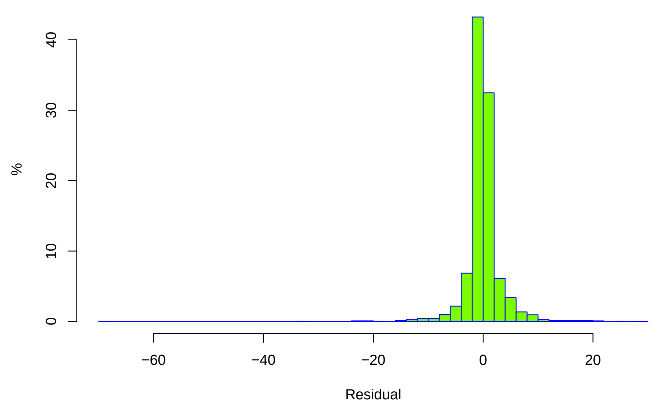

Figure 5: RRMSE's Comparison for Different Estimators for Government Functions with Sampling Rates Less than 10%

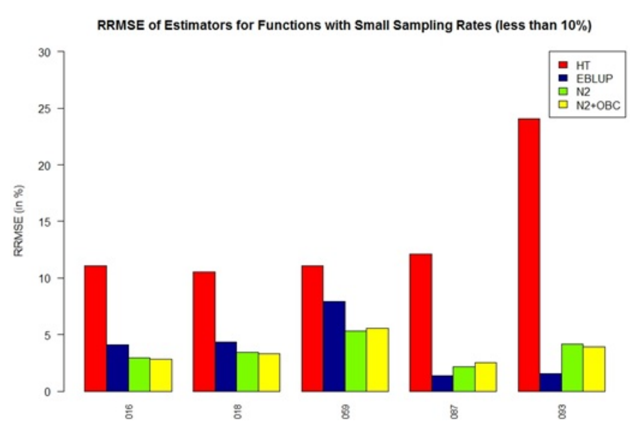

Table 1: Relative Root Mean Squared Error of Estimators for Different Areas (in California)

Function Code ⋆	N_m	med(f_m) (%)	RRMSE (%) HT	EBLUP	N2	N2+OBC
001	207	19.81	2.94	0.52	1.03	1.18
005	172	21.51	0.95	0.67	0.57	0.43
012	1,082	21.16	1.20	1.30	1.30	1.27
016	396	7.83	11.09	4.09	2.96	2.85
018	395	7.85	10.54	4.35	3.43	3.30
023	537	23.09	1.95	0.84	2.27	2.41
024	716	13.97	3.61	0.83	2.92	3.02
025	274	33.94	0.88	0.40	0.58	0.62
029	539	22.63	3.23	1.20	1.11	1.21
032	363	17.63	1.16	0.83	0.58	0.62
040	193	20.73	1.33	0.63	2.55	2.59
044	623	19.74	3.82	1.69	1.93	2.09
050	505	22.18	4.51	3.50	2.85	2.66
052	597	14.24	4.04	0.80	0.79	0.89
059	736	7.61	11.06	7.91	5.32	5.53
061	687	18.34	10.75	3.00	0.92	0.99
062	453	26.05	2.10	0.66	1.15	1.24
079	249	23.29	0.78	0.68	0.19	0.21
080	672	14.29	6.24	1.81	7.29	7.52
081	385	23.64	6.69	2.48	1.67	1.53
087	151	7.28	12.14	1.35	2.18	2.55
089	1,154	11.61	5.22	2.09	0.98	0.99
091	897	11.59	3.81	2.19	4.44	4.56
092	185	13.51	2.46	0.46	3.36	3.50
093	130	1.54	24.06	1.54	4.16	3.89
094	276	17.03	2.39	1.19	0.56	0.51
112	1,205	19.09	1.71	1.50	0.57	0.56
124	547	16.27	6.44	1.49	5.98	6.33
162	456	25.44	2.17	1.26	1.34	1.47

(⋆) function code description is in the Appendix

Appendix.

Function Code	
Air Transportation	001
Correction	005
Elementary and Secondary - Instruction	012
Higher Education - Other	016
Higher Education - Instructional	018
Financial Administration	023
Firefighters	024
Judicial & Legal	025
Other Government Administration	029
Health	032
Hospitals	040
Highways	044
Housing & Community Development	050
Libraries	052
Natural Resources	059
Parks & Recreation	061
Police Protection - Officers	062
Public Welfare	079
Sewerage	080
Solid Waste Management	081
Water Transport & Terminals	087
All Other & Unallocable	089
Water Supply	091
Electric Power	092
Gas Supply	093
Transit	094
Elementary & Secondary Schools - Other	112
Fire - Other	124
Police - Other	162

References

[1] Battese, G. E., Harter, R. M., and Fuller, W. A. (1988). An Error-components Model for Prediction of County Crop Areas Using Survey and Satellite Data, *Journal of the American Statistical Association, 83, 28-36*

[2] Bellow, M. E., and Lahiri, P.S. (2011). An Empirical Best Linear Unbiased Prediction Approach to Small-Area Estimation of Crop Parameters, *American Statistical Association - JSM 2011*

[3] Bell, W. R., and Huang, T. E. (2006). Using the t-distribution to deal with outliers in small area estimation. *Proceedings of Statistics Canada Symposium on Methodological Issues in Measuring Population Health. 2006.*

[4] Chambers, R., Chandra, H., Salvati, N., and Tzavidis, N. (2014) Outlier Robust Small Area Estimation, *Journal of Royal Statistical Society. Series B*

[5] Chandra, H. (2006). Small Area Estimation for Business Surveys, *American Statistical Association - JSM 2006.*

[6] Fay, R.E., and Herriot, R.A.(1979). Estimates of Income for Small Places: an Application of James-Stein Procedure to Census Data, *Journal of American Statistical Association, 74, 269- 277*

[7] Fellner, W. H. (1986). Robust Estimation of Variance Components, *Technometrics, 28, 51-60.*

[8] Gershunskaya, J. and Lahiri, P., (2008). Robust Estimation of Monthly Employment Growth Rates for Small Areas in the Current Employment Statistics Survey. *Proceedings of the Section on Survey Research Methods, American Statistical Association*

[9] Gershunskaya, J. (2010). Robust Small Area Estimation Using a Mixture Model. *Proceedings of the Section on Survey Research Methods, American Statistical Association – JSM 2010*

[10] Gershunskaya, J. and Lahiri, P., (2011). Robust Small Area Estimation Using a Mixture Model. *Int. Statistical Inst.: Proc. 58th World Statistical Congress, 2011, Dublin (Session IPS062)*

[11] Huber, P.J., and Ronchetti, E.M. (2009) Robust Statistics, *2nd Edition, John Wiley & Sons, Inc.*

[12] Rao, J.N.K. (2003). Small Area Estimation*, New-York, John Wiley & Sons, Inc.*

[13] Sinha, S.K., and Rao, J.N.K. (2008). Robust Methods for Small Area Estimation. *Proceedings of the American Statistical Association, Survey Research Methods Section, Alexandria, VA: American Statistical Association, 27-38.*

[14] Staudenmayer, J., Lake E.E., and Wand, M.P. (2009). Robustness for general design mixed models using the t-distribution. *Statistical Modeling 9.3 (2009): 235-255.*

[15] De Veaux, R., and Krieger, A. (1990). Robust Estimation of Normal Mixture. *Statistics & Probability Letters 10 (1990): 1-7. North-Holland.*