Small Area Estimation in Government Surveys (U.S. Census Bureau) [1]

Bac Tran
Bac.Tran@Census.gov
US Census Bureau, 4600 Silver Hill Road, Washington, DC 20233

Abstract

The Annual Survey of Public Employment & Payroll (ASPEP), conducted by the U.S. Census Bureau, provides statistics on the number of federal, state, and local government civilian employees and their gross payrolls. The universe of ASPEP is about 90,000+ state and local government units. Every five years (year ending with 2 and 7, e.g., (2007 and 2012) Census Bureau conducts a Census of Governments, Survey of Public Employment & Payroll (CoG:E). Between censuses, Census Bureau conducts the ASPEP, a nationwide sample survey covering all state and local governments in the United States. The ASPEP survey is designed to produce reliable estimates, for example, the number of full-time and part-time employees and payroll at the national level for large domains (e.g., government functions such as elementary and secondary education, higher education, police protection, financial administration, judicial and legal, etc., at the national level, and states aggregates of all function codes). However, it is also required to estimate the parameters for individual function codes within each state. This requirement prompted us to develop a methodology that employs Small Area Estimation (SAE) using unit-level covariate models in order to borrow strength from previous census data as an alternative to collecting expensive additional data for small cells. In this paper we summarize our applications of the estimators over the years for the ASPEP. The outlier treatments (Trinh & Tran, JSM 2016 & 2017) will also be discussed in this research to improve the quality of the estimates. The data we used in this research are the two CoG:E of the years 2007 and 2012.

Keywords: Governmental Units, Monte Carlo simulation, Small Area Estimation, Hierarchical Bayes, Empirical Best Linear Predictor

1. Introduction

Over the last few decades, the U.S. Census Bureau has pioneered in developing innovative small area methodologies in different programs. In one of the most cited papers in small area estimation (SAE) literature, Fay and Herriot (1979) developed a parametric empirical Bayes method to estimate per-capita income of small places with population less than 1,000 and demonstrated, using the Census data, that their method was superior to both direct design-based and synthetic methods. More recently, researchers at the U.S. Census Bureau implemented both empirical and hierarchical Bayes methodologies in the context of Small Area Income and Poverty Estimates (SAIPE) and Small Area Health Insurance Estimates (SAHIE) programs; see Bell et al. (2007) and Bauder et al. (2008).

Besides the Census Bureau's well-known SAIPE and SAHIE programs, researchers in the ESMD are actively pursuing state-of-the-art small area estimation techniques to improve the current estimation methodologies for small areas. In this paper we'd like to give overview of the estimators we used to estimate the parameters of the ASPEP over the year from different design-based estimators to Bayesian method, and lastly performed treatments to the outliers in both design-based and Bayesian approach.

[1]*Disclaimer: Any view expressed are those of the author and not necessarily those of the U.S. Census Bureau*

2. Annual Survey of Public Employment & Payroll (ASPEP)

The ASPEP population includes the civilian employees of all the Federal Government agencies (except the Central Intelligence Agency, the National Security Agency, and the Defense Intelligence Agency), all agencies of the 50 state governments, and $90{,}000^{+}$ local governments (i.e., counties, municipalities, townships, special districts, and school districts) including the District of Columbia. The survey measures the number of federal, state, and local civilian government employees and their gross payrolls for the pay period including March 12 each calendar year.

The survey provides state and local government data on full-time and part-time employment, part-time hours worked, full-time equivalent employment, and payroll statistics by governmental function (i.e., elementary and secondary education, higher education, police protection, fire protection, financial administration, central staff services, judicial and legal, highways, public welfare, solid waste management, sewerage, parks and recreation, health, hospitals, water supply, electric power, gas supply, transit, natural resources, correction, libraries, air transportation, water transport and terminals, other education, state liquor stores, social insurance administration, and housing and community development).

The survey provides Federal Government data on total employees, full-time employees, and total March payroll by governmental function. There is no detail available for part-time employment, part-time hours worked, full-time equivalent, or full-time or part-time employee payrolls. Three functions apply only to the Federal Government and have no counterpart at the state and local government levels: national defense and international relations, postal service, and space research and technology.

3. Estimators

Different estimators were used, and also researched in ASPEP from 2007 to 2017. Specifically,

a. Direct estimator- Horvitz-Thompson
b. Decision-based estimator
c. Synthetic estimator
d. Structure Preserving Estimator (SPREE)
e. Composite estimator
f. Empirical Best Linear Predictor- Unit/Area Covariates (EBLUP)
g. Benchmarking with EBLUP
h. Parametric Bootstrap Mean Square Error Estimates in Different Small Areas in ASPEP
i. Mixture models- Outliers Treatments (Design-based)
j. Bayesian approaches
    1. Bayesian version of EBLUP with types of government unit as a fixed effect
    2. Outlier treatments (t-distribution for errors term)
    3. Mixture models (two normal distributions for errors terms)

The readers can find the details of all items from a. to e. presented in JSM 2010-2013 by Tran et al; f. in JSM 2014 (Tran & Dumbacher); g. in JSM 2015 (Tran & Winters); h. in JSM 2016 (Tran); i. in JSM 2016 (Giang & Tran). j in JSM 2017 (Giang & Tran).

The results showed that EBLUP outperformed all of the estimators from a. to e. It also showed that the unit-level covariates outperformed the area-level covariates (Tran & Winters JSM 2015). Briefly, EBLUP performed very well and was implemented in production for the years of 2014-2015. The concerns on outliers were raised when there was lack of resources to do editing and imputing. The robust estimations were considered and the applications of t-distribution for errors terms and mixture of two normal distributions: both in design-based and Bayesian approaches, were studied, applied and evaluated against the EBLUP.

In this paper, we briefly review the EBLUP, discuss the Bayesian approach and then compare their performances.

3.1 EBLUP Estimators (area-level and unit-level models)

In this paper, the variable of interest is the number of full-time employees. Our data is skewed; therefore, we transformed the variable in a log scale (see Figure 2). We proposed two models: area-level model and unit-level on the auxiliary variable (see model (2) and model (5) below).

<u>Area-level Model</u>

Let y_{ij} denote the number of full-time employees for the j[th] governmental unit within the i[th] small area ($i = 1,\cdots,m;\ j = 1,\cdots,N_i$). The small area in this paper refers to the cell (state, function). In this paper, we are interested in estimating the total number of full-time employees for the i[th] small area given by $Y_i = \sum_{i=1}^{N_i} y_{ij}\ (i = 1,\cdots,m).$ An estimator of Y_i is given by:

$$\hat{Y}_i^{EB} = N_i \left[f_i \bar{y}_i + (1 - f_i)\hat{\bar{Y}}_{ir} \right] \tag{1}$$

where $\bar{y}_i = n_i^{-1} \sum_{j=1}^{n_i} y_{ij}$ is the sample mean; $f_i = n_i / N_i$, N_i and n_i are the sampling fraction, number of government units in the population and sample for area *i*, respectively; $\hat{\bar{Y}}_{ir}$ is a model-dependent predictor of the mean of the non-sampled part of area *i* $(i = 1,\cdots,m)$.

In this paper, we obtain $\hat{\bar{Y}}_{ir}$ using the following nested error regression model on the logarithm of the number of full-time employees at the government unit level:

$$\log(y_{ij}) = \beta_0 + \beta_1 \log(\bar{X}_i) + v_i + \varepsilon_{ij}, \tag{2}$$

$$v_i \overset{iid}{\sim} N(0,\tau^2) \text{ and } \varepsilon_{ij} \overset{iid}{\sim} N(0,\sigma^2), \tag{3}$$

where $\bar{X}_i$ is the average number of full-time employees for the *i*[th] small area obtained from the previous Census; β_0 and β_1 are unknown intercept and slope, respectively; v_i are small area specific random effects. The distribution of the random effects describes deviations of the area means from values $\beta_0 + \beta_1 \log(\bar{X}_i)$; ε_{ij} are errors in individual observations $(j = 1,...,N_i;\ i = 1,...,m)$. The random variables v_i and ε_{ij} are assumed to be mutually independent. We assume that sampling is non-informative for the distribution of measurements y_{ij} $(j = 1,...,N_i;\ i = 1,...,m)$. A similar model without logarithmic transformation can be found in Battese et al. (1988). The logarithmic transformation is taken to reduce the extent of heteroscedasticity in the employment data. Similar model using unit level auxiliary information was considered by Bellow and Lahiri [5] in the context of estimating total hectare under corn for U.S. counties. We use the following model-based predictor of $\bar{Y}_{ir}$:

$$\hat{\bar{Y}}_{ir} \approx \exp\left[\hat{\beta}_0 + \hat{\beta}_1 \log(\bar{X}_i) + \hat{v}_i + \frac{1}{2}(\hat{\sigma}^2 + \hat{\delta}_i^2) \right] \tag{4}$$

where $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{v}_i, \hat{\sigma}^2$, *and* $\hat{\delta}_i^2$ (standard error of $\hat{v}_i$) are obtained by fitting (2) using PROC MIXED of SAS. We obtain our estimate of total number of full-time employees in area *i* using equations (1) and (4).

Unit-level Model

Besides area-level (model 2), we also performed the unit-level (X_{ij}) model as below.

$$\log(y_{ij}) = \beta_0 + \beta_1 \log(X_{ij}) + v_i + \varepsilon_{ij}, \tag{5}$$

$$v_i \overset{iid}{\sim} N(0, \tau^2) \text{ and } \varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2), \tag{6}$$

After estimating the models parameters, the estimate will be obtained by two different ways: simple back transformed, and log-normal back transformed given as follows:

Simple Back Transformation

$\hat{Y}_i^{EB} = \sum_{i \in S} y_i + \sum_{i \notin S} \exp(\hat{\beta}_0 + \hat{\beta}_1 \log(X_{ij}) + \hat{v}_i)$ (simple)

Log-Normal Back Transformation

$\hat{Y}_i^{EB} = N_i(\mathrm{f_i}\bar{\mathrm{y}}_\mathrm{i} + (1 - \mathrm{f_i})\hat{\bar{Y}}_{ir}$, where

$\hat{\bar{Y}}_{ir} = \hat{\alpha}_{ir}\exp(\hat{v}_i + \frac{1}{2}(\hat{\sigma}^2 + \hat{\delta}_i^2))$, and $\hat{\alpha}_{ir} = (N_i - n_i)^{-1} \sum_{j \notin S_i} \exp(\hat{\beta}_0 + \hat{\beta}_1 \log(\mathrm{X_{ij}}))$

3.2 Robust Estimation

Hierarchical Bayes with t-distribution and types of government unit as fixed effect

$\boldsymbol{Model}$: $(y_{ijk}) = \beta_0 + \beta_1 x_{ijk} + \alpha_j + u_i + e_{ijk}$
y_{ijk}, x_{ijk} *are the number of full time employees from survey data and census year in* log *scale, respectively, where* $i =$ *area* i *(function code),* $j =$ *type of government,* $k =$ *unit* k^{th},
$i = 1,2,..,29$ *areas,* $j = 1,2,3,4$ *type of government,* $k = 1,2,..,N_i$

α_j *is fixed effect of government type* j^{th}
u_i *is random effect of function code* i^{th}

$e_{ijk} | \sigma_e^2 \overset{iid}{\sim} N(0, \sigma_e^2)$
$u_i | \sigma_u^2 \overset{iid}{\sim} N(0, \sigma_u^2)$
$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 50 & 0 \\ 0 & 50 \end{bmatrix}\right)$
$\sigma_u^2 \sim inverse\ gamma(0.01, 0.01)$
$\sigma_e^2 \sim inverse\ gamma(0.01, 0.01)$
$flat\ prior\ on\ \alpha_j$
$\log(y_{ijk}) | u_i, \alpha_j \sim t(mean = \beta_0 + \beta_1 \log(x_{ijk}) + \alpha_j + u_i, \sigma_e^2, df = 4)$

Mixture of two normal distributions for errors term (Design-based)

From hereafter for simplicity the index for state is dropped.
$\boldsymbol{i = area\ i\ (function\ code)\ and\ j\ is\ the\ unit\ j^{th}}$.

This model was proposed by Gershunskaya and Lahiri [8]. It is also call **N2** estimator. The model was specified as below.

$\boldsymbol{Model}$: $y_{ij} = x_{ij}^T \beta + u_i + \epsilon_{ij}$

$u_i \overset{iid}{\sim} N(0, \tau^2), \epsilon_{ij} | z \overset{iid}{\sim} (1 - z) N(0, \sigma_1^2) + z N(0, \sigma_2^2),$
$z | \pi \sim Bin(1, \pi)$
$u_i\ and\ \epsilon_{ij}\ are\ mutually\ independent$

$The\ parameter\ \theta = (\sigma_1, \sigma_2, \tau, \pi, \beta) is\ estimated\ by\ an\ EM\ algorithm$
$(See\ Giang\ \&\ Tran\ JSM\ 2016)$

Mixture of two normal distributions for errors term (Hierarchical Bayes)

$y_{ij} = x_{ij}^T \beta + u_i + \epsilon_{ij}$

$u_i | \sigma_u^2 \overset{iid}{\sim} N(0, \sigma_u^2)$

$z | \pi \sim Bin(1, \pi)$

$\pi = \frac{1}{(1 + e^{-z})}$
$\epsilon_{ij} | z, \sigma_1^2, \sigma_2^2, z \overset{iid}{\sim} (1 - z) N(0, \sigma_1^2) + z N(0, \sigma_2^2)$
$\sigma_u^2 \sim inverse\ gamma$
$\sigma_1^2 \sim inverse\ gamma$
$\sigma_2^2 \sim inverse\ gamma$

$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 50 & 0 \\ 0 & 50 \end{bmatrix}\right)$

$likelihood$: $y_{ij} | \sigma_1^2,\ \sigma_2^2,\ u_i \sim \pi * N(0, \sigma_1^2) + (1 - \pi) * N(0, \sigma_2^2)$
$(See\ Giang\ \&\ Tran\ JSM\ 2017)$

4. Results

In this paper, the universe is the intersection of the two census data, 2007 and 2012, i.e., government units that overlap between the 2007/2012 Censuses of Governments: Employment reporting strictly positive numbers of full-time employees. We developed a design-based Monte Carlo simulation experiment in which we draw repeated samples (1,000 of them) from the universe using the ASPEP sampling design. In each replicate we performed estimations in 3.1 and 3.2 which produced the estimates of full-time employees for state and local that contained 29 small areas (functions, see Appendix). The average of the RRMSEs from 1,000 replicates was compared with simulated true RRMSE for each small area. For simplicity, we showed the results for the biggest state and local data for California. Table 1 shows the relative root mean squared errors (RRMSE) of six different estimators: EBLUP, EBLUP with fixed effect (government type), hierarchical Bayes with t-distributed errors, hierarchical Bayes with t-distribution for errors terms and government type as fixed effect, mixture models with two normal distributions for errors terms (design-based-N2Design), and hierarchical Bayes with mixture of two normal distributed errors. Table 2 shows the number of times an estimator performs the best among rival estimators in terms of RRMSE. Table 3 shows the average sample sizes and average sampling rates in different small areas (functions).

Table 1: Relative Root Mean Squared Errors (RRMSE) of Six Different Estimators

Function	EBLUP	EBFixedType	HB	HBFixedType	N2Design	HBMixture
001	0.52%	0.44%	0.52%	0.61%	1.03%	0.29%
005	0.67%	0.64%	0.52%	0.48%	0.57%	0.57%
012	1.30%	1.67%	1.29%	0.98%	1.30%	1.65%
016	4.09%	4.13%	2.79%	2.23%	2.96%	2.53%
018	4.35%	3.19%	3.50%	2.76%	3.43%	2.50%
023	0.84%	0.81%	0.68%	0.68%	2.27%	0.97%
024	0.83%	2.80%	0.72%	0.55%	2.92%	1.20%
025	0.40%	0.47%	0.42%	0.46%	0.58%	0.58%
029	1.20%	1.66%	1.14%	1.07%	1.11%	1.49%
032	0.83%	0.69%	0.69%	0.57%	0.58%	0.35%
040	0.63%	0.83%	0.42%	0.43%	2.55%	0.78%
044	1.69%	1.09%	0.94%	0.91%	1.93%	0.42%
050	3.50%	0.94%	1.94%	1.37%	2.85%	0.69%
052	0.80%	0.51%	0.64%	0.62%	0.79%	0.93%
059	7.91%	2.51%	3.41%	1.98%	5.32%	3.92%
061	3.00%	0.81%	1.71%	1.54%	0.92%	1.55%
062	0.66%	2.71%	0.33%	0.45%	1.15%	0.39%
079	0.68%	0.72%	0.29%	0.28%	0.19%	0.25%
080	1.81%	2.02%	1.13%	1.36%	7.29%	1.25%
081	2.48%	1.67%	1.83%	1.73%	1.67%	1.35%
087	1.35%	1.39%	1.19%	1.03%	2.18%	1.18%
089	2.09%	2.33%	2.67%	2.36%	0.98%	1.98%
091	2.19%	2.96%	1.25%	0.95%	4.44%	0.38%
092	0.46%	0.60%	0.43%	0.45%	3.36%	0.51%
093	1.54%	1.65%	1.70%	1.76%	4.16%	1.80%
094	1.19%	1.19%	0.92%	0.86%	0.56%	1.07%
112	1.50%	1.17%	1.36%	0.99%	0.57%	0.23%
124	1.49%	4.15%	2.03%	2.43%	5.98%	2.74%
162	1.26%	1.31%	0.68%	0.73%	1.34%	0.50%

Table 2: Number of Times an Estimator Perform the Best Among Rival Estimators in terms of RRMSE

Function	flg_EBLUP	flg_EBFixedType	flg_HB	flg_HBFixedType	flg_N2Design	flg_HBMixture
001	.	.	.	.	.	1
005	.	.	.	1	.	.
012	.	.	.	1	.	.
016	.	.	.	1	.	.
018	.	.	.	.	.	1
023	.	.	1	.	.	.
024	.	.	.	1	.	.
025	1	.	.	.	.	.
029	.	.	.	1	.	.
032	.	.	.	.	.	1
040	.	.	1	.	.	.
044	.	.	.	.	.	1
050	.	.	.	.	.	1
052	.	1	.	.	.	.
059	.	.	.	1	.	.
061	.	1	.	.	.	.
062	.	.	1	.	.	.
079	.	.	.	.	1	.
080	.	.	1	.	.	.
081	.	.	.	.	.	1
087	.	.	.	1	.	.
089	.	.	.	.	1	.
091	.	.	.	.	.	1
092	.	.	1	.	.	.
093	1	.	.	.	.	.
094	.	.	.	.	1	.
112	.	.	.	.	.	1
124	1	.	.	.	.	.
162	.	.	.	.	.	1
	3	**2**	**5**	**7**	**3**	**9**

Table 3: Average Sample Size and Average Sampling Rate

Function	N	AVG(n)	median_n	n_positive	med(n_positive)	Sampling Rate
001	207	61.09	61	40.72	41	19.81%
005	172	55.04	55	36.97	37	21.51%
012	1082	233.78	234	229.04	229	21.16%
016	396	109.63	110	31.33	31	7.83%
018	395	109.62	110	31.32	31	7.85%
023	537	124.57	125	123.97	124	23.09%
024	716	115.58	116	100.54	100	13.97%
025	274	101.97	103	92.58	93	33.94%
029	539	124.58	125	122.52	122	22.63%
032	363	84.24	85	64.21	64	17.63%
040	193	61.94	62	40.41	40	20.73%
044	623	127.85	128	123.30	123	19.74%
050	505	122.49	123	111.56	112	22.18%
052	597	157.6	158	84.71	85	14.24%
059	736	79.88	80	55.81	56	7.61%
061	687	133.03	133	125.95	126	18.34%
062	453	121.87	122	117.88	118	26.05%
079	249	76.26	77	57.78	58	23.29%
080	672	113.09	113	95.58	96	14.29%
081	385	105.71	106	91.15	91	23.64%
087	151	35.57	36	11.24	11	7.28%
089	1154	144.79	145	133.70	134	11.61%
091	897	117.83	118	103.52	104	11.59%
092	185	47.22	47	24.94	25	13.51%
093	130	25.16	25	2.08	2	1.54%
094	276	68.63	69	46.97	47	17.03%
112	1205	237.27	237	229.57	230	19.09%
124	547	107.38	108	88.98	89	16.27%
162	456	120.51	121	115.35	116	25.44%

5. Conclusion

As we can see hierarchical Bayes model assuming t-distributed errors with fixed effect as government type (7 times out of 29) and hierarchical Bayes assuming Mixture of Normal distributed errors (9 times out of 29) perform better than the other estimators. In practice, we will create a 'hybrid' estimator, which is a combination of all the estimators where they perform better than the other ones.

Function	Description
000	Totals for Government
001	Airports
002	Space Research & Technology (Federal)
005	Correction
006	Nat Defense & International Relations (Federal)
012	Elementary and Secondary - Instruction
014	Postal Service (Fed)
016	Higher Education - Other
018	Higher Education - Instructional
021	Other Education (State)
022	Social Insurance Administration (State)
023	Financial Administration
024	Firefighters
025	Judicial & Legal
029	Other Government Administration
032	Health
040	Hospitals
044	Streets & Highways
050	Housing & Community Development (Local)
052	Local Libraries
059	Natural Resources
061	Parks & Recreation
062	Police Protection - Officers
079	Welfare
080	Sewerage
081	Solid Waste Management
087	Water Transport & Terminals
089	Other & Unallocable
090	Liquor Stores (State)
091	Water Supply
092	Electric Power
093	Gas Supply
094	Transit
112	Elementary and Secondary - Other Total
124	Fire - Other
162	Police-Other

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, 28-362.

[2] 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.

[3] Bauder, M., Riesz, S., and Luery, D. (2008) “Further Developments in a Hierarchical Bayes Approach to Small Area Estimation of Health Insurance Coverage: State-Level Estimates for Demographic Groups,” Proceedings of the Section on Survey Research Methods, Alexandria, VA: American Statistical Association, 1726-1733.

[4] Bell, W., Basel, W., Cruse, C, Dalzell, L., Maples, J., O’Hara, B., and Powers, D. (2007) ''Use of ACS Data to Produce SAIPE Model-Based Estimates of Poverty for Counties," Census Report.

[5] Bellow, M. and Lahiri, P. (2010) "Empirical Bayes Methodology for the NASS County Estimation Program," *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

[6] Chakraborty, A., Datta, S. and Mandal, A. (2016) "A Two-component Normal Mixture Alternative to the Fay-Herriot Model." Statistics in Transition new series 17.1 (2016): 67-90.

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

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

[9] 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)

[10] Lange, K.L., Little, R.J.A., and Taylor, J.M.G. (1989). "Robust Statistical Modeling Using the t-Distribution. Journal of the American Statistical Association," 84(408), 881-896.

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

[12] Stefanski, L. (1991), "A Normal Scale Mixture Representation of the Logistic Distribution. Statistics and Probability Letters," Vol 11, Issue 1, Pages 69-70. Elsevier

[13] 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.

[14] Trinh, G. and Tran, B. (2017). ""Robust Estimation for the Annual Survey of Public Employment & Payroll Using Mixture of Linear Mixed-effects Models with the MCMC Procedure." American Statistical Association, JSM2017

[15] Trinh, G. and Tran, B. (2016). "Outliers in Annual Survey of Public Employment & Payroll Small-area Approach," American Statistical Association, JSM2016 Conference Proceedings.

[16] Tran, B. and Winters, F (2015). "An Evaluation of Different Small Area Estimators and Benchmarking for the Annual Survey of Public Employment & Payroll." American Statistical Association, JSM2015 Conference Proceedings.

[17] Tran, B. and Dumbacher, B. (2014) "An Evaluation of Different Small Area Estimators for the Annual Survey of Public Employment and Payroll," American Statistical Association, JSM2014 Conference Proceedings.

[18] Tran, B. and Hogue, Carma (2013). "Small Area Estimation for Government Surveys." American Statistical Association, JSM2013 Conference Proceedings.

[19] Tran, B. and Yang, C. (2011). "Application of Small Area Estimation for Annual Survey of Employment and Payroll," American Statistical Association, JSM2011 Conference Proceedings.