An Evaluation of Different Small Area Estimators for the Annual Survey of Public Employment and Payroll

Bac Tran[1], Brian Dumbacher[1]
[1]U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233
Bac.Tran@census.gov, Brian.Dumbacher@census.gov

Abstract

The Governments Division of the U.S. Census Bureau employs small area estimation techniques for the Annual Survey of Public Employment and Payroll (ASPEP). ASPEP provides statistics on the number of federal, state, and local government civilian employees and their gross payrolls. Different small area estimators can be produced using the ASPEP data and auxiliary information from the preceding Census of Governments. We develop a design-based Monte Carlo simulation experiment in which we draw repeated samples from the 2007 Census of Governments data using the ASPEP sample design. We compute a wide range of estimates that use the generated sample and the 2002 Census of Governments data. We then compare simulated design-based biases, variances, mean squared errors and coverage probabilities of these estimators. We repeat the experiment using the 2012 Census of Governments data in order to understand if these properties change over years. The estimators covered under our simulation study include Horvitz-Thompson, Structure PREserving Estimation (SPREE), traditional composite and empirical Bayes methods.

Key Words: Government Units; Monte Carlo Simulation; Composite Estimator; Horvitz-Thompson; Empirical Bayes; SPREE

1. Introduction

Over the last few decades, the U.S. Census Bureau has pioneered 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 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 Governments Division are actively pursuing state-of-the-art small area estimation

---

This report is released to inform interested parties of research and to encourage discussion. Any views expressed on statistical, methodological, technological, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

techniques to improve the current estimation methodologies for small areas. Some results on the ASPEP estimation were presented at the 2013 SAE conference in Thailand. There are a large number of small area estimators available in the literature. These estimators typically use either implicit or explicit models to combine survey data with different administrative and Census records. The properties of such estimators are usually studied using the model used to derive the estimator. However, the design-based properties of small area estimators, which are most appealing to survey practitioners, are largely unknown. In this paper, we show different small area estimators that can be produced using the ASPEP data and auxiliary information from the preceding Census of Governments. We develop a design-based Monte Carlo simulation experiment in which we draw repeated samples from the 2007 Census of Governments data using the ASPEP sampling design and compute a wide range of estimates that use the generated sample and the 2002 Census of Governments data. We then compare simulated design-based biases, variances, and mean squared errors of these estimators. We repeat the experiment using the 2012 Census of Governments data in order to understand if these properties change over years. The estimators covered under our simulation study include Horvitz-Thompson, SPREE, traditional composite, and empirical Bayes.

The Governments Division of the U.S. Census Bureau conducts a census 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 Governments Division also conducts the Annual Survey of Public Employment and 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 conduct multiple activities. Activities in ASPEP are designated by function codes (see Appendix). School districts cover only education functions. Special districts usually perform only one function, but can perform two or three functions, like natural resources or sewerage. ASPEP is the only source of public employment data by program function and full-time and part-time employment. 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, see http://www.census.gov/govs/apes .

In 2009, ASPEP was redesigned and the old sample design was replaced by a stratified probability proportional-to-size (PPS) with modified cut-off sample design in order to reduce sample size and respondent burden for small townships and special districts. At the same time, the goal was to improve the precision of the estimates and data quality. The sample design was implemented in multiple steps. First, large governments were made initial certainties. Next, in the first phase of the design, a state-by-governmental type stratified PPS sample was selected, where size was taken as the total payroll (the sum of full-time pay and part-time pay) from the employment portion of the 2007 Census of Governments. In the second phase, a cut-off point was constructed to distinguish small and large government units in city, township, and special district strata. Lastly, the strata with small-size government units were subsampled using a simple random sampling design. In 2009, we selected 1,200 out of 2,000 small-size units.

Five years later in 2014, a new sample for ASPEP was selected based on the 2012 Census of Governments. The sample design was changed slightly. Initial certainty criteria were not used, and more sample was allocated to school district strata of small states. Systematic PPS sampling was performed in all strata after sorting by population (for general-purpose governments), enrollment (for school districts), and function (for special districts). In the second phase of the design, we again used modified cut-off sampling to select a subsample of small-size units.

The ASPEP is designed to produce reliable estimates of the number of full-time and part-time employees and payroll at the national level and for large domains (e.g., government functions such as elementary and secondary education, higher education, police protection, fire protection, financial administration, judicial and legal, etc., at the national level, and state aggregates of all function codes). However, it is also required to estimate the parameters for individual function codes within each state. This requirement leads us to explore small area estimation methodology that borrows strength from previous Census data instead of collecting expensive additional data for small cells. We refer to Rao (2003) and Jiang and Lahiri (2006) for a comprehensive account of small area estimation theory and applications. In Section 2, we briefly describe our method. In Section 3, we present our findings from our data analysis.

2. Estimation Methods

2.1 Empirical Bayes Models

In this paper, the variable of interest is the number of full-time employees. Our data are skewed; therefore, we transformed the variable on a logarithmic scale (see Figure 2). We propose two unit-level empirical Bayes models: a model with an area-level covariate and a model with a unit-level covariate, where the covariates are based on an auxiliary variable from the previous Census [see model (2) and model (5) below].

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 areas in this paper refer to the cells (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 $Y_i = \sum_{j=1}^{N_i} y_{ij}\ (i = 1, \cdots, m)$. An estimator of Y_i is given by:

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

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 number of units in the 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:

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

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

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 the logarithmic transformation can be found in Battese et al. (1988). The logarithmic transformation is used to reduce heteroscedasticity in the employment data. A similar model using unit-level auxiliary information was considered by Bellow and Lahiri (2012) in the context of estimating total hectare under corn for U.S. counties. We use the following model-based predictor of $\bar{Y}_{ir}$:

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

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

Besides the area-level covariate model [model (2)], we also consider the following:

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

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

where X_{ij} is the number of full-time employees from the previous Census for unit *j* in small area *i*.

2.2 Other Methods

2.2.1 Direct Horvitz-Thompson Estimator

The Horvitz-Thompson (HT) estimator for estimating the total in small area *i* is:

$$\hat{t}_i = \sum_{j \in S} w_{ij} y_{ij}\ , \quad (7)$$

where $w_{ij} = 1/\pi_{ij}$ is the weight and π_{ij} is the inclusion probability for unit *j* in small area *i*.

2.2.2 Decision-Based Estimator

The decision-based (DB) method is used to calculate the synthetic estimate in each cell by providing a stable state total as a reliable estimator in a large area covering all small areas, states by type of government and by function code for special districts. In other

words, it is used for estimating the aggregates. DB is a process of testing the possibility of combining the strata in order to get a better estimate of the total. This method strengthens the statistical models for the area of estimation. The state total is estimated by a single stratum weighted regression (GREG) estimator specified below:

$$\hat{t}_{y,GREG} = \hat{t}_{y,\pi} + \hat{b}(t_x - \hat{t}_{x,\pi}), \tag{8}$$

where $t_x = \sum_{i \in U} x_i$, $\hat{t}_{x,\pi} = \sum_{i \in S} \frac{x_i}{\pi_i}$, $\hat{t}_{y,\pi} = \sum_{i \in S} \frac{y_i}{\pi_i}$, $\hat{b} = \frac{\sum_{i \in S} (x_i - \bar{x})(y_i - \bar{y})/\pi_i}{\sum_{i \in S} (x_i - \bar{x})^2/\pi_i}$,

π_i is the inclusion probability, and x_i is the auxiliary data from the Employment portion of the Census of Governments for government unit *i*. The slope $\hat{b}$ was obtained by the DB process (Cheng et al., 2009). The DB method improved the precision of estimates and reduced the mean square error of weighted survey total estimates. The idea is to test the equality of linear regression lines to determine whether we can combine data in different substrata. The null hypothesis is $H_0 : b_1 = b_2$, that is, the equality of the frame population regression slopes for two substrata. In large samples, $\hat{b}$ is approximately normally distributed, $\hat{b} \sim N(b, \Sigma)$. Under the null hypothesis, with two sub-strata U_1 and U_2 (large and small) and samples S_1 and S_2 of sizes n_1 and n_2, respectively, we have $\hat{b}_1 - \hat{b}_2 \sim N(0, \Sigma_{1,2})$, where $\hat{b}_1 \sim N(b, \Sigma_1)$, $\hat{b}_2 \sim N(b, \Sigma_2)$, and $\Sigma_{1,2} = \Sigma_1 + \Sigma_2$. Therefore, the test statistic is

$$(\hat{b}_1 - \hat{b}_2)\Sigma_{1,2}^{-1}(\hat{b}_1 - \hat{b}_2) \sim \chi_1^2. \tag{9}$$

Our research showed that it was unnecessary to test the hypothesis for the intercept equality because our data analysis showed that we never rejected the null hypothesis of equality of intercepts when we could not reject the null hypothesis of equality of slopes.

The critical value for a test based on (9) was obtained from a chi-squared distribution with one degree of freedom. The test was performed with a significance level of α = 0.05. If we could not reject the null hypothesis, then the slopes estimated in the sub-strata using S_1 and S_2 were accepted as the same, and the DB estimator is equal to the GREG estimator for the union of two sample sets, that is, for $S = S_1 \cup S_2$. Otherwise, the DB estimator is the sum of two separate GREG estimators of stratum totals. That is,

$$\hat{t}_{y,DB} = \begin{cases} \hat{t}_{y,greg} & \text{if } H_0 \text{ is not rejected} \\ \sum_{h=1}^{2} \hat{t}_{y,greg}^{h} & \text{if } H_0 \text{ is rejected,} \end{cases} \tag{10}$$

where $\hat{t}_{y,greg}$ denotes the GREG estimator from the combined stratum from sample *S* and $\hat{t}_{y,greg}^{h}$ denotes the GREG estimator from substratum *h* from sample S_h. DB produces 51 (50 states and Washington, DC) totals for each key variable.

2.2.3 Synthetic Estimator

Synthetic estimation assumes small areas have the same characteristics as large areas and that there is a reliable estimate for large areas. There are many advantages of synthetic estimates. They are accurate, simple, intuitive, aggregated estimates that can be applied to all sample designs and borrow strength from similar small areas. Synthetic estimation can even provide estimates for areas with no sample from the sample survey and does not need a study model.

The general idea for synthetic estimation is that if we have a reliable estimate for a large area and this large area covers many small areas, then we can use this estimate to produce an estimate for a small area. The key element for calculating the synthetic estimate for a small area (state by function code) is to estimate the proportion of that small area of interest within the large state area. This estimate for the small area is known as the synthetic estimate.

The synthetic estimator for function *f* of state *i* is:

$$\hat{t}_{if}^{syn} = \frac{x_{if}}{\sum_{f} x_{if}} \hat{t}_{i}^{DB} \tag{11}$$

where x_{if} is the total of the auxiliary variable obtained from the Employment component of the Census of Governments. In our research x_{if} is the number of full-time employees in the previous Census in state *i* and function *f*. Also, $\hat{t}_{i}^{DB}$ is the DB estimate of the total for state *i* from equation (10).

2.2.4 Composite Estimator

To balance the potential bias of the synthetic estimator, $\hat{t}_{if}^{syn}$, against the instability of the design-based direct estimate, $\hat{t}_{if}^{HT}$, we take a weighted average of the two estimators. This composite estimation methodology was applied on the PPS sample for each cell (state by function). Generally, it has the form below:

$$\hat{t}_{if}^{composite} = \hat{\phi}_i \hat{t}_{if}^{HT} + (1 - \hat{\phi}_i)\, \hat{t}_{if}^{syn} \tag{12}$$

$$\text{where } \hat{\phi}_i = 1 - \frac{\sum_f \widehat{var}(\hat{t}_{if}^{HT})}{\sum_f (\hat{t}_{if}^{syn} - \hat{t}_{if}^{HT})^2} \quad \text{(see Rao 4.4.3)} \tag{13}$$

This estimator could give negative weight. In that case we set $\hat{\phi}_i = 0.5$.

2.2.5 Structure Preserving Estimation

Structure PREserving Estimation (SPREE) is a synthetic estimation method that uses the method of iterative proportional fitting to adjust the cell counts of a multi-way table in such a way that the adjusted counts agree with the specified margins. For detailed procedures, please see Rao (2003). In our research, we construct a two-way table (state by function with a dimension of 49 by 29) of full-time employee counts from Census of Governments data. Between two Censuses we collect survey data. Therefore, the margins are updated by SPREE. The margins could be obtained by reliable direct survey

estimates, or by the decision-based total estimates. In our research we use direct survey estimates.

3. Analysis and Results

We include three Censuses of Governments in this study: 2002, 2007, and 2012, from which we create two universes: 2002 data intersected with 2007 data and 2007 intersected with 2012 data. For simplicity, we make the intersections contain positive values of the variable of interest (full-time employees) so we can apply the logarithmic transformation. We apply production sample designs to select 200 replicated samples from each Census of Governments data, 2007 and 2012. With each replicate we estimate the full-time employee totals for states and functions (49 states and 29 functions) using the estimators: HT, SPREE, empirical Bayes with an area-level covariate [model (2)], empirical Bayes with a unit-level covariate [model (5)], and composite. We compute $\widehat{mse}$, $\widehat{var}$, and $\widehat{bias}$ for each estimator using the 200 replicates. The analysis covers 49 states and excludes Washington, DC and Hawaii because we collect data from all of their governments.

Figure 1 and Figure 2 show the data by function code for California before and after the log transformation, respectively.

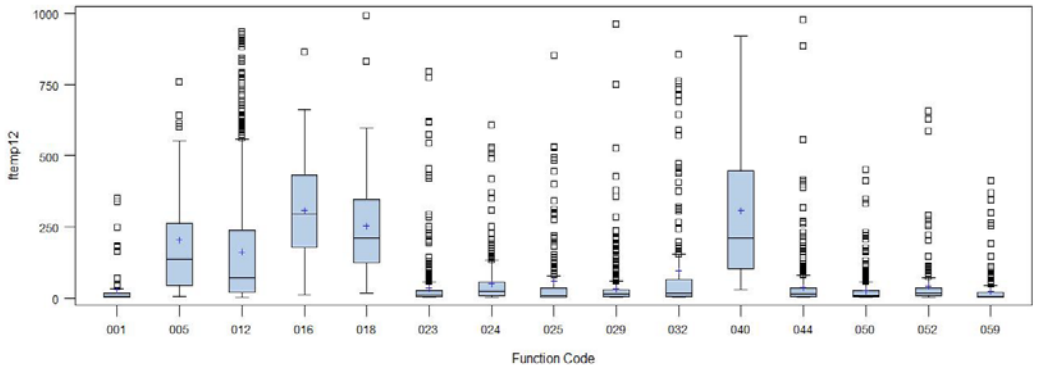

Figure 1: The Data before Log Transformation – California
Source: U.S. Census Bureau, 2012 Census of Governments: Employment

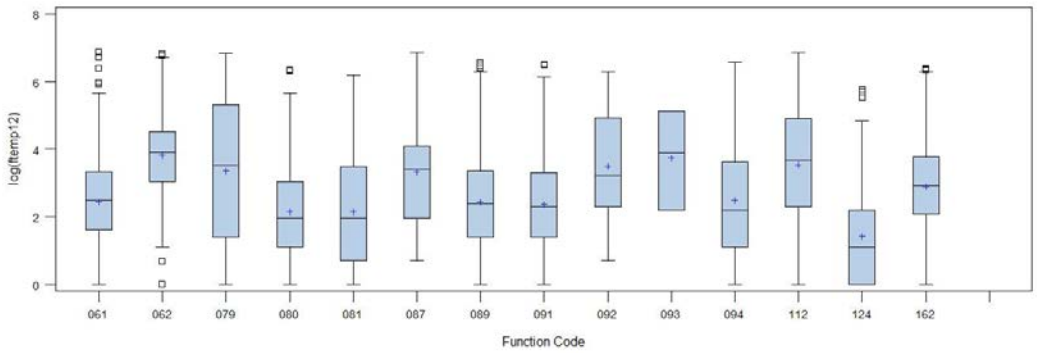

Figure 2: The Data after Log Transformation – California
Source: U.S. Census Bureau, 2012 Census of Governments: Employment

Figure 3 shows the distribution of the residuals after the log transformation. As we can see, the normality assumption in the model is satisfied very well.

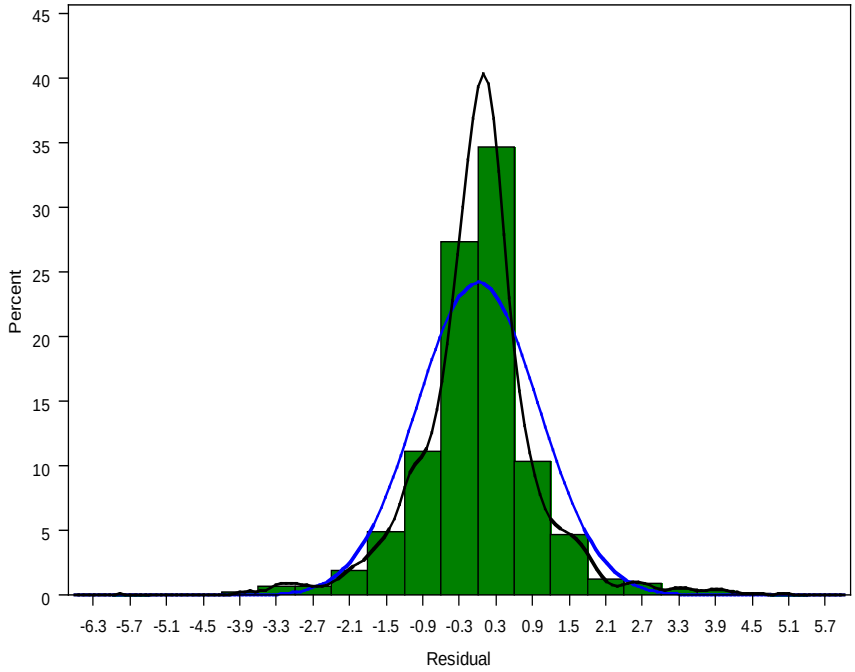

Figure 3: Normality of the Residuals after Log Transformation – California
Source: U.S. Census Bureau, 2012 Census of Governments: Employment

Table 1 and Table 2 show the comparisons of the $\widehat{mse}$ and $\widehat{bias}$ (actually, estimated relative bias) of five estimators for 2002/2007 and 2007/2012, respectively. The figures in the table show the number of times an estimator outperforms the others, i.e., smaller $\widehat{mse}$, and smaller $\widehat{bias}$.

Table 1: Comparisons for 2002-2007 Census of Governments Data (1,222 cells = state by function)

$\widehat{mse}$					$\widehat{bias}$				
Number of times the model outperforms the others (the larger the better)					*Number of times the model has larger bias vs the others (the smaller the better)*				
EB-Unit	EB-Area	Composite	HT	SPREE	EB-Unit	HT	EB-Area	Composite	SPREE
507	317	273	66	59	33	75	79	266	769

Source: U.S. Census Bureau, 2002 and 2007 Censuses of Governments: Employment

Table 2: Comparisons for 2007-2012 Census of Governments Data (1,231 cells = state by function)

$\widehat{mse}$					$\widehat{bias}$				
Number of times the model outperforms the others (the larger the better)					*Number of times the model has larger bias vs the others (the smaller the better)*				
EB-Unit	EB-Area	Composite	SPREE	HT	EB-Unit	HT	EB-Area	Composite	SPREE
704	306	161	45	15	26	48	83	165	909

Source: U.S. Census Bureau, 2007 and 2012 Censuses of Governments: Employment

As we can see, when using the mean square error or the bias, the empirical Bayes model with the unit-level covariate [model (5)] outperforms the other four estimators. This outperformance is consistent across data from three Censuses of Governments. The research in this paper uses real production sample designs for ASPEP. In our production, we will use a mixture of estimators where they show they perform best.

4. Conclusion

The HT estimator performed poorly in cells where sample sizes are relatively small. Our proposed estimator [model (5)] dominantly outperformed the other four estimators on small areas and on some large areas as well. Thanks to this research, in production we will select a mixture of the estimators where they perform best.

References

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.

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.

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.

Bellow, M., and Lahiri, P. (2010) "Empirical Bayes Methodology for the NASS County Estimation Program," *Proceedings of the Section on Survey Research Methods*, Alexandria, VA: American Statistical Association, 343-355.

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, 269-277.

Jiang, J., and Lahiri, P. (2006) "Mixed Model Prediction and Small Area Estimation (with discussions)," *Test*, 15, 1-96.

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

Shen, W., and Louis, T. (1998) "Triple-Goal Estimates in Two-Stage Hierarchical Models," *Journal of the Royal Statistical Society*, Series B, 60, 455–471.

Appendix

Function	Description
000	Total
001	Airports
002	Space Research & Technology (Federal)
005	Corrections
006	National Defense & International Relations (Federal)
012	Elementary and Secondary Education - Instruction
014	Postal Service (Federal)
016	Higher Education - Other
018	Higher Education - Instructional
021	Other Education (State)
022	Social Insurance Administration (State)
023	Financial Administration
024	Fire Protection - Firefighters
025	Judicial & Legal
029	Other Government Administration
032	Health
040	Hospitals
044	Streets & Highways
050	Housing & Community Development (Local)
052	Libraries
059	Natural Resources
061	Parks & Recreation
062	Police Protection - Officers
079	Public 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 Education - Other
124	Fire Protection - Other
162	Police Protection - Other