A MODEL-BASED APPROACH: COMPOSITE ESTIMATORS FOR SMALL AREA ESTIMATION

Kung-Jong Lui, Centers for Disease Control
W. G. Cumberland, University of California, Los Angeles
Kung-Jong Lui, 1600 Clifton Road, N.E., Koger, F-36, Atlanta, GA 30333

1. Introduction

In small area estimation, samples designed to provide estimates for large geographic areas are often used to provide estimates for small areas as well. In such cases the sample in a small area may be unrepresentative or too small to produce reliable estimates. The composite estimator, a weighted sum of two component estimators, can have a mean-squared-error (MSE) which is smaller than that of either component estimator when an appropriate weighting scheme is used (Schaible 1979). This technique has been frequently applied to combine the simple direct and the synthetic estimators (Schaible, Brock, and Schnack 1977, and Royall 1973). However, finding the optimal weight has generally been an insolvable problem in small area estimation. Although Schaible (1979) mentioned two conditions that might help in finding the optimal weight, those conditions turned out to be unrealistic. Schaible stated that the optimal weight could be found when the two components were independent and unbiased estimators of the domain total. These assumptions are difficult to evaluate with respect to a sampling plan when one of the two components is the synthetic estimator and the other is the simple expansion estimator. The second condition Schaible mentioned that would allow the optimal weight to be found approximately was to assume that the covariance (with respect to the sampling plan) of the expansion and the synthetic estimators was small relative to the MSE of either of these components. Again, this condition is difficult to evaluate.

In this paper, we discuss the composite estimate of the uniform minimum variance unbiased (UMVU) estimators under some Bayesian covariate models (Lui and Cumberland 1987), and derive the corresponding optimal weight in explicit form without making the assumptions mentioned previously. We will then show how to estimate the optimal weight and suggest a test to help in deciding how to best apply the results to small domain estimation. Finally, a discussion of the composite of the simple direct estimator and the modified synthetic estimator will be provided.

2. Composite Estimators

We suppose that the finite population is divided into I mutually exclusive sub-areas labeled $i = 1, \ldots, I$ for which we wish to produce estimates. Within each subdomain, units are further classified into J subgroups (for example, socioeconomic class, age, etc.); these are labeled $j = 1, \ldots, J$. The cell sizes N_{ij} resulting from this cross-classification are assumed to be known. Let y_{ijk} ($k = 1, 2, \ldots, N_{ij}$) be the measurement on the k_{th} individual in the ij_{th} cell and

$$T_i = \sum_{j=1}^{J} \sum_{k=1}^{N_{ij}} y_{ijk},$$

the total for the i_{th} subdomain. The primary focus is to estimate the T_i 's.

Letting s_{ij} denote the n_{ij} sampled units in the ij_{th} cell, we use $\sum_{k \varepsilon s_{ij}} y_{ijk}$ to denote the sample sum and $\bar{y}_{ij.}$ to represent the average for the sampled units in cell ij. Similarly, let $\bar{y}_{.j.} = \sum_i \sum_{k \varepsilon s_{ij}} y_{ijk}/n_{.j}$, where $n_{.j} = \sum_i n_{ij}$.

The composite estimator of any two estimators $\hat{T}_i^{(1)}$ and $\hat{T}_i^{(2)}$ for T_i is defined as

$$\gamma \ \hat{T}_i^{(1)} + (1-\gamma)\ \hat{T}_i^{(2)}, \text{ where } 0 \leq \gamma \leq 1.$$

The optimal weight $\gamma^\star$ for the composite estimator to have the minimal MSE is given by

$$\gamma^\star = [E(\hat{T}_i^{(2)} - T_i)^2 - E(\hat{T}_i^{(1)} - T_i)(\hat{T}_i^{(2)} - T_i)]/$$

$$[E(\hat{T}_i^{(1)} - T_i)^2 + E(\hat{T}_i^{(2)} - T_i)^2 - 2E(\hat{T}_i^{(1)} - T_i)(\hat{T}_i^{(2)} - T_i)],$$

if $0 \le \gamma^{\star} \le 1$, otherwise the optimal weight is equal to $\delta_{\{\gamma^{\star}>1\}}$, where δ is an indicator variable. (1) The composite estimator, $\gamma\ \hat{T}_i^{(1)} + (1-\gamma)\ \hat{T}_i^{(2)}$, has MSE smaller than $\hat{T}_i^{(2)}$ if $\gamma < 2\ \gamma^{\star}$. By symmetry, if $(1-\gamma) < 2\ (1-\gamma^{\star})$, then the composite estimator has MSE smaller than $\hat{T}_i^{(1)}$.

Traditionally, with respect to the distribution derived from the sampling plan, an attempt to find the optimal weight for the composite estimator is usually unsuccessful. This is because the formula for the optimal weight involves terms, such as MSE's of the estimators, that are difficult to evaluate with respect to the sampling plan. In the following discussion, on the basis of the super-population model-based approach (Royall 1970), expectation is taken with respect to the distribution given in a model rather than the sampling plan.

3. Bayesian Covariate Model

Consider the model that is the implicit assumption of the most common estimator, the synthetic estimator, for small area estimation:

$$y_{ijk} = b_j + \epsilon_{ijk} \qquad (2),$$

where ϵ's are independent, normally distributed with mean 0 and variance σ^2. Furthermore, assuming that we have some prior knowledge, let b_j be independently distributed with $N(\beta_0, \sigma_b^2)$ and be independent of ϵ's. Note that if $\sigma_b^2 = 0$, then we get one of the simple least-squares models that have been considered by Holt, Smith, and Tomberlin 1979). In the following discussion, we assume that the ratio of $\kappa = \sigma_b^2 / \sigma^2$, which can be interpreted as the relative confidence of the prior knowledge to the current information, can be assigned by investigators, or is known. Methods to estimate this parameter can be found elsewhere (Ghosh and Meeden 1986, Dempster and Raghunathan 1986). When β_0 is known, the Bayesian estimator for T_i has been presented in our previous report (Lui and Cumberland 1987) and it can be easily proved that Bayesian estimator always has a smaller mean-squared-error than the corresponding least-squares estimator. However, if β_0 is unknown, the corresponding empirical Bayes estimator does not necessarily outperform the corresponding least-squares estimator (Lui and Cumberland 1987).

When β_0 is unknown, the UMVU estimator of T_i is given by

$$\hat{T}_i^{gs1} = \sum_j \sum_{k \in s_{ij}} y_{ijk} + \sum_j \sum_{k \notin s_{ij}} [(1-\lambda_j)\ \bar{y}_w + \lambda_j\ \bar{y}_{.j.}], \quad (3)$$

where $\bar{y}_w = \sum_j \lambda_j\ \bar{y}_{.j.} / \sum_j \lambda_j$

$$\lambda_j = n_{.j}\ \kappa / (n_{.j}\ \kappa + 1).$$

The prediction variance of $\hat{T}_i^{gs1}$ is given by

$$V(\hat{T}_i^{gs1} - T_i) = \sum_j (N_{ij} - n_{ij})\ \sigma^2 +$$

$$\sum_j (N_{ij} - n_{ij})^2 \lambda_j \sigma^2 / n_{.j} + (\sum_j (N_{ij} - n_{ij})(1-\lambda_j))^2 \sigma_b^2 / (\sum_j \lambda_j).$$

Assuming we have some prior covariate information related to b_j and assuming that b_j is normally distributed as $N(\beta_0 + \beta_1 x_j, \sigma_b^2)$, it is easy to show that the UMVU estimator of T_i under this covariate model when β_0 and β_1 are unknown is given by

$$\hat{T}_i^{cs1} = \sum_j \sum_{k \in s_{ij}} y_{ijk} + \sum_j \sum_{k \notin s_{ij}} [(1-\lambda_j)(\hat{\beta}_0 + \hat{\beta}_1 x_j) + \lambda_j \bar{y}_{.j.}],$$

where $\hat{\beta}_0 = \bar{y}_w - \hat{\beta}_1 \bar{x}_w$,

$$\hat{\beta}_1 = \sum_j \lambda_j\ (x_j - \bar{x}_w)(\bar{y}_{.j.} - \bar{y}_w) / \sum_j \lambda_j\ (x_j - \bar{x}_w)^2,$$

and $\bar{x}_w = \sum_j \lambda_j\ x_j \ / \sum_j \lambda_j$.

The prediction variance of $\hat{T}_i^{cs1}$ is given by

$$V(\hat{T}_i^{cs1} - T_i) = V(\hat{T}_i^{gs1} - T_i) +$$
$$(\sum_j (N_{ij} - n_{ij})(1-\lambda_j)(x_j - \bar{x}_w))^2 \sigma_b^2 / \sum_j \lambda_j (x_j - \bar{x}_w)^2.$$

$\hat{T}_i^{gs1}$, though biased under the covariate model, can be a better estimator with respect to the MSE than the UMVU estimator $\hat{T}_i^{cs1}$. This occurs when the coefficient of variation for $\hat{\beta}_1$ is large. Determining the choice between $\hat{T}_i^{gs1}$ and $\hat{T}_i^{cs1}$ leads us to consider a weighted average of these two estimators that could have a MSE smaller than either $\hat{T}_i^{gs1}$ or $\hat{T}_i^{cs1}$. From the formula (1), we have the optimal weight $\gamma^{\star}$ equal to $CV^2(\hat{\beta}_1)/(CV^2(\hat{\beta}_1)+1)$, where $CV^2(\hat{\beta}_1)=V(\hat{\beta}_1)/(E(\hat{\beta}_1))^2$. Note that the condition $\gamma < 2\gamma^{\star}$ is automatically satisfied if $CV^2(\hat{\beta}_1) > 1$. Therefore, $\gamma\ \hat{T}_i^{gs1} + (1-\gamma)\ \hat{T}_i^{cs1}$ always has a MSE smaller than $\hat{T}_i^{cs1}$, if the coefficient of variation of $\hat{\beta}_1$ is greater than 1. Conversely, if $CV^2(\hat{\beta}_1) < 1$, then the composite estimator always has a MSE smaller than $\hat{T}_i^{gs1}$. Furthermore, we can easily show that the composite estimator always has a MSE smaller than either of its components if

a. $CV^2(\hat{\beta}_1) > 1$ and $\gamma \geq (CV^2(\hat{\beta}_1)-1)/(CV^2(\hat{\beta}_1)+1)$,
or if
b. $CV^2(\hat{\beta}_1) < 1$ and $\gamma \leq 2CV^2(\hat{\beta}_1)/(CV^2(\hat{\beta}_1)+1)$,
or if
c. $CV^2(\hat{\beta}_1) = 1$.

Note that if $CV^2(\hat{\beta}_1) \to \infty$, then the composite estimator with the optimal weight converges to $\hat{T}_i^{gs1}$. This implication is quite reasonable, because if the coefficient of variation $CV(\hat{\beta}_1)$ is very large, then using the information about β_1 in the estimator $\hat{T}_i^{cs1}$ might lead to a worse estimator than the estimator $\hat{T}_i^{gs1}$, which does not use this information.

Finally, the MSE of the composite estimator with the optimal weight $\gamma^{\star}$ is given by

$$E(\gamma^{\star}\hat{T}_i^{gs1} + (1-\gamma^{\star})\hat{T}_i^{cs1} - T_i)^2 = V(\hat{T}_i^{gs1}-T_i)$$

$$+ V(\hat{\beta}_1)[E(\hat{T}_i^{gs1}-T_i)]^2/[V(\hat{\beta}_1)+\beta_1^2],$$

where $V(\hat{T}_i^{gs1}-T_i)$, $V(\hat{\beta}_1)$, β_1, and the bias $E(\hat{T}_i^{gs1}-T_i)$ can be easily estimated if κ, the relative size σ_b^2 to σ^2, is assumed to be known.

We can generalize the above results to include any p-covariate ($p \leq J-1$). Let $b_j \sim N(\beta_0 + \beta'\underline{x}_j, \sigma_b^2)$, where $\underline{\beta}' = (\beta_1, \beta_2, \ldots, \beta_p)$, and $\underline{x}_j' = (x_{j1}, x_{j2}, \ldots, x_{jp})$. In this p-covariate model, the UMVU estimator $\hat{T}_i^{csp}$ of T_i is given by

$$\hat{T}_i^{csp} = \sum_j \sum_{k\epsilon s_{ij}} y_{ijk} + \sum_j \sum_{k\notin s_{ij}} [(1-\lambda_j)(\hat{\beta}_0 + \hat{\beta}_1 x_{j1} + \ldots + \hat{\beta}_p x_{jp}) + \lambda_j\ \bar{y}_{.j.}],$$

where $\hat{\beta}_0 = \bar{y}_w - \sum_{t=1}^{p}\hat{\beta}_t\bar{x}_{wt}$ and $\hat{\underline{\beta}} = (\underline{X}_D'\underline{\Lambda}\ \underline{X}_D)^{-1}\underline{X}_D'\underline{\Lambda}\ \underline{\bar{y}}$

$$\underline{X}_D = (x_{jt} - \bar{x}_{wt})_{J\times P},\quad \bar{x}_{wt} = \sum_j \lambda_j x_{jt} \Big/ \sum_j \lambda_j$$

$$\underline{\Lambda} = (\mathrm{diag}(\lambda_j))_{J\times J},\quad \underline{\bar{y}}' = (\bar{y}_{.1.}, \bar{y}_{.2.}, \ldots, \bar{y}_{.J.}).$$

The prediction variance of $V(\hat{T}_i^{csp}-T_i) =$

$$V(\hat{T}_i^{gs1} - T_i) + [\underline{L}_i'(\underline{X}_D'\ \underline{\Lambda}\ \underline{X}_D)^{-1}\underline{L}_i]\ \sigma_b^2,$$

where $\underline{L}_i' = (\sum_j (N_{ij}-n_{ij})(1-\lambda_j)(x_{j1}-\bar{x}_{w1}), \ldots, \sum_j (N_{ij}-n_{ij})(1-\lambda_j)(x_{jp}-\bar{x}_{wp}))$.

In fact, every argument in the univariate case can be carried through simply replacing $CV^2(\hat{\beta}_1)$ in the univariate case with $CV^2(\underline{L}_i'\hat{\underline{\beta}})$.

4. Least-Squares Model Related to The Simple Direct And The Synthetic Estimators

Considering the most commonly used composite estimator in small area estimation, a weighted sum of a simple direct and a synthetic estimator leads to the estimator $\gamma\ \hat{T}_i^D + (1-\gamma)\ \hat{T}_i^{MS}$, a linear combination of the simple direct estimator and the modified synthetic estimator (Holt, Smith, and Tomberlin 1979).

The simple direct estimator,

$$\hat{T}_i^D = \sum_j N_{ij}\ \bar{y}_{ij.},$$

is the UMVU estimator of T_i under the model

$$y_{ijk} = \mu_{ij} + \epsilon_{ijk}$$

where ϵ's are independent, normally distributed with mean 0 and variance σ^2 and μ_{ij} are fixed unknown constants.

The modified synthetic estimator,

$$\hat{T}_i^{MS} = \sum_j \sum_{k\epsilon s_{ij}} y_{ijk} + \sum_j \sum_{k\notin s_{ij}} \bar{y}_{.j.},$$

is the UMVU estimator of T_i under the above model when $\mu_{1j} = \mu_{2j} = \ldots = \mu_{Ij}$. It is easy to show that $E(\hat{T}_i^D - T_i)(\hat{T}_i^{MS} - T_i) = V(\hat{T}_i^{MS} - T_i)$ under the non-restricted model. Therefore, we find the optimal weight from formula (1) is

$$\gamma^{\star} = [\sum_j (N_{ij} - n_{ij})^2(1/n_{ij} - 1/n_{.j})\sigma^2]/$$
$$[\sum_j (N_{ij} - n_{ij})^2(1/n_{ij} - 1/n_{.j})\sigma^2 + (\sum_j (N_{ij} - n_{ij})(\mu_{ij} - \bar{\mu}_{.j}))^2],$$

where $\bar{\mu}_{.j} = \sum_i n_{ij}\mu_{ij}/n_{.j}$.

A reasonable estimate $\hat{\gamma}^{\star}$ of $\gamma^{\star}$ can be obtained by substituting the UMVU estimators

$$\hat{\mu}_{ij} = \bar{y}_{ij.}, \ \hat{\bar{\mu}}_{.j} = \sum_i n_{ij}\bar{y}_{ij.}/n_{.j} \text{ and}$$

$$\hat{\sigma}^2 = \sum_j \sum_i \sum_{k \in s_{ij}} (y_{ijk} - \bar{y}_{ij.})^2/(n_{..} - J)$$

for μ_{ij}, $\bar{\mu}_{.j}$, and σ^2 in $\gamma^{\star}$ respectively.
A measure of error $E(\gamma\ \hat{T}_i^D + (1-\gamma)\ \hat{T}_i^{MS} - T_i)^2$ for $\gamma\ \hat{T}_i^D + (1-\gamma)\ \hat{T}_i^{MS}$ can be obtained from easily accessible estimates of $E(\hat{T}_i^D - T_i)^2$, $E(\hat{T}_i^{MS} - T_i)^2$, and $E(\hat{T}_i^D - T_i)(\hat{T}_i^{MS} - T_i)$, which can be derived from the model-based approach (Holt, Smith, and Tomberlin 1979).

5. Testing of $CV^2(\hat{\beta}_1)$

We have shown that $CV^2(\beta_1) > 1$ ($CV^2(\beta_1) < 1$) implies that the composite estimator $\gamma\ \hat{T}_i^{gs1} + (1-\gamma)\ \hat{T}_i^{cs1}$ always has a MSE smaller than $\hat{T}_i^{cs1}$ ($\hat{T}_i^{gs1}$). We must, however, decide whether $CV^2(\beta_1) > 1$ or $CV^2(\beta_1) < 1$, because usually the values of β_1 and σ_b^2 are unknown. One method for deciding whether $CV^2(\beta_1) > 1$ is hypothesis testing, which requires that we find the distribution of the estimator of $CV^2(\hat{\beta}_1)$. We can easily rewrite the model assumption given in section 3, with $b_j \sim N(\beta_0 + \beta_1 x_j, \sigma_b^2)$, into the matrix form:

$$\begin{pmatrix} \underline{Y}_s \\ \underline{Y}_{\bar{s}} \end{pmatrix} = \begin{pmatrix} \underline{X}_s \\ \underline{X}_{\bar{s}} \end{pmatrix} \underline{B} + \begin{pmatrix} \epsilon_s \\ \epsilon_{\bar{s}} \end{pmatrix}, \text{ where } \begin{pmatrix} \epsilon_s \\ \epsilon_{\bar{s}} \end{pmatrix}$$

is distributed with $N(\underline{0}, \sigma^2 I)$, $\underline{B} = (b_1, \ldots, b_J)'$ is distributed with $N(\underline{A}\ \underline{\beta}, \sigma_b^2 I)$, and is independent of

$$\begin{pmatrix} \epsilon_s \\ \epsilon_{\bar{s}} \end{pmatrix}, \ \underline{A}' = \begin{pmatrix} 1, & 1, & \ldots & 1 \\ x_1, & x_2, & \ldots & x_J \end{pmatrix}_{2 \times J}, \ \underline{\beta}' = (\beta_0, \beta_1), \text{ and}$$

where s and $\bar{s}$ denote the sampled and non-sampled units respectively.

Applying the standard results for the general linear model (Graybill 1976), we get the UMVU estimator of σ^2,
$\hat{\sigma}^2_{cs1} = \underline{Y}_s'[\underline{V}_s^{-1} - \underline{V}_s^{-1}\ \underline{X}_s^{(2)}(\underline{X}_s^{(2)'}\underline{V}_s^{-1}\underline{X}_s^{(2)})^{-1}\underline{X}_s^{(2)'}\underline{V}_s^{-1}]\underline{Y}_s/(n_{..}-2)$,
where $\underline{Y}_s'$ is $1 \times n_{..}$ vector of measurements on the sampled individuals, $\underline{V}_s = (\mathrm{diag}(\kappa\ \underline{1} + \underline{I})_{n_{.j} \times n_{.j}})_{n_{..} \times n_{..}}$ is a block diagonal matrix, and $\underline{X}_s^{(2)} = \underline{X}_s\ \underline{A}$. It is easy to show that $(n_{..} - 2)\ \hat{\sigma}^2_{cs1}/\sigma^2 \sim \chi^2\ (n_{..} - 2)$ which implies that $(n_{..} - 2)\ \kappa\ \hat{\sigma}^2_{cs1}/\sigma_b^2 \sim \chi^2\ (n_{..} - 2)$. Also, the UMVU estimator $\hat{\beta}_1$ of β_1 is

$$\hat{\beta}_1 = (0,1)\ (\underline{X}_s^{(2)'}\underline{V}_s^{-1}\ \underline{X}_s^{(2)})^{-1}\ \underline{X}_s^{(2)'}\underline{V}_s^{-1}\ \underline{Y}_s,$$

and $\hat{\beta}_1$ and $\hat{\sigma}^2_{cs1}$ are independent. Thus we have $(\hat{CV}^2(\hat{\beta}_1))^{-1} = \hat{\beta}_1{}^2\ (\hat{\sigma}_b^2/(\sum_j \lambda_j\ (x_j - \bar{x}_\omega)^2)^{-1}$, which is distributed as $F(1/[2CV^2(\beta_1)];\ 1,\ n_{..} - 2)$, a noncentral F-distribution, where $\hat{\sigma}_b^2 = \kappa\ \hat{\sigma}^2_{cs1}$.

We can use this resulting distribution to test the hypothesis $H_0 : CV^2(\beta_1) = 1$ versus $(CV^2(\beta_1))^{-1} > 1$. Under the null hypothesis, the distribution of the test statistic is a F-distribution with noncentrality parameter 0.5, leading to a simple test of the null hypothesis. When we reject the null hypothesis, we prefer using $\gamma\ \hat{T}_i^{gs1} + (1-\gamma)\ \hat{T}_i^{cs1}$ (or in particular, $\hat{T}_i^{cs1}$) to using $\hat{T}_i^{gs1}$. This result is consistent with the fact that when $(CV^2(\beta_1))^{-1}$ is very large, (implying that the prior knowledge about β_1 is very precise), the estimator $\hat{T}_i^{cs1}$ using the covariate information will be more accurate than the estimator $\hat{T}_i^{gs1}$ which ignores it. This test statistic can be easily generalized for the P-variate case. The test statistic for testing hypothesis $CV^2(\underline{L}_i'\underline{\beta}) = 1$ has a noncentral F-distribution with parameter equal to 0.5 and degrees of freedom 1 and $n_{..} - P - 1$.

6. Discussion

From the traditional point of view, Schaible (1979) pointed out two major problems in using the composite estimator. The first problem concerns how to estimate the optimal weight, given two estimators. The difficulty stems from the near impossibility of calculating, with respect to sampling plan, the MSE of small area estimators such as the synthetic estimator. Although Schaible (1979) suggested several different methods to estimate the optimal weight corresponding to different possible

MSE's assumed for the population, these methods all depend on the true value T_i, and hence cannot be applied to estimate T_i in practice. The second problem, common to all small area estimators, is how to provide a measure of error of a composite estimator for a given small area. Using the model-based approach, however, we can estimate the optimal weight explicitly and provide the measure of error of the composite estimator for each small area. Therfore, the results presented here should be useful for survey statisticians or epidemiologists in estimating the local area characteristics.

References

Dempster, A. P. and Raghunathan, T. E. (1986): A Method of Estimating the Interval Population of Counties. Journal of the American Statistical Association, 51, pp. 587-590.

Ghosh, M. and Meeden, G. (1986): Empirical Bayes Estimation in Finite Population Sampling. Journal of the American Statistical Association, 81, pp. 1058-1062.

Graybill, F. A. (1976): Theory and Application of the Linear Model. North Scituate: Duxburg.

Holt, D., Smith, T. M. F., and Tomberlin, T. J. (1979): A Model-Based Approach to Estimation for Small Subgroups of a Population. Journal of the American Statistical Association, 74, pp. 405-410.

Lui, K.-J. and Cumberland W. G. (1987): A Bayesian Approach to Small Domain Estimation. Proceedings of the American Statistical Association, Survey Statistics Section, pp. 347-352.

Royall, R. M. (1970): On Finite Population Sampling Theory under Certain Linear Regression Models. Biometrika, 57, pp. 377-387.

Royall, R. M. (1973): Discussion of Two Papers on Recent Developments in Estimation of Local Areas. Proceedings of the American Statistical Association, Social Statistics Section, pp. 43-44.

Schaible, W. L. (1979): A Composite Estimator for Small Area Statistics. In Synthetic Estimates for Small Areas: Statistical Workshop Papers and Discussion, edited by J. Steinberg, NIDA Research Monograph, 24, pp. 36-59.

Schaible, W. L., Brock, D. B., and Schnack, G. A. (1977): An Empirical Comparison of the Simple Inflation, Synthetic and Composite Estimators for Small Area Statistics. Proceedings of the American Statistical Association, Social Statistics Section, pp. 1017-1021.