Comparison of Small Area Models in SAIPE

Tapabrata Maiti[1,2] and Eric V. Slud[1,3]
[1] Census Bureau, [2] Iowa State University, [3] University of Maryland

May 17, 2002

Abstract. The Small Area Income and Poverty Estimation (SAIPE) project is an ongoing Census project to estimate numbers of poor school-age children by state, county, and ultimately school district, based upon Current Population Survey (CPS) and IRS data, together with information from the latest decennial census. The project and its mandated governmental role are summarized in Citro & Kalton (1999). The current county-level methodology, summarized by Bell (1997), relies on a Fay-Herriot (1979) model fitted to log-counts (by county) of related school-age children in CPS-sampled households, and discards data from those sampled counties with no sampled poor children. The present paper compares SAIPE small-area estimation by analogous Fay-Herriot models for logarithms of county child poverty *rates* with a unit- (i.e., individual-) level logistic- regression model with county-level random effects. This comparison is done in two ways. First, it is shown via simulation that the effect of fitting a Fay-Herriot model on what amounts to a truncated dataset (i.e., a dataset omitting cases with quantitative normal responses below a certain threshold) introduces systematic trends relating bias to response-level which degrades not only the quality of small-area estimates (SAE's) but renders especially unreliable the available estimates (Prasad and Rao 1990) of mean-squared error of the SAE's. Our second approach to understanding the differences in performance between the linear (Fay-Herriot) and generalized-linear (GLM, or mixed logistic regression) SAE methods, is grounded in data analyses using SAIPE datasets from 1994 and 1990, using 1990 decennial census data as an external standard of truth for the county-level child poverty rates being estimated. It is shown that the GLM method and a weighted variant fit the data better than the (truncated) linear models currently used, when judged by the internal evidence of the 1994 and 1990 datasets. However, a SAIPE Fay-Herriot fitting method for 1990 data, with particularly small PSU variance and a particularly large coefficient for the 1980 Census log child-poverty rate, performs excellently in matching to the 1990 Census log-rate in CPS-sampled counties. Still, for counties in which there was no CPS sample, GLM SAE's perform better than those from linear models on 1990 data with Census '90 as the standard of truth, with respect to various loss criteria.

Key words: empirical best linear unbiased predictor (EBLUP), mean-squared errors, mixed effects linear model, mixed effect logistic regression, loss function, small area estimation, weighted likelihood.

This paper reports on research and analysis undertaken for the Census Bureau by Tapabrata Maiti and Eric Slud, and is released to inform interested parties and to encourage discussion. Results and conclusions expressed are those of the authors and have not been endorsed by the Census Bureau.

1 Introduction

As summarized by Citro & Kalton (1999) and Bell (1997) the Small Area Income and Poverty Estimates (SAIPE) project at the Census Bureau has developed methods for estimating poverty and income statistics at the county and state level. At the county level, these methods rely on a mixed-effects linear model which is applied to the logarithms of the observed numbers (e.g. of poor children 5–17) in counties for which CPS samples were taken and in which the sample contained a nonzero number of poor children. Those sampled counties without poor children 5–17 in-sample are dropped from the analysis, a bothersome aspect of the methodology. It would be desirable instead to model the essential discreteness of the response-counts by some sort of unit-level model.

Based upon the methods proposed in Slud (1998, 2000b) of estimating mixed-effect logistic regression models via mixed-nonlinear-regression software or by maximization of an approximate log-likelihood, we describe here a mixed-effect unit-level logistic regression model for SAIPE data. This model would make use of all of the SAIPE data. In order to compare this new method to the one which is now in use, Slud (2000, 2002) has previously conducted several simulation experiments, with model parameters progressively becoming more realistic for the SAIPE application, to document the effect on small-area point estimates and on estimates of mean-squared error of estimation using both the truncated Fay-Herriot model and the mixed logistic *unit-level* regression model. Briefly, those experiments found that the mixed-logistic method had 10–30% lower empirical mean-squared error than the linear-model method when the true model was mixed logistic, and even when the true (unit-level) model had an exponential link — a setting more closely resembling the log-linear mean-rate specification of the Fay-Herriot model standardly used in SAIPE — the MSE of the mixed logistic estimators was not more than 5–10% worse *and was often better* than that the linear-model estimators. Part of the explanation for this was that even with an exponential link for unit-level mean child-poverty rate, the log-linear model analysis was misspecified because, after discarding sampled PSU's with zero counts of child poor, it was based on a randomly left-truncated dataset. However, the simulation experiments showed a small systematic bias of SAE's based on the log-linear model, as a function of actual child-poverty rate, when the mixed logistic model holds. Similarly, the GLM-based SAE's showed a small systematic bias versus the relatively unbiased linear-model SAE's, as a function of actual PSU response-rate, when the true link was exponential. However, both of these biases were small enough (with magnitudes of order 0.01–0.02) to be overwhelmed by the variability due to PSU random effects in actual SAIPE datasets. On the other hand, PSU-level MSE estimates for SAE's constructed under each of the two models — while accurate within 20% or so when the link was properly specified, exponential for the log-linear model and logistic for the GLM — turned awful (often with errors of 50% or more) for the model with wrongly specified link. This raises a serious concern that the functional form of the average child-poverty rate as a function a linear combination of CPS and IRS predictors be carefully ascertained and checked in actual SAIPE data-analyses. There is also a remaining worry about the extent to which the standard SAIPE approach of using a complete-data (and therefore misspecified) model on a truncated dataset can be causing bias in SAE's and MSE estimates.

To address the latter concern, in this paper we first describe in Section 2 a simulation study of the performance of small-area estimators under a Fay-Herriot model, with PSU's, covariates, and response-rates like those of SAIPE, under either 20% or 30% lower truncation of a normally distributed quantitative response meant to mimic the PSU-level observed log child-poverty rate. We study the biases in the truncated-data SAE's (as a function of true response rate) and in the corresponding estimators of MSE adapted from those provided by Prasad and Rao (1990), finding that the MSE estimators are unreliable for the truncated data.

We then take a second approach, in Section 3, to comparing the performance of SAE's based on log-linear and GLM models in SAIPE, based on careful data analysis of SAIPE CPS/IRS datasets from 1994 and 1990. First, we look at the internal evidence in these datasets concerning the relative adequacy of fit of the two types of model. In this effort, we calculate several different 'loss-functions' — weighted and unweighted sums of squared, absolute, and absolute relative errors — to aggregate the discrepancies between fitted county-level child poverty rates and the CPS estimates. We also break down the losses into the contributions arising from (CPS-sampled) counties in which there were sampled poor children from those where there were not.

Next, by appealing to 1990 decennial census data as the external standard of truth for the county-level child poverty rates, we re-calculate the corresponding losses for CPS-sampled counties, according to whether or not there were sampled poor children, and finally for the non-CPS-sampled counties. We devote particular attention to the non-sampled counties to show that, to the (incomplete) extent that the Census measurement of child poverty rates accords conceptually with that defined in SAIPE, there is some advantage for the GLM methodology according to virtually all loss measures. We pull together our recommendations based on these findings (together with those of the simulations in Section 2) in Section 4.

Along the way, we investigate in Section 3.3 several possible explanations of the different behavior of the log-linear and GLM estimating methods. We find most convincing an explanation based on the different ways in which the Fay-Herriot aggregate-level method and the mixed-logistic method weigh counties — respectively, roughly equally or in proportion to size — in their respective Maximum-likelihood estimators of fixed-effect coefficients.

Acknowledgements. We are grateful to the Census Bureau for supporting this research, to William Bell for his advice and comments, and to Robin Fisher for helping us with access to and interpretation of the SAIPE datafiles.

2 Effect of Data Truncation in a linear model

We begin by studying the effect of an unavoidable and potentially unfavorable aspect of SAIPE county-level methodology, namely the truncation of the CPS dataset by treating counties in which there were sampled children but no observed *poor* children as though they had no sample at all. We operate within the framework of a Fay-Herriot linear model for logarithms of within-sample county child poverty rates, and we idealize it by studying through simulations the comparative performance of small-area estimates of means and transformed means within Fay-Herriot models in which responses are left-truncated.

2.1 Model and Estimators, for complete and truncated data

Consider the following general models, first for complete and then for truncated data. For each PSU (county) indexed by $i = 1, \ldots, m$, assume that sample-sizes n_i and p-dimensional vectors x_i of predictor variables are known, and that response-variables

$$y_i = x_i^{\text{tr}} \beta + u_i + e_i \quad , \qquad u_i \sim \mathcal{N}(0, \sigma_u^2) \quad , \qquad e_i \sim \mathcal{N}\left(0, \frac{v_e}{n_i}\right) \tag{1}$$

are observed (whenever $n_i > 0$), where $\beta \in \mathbf{R}^p$ is a vector of unknown fixed-effect coefficients, and u_i, e_i are respectively PSU random effects and sampling errors, independent of each other within and across PSU's. Ordinarily, σ_u^2 is unknown and estimated, while v_e is known. This is a standard Fay-Herriot (1979) model, used for small-area estimation, i.e., for estimation of the quantities

$$\vartheta_i = x_i^{\text{tr}} \beta + u_i \quad , \qquad i = 1, \ldots, m$$

However, in SAIPE it also makes sense, via consideration of the 'bivariate model' of Bell (1997) using auxiliary decennial census data (summarized also in Citro and Kalton 1999), to treat σ_u^2 as known and v_e as unknown and estimated.

In idealizing SAIPE, we view y_i as the observed log child-poverty rate for the i'th PSU, and define the rate itself by exponentiating:

$$\vartheta_i^* = \exp(\vartheta_i) \equiv \exp(x_i^{\text{tr}} \beta + u_i)$$

In the context of SAIPE, one would discard data from PSU's with no child poor, which we idealize by defining a *truncated dataset*

$$(\tilde{y}_i \; : \; 1 \le i \le m \, , \; y_i > \tau)$$

with corresponding sample-sizes

$$\tilde{n}_i \equiv \begin{cases} n_i & \text{if} \quad n_i > 0 \quad \text{and} \quad y_i > \tau \\ 0 & \text{if} \quad n_i = 0 \quad \text{or} \quad y_i \le \tau \end{cases}$$

Here the threshold-parameter τ is regarded as known, analogous to the level 0 for observed count of poor children.

We study through simulation the behavior of estimators of ϑ_i and ϑ_i^* (*small-area estimators*, or SAE's), through their empirical Mean-Squared Errors (MSE's) and statistical estimators (*mse*'s) of these MSE's. The estimators for ϑ_i based on the complete data $\{y_j, n_j\}_{j=1}^m$ above are the standard EBLUP estimators (*cf.* Prasad and Rao 1990, Ghosh and Rao 1994)

$$\hat{\vartheta}_i \;=\; x_i^{\rm tr}\,\hat{\beta} \,+\, \hat{\gamma}_i\,(y_i - x_i^{\rm tr}\hat{\beta})$$

where $(\hat{\beta}, \hat{\sigma}_u^2)$ or $(\hat{\beta}, \hat{v}_e)$ are the maximum likelihood estimators in the model (1), and $\hat{\gamma}_i = \hat{\sigma}_u^2/(\hat{\sigma}_u^2 + v_e/n_i)$ or $\hat{\gamma}_i = \sigma_u^2/(\sigma_u^2 + \hat{v}_e/n_i)$, depending on which variance component is being estimated. We follow the convention that $\hat{\gamma}_i \equiv 0$ (so that $\hat{\vartheta}_i = x_i^{\rm tr}\,\hat{\beta}$) when $n_i = 0$.

In our simulation study, we present results with the usual Fay-Herriot model (cited later as **LmA**), i.e., with v_e assumed known but σ_u^2 unknown. We have also performed the simulation study using the current approach of the Census bureau i.e., treating σ_u^2 as known but v_e unknown within model (1), an analysis method we cite below as **LmB**. The detailed results of the **LmB** simulation are not presented because their general behavior is the same as for **LmA**. In either model, the (complete-data) estimators for the exponentiated small-area parameters $\vartheta_i^* = \exp(\vartheta_i)$ are given by the approximately bias-corrected formula

$$\hat{\vartheta}_i^* \;=\; \exp\left(x_i^{\rm tr}\,\hat{\beta} \,+\, \hat{\gamma}_i\,(y_i - x_i^{\rm tr}\hat{\beta}) + \frac{1}{2}\,\hat{\sigma}_u^2\,(1 - \hat{\gamma}_i)\right) \tag{2}$$

(This formula, used in SAIPE in the **LmB** setting with $\hat{\sigma}_u^2$ replaced by σ_u^2, ignores the variability of the parameter-estimators but corrects the bias.)

An approximate expression for MSE estimates is given (along with definitions of the component expressions g_{ki}, $k = 1, \ldots, 4$) by Datta and Lahiri (2001), in the form

$$mse(\hat{\vartheta}_i) = \begin{cases} g_{1i}(\hat{\sigma}_u^2) + g_{2i}(\hat{\sigma}_u^2) + 2g_{3i}(\hat{\sigma}_u^2) - g_{4i}(\hat{\sigma}_u^2) & \text{if} \quad n_i > 0 \\ \hat{\sigma}_u^2 \;+\; x_i^{\rm tr}\,\widehat{\rm Var}(\hat{\beta})\,x_i & \text{otherwise} \end{cases} \tag{3}$$

where $\widehat{\rm Var}(\hat{\beta})$ denotes a large-sample estimator of (asymptotic) variance, which could be given in a robustified (misspecified-model 'sandwich-estimator') form if desired. The significance of each of the g_{ki}'s, $k = 1, 2, 3,$ is nicely described in Prasad and Rao (1990). The g_{1i} term is the posterior variance of the Bayes estimator (or BLUP) and is of order $\mathcal{O}(1)$. The g_{2i} term accounts for the variability of $\hat{\beta}$ and is of order $\mathcal{O}(m^{-1})$. The term g_{3i} accounts for variability of the variance-component estimator ($\hat{\sigma}_u^2$ or $\hat{v}_e$) and is also of order $\mathcal{O}(m^{-1})$. The last term, g_{4i}, is an asymptotic bias adjustment for the ML estimate of σ_u^2 derived in Datta and Lahiri (2001), and does not arise when ANOVA- or moments-based estimates of σ_u^2 are used. Equation (3) is an estimator of an approximate MSE which is correct up to order of magnitude $\mathcal{O}(m^{-1})$ under the Fay-Herriot model. So the neglected terms are of order $o(m^{-1})$ and hence (3) is a *second-order corrected* MSE estimator.

The corresponding MSE estimators for the exponentiated small-area estimates $\hat{\vartheta}_i^*$ are approximated by a first-order Taylor expansion, yielding

$$mse(\hat{\vartheta}_i^*) = (\hat{\vartheta}_i^*)^2 \cdot mse(\hat{\vartheta}_i)$$

2.2 Design of the Simulation

A simulation study has been performed as follows to check the performance of the estimation strategy described above. In addition to the parameters $(\beta,\, \sigma_u^2,\, v_e,\, \{n_i,\, x_i\})$, we fix throughout each run of the simulation a quantile $p \in (0,1)$, equal to 0.2 or 0.3 in the results below.

- Step I. Generate according to (1) the variables $u_i,\, y_i,\ i = 1, \ldots, m$, treat the data set $\{y_i,\, x_i\}_{i=1}^m$ as the complete data set, denoted S_1.
- Step II. Define the *left-truncated dataset* as the subset $S_2 = \{(y_i,\, x_i) : 1 \le i \le m,\ y_i > \tau\}$ of S_1, where τ is defined to satisfy

$$p \;=\; \frac{1}{m} \sum_{i=1}^{m} P(\, y_i \le \tau\,)$$

  or equivalently,

$$p \;=\; \frac{1}{m} \sum_{i=1}^{m} \Phi\left(\frac{\tau - x_i^{\rm tr}\beta}{\sqrt{\sigma_u^2 + v_e/n_i}} \right) \tag{4}$$

- Step III. Compute the point estimates $\{\hat{\vartheta}_i,\, \hat{\vartheta}_i^*\}$ and the MSE estimates $\{mse(\hat{\vartheta}_i),\, mse(\hat{\vartheta}_i^*)\}$ for both the data sets S_1 and S_2. Note that $\tilde{n}_i = 0$ and $\hat{\gamma}_i = 0$ in the S_2 estimations for all indices i which are in S_1 but not in S_2.

- Step IV. Repeat Steps I–III a large number B of times.
- Step V. For each estimator t_i of T_i (where $T_i = \vartheta_i,\ \vartheta_i^*,\ mse(\hat{\vartheta}_i)$), and $mse(\hat{\vartheta}_i^*)$, calculate the following:

$$\text{(i) Simulated MSE:} \qquad SMSE(t_i) \ = \ \frac{1}{B}\sum_{b=1}^{B}\left(t_i^{(b)} - T_i^{(b)}\right)^2$$

$$\text{(ii) Simulated bias:} \qquad SB(t_i) \ = \ \frac{1}{B}\sum_{b=1}^{B}\left(t_i^{(b)} - T_i^{(b)}\right)$$

In addition, we calculate for each simulation run of B iterations the average of each of the parameter-estimates $\hat{\beta}$ and $\hat{\sigma}_u^2$ or $\hat{v}_e$ (and therefore the corresponding biases).

2.3 Simulation results

For our presentation we used SAIPE county predictors x_i for the year 1994, as described in detail in Section 3 below. Since 304 counties (= 20.4% of the $m = 1488$ PSU's in the 1993–95 3-year aggregated CPS sample) do not have any sampled poor children aged 5-17, and there is interest in extending the SAIPE methodology to CPS sample data from single years (without 3-year aggregation, i.e., with smaller samples, which would result in more 0-counts), we chose $p = .2$ or $.3$ as the truncation quantile for our simulations.

In Figure 1 we plot the average MSE and the corresponding average *mse* for all the 1488 counties. Figures 1(a) and (c) show the Prasad-Rao (1990) type of MSE estimates tracking the MSE's very closely. However, Figures 1(b) and 1(d) show the poor performance of *mse* in the presence of truncation. In the truncated case, the Prasad-Rao type *mse* severely underestimates MSE, and Figure 2 displays the percent underestimation. Note that the underestimation increases as the fixed effect parameters approach the extremes (respectively -3.9 and -0.45) of values of $x_i^{\text{tr}}\beta$. A similar pattern is observed for exponentiated SAE's and is depicted in Figure 3 and 4. Note that Figures 2 and 4, as well as 5 and 6 below, display results only for the case $p = 0.2$. The corresponding results for $p = 0.3$ were calculated but are not shown because they were very similar to the case $p = 0.2$.

To investigate the reason behind the underestimation, we plot MSE against the sum of *mse* and the square of the bias of the point estimate.

Figure 5 shows that, although there is a large bias in the SAE, the sum of squared bias and *mse* is very close to the MSE, at least for the linear scale of measurement. That is, the *mse* continues to estimate the *variance* of the SAE well, rather than the MSE. Figure 6 shows the situation for the exponential scale. Note that, in this case, *mse* plus squared SAE bias is no longer very close to the MSE. This discrepancy between the behavior of MSE estimates in the two different scales arises in part because the estimator in exponential scale corrects for the bias due to the nonlinear transformation, without adjustment of the variance. A close examination reveals that the term $g_{1i}(\hat{\sigma}_u^2) = \hat{\sigma}_u^2 v_e/(v_e + n_i\hat{\sigma}_u^2)$ is the dominant term contributing to *mse* (about 99% of it for most of the counties), and all other terms contribute minimally. From our simulation study, we notice that there is a systematic negative bias in the estimation of σ_u^2 for the truncated data, in turn leading to an underestimated *mse* . Our numerical finding was that, if the estimate of σ_u^2 for the truncated data set were replaced by its true value, then the $g_{1i}(\hat{\sigma}_u^2)$ values under truncated and complete data are very close. Thus, underestimation of σ_u^2 seems to be the primary cause of bad performance of *mse*'s. Note that the asymptotic bias-adjusted term $g_{4i}(\hat{\sigma}_u^2)$ adjusts the bias due to ML estimation from the complete data set but does not accommodate the bias due to truncation.

We have mentioned above that truncation causes biases in the small area estimates. As an illustration, Table 1 reports summary statistics across PSU's of the distribution of simulated biases of the SAE's for complete and truncated data, for both types of Fay-Herriot models (**LmA** and **LmB**) on the 'linear' measurement scale (ϑ_i, before exponentation). In light of the discussion above, the Prasad-Rao (1990) estimation strategy for SAE's will be properly applicable to SAIPE data only after modification of the SAE estimators themselves to adjust for truncation bias, and of the *mse*'s to account for nonlinear transformation and possibly also for truncation. Both of these modifications require some further research, which we are undertaking.

In the rest of this article, we study an alternative methodology for producing SAE's without the need for truncation, but with a binomial-normal generalized linear model with random effect. These alternative SAE's are compared with those arising from the truncated Fay-Herriot model, for actual SAIPE data.

Table 1. SAE biases on linear measurement scale ϑ_i for two Fay-Herriot analysis methods, for complete (Table 1a) and truncated (Table 1b) data.

Table 1a: SAE biases for complete data.

Model	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
LmA	$-.0155$	$-.00354$	2.19×10^{-5}	-1.06×10^{-4}	.00346	.0151
LmB	$-.0205$	$-.00325$	1.84×10^{-5}	9.08×10^{-5}	.00330	.0187

Table 1b: SAE biases under (20%) truncation.

Model	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
LmA	-.1006	.01352	.04658	.05329	.09046	.31030
LmB	-.1001	.01308	.04647	.05623	.09518	.32950

3 SAIPE Data Analyses: Truncated Fay-Herriot vs. GLM

As described in Bell (1997) and Citro & Kalton (1999), a large number of exploratory analyses have been performed on SAIPE data, with a view to choosing the best available model specification and set of predictors from the variables derived from IRS county-aggregated data, CPS sample data, and long-form county aggregates from the most recent decennial census. The predictors which have been agreed upon vary a bit according to whether the response variable is chosen to be the log-count of child poor or log-rate of child poverty. For purposes of comparability between log-linear and GLM unit-level models, we restrict attention throughout to linear (Fay-Herriot-type) models of log child-poverty *rate*. Then the predictors are:

LTAXRT = logarithm of the current-year IRS-estimated Child Poverty Rate for the county;

LSTMPRT = logarithm of the current-year county Food-Stamp Participation rate;

LFILRT = logarithm of current-year county IRS Child Tax-Exemptions divided by current-year Population Estimate;

LCPRT = logarithm of county Poverty Rate for residents aged 5–17 from the latest decennial census.

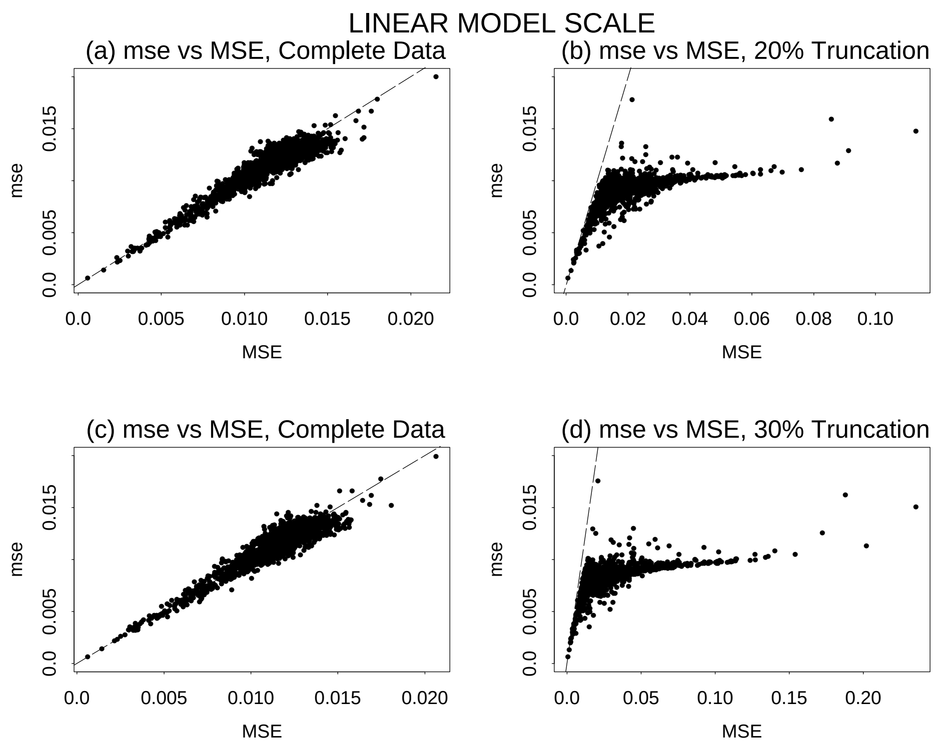

Figure 1: Scatterplots of *mse* versus MSE of linear-scale SAE's of ϑ_i, for four simulations with complete and truncated (20% or 30%) data. Each point plotted corresponds to one single PSU. In each plot, the dashed line is the 45° line $mse =$ MSE. (a) Case of complete data. (b) Subset of 20%-truncated dataset from case a. (c) Independently generated dataset, case of compete data. (d) Subset of 30%-truncated dataset from case c.

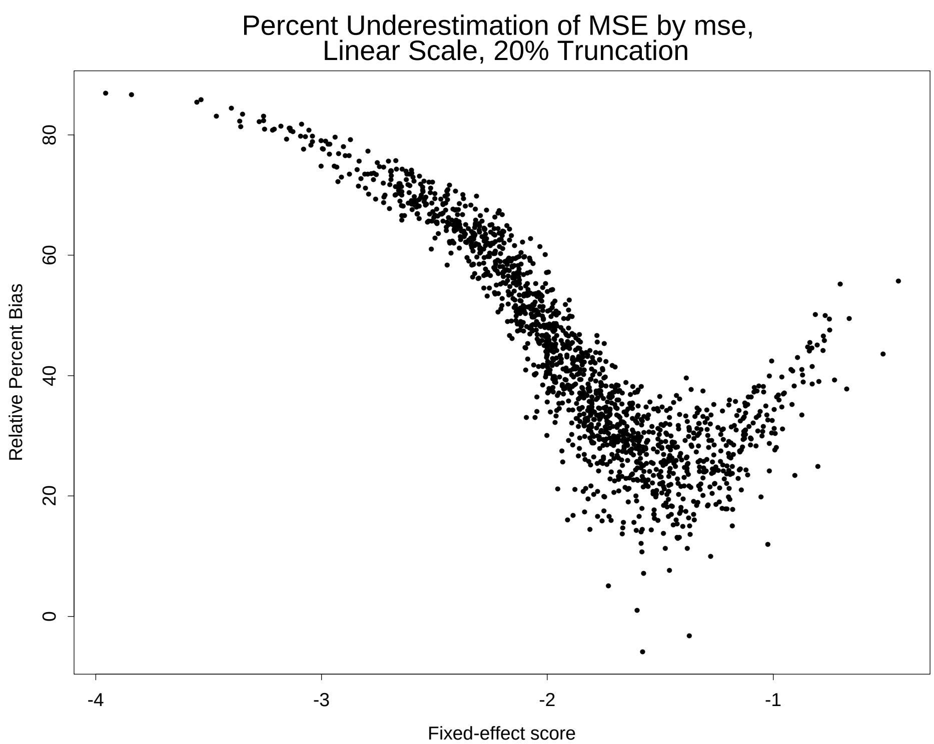

Figure 2: Scatterplot of relative bias of *mse* for linear-scale SAE's $\hat{\vartheta}_i$ plotted, by PSU, versus fixed-effect scores $x_i^{\rm tr}\beta$. The relative bias plotted on the y-axis for the i'th PSU is $100(1 - mse(\hat{\vartheta}_i)/MSE(\hat{\vartheta}_i))$.

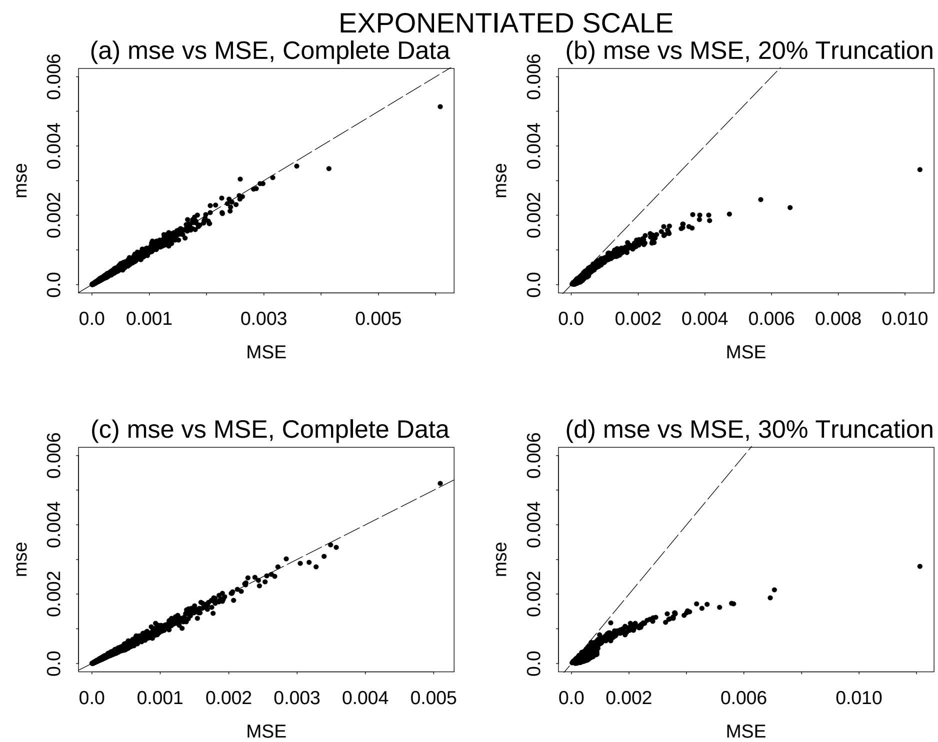

Figure 3: Scatterplots of *mse* versus MSE of exponentiated-scale SAE's of ϑ_i, for four simulations with complete and truncated (20% or 30%) data. Each point plotted corresponds to one single PSU. In each plot, the dashed line is the 45° line $mse = \text{MSE}$. (a) Case of complete data. (b) Subset of 20%-truncated dataset from case a. (c) Independently generated dataset, case of compete data. (d) Subset of 30%-truncated dataset from case c.

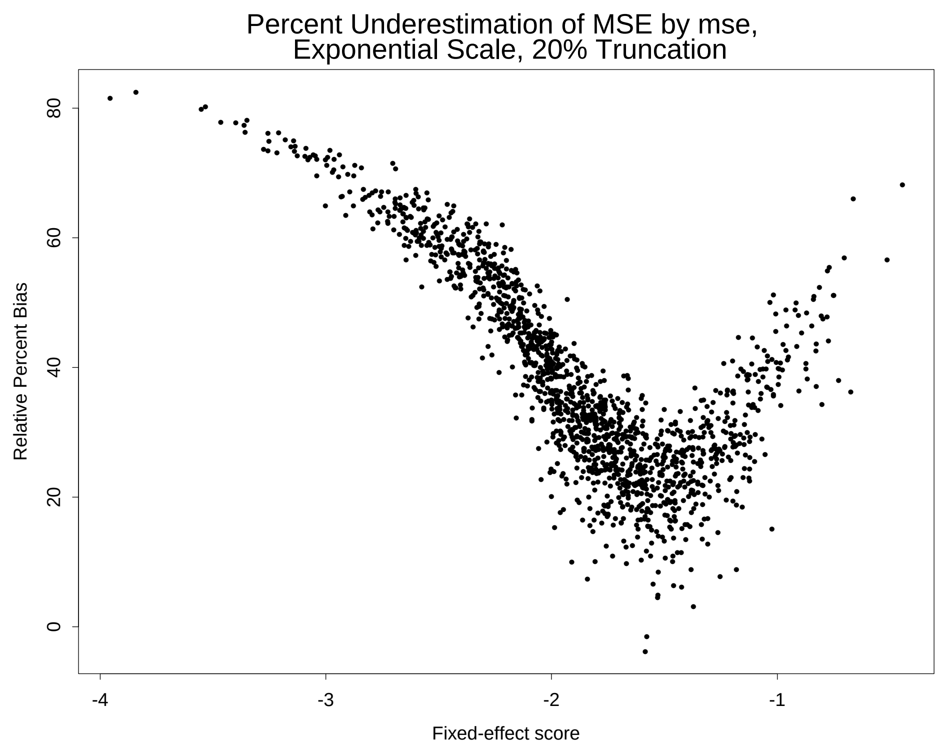

Figure 4: Scatterplot of relative bias of mse for exponentiated-scale SAE's $\hat{\vartheta}_i^*$ plotted, by PSU, versus fixed-effect scores $x_i^{\mathrm{tr}}\beta$. The relative bias plotted on the y-axis for the i'th PSU is $100(1 - mse(\hat{\vartheta}_i^*)/MSE(\hat{\vartheta}_i^*))$.

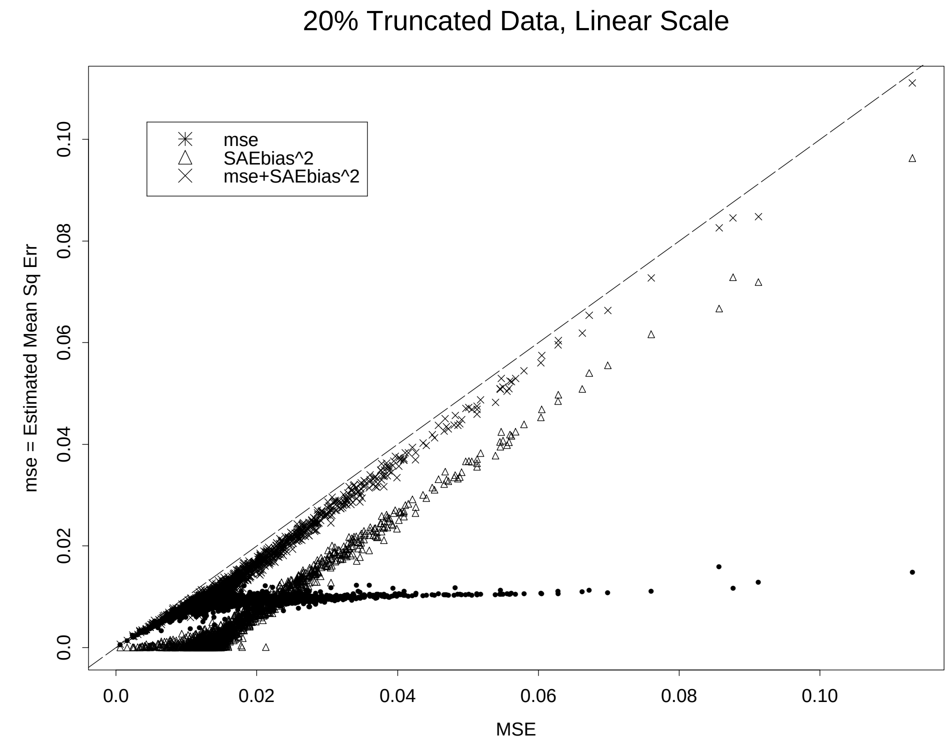

Figure 5: Scatterplots on a single set of axes of three quantities related to *mse* for the linear-scale SAE's $\hat{\vartheta}_i$ versus $MSE(\hat{\vartheta}_i)$. On the y-axis are successively plotted: *mse* (solid dot), squared bias of SAE (hollow triangle), and the sum of *mse* and SAE-bias squared ($\times$ symbol). Dashed line is the $45°$ line.

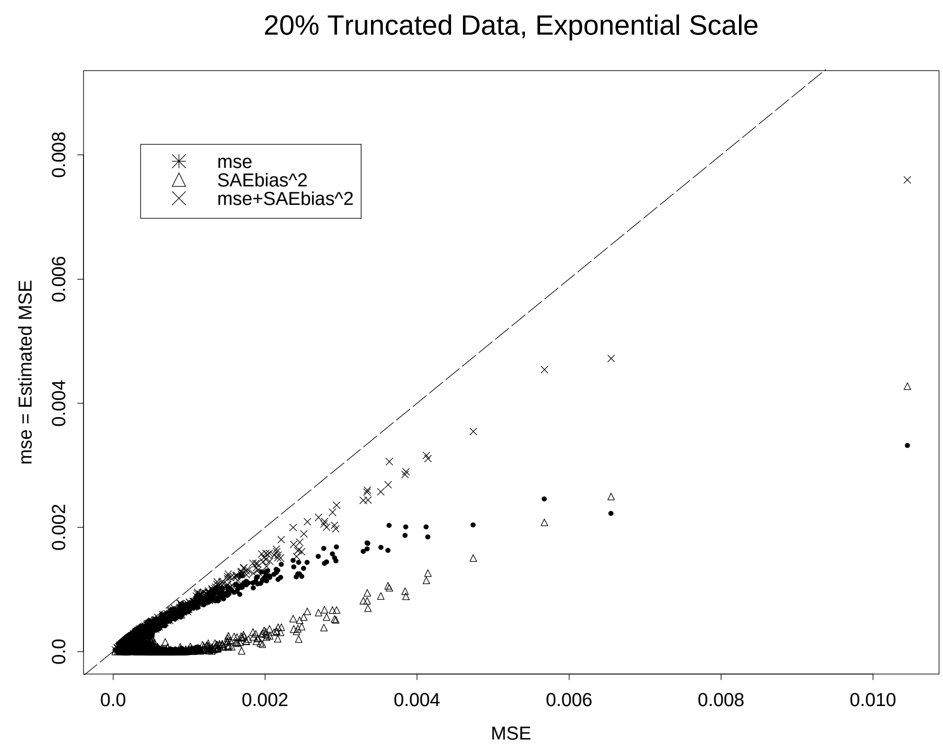

Figure 6: Scatterplots on a single set of axes of three quantities related to *mse* for the exponentiated-scale SAE's $\hat{\vartheta}_i^*$ versus $MSE(\hat{\vartheta}_i^*)$. On the y-axis are successively plotted: *mse* (solid dot), squared bias of SAE (hollow triangle), and the sum of *mse* and SAE-bias squared ($\times$ symbol). Dashed line is the 45° line.

The response-variable is the logarithm of the CPS-estimated fraction of the county 5–17 year-olds related to the principal householder of households in poverty. This response, in the current SAIPE methodology, is based upon a 3-year aggregation of CPS samples (including those of the year before and the year after which a county-level child-poverty rate is desired). Similarly, the numbers of sampled related county children and the subset of poor children would be aggregated over a 3-year window of CPS samples. Thus, in the data referred to below as the '1994 SAIPE dataset', CPS samples for the years 1992–1994 are aggregated for the purpose of producing county-level child poverty estimates for the middle of the 'current' year 1993. Similarly, the '1990 SAIPE dataset' refers to CPS aggregates over 1988–1990, for estimation as of mid-1989. The predictor variables used in the simulation of Section 2 above were precisely the 1488 four-dimensional vectors $x_i^{tr} = (1, \text{LTAXRT}, \text{LSTMPRT}, \text{LFILRT}, \text{LCPRT})$ for 1992–1994 CPS-sampled counties, and the parameters (given only to one significant figure, and including intercept as first fixed-effect) $\beta = (0.0, 0.2, 0.3, -0.1, 0.4)$, $\sigma_u^2 = 0.01$, $v_e = 10$ used in the simulations were the ones fitted in 1997 to the log-rate responses for the 1994 SAIPE dataset. The Fay-Herriot model currently used by SAIPE is equation (1) above, with n_i denoting the number of sampled households in CPS. However, since the relevant households for SAIPE are really only the ones with related children, in what follows we use the same model with n_i denoting instead the (estimated) number of CPS sampled children.

For statistical analysis with the mixed logistic unit-level model, numbers of related children and related poor children are required for each sampled county. During much of the present research, these numbers were not available, and we used instead the CPS 3-year sample-weighted estimates of children per HU multiplied by the actual sampled number of HU's, which were fractional, not integers. We have now re-done all analyses using the exact numbers of CPS-sampled children and poor children, and the results are essentially identical. However, the Tables and Figures displayed in this paper were calculated using the sample-weighted CPS estimates.

3.1 Internal Evidence from 1994 & 1990 SAIPE Data

In the SAIPE methodology based upon Fay-Herriot models, the county-level random effect variance σ_u^2 in model (1) is assumed known, from a previous modelling stage using Census data, described by Bell (1997) and Citro & Kalton (1999). Usually, Fay-Herriot (1979) models assume instead that the quantity v_e represents sampling error and is known for that reason. However, it is not true, strictly speaking, that either of these variance parameters can be regarded as known precisely. In the present study, we first estimate county-level random-effect variance σ_u^2 within the mixed-effect logistic regression model

$$y_i \sim Binom(n_i, \pi_i) \quad , \quad \log\left(\frac{\pi_i}{1-\pi_i}\right) = x_i^{tr}\, b + u_i \tag{5}$$

where x_i, u_i have the same meanings as in (1), and the unknown fixed-effect coefficients are now b. This analysis is referred to below as the Generalized Linear Model, mixed-logistic, or simply **Glm** analysis method. The parameters (b, σ_u^2) are estimated via Maximum-likelihood, using the accurate numerical log-likelihood approximations described in Slud (2000b). The random-effect variance σ_u^2 is then fixed and treated as known, and (using only the data on PSU's with positive number of CPS-sampled child poor) β, v_e are estimated via Maximum Likelihood within the Fay-Herriot model (1), an analysis method which we refer to as Linear Model B, or simply **LmB**. Next, fixing this value v_e estimated within **LmB** as though known, we estimate β, σ_u^2 via maximum likelihood within model (1), still using only data on PSU's with postive number of sampled child poor. This latter analysis is referred to as Linear Model A, or **LmA**. Thus the σ_u^2 numbers match by design for **Glm** and **LmB** in Table 2 below, and the v_e numbers match for **LmB** and **LmA**, and these models were fitted in the order **Glm**, **LmB**, **LmA**. Because of its similarity to the SAIPE Fay-Herriot analysis with fixed (assumed known) σ_u^2, **LmB** is taken to be the main linear-model competitor to **Glm**. However, for completeness we also include some estimates and comparisons made with another Fay-Herriot style model like **LmB** (fitted to the same dataset consisting of counties with no child-poor) but with σ_u^2 fixed at the value derived from earlier SAIPE analyses of Census data, as part of the 'bivariate model' expounded by Bell (1997). For analyses based upon either the 1994 or the 1990 SAIPE dataset, the σ_u^2 value fixed was the same (0.01 up to one significant figure). In each case, we re-fitted the model with this fixed σ_u^2 value by exactly the same method, on exactly the same data, as for **LmB**: we refer to this form of Fay-Herriot linear-model analysis as **LmS**.

The form of small-area estimators $\hat{\vartheta}_i^*$ generated from the linear models were given in formula (2) above. The small-area estimators based on the

Generalized Linear Models have the form

$$\hat{\vartheta}_i^* \;=\; \int \frac{e^{(\gamma+\sigma z)(y_i+1)}}{(1+e^{(\gamma+\sigma z)})^{n_i+1}}\,\phi(z)\,dz \Big/ \int \frac{e^{(\gamma+\sigma z)\,y_i}}{(1+e^{(\gamma+\sigma z)})^{n_i}}\,\phi(z)\,dz \qquad (6)$$

where values $\gamma = x_i^{tr}\,\hat{\beta}$ and $\sigma = \hat{\sigma}_u$ are substituted; where ϕ is the standard normal density; and where the values y_i, n_i are taken to be 0 in PSU's with no sample. These SAE's are as given by Slud (1999), and are *not* bias-corrected.

We begin our description of the 1994 SAIPE data-analysis by showing in Table 2 the estimators of fixed-effect coefficients and variance components for all of the modelling approaches described above.

Table 2. Coefficient and variance-component estimates in 1994 SAIPE dataset for all of the models discussed in the text. The model labelled **Saipe** is the log-rate model fitted (as one of the many alternatives mentioned in Citro & Kalton 1999) on (an earlier version of) the '94 SAIPE data.

Parameter	**Glm**	**LmA**	**LmB**	**LmS**	**Saipe**	**GlmW**
Intercept	.726	-.108	-.104	.032	.0	1.214
LTAXRT	.390	.390	.394	.278	.2	.722
LSTMPRT	.406	.211	.210	.294	.3	.520
LFILRT	-.318	-.357	-.359	-.370	-.1	-.881
LCPRT	.441	.296	.294	.369	.4	.371
σ_u^2	.549	.404	.549	.01	.01	1.340
v_e	0	.762	.762	10	10	0

The first major point to be derived from careful examination of the SAIPE 1994 model fit is the rather different behavior of residuals for child-poverty rates for CPS-sampled counties with and without child poor. This is hardly surprising for **LmA**, **LmB**, and **LmS** because these models are fitted using only the data from counties with sampled child poor. But there is a systematic bias between the fitted child poverty rates in the **Glm** analysis and those provided in the truncated-data Fay-Herriot models, which clearly distinguishes those sampled counties with and without sampled child poor. First, in Figure 7 it is clear that all estimated rates are over-estimates in counties with no sampled child-poor, but it is slightly less easy to anticipate that relatively few counties with child poor would have estimated rates

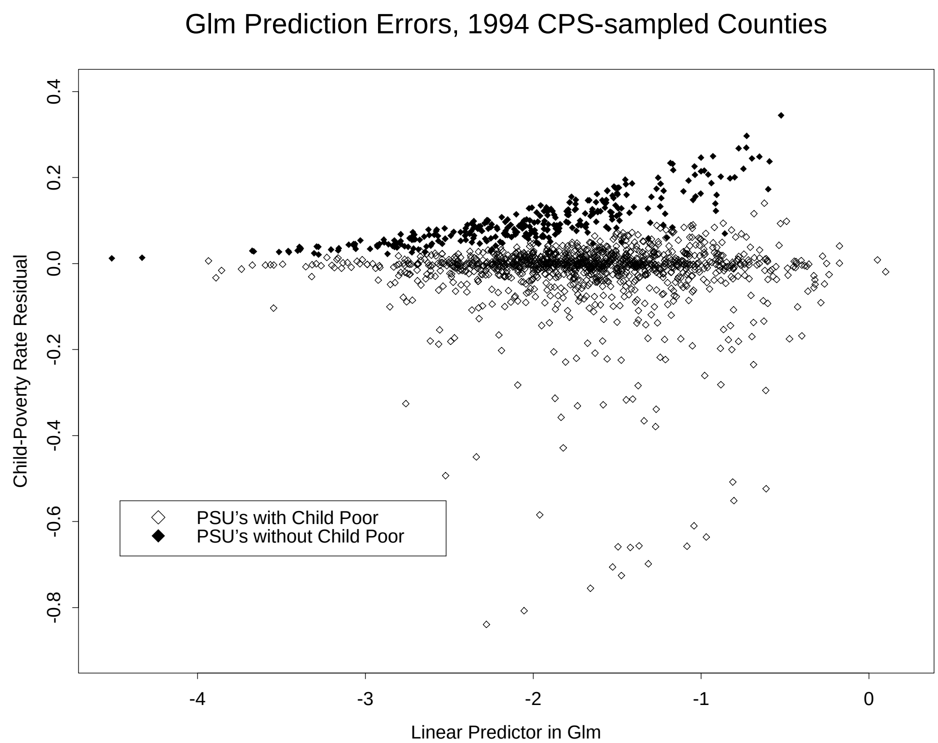

Figure 7: Scatterplot of residuals between SAE's and CPS-estimated child poverty rates based on 1994 SAIPE county data, where counties *with* sampled poor children are plotted with hollow triangles, and counties *without* sampled poor children are plotted with solid triangles.

which are too large. (The same pattern obtains in the 1990 SAIPE dataset.) A more meaningful distinction between predictability of child poverty rates by county is given by Figure 8. There we plot versus the 1990 Census child-poverty rate the county-by-county differences between the *logit* of estimated 1994 child-poverty rates by **Glm** and by **LmB**. (That is, the y-ordinate for the i'th county is $\log((\hat{\vartheta}_i^*)^{Glm}/(1-(\hat{\vartheta}_i^*)^{Glm})) - \log((\hat{\vartheta}_i^*)^{LmB}/(1-(\hat{\vartheta}_i^*)^{LmB}))$.) Figure 8 shows that there are systematic differences between the way in which the **Glm** and **LmB** models predict within sampled counties which happened to have no sampled poor (related, school-age) children as compared with those that did have sampled poor children. While both the fitted **Glm** and **LmB** child-poverty-rate values had average logit of -1.5

among counties with poor children (slightly greater than the counties' average logit child-poverty rate found in the 1990 decennial census), the **Glm** had average logit rate of -2.36 (much lower than the 1990 census value of -1.97) among the counties which in 1993-5 had CPS sample but no sampled poor children, while the **LmB** average logit estimated rate had the much higher value of -1.52 (more or less the *same* as the **LmB** and **Glm** average logit rate on counties with sampled poor children). For application of model predictions to non-sampled counties, it is much more plausible to use a method which tends to give lower child-poverty-rate predictions to the sampled counties without sampled poor children. On this account, we might expect **Glm** SAE's to outperform those from **LmB** on the non-sampled counties.

The figures analogous to Figures 7 and 8, based on fitting-method **LmA** in place of **LmB** or on the 1990 SAIPE dataset in place of the 1994 dataset, look very similar to Figures 7 and 8, and are not shown. As a further indication of the same behavior cited above, we show in Figure 9 the scatter-plot of logit **Glm** estimated child-poverty rates versus the logit **LmA** estimated rates for CPS-sampled (1988-1990) counties, both fitted on the 1990 SAIPE data. In this picture, the counties without sampled poor children are plotted as solid diamonds, and those with poor children are plotted as hollow diamonds. Again, it is very clear that the two sets of counties are effectively distinguished by their estimated rates under **Glm** versus **LmA**.

We now undertake a series of comparisons between the fits to the 1990 and 1994 CPS sample-weighted child-poverty rate estimates using only the 1990 and 1994 SAIPE data, respectively. These comparisons are *internal*, examining fit based on the same dataset used in fitting, and are expressed through various *loss criteria*, defined as follows:

$$\begin{aligned} SSQ &= \sum_i (\hat{\vartheta}_i^* - \vartheta_i^{*0})^2 \\ WtSSQ &= \sum_i n_i\,(\hat{\vartheta}_i^* - \vartheta_i^{*0})^2 \\ Abs &= \sum_i |\hat{\vartheta}_i^* - \vartheta_i^{*0}| \\ WtAbs &= \sum_i n_i\,|\hat{\vartheta}_i^* - \vartheta_i^{*0}| \end{aligned}$$

Here $\hat{\vartheta}_i^*$ denotes one of the SAE's under study for county child-poverty rate, and ϑ_i^{*0} denotes a county rate to be used as standard of truth. In the present subsection, the standard is the CPS-estimated child-poverty rate. In the next subsection, the standard of truth is the county rate estimated from the 1990 census. Throughout, n_i denotes the (CPS-estimated) number of related school-age children sampled in the i'th county. The summation-range for county-indices i is first of all the set of m sampled counties, but the same loss-criteria are also separately calculated for the sets of counties with and without sampled poor children. These losses are calculated for the estimation methods described above, with one additional method denoted **GlmW**. This method is based upon parameter estimation in the full dataset, by the same maximum-likelihood calculation for model (5) done in **Glm**, except that the log-likelihood term for county i is downweighted by a factor $1/n_i$. There are two motivations for introducing this weighted alternative to **Glm**. First, we note that in the likelihood which is maximized for the Fay-Herriot model (1), each county enters with equal weight, the only distinction between counties based on sample-size being made through the sampling-error variance term v_e/n_i. On the other hand, the **Glm** model (5) is a *unit-level* model, which (due to the identical covariates used over each county) enters n_i log-likelihood terms for county i, thereby giving larger counties greater influence in determining the fixed-effect coefficients and variance components than is done in the Fay-Herriot model MLE's. A second reason for expecting that **GlmW** might have value in the SAIPE context is the possibility, discussed more fully in subsection 3.2 below, that the sampled and unsampled sets of counties may satisfy slightly different models (of either the Fay-Herriot or mixed-effect logistic type), and since the large counties are always sampled in CPS while the small ones rotate in and out of the sample over time, an analysis method like **Glm** which (by virtue of its definition as a unit-level model) weights the log-likelihood contributions of large counties too heavily, might distort the fitted parameters in a way which would cause poorer performance of SAE estimates on *un*sampled counties. Generally speaking, the SAE's for large sampled counties are always reasonably good, for any of the models, since they resemble the direct sample-based estimators.

Tables 3 and 4 respectively show the loss-values for the 1994 and 1990 SAIPE datasets, due to discrepancies between SAE's of county child-poverty rates and the CPS-estimated values. The immediate conclusion, by all loss-criteria, is that the **Glm** method performs better than **LmA** and **LmB** on the counties with no child poor, while **LmB** is uniformly the best of the three methods for the counties with sampled poor children. Similarly, **LmS** performs uniformly worse than **LmB** and **LmA**, in counties with poor children, but better in counties without them. On the set of all counties, **Glm** performs better than **LmA** and **LmB** with respect to the squared-discrepancy loss criteria, but worse with respect to the absolute-error criteria. However, **GlmW** clearly outperforms all other methods in this internal comparison, by every criterion, in every case.

Table 3. Loss-criterion values for 5 SAE methods versus CPS (weighted sampled-based) estimates, from 1994 SAIPE data, for all 1488 sampled counties, for all 1184 counties with sampled poor children, and for all 304 counties without sampled poor children.

	Glm	**LmA**	**LmB**	**LmS**	**GlmW**
All counties					
SSQ	15.85	17.50	18.04	40.84	9.69
WtSSQ	91.12	142.44	158.39	682.54	42.39
WtAbs	1221.5	857.2	856.4	4942.7	726.1
Abs	79.96	75.17	75.45	163.24	55.14
With child poor					
SSQ	11.91	5.44	4.00	32.68	7.30
WtSSQ	66.08	14.32	9.18	597.22	30.14
WtAbs	949.0	209.4	157.3	4422.3	547.4
Abs	49.78	20.15	16.08	118.76	32.90
No child poor					
SSQ	3.94	12.05	14.04	8.16	2.39
WtSSQ	25.05	128.12	149.21	85.31	12.25
WtAbs	272.5	647.8	699.0	520.3	178.8
Abs	30.17	55.01	59.37	44.48	22.24

3.2 Explaining Patterns of Difference between the SAE's

Why do the log-linear (Fay-Herriot) estimating methods perform better than the mixed-effect logistic on the counties with sampled poor children, and worse on the others ? There are several important differences between these two classes of models, and it is desirable to understand which differences have impact on the performance of SAIPE estimation.

The differences between the (log-) linear model methods and the **Glm** methods fall into the following categories: (i) aggregate-PSU versus unit-level specification, (ii) truncation of dataset (deletion of sampled counties with 0-counts of poor children) in the **Lm** model, and (iii) flexibility of specification in incorporating external information. We discuss each of these categories in turn.

Table 4. Loss-criterion values for 5 SAE methods versus CPS (weighted sampled-based) estimates, from 1990 SAIPE data, for all 1259 sampled counties, for all 1028 counties with sampled poor children, and for all 231 counties without sampled poor children.

	Glm	**LmA**	**LmB**	**LmS**	**GlmW**
All counties					
SSQ	5.64	8.36	8.91	22.11	3.29
WtSSQ	58.33	108.39	117.92	582.5	27.81
WtAbs	1108.5	747.1	752.3	4802.7	726.3
Abs	47.72	44.70	45.32	117.79	33.61
With child poor					
SSQ	3.66	0.95	0.73	16.89	2.22
WtSSQ	40.67	4.46	3.18	512.9	19.50
WtAbs	872.0	142.2	116.4	4331.7	566.0
Abs	29.46	7.66	6.40	88.10	20.44
No child poor					
SSQ	1.99	7.41	8.18	5.21	1.08
WtSSQ	17.66	103.93	114.74	69.60	8.31
WtAbs	236.5	604.9	635.8	471.0	160.3
Abs	18.26	37.04	38.92	29.68	13.17

(i). The unit- or aggregate-level specification has an immediate impact on the effective weighting of PSU's in the respective Maximum Likelihood estimation of parameters. While the Fay-Herriot model allows PSU sample-size n_i to enter the variance explicitly, it weights each PSU equally; by contrast, the unit-level models give equal weight to each sample unit, so that there are n_i equal log-likelihood contributions in the i'th PSU.

(ii). The specifications of all model types implicitly or explicitly involve a nonlinear transformation of the true and sampled counts of related school-age children and related school-age poor children. The *log* transformation of the **Lm** methods dictates analysis without the sampled counties having 0 counts of poor children. The binomial response and logistic link-function of the **Glm** models implicitly determine the form of the variance of the estimands ϑ_i and ϑ_i^* as functions of the covariates.

(iii). The **Lm** methods allow aggregate-level information on variance-components to fix or restrict parameter-values if available, and allow the possibilty of generalized variance function models to change the specific parametric form of the variance components. The **Glm** (logistic) variance-structure is determined by the independence of unit responses and the link-function, and is not easy to change except through alternative links or further modelling of actual household- or cluster- dependence among CPS-sampled children within county.

To check whether (i) accounts in significant degree for the differences in performance of **Glm** and the **Lm** methods in Tables 2–3 (with CPS-estimated rates as the standard), we included **GlmW** in those and all further comparisons. This was done because the downweighting of counties with large sample in **GlmW** accords with the equal county-level weighting in the Fay-Herriot log-likelihood. The loss-values of **GlmW** accord much more closely with those of **LmA** and **LmB** than do those of **Glm**, with respect to the census standard (Tables 5 and 6) below, but **not** in the comparisons using CPS-estimated rates as standard. On the other hand, in the latter comparisons, the **GlmW** method shows closer agreement to CPS-estimated rates than **Glm**, both with weighted and unweighted loss criteria.

We cannot say simply that the downweighting brings **GlmW** closer to the **Lm** methods. The effect of truncation (item (ii) above) in the **Lm** parameter estimations has no counterpart in the GLM analyses. It seems likely, however, that there is an interaction (in the SAIPE 1990 data) between the effect on GLM SAE's of the different weightings and the difference between Census- and CPS- estimates of the child poverty rates. Indeed, the correlation between the 1259 values $\vartheta^{*0}_{i,Census} - \vartheta^{*0}_{i,CPS}$ and the SAE differences $\hat{\vartheta}_i^{*,Glm} - \hat{\vartheta}_i^{*,GlmW}$ is calculated to have the astonishingly high value of 0.917. The scatterplot of the close countywise relationship between these two differences is given in Figure 10, where the relationship is also seen not to differ much between those counties with sampled child poor and those without. However, further plots (not shown) and correlations indicate that neither of the two quantities plotted has a monotone relationship with number of sampled children or with the predictor variables used in the models.

Our tentative conclusion is that the pattern of different performance between Fay-Herriot and GLM SAE methods is due to the truncated dataset used by the former, to the implicit differences in PSU-weighting used by the two methods, to the systematic differences between counties with moderate and low child-poverty rates, and to the systematic countywise differences in Census and CPS definitions of child-poverty rates. But we have not been able to parse the relative importance of these different effects.

3.3 Model-based estimators versus 1990 Census estimates

We turn now to external comparisons between the various model-based SAE's described above, with special reference to the 1990 SAIPE dataset where the 1990 decennial census estimates of county child-poverty rates can be used as a standard of truth. As pointed out by Bell (1997) and in Citro & Kalton (1999), the definition of child-poverty rate within the decennial census does differ somewhat from that of the CPS (in the ages of the children considered, the CPS criterion that sampled children must related to the principal householder, etc.), but the consensus seems to be that the corresponding rates are close enough that the SAE's can be judged essentially by how close they come (both on sampled and unsampled counties) to the census estimates.

Since the census child-poverty rates are never 0, unlike those estimated from CPS samples, we supplement the loss-criteria already considered with

$$AbsRel \;=\; \sum_i \, |\hat{\vartheta}^*_i / \vartheta^{*0}_i \,-\, 1|$$

Consider first the counties in the 1990 SAIPE dataset, i.e., those sampled by CPS from 1988-1990. Table 5 displays the 5 loss-criteria for each of the 5 methods of constructing SAE's of county child poverty rates, with respect to the 1990 census estimated values. Now **Glm** performs better than **LmA** and **LmB** in counties with sampled poor children, and worse in counties without them. This pattern is exactly the reverse of the one shown in Table 4. However, **Glm** is now better than **LmA** and **LmB** with respect to all loss-criteria, on the set of all sampled counties. Moreover, **Glm** now uniformly outperforms **GlmW** in all cases. However, in this external comparison of sampled counties,(with census as standard, the SAIPE-style method **LmS** (fitted like **LmB**, but with county random-effect variance determined in Bell's 1997 bivariate model from 1980 decennial census data) is strikingly better than all the other methods, by all loss-criteria.

Finally, we turn to the performance of the various SAE methods on the set of counties which were not sampled by CPS in the period 1988-90, using the 1990 census-estimated rates as the standard. Table 6 gives the results.

In the non-sampled counties, as in the sampled counties with census-estimated rates as the standard, **LmS** is again the best of the **Lm** methods, and **Glm** clearly outperforms **GlmW**. But now, **Glm** shows an advantage over **LmS**: more than 30% by the SSQ and WtSSQ loss-criteria, and 8–12% by the other criteria.

Table 5. Loss-criterion values for 5 SAE methods, from 1990 SAIPE data, with respect to Census 1990 estimates: for all 1259 sampled counties, for all 1028 counties with sampled poor children, and for all 231 counties without sampled poor children.

	Glm	**LmA**	**LmB**	**LmS**	**GlmW**
All counties					
SSQ	8.63	14.04	14.82	1.37	11.87
WtSSQ	409.1	552.7	564.7	60.42	491.9
WtAbs	4397.4	5001.3	5055.5	1629.0	4814.7
Abs	77.65	95.55	98.43	27.86	91.81
AbsRel	509.5	666.1	689.9	179.9	611.1
With child poor					
SSQ	7.65	13.55	14.17	1.09	10.11
WtSSQ	389.6	545.9	555.4	57.17	460.2
WtAbs	4163.5	4857.9	4885.2	1561.7	4505.9
Abs	66.16	87.03	88.32	23.00	75.67
AbsRel	423.0	577.5	586.7	139.28	488.7
No child poor					
SSQ	0.98	0.49	0.65	0.28	1.76
WtSSQ	19.43	6.81	9.28	3.25	31.66
WtAbs	233.89	143.39	170.36	67.22	308.75
Abs	11.48	8.52	10.11	4.86	16.14
AbsRel	86.53	88.51	103.22	40.59	122.42

The excellent performance of **LmS** SAE's with respect to a decennial-census standard of truth, and rather poor performance with respect to current CPS data, has a rather clear explanation. It turns out that **LmS** on the 1990 data has a much larger coefficient for the predictor LCPRT (the log child-poverty rate derived from the previous decennial census) than do the other models. For example, the coefficient in **LmB** was 0.252 with estimated standard deviation 0.089, while for **LmS** the coefficient is 0.387 with estimated standard deviation 0.047. Moroever, this feature of the **LmS** model is not accidental. Since it is prescribed to have a very small σ_u^2 $(= 0.01)$ compared to the other models (such as 0.553 for **LmB**), its parameter-estimates give large v_e estimates and large influence to the PSU's with large n_i, where the direct CPS child-poverty rate estimates are most likely to be in close accord with recent decennial estimates. This point is worth dwelling on: the correlation between county-level child-poverty rates for the 1980 and 1990 censuses, for counties which were CPS-sampled in 1988-90, is 0.88, while the correlation between 1990 CPS unweighted child-poverty rates and those from the 1990 census is only 0.50, while the correlation between census and CPS 1990 rates for counties with at least 50 children in sample in 1988-90 (of which there were 500) was 0.73. Thus, the heavy influence of large counties in **LmS** due to its large v_e and small σ_u^2 values, result in high correlation with census-estimated rates. However, in smaller counties and counties where current CPS data is not in conformity with the last censal child-poverty rates, **LmS** is not reliable. This comment reinforces the good behavior of **Glm** with respect to either censal or CPS standards of truth.

Without drawing definitive conclusions from these results, we do find systematic differences between the precise models applicable to the sampled and nonsampled counties. These have been reflected by the contrasting performance of the various models with respect to county weighting, size of county random-effect variance, and data-truncation, within different sets of counties with respect to the different (CPS or census) standards used. If the dual objective of analysis is at the same time to reflect accurately the CPS sample-weighted estimates of child poverty rate within sampled counties, and also to track closely with the most current decennial-census estimates, then **Glm** should perhaps be the recommended method. However, if a hybrid method can be acceptable, *and* if close agreement with the most current decennial census rate-estimates is deemed more important than agreement with current CPS sample-based estimates, then the Tables indicate that **LmS** be used on CPS-sampled counties and **Glm** on the non-sampled ones.

Table 6. Loss-criterion values for 5 SAE methods, fitted from 1990 SAIPE data, with respect to Census 1990 estimates for all 1870 counties not sampled by CPS in 1988–90.

	Glm	**LmA**	**LmB**	**LmS**	**GlmW**
SSQ	3.26	5.83	8.02	4.79	5.04
WtSSQ	7683	16463	24465	11054	13005
WtAbs	183377	293541	373105	208820	228314
Abs	310.3	498.6	594.3	336.24	321.2
AbsRel	55.50	81.55	99.25	63.76	67.03

4 Conclusions and Recommendations

The main conclusions of this research are as follows:

(1) In the SAIPE setting, SAE biases within the Fay-Herriot log-rate model due to analysis with truncated datasets are small but systematic. The mse formulas developed by Prasad and Rao (1990), adapted to the first-order bias-correction for the exponential transformation from the log-rate model to the rate SAE's, can reasonably estimate the variance but not the MSE of the county estimates.

(2) The SAE's based on the weighted GLM method **GlmW** generally outperform all competing linear-model (Fay-Herriot) and **Glm** competitors with respect to internal loss-function criteria, on the complete set of CPS-sampled counties.

(3) Choosing a large county random-effect to resemble in order of magnitude the random effect in the **Glm**, as is done artificially in **LmA** and **LmB**, leads to SAE methods which fit the CPS rate-estimates data (both in counties with and without sampled poor children) better than the SAIPE method which estimates a much smaller county random effect from the residuals of an analogous linear model fitted to the most recent decennial census log rates. However, the SAIPE method yields *much* better agreement than **LmA** and **LmB** with current census child-poverty rates in a decennial year, both in CPS-sampled and non-sampled counties.

(4) The **LmS** (current SAIPE) method yields much better agreement with the external Census standard than **Glm** in CPS-sampled counties, both in counties with and without sampled poor children, and **Glm** is much better than the weighted version **GlmW** and the linear model methods in

this setting. However, on the counties which were *not* sampled by CPS, **Glm** shows closer agreement with Census rates than all linear-model methods.

(5) Although the current SAIPE estimation method (essentially **LmS**) shows the closest agreement with current censal estimated child-poverty rates in the CPS-sampled counties, we have found this to be due primarily to **LmS**'s heavy reliance on previous censal estimates together with the higher county-by-county correlation between successive censal estimates than between censal and direct CPS estimates. Since a key objective of SAIPE estimation is to allow recognition of county-level changes in child-poverty rates after the previous decennial census, while maintaining good agreement with decennial-census estimates, we recommend the **Glm** method over **LmS** and the other linear-model methods.

(6) There is currently no good approach to MSE estimation in the SAIPE setting, due to the likely model misspecifications. Currently available methods of estimating MSE, based on the 'delta method', assume the specified GLM is valid. While these methods can be 'robustified' using sandwich variance estimators, they address only the SAE variance and cannot correct for model misspecification bias. Reliable MSE estimation may be easier to develop for the **Glm** estimation methodology than for the other methods, due to **Glm**'s slightly better fit and robustness.

5 REFERENCES

Bell, W. (1997) Models for county and state poverty estimates. Preprint, Census Statistical Research Division.

Citro, C. and Kalton, G., eds. (1999) **Small-Area Estimates of School-Age Children in Poverty**, Interim Report 3 (National Research Council), Washington DC: Nat. Acad. Press.

Datta, G. and Lahiri, P. (2001) A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems. *Statistica Sinica* **10**, 613-627.

Fay, R. and Herriot, R. (1979) Estimates of income for small places: an application of James-Stein procedures to census data. *JASA* **74**, 341-53.

Ghosh, M., Natarajan, K. Stroud, T. and Carlin, B. (1998) Generalized linear models for small-area estimation. *JASA* **93**, 273-82.

Ghosh, M., and Rao, J.N.K. (1994) Small Area Estimation: an appraisal. *Statist. Sci.* **9**, 55-93.

Prasad, N. and Rao, J.N.K (1990) The estimation of mean-squared errors of small-area estimators. *Jour. Amer. Statist. Asssoc.* **85**, 163-171.

Slud, E. (1999) Models for simulation and comparison of SAIPE analyses. Census Bureau preprint, posted at SAIPE web-site (see below).

Slud, E. (2000) Comparison of aggregate versus unit-level models for small-area estimation. *ASA Proc. Surv. Res. Meth.*, 563-8.

Slud, E. (2000b) Accurate calculation and maximization of log-likelihood for mixed logistic regression. Census Bureau preprint, posted at SAIPE web-site: **ftp.census.gov/hhes/www/saipe/tecrep.html**

Slud, E. (2002) Simulation Studies of Bias in Small-Area Point and MSE Estimates in SAIPE. Census Bureau preprint.

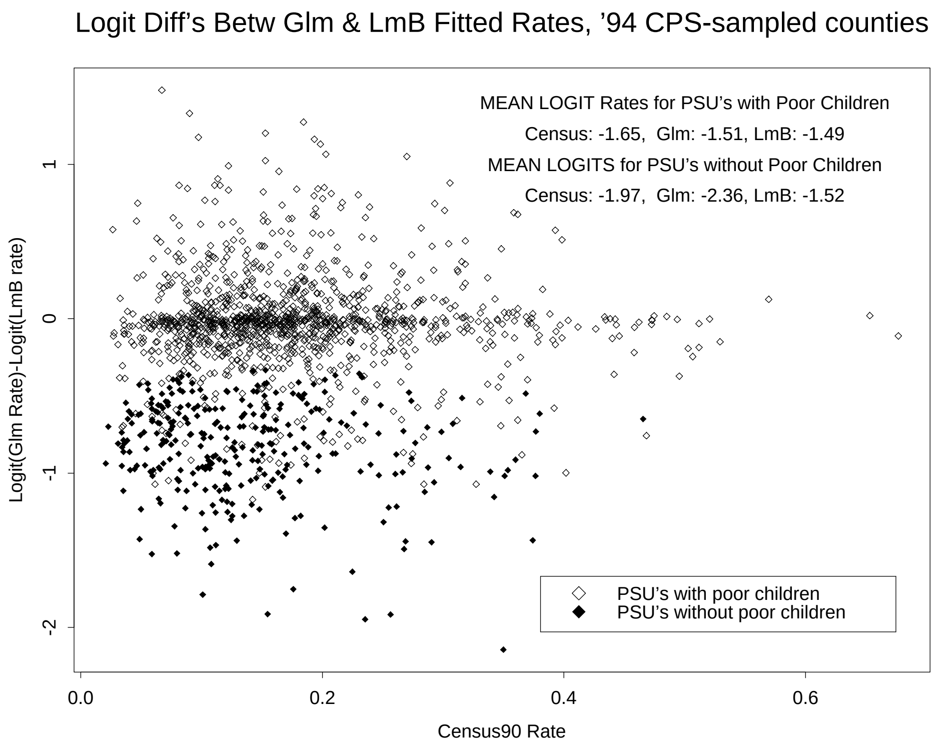

Figure 8: Scatterplot of differences between SAE's of 1994 child-poverty rates from **Glm** and from **LmB**, plotted for each county versus county child-poverty rate from the 1990 census. Counties *with* sampled poor children are plotted with hollow triangles, and counties *without* sampled poor children are plotted with solid triangles.

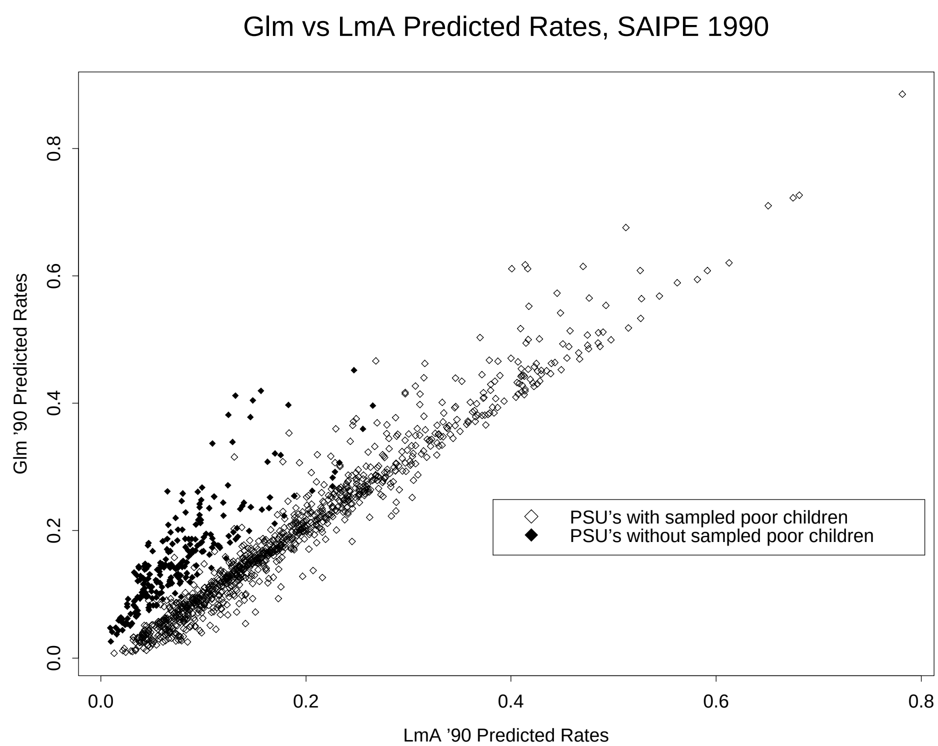

Figure 9: Scatterplot of Child Poverty-rate Predictors for **Glm** versus **LmA** based on 1990 SAIPE dataset. Counties *with* sampled poor children are plotted with hollow triangles, and counties *without* sampled poor children are plotted with solid triangles.

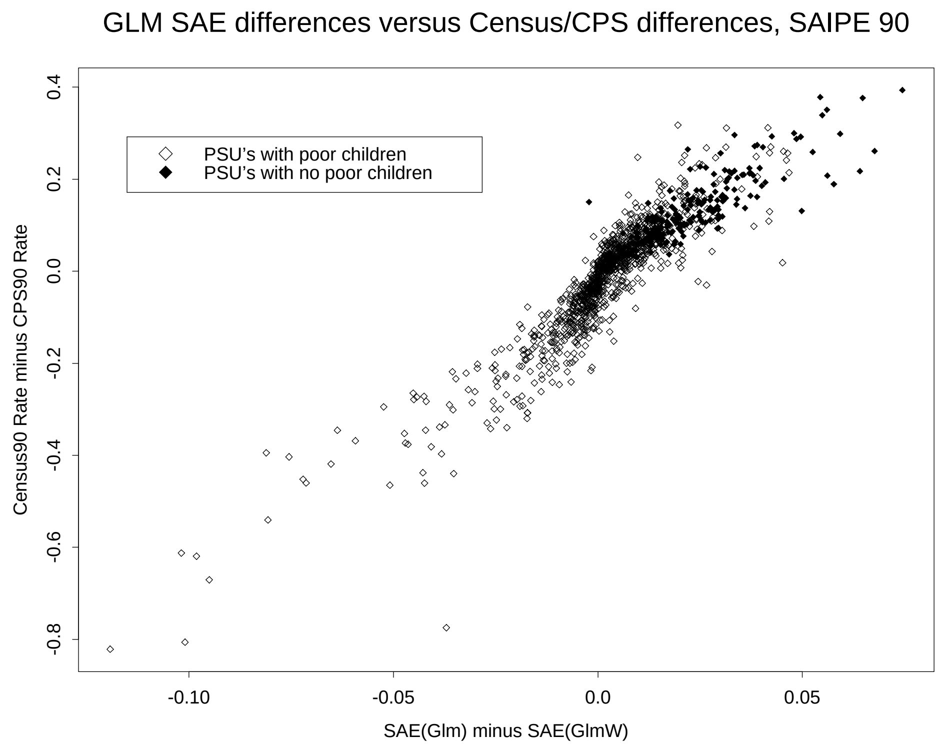

Figure 10: Scatterplot of differences between Census- and CPS- estimated Child Poverty-rates in 1990 versus differences between Predictors from **Glm** and **GlmW**. Counties *with* sampled poor children are plotted with hollow triangles, and counties *without* sampled poor children are plotted with solid triangles.