Methodology for Testing for a Rise in Child Poverty Rate of Five Percent or Greater: 2003 to 2004

David S. Powers and William R. Bell
March 26, 2007

1 Background

The U.S. Census Bureau's Small Area Estimates Branch annually provides the Administration for Children and Families in the Department of Health and Human Services with model-based estimates of the number of children ages 0 to 17 in poverty. These estimates are used to determine if any states had an increase of 5 percent or greater in child poverty rate between two consecutive years. This document addresses change between 2003 and 2004.

The data presented help identify states for which the following equivalent statements are true:

$$\frac{(\text{2004 Poverty Rate}) - (\text{2003 Poverty Rate})}{\text{2003 Poverty Rate}} \geq 0.05$$

$$(\text{2004 Poverty Rate}) - (\text{2003 Poverty Rate}) \geq 0.05 \cdot (\text{2003 Poverty Rate})$$

$$(\text{2004 Poverty Rate}) - 1.05 \cdot (\text{2003 Poverty Rate}) \geq 0$$

To statistically test these statements, standardized normal z-statistics are compared to critical values at the 10 percent significance level. This document summarizes state and national test results and describes the derivation of z-statistics and change variance estimates.

The poverty estimates used in this analysis are from the Small Area Income and Poverty Estimates (SAIPE) program. The SAIPE program produces model-based estimates of poverty that match official poverty at the national level as measured by the Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS). Documentation of the methods used to produce state poverty estimates for 2003 and 2004 is available on the SAIPE program's website, www.census.gov/hhes/www/saipe/documentation.html.

Within this document, "change estimate" refers to the 2004 poverty rate estimate for children ages 0 to 17 minus the 2003 poverty rate estimate for children ages 0 to 17. Accordingly, "change variance estimate" refers to the variance of this quantity, and "*z*-statistic" refers to the ratio of "change estimate" to the square root of "change variance estimate." Terminology for the 5 percent or greater data corresponds: "1.05 change estimate" refers to the 2004 poverty rate estimate for children ages 0 to 17 minus 1.05 times the 2003 poverty rate estimate for children ages 0 to 17, "1.05 change variance estimate" refers to the variance of this quantity, and "1.05 *z*-

statistic" refers to the ratio of "1.05 change estimate" to the square root of "1.05 change variance estimate."

Section 2 below describes the type of hypothesis tests used to assess year-to-year change in the child poverty rates. *Sections 3 and 4* present state and national results for the hypothesis tests. *Section 5* presents mathematical details behind the poverty rate estimation and change variance estimation.

State poverty estimates, standard errors and z-statistics are shown in *Table 1* and *Table 2* at the end of the document.

2 Hypothesis Tests

The change estimate and change variance estimate can be used to test whether there is statistically significant evidence that the child poverty rate has increased. Likewise, the 1.05 change variance estimate and the 1.05 change estimate can be used to test whether there is statistically significant evidence that the child poverty rate has increased by 5 percent or greater. The 1.05 *z*-statistics are created for the one-tailed hypothesis test as follows:

Null Hypothesis: Poverty rate has *not* increased by 5 percent or greater

$$(\text{2004 Poverty Rate}) - 1.05 \cdot (\text{2003 Poverty Rate}) < 0$$

Alternative Hypothesis: Poverty rate has increased by 5 percent or greater

$$(\text{2004 Poverty Rate}) - 1.05 \cdot (\text{2003 Poverty Rate}) \geq 0$$

Test Statistic (the 1.05 *z*-statistic):

$$z = \frac{((\text{2004 Poverty Rate Estimate}) - 1.05 \cdot (\text{2003 Poverty Rate Estimate}))}{\sqrt{\text{Var}((\text{2004 Poverty Rate Estimate}) - 1.05 \cdot (\text{2003 Poverty Rate Estimate}))}} \tag{1}$$

Under the SAIPE program's state models, z can be compared to critical values from the standard normal distribution.

A single one-tailed test would be appropriate to test for an increase of 5 percent or greater in a particular state. However, since we are testing for an increase of 5 percent or greater in all 50 states and Washington, D.C., applying one-tailed tests separately for each state would be inappropriate. In particular, if no state had an increase of 5 percent or greater and we performed this test separately for each state, then the probability we would conclude one or more states had a 5 percent or greater increase may be larger than the stated significance level. This is referred to as the problem of "multiple comparisons."

In order to test whether there has been a child poverty rate increase of 5 percent or greater in any of the 51 states, we follow the Bonferroni approach. The Bonferroni approach addresses the problem of multiple comparisons by using a critical value such that, if all the null hypotheses for a set of tests were true, the probability that one or more of these tests would yield a statistically significant result will be no larger than the specified significance level.

3 State Results

We use a 10 percent significance level for our hypothesis tests. For a set of 51 tests (for the 50 states and Washington, D.C.) with the standard normal *z*-statistic, the Bonferroni one-tailed critical value is 2.88. If any state has a 1.05 *z*-statistic greater than or equal to 2.88, then there is evidence that the child poverty rate for that state increased by 5 percent or greater. We find that no state has a 1.05 *z*-statistic greater than or equal to 2.88 when comparing 2003 and 2004 child poverty rate estimates. Thus, using the Bonferroni test, *we do not find statistical evidence that any state has a child poverty rate increase of 5 percent or greater between 2003 and 2004.*

As described, the Bonferroni approach is appropriate for answering the question, "Is there evidence that *any* state had a child poverty rate increase of 5 percent or greater?" A different critical value would be appropriate to test for evidence of a child poverty rate increase of 5 percent or greater in a *particular* state that was selected in advance, i.e., not selected based on looking at the results for all the states. The critical value when an individual state is selected in advance is 1.28, the cutoff for a single one-tailed test with a 10 percent significance level.

If we ignore multiple comparisons issues, we find that two states have evidence of an increase in child poverty rate of 5 percent or greater from 2003 to 2004: Wisconsin (with a 1.05 z-statistic of 1.49 > 1.28) and Michigan (with a 1.05 z-statistic of 1.31 > 1.28). Therefore, someone particularly interested in the result for Wisconsin (or Michigan) could find evidence at the 10 percent significance level that the child poverty rate in Wisconsin (or Michigan) increased by 5 percent or more from 2003 to 2004. *However*, this assumes it was decided in advance to examine the specific result for Wisconsin (or Michigan) and that Wisconsin (or Michigan) was not picked by looking at the results for all states and selecting the one or two with the largest 1.05 z-statistics.

Since we are interested in examining *all states* for evidence of whether *any state* had an increase in child poverty rate of 5 percent or greater, the Bonferroni approach, which takes into account multiple comparisons issues, is more appropriate than individual state one-tailed tests. Using this method, the 1.05 z-statistics, collectively, are judged not to be statistically significant. Therefore, as noted above, there is no statistically significant evidence that any state had a 5 percent or greater increase in child poverty rate between 2003 and 2004.

Two changes were made to the SAIPE program's state models for 2004, as described on the SAIPE program's webpage, www.census.gov/hhes/www/saipe/techdoc/methods/04change.html. The model changes for 2004 have the result of enhancing the differences between the 2003 and 2004 SAIPE estimates for some states, and diminishing them for other states, relative to what would be the case if the same models were used for both years. This result is typical when making model changes and tends to increase the likelihood of finding estimated changes that are statistically significant. Despite the effects of the model changes for 2004, we find no statistical evidence that any state had a rise in child poverty rate of 5 percent or greater between 2003 and 2004 using the Bonferroni method.

Data for each state are presented in Tables 1 and 2. Table 1 contains poverty rate estimates, and Table 2 contains standard errors and *z*-statistics. The critical value of 2.88 should be used when checking for statistically significant evidence (at the 10 percent level) that any state had a child poverty rate increase, or an increase of 5 percent or greater, and the critical value of 1.28 should be used by individual states examining their own results separately.

4 National Results

In addition to state-level child poverty rates, we also consider the child poverty rate at the national level. The official national poverty estimates are direct estimates from the ASEC, and the SAIPE program does not model these national estimates.

Standard errors for the 2003 estimates are computed using the methods described in U.S. Census Bureau's "Source and Accuracy of Estimates for Income, Poverty and Health Insurance Coverage in the United States: 2003," available at www.census.gov/hhes/www/income/p60_226sa.pdf; specifically, formula (2) on page 7 and Table 2 on page 6 are used. Standard errors for the 2004 estimates are computed using the methods described in U.S. Census Bureau's "Source and Accuracy of Estimates for Income, Poverty and Health Insurance Coverage in the United States: 2004," available at www.census.gov/hhes/www/income/p60_229sa.pdf; specifically, formula (2) on page 10 and Table 3 on page 7 are used. These estimates and standard errors are the latest available as of the posted release date and are not subsequently updated.

The first line of Table 1 refers to the United States as a whole. We see from Table 1 that, while the estimated U.S. child poverty rate increased between 2003 and 2004, the estimated percent increase was only 1.1 percent, which is considerably less than 5 percent. Hence, the 1.05 *z*-statistic does not reject the null hypothesis (1.05 *z*-statistic = -2.28 < 1.28), and we do not find evidence (at the 10 percent level of significance) of a 5 percent or greater increase in the national child poverty rate between 2003 and 2004.

To compute the 1.05 *z*-statistic at the national level, we use equation (1) for *z* shown in Section 2. In this case, the variance in the denominator is computed as:[1]

$$\begin{aligned}&\mathrm{Var}((\text{2004 Poverty Rate Estimate}) - 1.05 \cdot (\text{2003 Poverty Rate Estimate}))\\ &\qquad = \quad s_x^{\,2} + (1.05 \cdot s_y)^2 - 2 \cdot r \cdot s_x \cdot (1.05 \cdot s_y) \ ,\end{aligned}$$

where

[1] This is the standard formula for *standard errors of differences* (squared). See formula (3) and the accompanying text on page 11 of U.S. Census Bureau's "Source and Accuracy of Estimates for Income, Poverty and Health Insurance Coverage in the United States: 2004," available at www.census.gov/hhes/www/income/p60_229sa.pdf.

s_x = standard error of 2004 poverty rate estimate for children ages 0-17

s_y = standard error of 2003 poverty rate estimate for children ages 0-17

r = correlation coefficient for year-to-year comparisons of ASEC poverty estimates or proportions

Here, r equals 0.45, the value in the upper-left cell of Table 6 on page 9 of U.S. Census Bureau's "Source and Accuracy of Estimates for Income, Poverty, and Health Insurance Coverage in the United States: 2004," available at www.census.gov/hhes/www/income/p60_229sa.pdf.

5 Mathematical Details

5.1 State Poverty Models

The SAIPE program's poverty models employ both direct survey-based estimates of poverty from the ASEC and regression predictions of poverty based on administrative records and Census 2000 data. The state poverty models are defined as follows:

$$\begin{aligned} y_i &= Y_i + e_i \qquad e_i \sim N(0, V_{ei}) \\ Y_i &= X_i\beta_i + u_i \quad u_i \sim N(0, \sigma_{ui}^2 I) \end{aligned} ,$$

where

y_i = vector of 51 state ASEC estimates of poverty ratios for a given age group and a given year,

Y_i = vector of "true" poverty ratios for a given age group and a given year,

X_i = matrix of predictor variables for a given age group and a given year; β_i is the corresponding vector of regression coefficients,

u_i = vector of model errors for a given age group and a given year, assumed independent across states; σ_{ui}^2 is their common variance,

e_i = vector of sampling errors for a given age group and a given year, assumed independent across states; V_{ei} is the diagonal matrix of the sampling error variances for each state for a given age group and a given year.

Poverty ratios for children ages 0-4 and poverty ratios for children ages 5-17 are modeled separately. The subscript i = 1, 2, 3, 4 indexes the four ASEC equations for the two years (2003 and 2004) and the two age groups (0-4 and 5-17) according to the following scheme:

i = 1: y_1 = 2003 ASEC estimated poverty ratio for children ages 0-4
i = 2: y_2 = 2003 ASEC estimated poverty ratio for children ages 5-17
i = 3: y_3 = 2004 ASEC estimated poverty ratio for children ages 0-4
i = 4: y_4 = 2004 ASEC estimated poverty ratio for children ages 5-17

The coefficient vector, β_i, and the model error variance, σ_{ui}^2, are estimated by Bayesian techniques, treating the estimated sampling error variances, V_{ei}, as known. (Estimation of σ_{ui}^2 and V_{ei} is discussed in *Section 5.4*.) The Bayesian techniques combine the regression predictions with the direct ASEC estimates, weighting the contribution of these two components on the basis of their relative precision, in order to obtain model-based estimates of child poverty ratios by state.

Additional documentation of the SAIPE program's poverty models and estimation procedures is available at www.census.gov/hhes/www/saipe/documentation.html.

5.2 Poverty Rates, Ratios and Universes

Poverty ratio estimates for children ages 0-4 are defined as the ASEC estimated number of children ages 0-4 in poverty divided by the ASEC estimated population ages 0-4. Likewise, poverty ratio estimates for children ages 5-17 are defined as the ASEC estimated number of children ages 5-17 in poverty divided by the ASEC estimated population ages 5-17. Poverty *rates* differ from poverty *ratios* in that they have different denominators. Poverty rates have as their denominator the ASEC *poverty universe estimate* (described below), whereas poverty ratios have as their denominator the ASEC *population estimate*.

Both the ASEC poverty universe and ASEC population exclude people living in military barracks or in institutional group quarters since the ASEC does not sample from these groups. The poverty universe also excludes children ages 0-14 who are not related to the householder by birth, marriage or adoption since for these children no one ages 15 or over answers ASEC income questions. For discussion of poverty measurement and ASEC definitions, see www.census.gov/hhes/www/poverty/definitions.html.

The SAIPE program's state models are run using poverty ratios rather than poverty rates. In computing poverty ratios, we use ASEC estimates in both the numerators and denominators (as opposed to using demographic population estimates in the denominators) because the positive correlation among these ASEC estimates reduces the variances of the resulting poverty ratio estimates.

We convert model-based estimates of poverty ratios for children ages 0-4 and children ages 5-17 into estimates of poverty rates for children ages 0-17 through the following steps:

1. Multiply the model-based estimates of poverty *ratios* for each combination (i) of age group and year by the corresponding demographic population estimates in order to obtain estimates of the number of children ages 0-4 and 5-17 in poverty in each state.

The demographic population estimates are from the Census Bureau's Population Estimates Program, and these estimates are adjusted to represent the population covered by the ASEC.

2. Multiply the estimated number in poverty in each state by a raking factor (defined in Section 5.3) for each combination (i) of age group and year so that the resulting state estimated numbers in poverty sum to the ASEC national estimate for that combination of age group and year.
3. For each state add the raked estimate of the number of children ages 0-4 in poverty to the raked estimate of the number of children ages 5-17 in poverty to get the raked estimate of the number of children ages 0-17 in poverty for a given year.
4. Form the estimated poverty *rates* for children ages 0-17 by dividing the estimated number of children ages 0-17 in poverty by the demographic poverty universe estimate for children ages 0-17 (children ages 0-4 plus children ages 5-17).

Note that in the first step we multiply the estimated poverty ratios by the demographic estimates of population rather than by the ASEC estimates of population. The demographic estimates of population are assumed to have no sampling error and are more accurate than population estimates constructed from ASEC sample data. Thus, while ASEC population estimates are suitable denominators for the modeled poverty ratios (due to their correlation with the poverty ratio numerators, as noted above), demographic population estimates are more appropriate for multiplying the model-based poverty ratio estimates to obtain the estimated numbers of children in poverty.

The modeled ASEC estimates use data from interviews conducted in February, March, and April of a given year (the survey year, SY) regarding income from the previous year (the income year, IY). The relevant population and poverty universe estimates to use as denominators in the poverty ratios and poverty rates are those for the survey year. Therefore, the estimated poverty ratios and poverty rates for 2003 use population and poverty universe estimates for 2004, and the estimated poverty ratios and poverty rates for 2004 use population and poverty universe estimates for 2005.

For further discussion of the denominators used in poverty rates, see www.census.gov/hhes/www/saipe/techdoc/inputs/denom.html.

5.3 Change Variance Estimates

This section describes mathematical details behind the computation of change variance estimates and 1.05 change variance estimates. The square roots of these variance estimates form the denominators of the z-statistics and 1.05 z-statistics used to assess possible changes in state child poverty rates.

We represent the state demographic population estimates in mathematical notation as:

N_{1k} = 2004 demographic population estimate for children ages 0-4 in state k,
N_{2k} = 2004 demographic population estimate for children ages 5-17 in state k,
N_{3k} = 2005 demographic population estimate for children ages 0-4 in state k,

$N_{4k} =$ 2005 demographic population estimate for children ages 5-17 in state k,

and we represent the state demographic poverty universe estimates as:

U_{1k} = 2004 demographic poverty universe estimate for children ages 0-4 in state k,
U_{2k} = 2004 demographic poverty universe estimate for children ages 5-17 in state k,
U_{3k} = 2005 demographic poverty universe estimate for children ages 0-4 in state k,
U_{4k} = 2005 demographic poverty universe estimate for children ages 5-17 in state k.

We define *scaling factors* for the two age groups in each year as:

$$r_{1k} = \frac{N_{1k}}{U_{1k} + U_{2k}} \;,\quad r_{2k} = \frac{N_{2k}}{U_{1k} + U_{2k}}$$
$$r_{3k} = \frac{N_{3k}}{U_{3k} + U_{4k}} ,\quad r_{4k} = \frac{N_{4k}}{U_{3k} + U_{4k}} ,$$

and we define *raking factors* for each combination (i) of age group and year as:

$$RF_i = \frac{\text{ASEC direct national estimate of number in poverty for age group and year combination } i}{\sum_k (\text{model-based estimate of number in poverty for state } k \text{ for age group and year combination } i)} .$$

The scaling factors weight the ages 0-4 and ages 5-17 estimated poverty ratios in proportion to the number of children in each age group, forming contributions to the ages 0-17 poverty rates. The raking factors scale the poverty ratio estimates so that the sum of the products of poverty ratios and demographic population estimates equals the national ASEC estimate of the number of children in poverty in each age group and year.

Letting R_i be a 51x51 diagonal matrix with the r_{ik} terms (scaling factors) on the diagonal, the vector of contributions to the 2003 ages 0-17 poverty rate from the ages 0-4 group and the ages 5-17 group are $R_1 \cdot Y_1$ and $R_2 \cdot Y_2$, respectively. The raked estimators of these products are $R_1 \cdot RF_1 \hat{Y}_1$ and $R_2 \cdot RF_2 \hat{Y}_2$. Likewise, the vector of contributions to the 2004 ages 0-17 poverty rate from the ages 0-4 group and the ages 5-17 group can be written as $R_3 \cdot Y_3$ and $R_4 \cdot Y_4$, respectively, and the raked estimators of these products are $R_3 \cdot RF_3 \hat{Y}_3$ and $R_4 \cdot RF_4 \hat{Y}_4$.

The error in the change estimate can then be written as:

$$[R_3(Y_3 - RF_3\hat{Y}_3) + R_4(Y_4 - RF_4\hat{Y}_4)] - [R_1(Y_1 - RF_1\hat{Y}_1) + R_2(Y_2 - RF_2\hat{Y}_2)], \qquad (2)$$

where $Y_i - RF_i\hat{Y}_i$ is the error in the raked poverty ratio estimate for combination (i) of age group and year. The diagonal of the variance matrix of this expression will be the change variance estimates. Similarly, the error in the 1.05 change estimate can be written as:

$$[R_3(Y_3 - RF_3\hat{Y}_3) + R_4(Y_4 - RF_4\hat{Y}_4)] - 1.05[R_1(Y_1 - RF_1\hat{Y}_1) + R_2(Y_2 - RF_2\hat{Y}_2)], \qquad (3)$$

and the diagonal of the variance matrix of this expression will be the 1.05 change variance estimates.

Bell (1999) determined that the vector of prediction errors, $Y_i - RF_i\hat{Y}_i$, for combination (i) of age group and year can be expressed as:

$$Y_i - RF_i\hat{Y}_i = A_i \cdot u_i + (A_i - I)\cdot e_i + A_i X_i \beta_i \ ,$$

where

$$A_i = (1 - RF_i)I + RF_i(I - H_i)(I - M_i) \ ,$$
$$H_i = \sigma_{ui}^2 \Sigma_i^{-1}, \quad \Sigma_i = \sigma_{ui}^2 I + V_{ei}, \quad \text{and } M_i = X_i(X_i'\Sigma_i^{-1}X_i)^{-1}X_i'\Sigma_i^{-1}.$$

The term $A_i X_i \beta_i$ can be rewritten as $(1 - RF_i)\times X_i\beta_i$. This is, fundamentally, a bias term that arises from raking state estimates to national totals under the model assumption that the regression function $X_i\beta_i$ produces unbiased estimates. (The raking factor, RF_i, also includes some random estimation error.) The model is, of course, an approximation, and the raking is done because it is believed to reduce possible bias arising from failure of the model assumptions. We therefore ignore this bias term in computing measures of error for the raked estimates and compute the covariance matrix based on just the first two terms, $A_i.u_i$ and $(A_i - I)\cdot e_i$.

Proceeding with the assumption that the $A_i X_i \beta_i$ term can be ignored, the errors in the change and 1.05 change estimates shown in formulas (2) and (3) can be re-written as:

$$\begin{aligned} &R_3[A_3 \cdot u_3 + (A_3 - I)\cdot e_3] + R_4[A_4 \cdot u_4 + (A_4 - I)\cdot e_4] \\ &\quad + \tilde{R}_1[A_1 \cdot u_1 + (A_1 - I)\cdot e_1] + \tilde{R}_2[A_2 \cdot u_2 + (A_2 - I)\cdot e_2] \end{aligned} \ , \qquad (4)$$

where $\tilde{R}_1$ and $\tilde{R}_2$ are $-1.05R_1$ and $-1.05R_2$, respectively, for the error in the 1.05 change estimate and are $-R_1$ and $-R_2$, respectively, for the error in the change estimate.

The covariance matrix of formula (4) can, then, be written as:

$$\sum_i \sum_j [\overline{R}_i \cdot (A_i - I)]Cov(e_i, e_j)[\overline{R}_j \cdot (A_j - I)]' + \sum_i \sum_j [\overline{R}_i \cdot A_i]Cov(u_i, u_j)[\overline{R}_j \cdot A_j]' \ , \qquad (5)$$

where, for the 1.05 change variance estimates:

$\overline{R}_i = -1.05R_i$ when i = 1 or 2, and $\overline{R}_i = R_i$ when i = 3 or 4,

and, for the change variance estimates:

$$\overline{R}_i = -R_i \text{ when i = 1 or 2, and } \overline{R}_i = R_i \text{ when i = 3 or 4.}$$

$[Cov(e_i, e_j)]$ and $[Cov(u_i, u_j)]$ are 51x51 matrices. They are diagonal matrices because we assume the sampling errors and model errors are uncorrelated across states and uncorrelated with each other. The 16 different *i, j* pairs correspond to the individual cells in the figure below:

		1	2	3	4
		04yr1	517yr1	04yr2	517yr2
1	04yr1	v11	v12	v13	v14
2	517yr1	v21	v22	v23	v24
3	04yr2	v31	v32	v33	v34
4	517yr2	v41	v42	v43	v44

There are 32=16+16 terms altogether in formula (5)'s summation. The change variance and 1.05 change variance estimates are then the 51 diagonal elements of formula (5) evaluated.

5.4 Covariances Needed for Change Variance Estimates

In order to estimate formula (5), we need to estimate the diagonal elements of $[Cov(e_i, e_j)]$ and $[Cov(u_i, u_j)]$. We do this through the following steps:

- Fit models for the sampling error variances of ASEC state estimates using direct estimates of ASEC sampling error variances and covariances;
- Compute averages of direct estimates of ASEC sampling error correlations;
- Fit models for the state ASEC estimates used to produce the state poverty ratio predictions;
- Treat pairs of ASEC state equations (by age group and year) jointly and use Bayesian techniques to estimate the correlation between model errors in the two equations; and
- Combine the estimated sampling error variances and correlations to obtain estimated sampling error covariances, and combine the estimated model error variances and correlations to obtain estimated model error covariances.

These steps are described in more detail below.

Sampling Error Variances

We estimate the sampling error variances, V_{ei}, for each age-group poverty ratio (0-4 and 5-17) by fitting sampling error models to directly-estimated ASEC sampling error covariance matrices for each state. We produce direct estimates of state ASEC sampling error variances using the VPLX program, which implements a successive difference replication method, as described in Fay and Train (1995). Separately for each age-group poverty ratio, we then fit sampling error models to these directly-estimated variances by maximum likelihood estimation assuming a Wishart distribution for the covariance matrices. Modeling the sampling error variances as such improves estimates of the sampling error variances, V_{ei}, for each age-group poverty ratio (and for median household income) for each state. Otto and Bell (1995) discuss the type of sampling error models used.

The sampling error models allow the sampling error variances (nonzero elements of the diagonal matrices, V_{ei}) to differ across states and years through a generalized variance function (GVF) that depends on the level of the poverty ratio estimates and on the ASEC state sample sizes. The models assume that the sampling error correlations between years (such as ρ_{e13} and ρ_{e24}) are constant across states. The models also assume that the sampling error correlations are stationary and thus depend only on the lag, $t - j$, between years t and j.[2] Also, because we use separate sampling error models for each age group, the fitted sampling error models do not provide estimates of sampling error correlations between the poverty ratios for different age groups.

For 2003, changes were made to the ASEC sample weighting used to produce the ASEC direct estimates, and so adjustments were made to the 2003 variance estimates. (For more discussion of this, see www.census.gov/hhes/www/saipe/techdoc/methods/03change.html.) For 2004, the direct variance estimates drew on samples from both the 1990 Census based CPS design and the newly introduced Census 2000 based design. This complicated the construction of replicate weights and required some approximations in the estimation of sampling error variances and covariances. While the replicate weights were believed to produce reasonable estimates of the sampling error variances for 2004, the estimates of the covariances and correlations between 2004 and earlier years were believed to be unreliable. (This issue is dealt with below.) Because the design changes for 2004 could have affected the 2004 sampling variances, the sampling error models used different design effect parameters for 2004 than for 2000-2003. (For more discussion of this, see the last paragraph of www.census.gov/hhes/www/saipe/techdoc/methods/04change.html.) The sampling error models were then fit to the estimated covariance matrices to produce the model-based sampling error variance estimates for 2003 and 2004.

Sampling Error Correlations

We estimate the sampling error correlations between the poverty ratios (ρ_{e13}, ρ_{e24}, ρ_{e14}, ρ_{e23}, ρ_{e12}, ρ_{e34}) by averaging the corresponding direct estimates over states and years. More specifically, we construct correlation matrices from direct sampling covariance matrices (discussed above) and then average these over the 50 states and Washington, D.C. When doing this in past years we assumed stationarity of the sampling error correlations, which, again, means assuming that the correlation between years t and j depends only on the lag, $t - j$.

[2] For reasons discussed in the next paragraph, this assumption is probably violated for the correlations of 2004 with earlier years, but the assumption has little effect on the fitted GVFs, which are the important results of this model fitting.

As noted above, the direct estimates of sampling error correlations between 2003 and 2004 were believed to be unreliable. To deal with this, we applied an adjustment factor to the average year-to-year correlations from previous years to approximate the correlations between 2003 and 2004. The adjustment factor was computed as the percent sample overlap that occurred from 2003 to 2004 divided by the typical sample overlap for adjacent years. For 2003-2004, 42 percent of the CPS sample housing units for 2003 were in rotation groups to also be interviewed in 2004, whereas in the other years (2000-2001, 2001-2002, and 2002-2003), roughly 50 percent of the CPS sample housing units for one year were also in sample in the next year. Dividing the 42 percent by 50 percent yields the adjustment factor of 0.84.

Given the stationarity assumption, for the cross-year correlations (ρ_{e13}, ρ_{e24}, ρ_{e14}, ρ_{e23}) we averaged across states the cross-year correlations that correspond to the same two age groups. We thus averaged these over 2000-2001, 2001-2002, 2002-2003 and applied the 0.84 adjustment factor. For the within-year correlation for 2003 (ρ_{e12}), we averaged across states and over 2000, 2001, 2002 and 2003, and for the within-year correlation for 2004 (ρ_{e34}), we averaged across states and used 2004 alone. In each case we used simple unweighted averages of the correlations.

Sixteen sampling error correlations are estimated. Four of the sixteen correlations are along the diagonal and equal one. Of the twelve remaining correlations, six correlations are equal to the other six correlations due to symmetry. Thus, effectively six correlations are estimated: ρ_{e13}, ρ_{e24}, ρ_{e14}, ρ_{e23}, ρ_{e12}, ρ_{e34}.

Model Error Variance

We estimate the model error variance, σ_{ui}^2, while fitting the state models to the ASEC direct poverty ratio estimates. We use a Bayesian approach in estimation of the state model, and we regard σ_{ui}^2 as estimated by its posterior mean. We use a noninformative (flat) prior for all model parameters.

Model Error Correlations

We estimate the model error correlations (ρ_{u12}, ρ_{u13}, ρ_{u14}, ρ_{u23}, ρ_{u24}, ρ_{u34}) using the Bayesian approach and treating each pair of ASEC state equations jointly. For each of the six possible distinct pairs of equations among the four ASEC state equations, we specify flat prior distributions for the regression coefficients and the model error variances, as was done when fitting the models one equation at a time. The prior for the model error correlation is taken to be uniform on the interval [-1,1]. We then take the posterior mean of the model error correlation as its point estimate. Note that although this model-fitting procedure produces new estimates of the other model parameters involved in each pair of equations (i.e., the regression parameters and model error variances), for calculation of the change estimates and their variances we leave these model parameters at their original Bayesian estimates obtained from fitting the single ASEC equations separately. This is done so that the results remain consistent with the SAIPE program's published estimates, which are produced by fitting single ASEC state equations at a time. This joint Bayesian treatment of two ASEC equations simultaneously is done using the WinBUGS package (Spiegelhalter, et al., 2003).

Unlike the sampling error correlations for which comparable time lags of cross-year and within-year model error correlations are averaged across years, for the model error correlations, estimates of the cross-year and within-year model error correlations are not averaged across years. Sixteen model error correlations are estimated. Four of the sixteen correlations are along the diagonal and equal one. Of the twelve remaining correlations, six correlations are equal to the other six correlations due to symmetry. Thus, effectively six correlations are estimated: ρ_{u13}, ρ_{u24}, ρ_{u14}, ρ_{u23}, ρ_{u12}, ρ_{u34}.

Covariances

Finally, in order to estimate the sampling error covariances in formula (5), we combine the estimated sampling error variances with the estimated sampling error correlations as follows:

$$Cov(e_i, e_j) = \rho_{eij}(V_{ei}V_{ej})^{1/2} \quad ,$$

where $(V_{ei}V_{ej})^{1/2}$ represents the matrix formed by taking the square root of each element of $V_{ei}V_{ej}$. Likewise, in order to estimate the model error covariances, we combine the estimated model error variances with the estimated model error correlations as follows:

$$Cov(u_i, u_j) = \rho_{uij}\sigma_{ui}\sigma_{uj}I \quad .$$

There are sixteen sampling error covariances, of which four are sampling error variances, and there are sixteen model error covariances, of which four are model error variances. These covariances are used to evaluate formula (5) at the end of Section 5.3.

References

Bell, William R. (1999), “Derivation of Dependence of Prediction Errors on Model and Sampling Errors,” Unpublished U.S. Census Bureau report.

Fay, Robert E. and Train, George F. (1995), “Aspects of Survey and Model-Based Postcensal Estimation of Income and Poverty Characteristics for States and Counties,” American Statistical Association, Proceedings of the Section on Government Statistics.

Otto, Mark C. and Bell, William R. (1995), “Sampling Error Modelling of Poverty and Income Statistics for States,” American Statistical Association, Proceedings of the Section on

Government Statistics, pp. 160-165.

Spiegelhalter, David, Andrew Thomas, Nicky Best, and Dave Lunn (2003), "WinBUGS v1.4: Bayesian Inference Using Gibbs Sampling User Manual," MRC Biostatistics Unit, Institute of Public Health, Cambridge, U.K.

U.S. Census Bureau (2004). "Source and Accuracy of Estimates for Income, Poverty and Health Insurance Coverage in the United States: 2003;" Posted on World Wide Web, August 26, 2004, <http://www.census.gov/hhes/www/income/p60_226sa.pdf>

U.S. Census Bureau (2005) "Source and Accuracy of Estimates for Income, Poverty and Health Insurance Coverage in the United States: 2004," Posted on World Wide Web, August 30, 2005, <http://www.census.gov/hhes/www/income/p60_229sa.pdf>

Table 1. Poverty Rate Estimates for Children Ages 0-17

FIPS code	name		ages 0-17 poverty rate 2003	ages 0-17 poverty rate 2004	% change poverty rate[1] '03-'04	change estimate[2] '03-'04	1.05 change estimate[3] '03-'04
00	United States		17.6	17.8	1.1	0.2	-0.7
01	Alabama		22.3	22.6	1.3	0.3	-0.8
02	Alaska		12.5	12.9	3.2	0.4	-0.2
04	Arizona		20.6	21.5	4.4	0.9	-0.1
05	Arkansas		23.6	22.7	-3.8	-0.9	-2.1
06	California		19.5	18.7	-4.1	-0.8	-1.8
08	Colorado		12.8	12.8	0.0	0.0	-0.6
09	Connecticut		10.5	11.4	8.6	0.9	0.4
10	Delaware		13.0	13.4	3.1	0.4	-0.3
11	District of Columbia		29.5	29.2	-1.0	-0.3	-1.8
12	Florida		19.4	17.3	-10.8	-2.1	-3.1
13	Georgia		19.1	19.4	1.6	0.3	-0.7
15	Hawaii		14.6	10.8	-26.0	-3.8	-4.5
16	Idaho		15.9	15.1	-5.0	-0.8	-1.6
17	Illinois		15.6	16.7	7.1	1.1	0.3
18	Indiana		13.7	15.7	14.6	2.0	1.3
19	Iowa		12.3	13.1	6.5	0.8	0.2
20	Kansas		13.8	14.6	5.8	0.8	0.1
21	Kentucky		21.1	22.2	5.2	1.1	0.0
22	Louisiana		26.6	27.4	3.0	0.8	-0.5
23	Maine		14.3	14.3	0.0	0.0	-0.7
24	Maryland		11.5	11.1	-3.5	-0.4	-1.0
25	Massachusetts		11.9	12.8	7.6	0.9	0.3
26	Michigan		15.0	17.3	15.3	2.3	1.6
27	Minnesota		10.2	10.6	3.9	0.4	-0.1
28	Mississippi		26.8	28.6	6.7	1.8	0.5
29	Missouri		16.5	18.5	12.1	2.0	1.2
30	Montana		19.9	18.5	-7.0	-1.4	-2.4
31	Nebraska		12.9	12.8	-0.8	-0.1	-0.8
32	Nevada		15.3	15.3	0.0	0.0	-0.8
33	New Hampshire		7.8	8.8	12.8	1.0	0.6
34	New Jersey		11.8	10.2	-13.6	-1.6	-2.2
35	New Mexico		25.8	23.8	-7.8	-2.0	-3.3
36	New York		20.2	20.8	3.0	0.6	-0.4
37	North Carolina		19.1	18.7	-2.1	-0.4	-1.4
38	North Dakota		13.4	13.5	0.7	0.1	-0.6
39	Ohio		15.4	16.8	9.1	1.4	0.6
40	Oklahoma		21.1	20.2	-4.3	-0.9	-2.0
41	Oregon		17.4	17.3	-0.6	-0.1	-1.0
42	Pennsylvania		14.9	16.0	7.4	1.1	0.4
44	Rhode Island		16.6	17.4	4.8	0.8	0.0
45	South Carolina		19.8	21.2	7.1	1.4	0.4
46	South Dakota		16.5	17.9	8.5	1.4	0.6
47	Tennessee		19.2	20.1	4.7	0.9	-0.1
48	Texas		22.8	22.7	-0.4	-0.1	-1.2
49	Utah		12.5	12.4	-0.8	-0.1	-0.7
50	Vermont		11.9	10.6	-10.9	-1.3	-1.9
51	Virginia		13.6	12.2	-10.3	-1.4	-2.1
53	Washington		15.3	15.8	3.3	0.5	-0.3
54	West Virginia		24.3	22.6	-7.0	-1.7	-2.9
55	Wisconsin		12.4	14.9	20.2	2.5	1.9
56	Wyoming		14.6	13.7	-6.2	-0.9	-1.6

[1] 100 × [(2004 Poverty Rate Estimate – 2003 Poverty Rate Estimate)/(2003 Poverty Rate Estimate)]
Percent changes may not be statistically significant. See Sections 3 and 4 for details.

[2] 2004 Poverty Rate Estimate – 2003 Poverty Rate Estimate

[3] 2004 Poverty Rate Estimate – 1.05 × (2003 Poverty Rate Estimate)

Source: U.S. Census Bureau, SAIPE program, www.census.gov/hhes/www/saipe/index.html.

Table 2. Standard Errors and *z*-statistics for Children Ages 0-17

FIPS code	name	change estimate[1] '03-'04	S.E. of change est. '03-'04	z-statistic[2] '03-'04	1.05 change estimate[3] '03-'04	S.E. of 1.05 change est. '03-'04	1.05 z-statistic[4] '03-'04
00	United States	0.2	0.30	0.61	-0.7	0.31	-2.28
01	Alabama	0.3	1.38	0.23	-0.8	1.40	-0.57
02	Alaska	0.4	1.21	0.36	-0.2	1.23	-0.16
04	Arizona	0.9	1.59	0.58	-0.1	1.62	-0.07
05	Arkansas	-0.9	1.50	-0.58	-2.1	1.53	-1.34
06	California	-0.8	1.01	-0.81	-1.8	1.04	-1.73
08	Colorado	0.0	1.27	-0.01	-0.6	1.30	-0.50
09	Connecticut	0.9	1.21	0.73	0.4	1.24	0.29
10	Delaware	0.4	1.28	0.28	-0.3	1.31	-0.23
11	District of Columbia	-0.3	2.56	-0.13	-1.8	2.63	-0.69
12	Florida	-2.1	1.18	-1.77	-3.1	1.21	-2.54
13	Georgia	0.3	1.44	0.20	-0.7	1.47	-0.45
15	Hawaii	-3.8	1.44	-2.63	-4.5	1.47	-3.09
16	Idaho	-0.8	1.40	-0.54	-1.6	1.43	-1.09
17	Illinois	1.1	1.10	0.98	0.3	1.12	0.27
18	Indiana	2.0	1.23	1.68	1.3	1.25	1.10
19	Iowa	0.8	1.19	0.75	0.2	1.21	0.23
20	Kansas	0.8	1.25	0.69	0.1	1.27	0.14
21	Kentucky	1.1	1.40	0.78	0.0	1.43	0.02
22	Louisiana	0.8	1.70	0.46	-0.5	1.74	-0.31
23	Maine	0.0	1.37	0.05	-0.7	1.39	-0.46
24	Maryland	-0.4	1.25	-0.32	-1.0	1.28	-0.76
25	Massachusetts	0.9	1.35	0.71	0.3	1.38	0.26
26	Michigan	2.3	1.15	1.98	1.6	1.17	1.31
27	Minnesota	0.4	1.17	0.39	-0.1	1.19	-0.05
28	Mississippi	1.8	1.74	1.02	0.5	1.78	0.24
29	Missouri	2.0	1.39	1.44	1.2	1.41	0.83
30	Montana	-1.4	1.52	-0.92	-2.4	1.55	-1.54
31	Nebraska	-0.1	1.27	-0.05	-0.8	1.29	-0.55
32	Nevada	0.0	1.37	-0.01	-0.8	1.40	-0.56
33	New Hampshire	1.0	1.22	0.82	0.6	1.24	0.49
34	New Jersey	-1.6	1.15	-1.37	-2.2	1.17	-1.85
35	New Mexico	-2.0	1.58	-1.29	-3.3	1.62	-2.06
36	New York	0.6	1.11	0.53	-0.4	1.14	-0.36
37	North Carolina	-0.4	1.28	-0.29	-1.4	1.31	-1.01
38	North Dakota	0.1	1.41	0.08	-0.6	1.44	-0.38
39	Ohio	1.4	1.17	1.26	0.6	1.20	0.59
40	Oklahoma	-0.9	1.50	-0.61	-2.0	1.54	-1.28
41	Oregon	-0.1	1.48	-0.03	-1.0	1.51	-0.61
42	Pennsylvania	1.1	1.12	0.97	0.4	1.14	0.29
44	Rhode Island	0.8	1.39	0.62	0.0	1.42	0.02
45	South Carolina	1.4	1.37	1.03	0.4	1.40	0.30
46	South Dakota	1.4	1.53	0.92	0.6	1.56	0.37
47	Tennessee	0.9	1.48	0.61	-0.1	1.50	-0.04
48	Texas	-0.1	1.12	-0.13	-1.2	1.15	-1.12
49	Utah	-0.1	1.38	-0.02	-0.7	1.41	-0.46
50	Vermont	-1.3	1.25	-1.05	-1.9	1.28	-1.50
51	Virginia	-1.4	1.21	-1.13	-2.1	1.23	-1.66
53	Washington	0.5	1.22	0.42	-0.3	1.25	-0.20
54	West Virginia	-1.7	1.53	-1.15	-2.9	1.56	-1.90
55	Wisconsin	2.5	1.24	2.02	1.9	1.26	1.49
56	Wyoming	-0.9	1.43	-0.63	-1.6	1.46	-1.12

[1] 2004 Poverty Rate Estimate – 2003 Poverty Rate Estimate

[2] $((\text{2004 Poverty Rate}) - (\text{2003 Poverty Rate}))/\sqrt{\text{Var}((\text{2004 Poverty Rate}) - (\text{2003 Poverty Rate}))}$

[3] 2004 Poverty Rate Estimate – 1.05 × (2003 Poverty Rate Estimate)

[4] $((\text{2004 Poverty Rate}) - 1.05\times(\text{2003 Poverty Rate}))/\sqrt{\text{Var}((\text{2004 Poverty Rate}) - 1.05\times(\text{2003 Poverty Rate}))}$

Source: Author's calculations. See Sections 3 and 4 for discussion of critical values.