An Empirical Study on Using ACS Supplementary Survey Data in SAIPE State Poverty Models

Elizabeth T. Huang and William R. Bell
U.S. Bureau of the Census, Washington D.C. 20233 -9100

Key Words: Small area estimation, Fay-Herriot model, bivariate model

1. Introduction

The Census Bureau's Small Area Income and Poverty Estimates (SAIPE) program produces poverty estimates for various age groups for states, counties, and school districts. For states the age groups are 0-4, 5-17, 18-64, and 65+. The state estimates come from a regression model with state random effects (Fay and Herriot 1979) applied to direct state estimates from the Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC, formerly known as the CPS March income supplement). The models borrow information from regression variables related to poverty that are constructed from administrative records data and from poverty estimates from the previous decennial census. Estimates are identified by the "income year" (IY), which refers to the year for which income is reported in the ASEC. Beginning with IY 2000, the CPS ASEC sample was expanded to produce estimates of health insurance coverage for the State Children's Health Insurance Program. This increased the sample size from about 60,000 households to about 98,000 households. Further information is available on the SAIPE web site at http://www.census.gov/hhes/www/saipe/documentation.html.

In recent years supplementary surveys fielded to test data collection procedures for the American Community Survey (ACS) have also provided state poverty estimates. The ACS asks essentially the same questions as the decennial census long form survey, and is proposed to replace the long form, but with the data collection spread continuously throughout the decade, rather than at a single point in time. When fully implemented, the ACS will have a national annual sample size of approximately 3 million addresses. The supplementary surveys have had sample sizes of about 700,000, significantly larger than the CPS ASEC. Further information on the ACS may be found at http://www.census.gov/acs.

The ACS procedures for collecting income data differ from those of the CPS ASEC. ACS collects income data continuously with a reference period of the previous 12 months (at the time income is reported) whereas the CPS ASEC collects income data in February–April with a reference period of the previous calendar year. Annual ACS state estimates use data collected over a full year, and thus involve income reports that cover different 12 month time frames extending over a period of nearly two years. The data collection differences mean that the ACS results would be expected to have some bias relative to what the CPS ASEC is estimating. Since the CPS ASEC provides the official direct poverty estimates at the national level, SAIPE strives to estimate poverty as defined by the CPS ASEC. ACS has the advantage, though, of a much larger sample than the CPS ASEC.

In this paper we report results of an empirical study investigating the potential benefits to the SAIPE state poverty models of using data from the ACS supplementary surveys. (Benefits of using the full production ACS data when it becomes available should be greater due to the larger sample size planned for the full production ACS.) We compared results in terms of prediction error variances from the current state poverty ratio models with results from a bivariate model that used data from both CPS ASEC and ACS. We did this for both IYs 2000 and 2001 using data from two ACS supplementary surveys: the Census 2000 Supplementary Survey (C2SS, for IY 2000), and the 2001 Supplementary Survey (SS01, for IY 2001). Our results suggest that use of the ACS supplementary survey data has potential to reduce prediction error variances from the models, but there are two qualifications. First, the results vary over states, with some states actually showing increased variances. Second, we tried alternative models for using the ACS supplementary survey data, and results varied across the alternative models. Models that made more restrictive assumptions yielded apparently greater improvements in prediction error variances. The validity of these results depends, though, on the more restrictive model assumptions holding. Therefore, we also examined statistical tests (chi-squared tests) of these restrictions. The most restrictive assumptions (such as assuming no difference between what CPS ASEC and ACS are estimating) were rejected.

Section 2 presents the alternative models we tried for the CPS ASEC and ACS state poverty ratios. Section 3 contains the empirical results including prediction error variance comparisons and results of the chi-squared tests. The prediction error variances for our models are posterior variances computed via a Bayesian approach discussed in Section 2. Finally, Section 4 summarizes our conclusions.

2. Alternative Models for State Poverty Ratios

To incorporate information from both CPS ASEC and ACS data we use a general bivariate regression model with random effects. Bell (2000) discussed this model in the context of county poverty models. Section 2.1 discusses the general bivariate model, and Section 2.2 some alternative (restricted) bivariate models, as well as the univariate model currently used in SAIPE production. Section 2.3 then discusses Bayesian treatment of the models – i.e., how we obtain posterior means and variances for the state poverty ratios.

2. 1 General Bivariate Model

For any given year and age group, let Y_{1i} and Y_{2i} be the true poverty ratios (number poor / population) for state i that are being estimated by the CPS ASEC and ACS, respectively, for i = 1,ÿ,51 (including the 50 states and the District of Columbia). Note that due to data collection differences between the CPS ASEC and ACS, this model assumes that Y_{1i} … Y_{2i}, in general. Also let y_{1i} and y_{2i} be the direct sample estimated poverty ratios for state i from the CPS ASEC and ACS, respectively. Then we have

$$y_{1i} = Y_{1i} + e_{1i}$$

$$y_{2i} = Y_{2i} + e_{2i},$$

where the sampling errors e_{1i} and e_{2i} are assumed to be independently distributed as $N(0, v_{ji}), j = 1,2$. Here the v_{ji} are assumed known, though in reality they are estimates of the actual sampling variances. In the case of CPS ASEC, the direct variance estimates are smoothed using a sampling error model (Otto and Bell 1995) to get the v_{1i}. In the case of ACS, we use the direct sampling variance estimates as the v_{2i}. Finally, we assume Cov(e_{1i}, e_{2i}) = 0, because the CPS ASEC and ACS are independent samples.

Our model for the true poverty ratios is:

$$Y_{1i} = \alpha_1 + x_i' \beta_1 + u_{1i}$$

$$Y_{2i} = \alpha_2 + x_i' \beta_2 + u_{2i}$$

where

(u_{1i}, u_{2i})N are independently and identically normally distributed with means zero, with Var(u_{1i}) = s_{11} and Var(u_{2i}) = s_{22}, and with Corr(u_{1i}, u_{2i}) = ρ, and

x_i' is a row vector of regression variables.

The regression variables in x_i' for IYs 2000 and 2001 include pseudo state poverty rates constructed from Internal Revenue Service (IRS) tax data, tax non-filer ratios constructed from IRS data and state population estimates, Supplementary Security Income (SSI) state participation rates (for age 65+ only) constructed from Social Security Administration data and state population estimates, and Census 2000 state poverty ratios. For more information see the SAIPE web site mentioned earlier.

Noninformative prior distributions for the model parameters are assumed as follows:

β = (α_1, β_1N, α_2, β_2N)N is assumed to be multivariate N(**0**, cI), with c large,
s_{11} and s_{22} are assumed to be Uniform (0, m_1) and Uniform (0, m_2), with m_1 and m_2 large, and
ρ is assumed to be Uniform (-1, 1).

The values of c, m_1, and m_2 were chosen to be sufficiently large so that the priors could effectively be regarded as flat on (! 4, +4) and (0, +4) as appropriate. We used c = 1,000 for all age groups and chose appropriate values for m_1 and m_2 separately for each age group (e.g., for age 5-17 we used $m_1 = m_2 = 20$).

2. 2 Alternative Models

Bivariate Model A is the general bivariate model discussed above with no restrictions on the model parameters (except $s_{11} > 0$, $s_{22} > 0$, and $|\rho| < 1$.)

Bivariate Model B1 assumes that the CPS ASEC and the ACS estimate the same state poverty ratio, that is, $Y_{1i} = Y_{2i}$. (For Model A this implies the constraints $\alpha_1 = \alpha_2$, $\beta_1 = \beta_2$, and $u_{1i} = u_{2i}$, which in turn imply that $s_{11} = s_{22}$ and ρ = 1.)

Bivariate Model B2 assumes that the CPS ASEC and ACS models have the same regression parameters ($\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$), but with different model errors (u_{1i} … u_{2i} , so s_{11} … s_{22} and ρ … 1, in general).

Bivariate Model C assumes that, excluding the intercepts, the regression coefficients in the CPS ASEC and ACS regression equations are the same ($\beta_1 = \beta_2$).

Univariate Models: If $\rho = 0$, then Model A reduces to separate univariate regression models and we fit the CPS ASEC and ACS equations separately.

The univariate model using the CPS ASEC equation is the current SAIPE state model. If $\rho \neq 0$ then a bivariate model has potential benefits compared to this univariate model.

2. 3 Bayesian Inference for the Models

For the bivariate Model A, we used Gibbs sampling via WinBUGs (Spiegelhalter, et al. 2003) to simulate 10,500 (first 500 discarded as burn in) sets of model parameters (ρ, s_{11}, s_{22}, $\boldsymbol{\beta}$). The posterior means and variances of Y_{1i} from the CPS ASEC equation were approximated by averaging results over the simulations of (ρ, s_{11}, s_{22}) to approximate the following formulas:

$$E(Y_{1i} \mid \mathbf{y}) = E_{\rho, s11, s22}\left[E(Y_{1i} \mid \mathbf{y}, \rho, s_{11}, s_{22})\right] \quad (1)$$

$$\begin{aligned} Var(Y_{1i} \mid \mathbf{y}) = {} & E_{\rho, s11, s22}\left[Var(Y_{1i} \mid \mathbf{y}, \rho, s_{11}, s_{22})\right] \\ & + Var_{\rho, s11, s22}\left[E(Y_{1i} \mid \mathbf{y}, \rho, s_{11}, s_{22})\right] \end{aligned} \quad (2)$$

where $\mathbf{y} = \{(y_{1i}, y_{2i}), i = 1,\ldots,51\}$ is the observed data. In (1) and (2) $E(Y_{1i} \mid \mathbf{y}, \rho, s_{11}, s_{22})$ and $Var(Y_{1i} \mid \mathbf{y}, \rho, s_{11}, s_{22})$ can be readily calculated from standard formulas that account for the effects of estimating the unknown βs. (See, e.g., Bell 2000.) $E_{\rho, s11, s22}$ and $Var_{\rho, s11, s22}$ were approximated by taking the sample mean and variance across the simulations of the terms as indicated. The analogous calculations were made to obtain the posterior means and variances of the Y_{2i}, the true poverty ratios in the ACS equation, but the results are not reported here.

For the other bivariate models, and the univariate model, we used the same set of simulated parameter values (ρ, s_{11}, s_{22}) obtained under Model A to compute posterior means and variances of Y_{1i}, using analogous formulas to (1) and (2) above, but that reflected the restrictions imposed by the various models. For example, results for the univariate model were obtained by replacing $\mathbf{y}$ by $y_1 = (y_{11},\ldots,y_{1,51})'$. The use of the same simulations of (ρ, s_{11}, s_{22}) for all models provides a better indication of differences between the posterior means and variances due to conditioning on different amounts of information (i.e., CPS ASEC alone versus CPS ASEC in combination with ACS) than we would obtain by using different simulations of the model parameters from their posterior distributions under the alternative models. When doing the latter, differences in $E(Y_{1i} \mid \mathbf{y})$ and $Var(Y_{1i} \mid \mathbf{y})$ from conditioning on different amounts of information are confounded by differences in the posterior distributions of the model parameters, particularly the variances.

3. Empirical Model Comparisons

Our primary interest is in whether using ACS data in conjunction with the CPS ASEC data can reduce prediction error (posterior) variances of Y_{1i}, the "true" poverty ratio as estimated by CPS ASEC? We present here three sets of empirical results relevant to this question. First, in Section 3.1 we compare posterior variances of Y_{1i} from Model A with those from the univariate (current production) model to see what improvements may result from use of the general bivariate model. Then, in Section 3.2 we present results of chi-squared tests of the restrictions imposed on the regression coefficients by Models B2 and C to see if these more restrictive models are consistent with the data. We find Model B2 is rejected (and by implication, so is the more restrictive Model B1), while Model C is not. Finally, in Section 3.3 we compare posterior variances of Y_{1i} from Model C with those from the univariate model to see what improvements may result if we are willing to use the more restrictive Model C.

To compare posterior variances we examine their relative percentage differences under an alternative model with those from the univariate model. For example, we compare posterior variances from Model A with those from the univariate model by computing

$$100 \times \frac{Var(Y_{1i}|\mathbf{y}, \text{ Model A}) - Var(Y_{1i}|\mathbf{y}, \text{ univ. model})}{Var(Y_{1i}|\mathbf{y}, \text{ univ. model})}.$$

3. 1 Posterior Variance Comparisons for Model A

The posterior means and standard deviations of the model parameters (ρ, s_{11}, s_{22}) from the Gibbs sampling via WinBUGs of 10,000 simulations from bivariate Model A are shown in Table 3.1.

Table 3.1 Posterior means and standard deviations of the parameters of Model A for IY 2000 and IY 2001

IY 2000

age	0-4	5-17	18-64	65+
ρ	0.53 (0.38)	0.29 (0.46)	0.34 (0.41)	-0.20 (0.47)
s_{11}	2.92 (2.23)	0.81 (0.70)	0.23 (0.19)	0.76 (0.71)
s_{22}	2.24 (1.21)	1.26 (0.58)	0.48 (0.17)	0.66 (0.30)

IY 2001

age	0 - 4	5 - 17	18-64	65+
ρ	0.17 (0.53)	0.54 (0.33)	0.49 (0.40)	!0.07 (0.53)
s_{11}	1.80 (1.70)	1.85 (1.21)	0.18 (0.16)	0.39 (0.40)
s_{22}	1.53 (0.97)	0.92 (0.45)	0.37 (0.13)	0.38 (0.21)

Notice that the standard deviations for s_{11} and s_{22} are rather large relative to the posterior means, and the posterior standard deviations for ρ are large relative to the width of the interval (-1,1). These results reflect considerable uncertainty about these model parameters. This uncertainty can also be seen from estimates of the posterior densities for (ρ, $s_{11,}$ s_{22}, **β**) plotted in Figure 1 for age 5-17 in IY 2000. In fact, in regard to ρ, for none of the age groups in IY 2000 or 2001 can we conclusively determine that ρ > 0. Plots of the posterior densities of the regression coefficients shown in Figure 1 also show considerable uncertainty, more so for the first four coefficients shown (which refer to the CPS ASEC equation) than for the last four coefficients (which refer to the ACS equation). The lower level of sampling error in the ACS estimates leads to more precise estimates of the ACS regression coefficients. Also note that the posterior densities of the regression coefficients appear reasonably normal. [Note: The notation in Figure 1 of beta[1],...,beta[8] corresponds to the regression parameters **β** = (α_1, β_1N, α_2, β_2N)N.]

Table 3.2 summarizes the comparisons of the posterior variances of the state poverty ratios Y_{1i} from Model A with those from the univariate model. We see, for both years and across all age groups, only small improvements in posterior variances on average from use of the bivariate model. The min values show that some states show more dramatic variance reductions than others, while the max values show that some states show substantial variance increases. Appendix A shows more detail from these results, presenting the frequency distribution of the percentage differences in posterior variances. These tables show that substantial variance increases for states are relatively rare, while small to moderate variance reductions predominate.

The best cases for variance reductions from the bivariate Model A are age 0-4 in IY 2000 and ages 5-17 and 18-64 in IY 2001. Even for these cases the average variance reductions are small, and for most individual states the percentage differences in posterior variances are small or moderate at best. For ages 5-17, 18-64, and 65+ in IY 2000, and for ages 0-4 and 65+ in IY 2001, it is difficult to claim an advantage from using

Table 3.2 Relative percent differences of the posterior variances from Model A and the univariate model

IY 2000

Age	Mean	Min	Max
0-4	- 6.67	-19.64	19.07
5-17	- 2.65	-12.14	17.29
18-64	- 4.31	-13.33	20.30
65+	- 1.61	- 9.54	23.63

IY 2001

Age	Mean	Min	Max
0-4	! 1.46	!10.84	34.05
5-17	! 7.08	!15.71	5.53
18-64	! 7.81	!19.38	27.06
65+	! 0.48	! 9.05	20.37

bivariate Model A. Also, the fact that the age group providing the best case for IY 2000 differs from those providing the best cases for IY 2001 is somewhat disturbing, as it does not suggest consistent improvements for any age group.

3.2 Chi-Squared Tests of Model Restrictions

Section 2.2 presented three alternative bivariate models (B1, B2, and C) all of which impose restrictions on the general bivariate Model A. The restrictions implied by Models B2 and C can be tested by testing the following null hypotheses:

H1: $\alpha_1 = \alpha_2$, $\beta_1 = \beta_2$ (Model B2)

H2: $\beta_1 = \beta_2$ (Model C)

Hypothesis H1 postulates equality of the regression coefficients in the CPS ASEC and ACS equations. Hypothesis H2 postulates this equality apart from the intercept terms. We test these hypotheses against the alternative hypothesis of Model A holding with no restrictions. While we will not explicitly test Model B1 (CPS ASEC and ACS estimate the same poverty ratios), note that if H1 is rejected so, by implication, is the more restrictive Model B1.

To test the hypotheses H1 and H2 we formulate chi-squared statistics using the posterior means and covariance matrices of the regression coefficients under Model A. From a Bayesian perspective, this is equivalent to seeing if the values under these null hypotheses lie within a given highest posterior density region. More broadly speaking, the chi-squared statistics check if the restrictions under H1 and H2 are reasonably consistent with the posterior distribution of the regression parameters under the general Model A.

More specifically, the chi-squared statistic for testing H2 is

$$\chi^2 = (b_1 - b_2)'[\mathrm{Var}\,(b_1 - b_2)]^{-1}(b_1 - b_2)$$

where b_1 and b_2 are the posterior means of β_1 and β_2, while $\mathrm{Var}(b_1 - b_2)$ is the posterior covariance matrix of $\beta_1 - \beta_2$. This statistic has three degrees of freedom for ages 0-4, 5-17, and 18-64, and four degrees of freedom for age 65+ (the difference being due to the additional inclusion of the SSI participation rate in x_i' for age 65+). For testing H1 an analogous statistic is used that also involves the intercepts, α_1 and α_2, and which has four degrees of freedom for ages 0-4, 5-17, and 18-64, and five degrees of freedom for age 65+. We compare χ^2 to five percent critical values from the chi-squared distribution; these are 7.8, 9.5, and 11.1 for three, four, and five degrees of freedom, respectively.

The results of the Chi-squared tests for IYs 2000 and 2001 are given in Table 3.3. Values that are significant at the five percent level are shown in bold.

Table 3.3 Chi-Squared statistics for testing hypotheses H1 and H2 for IYs 2000 and 2001

IY 2000

Age	0-4	5-17	18-64	65+
H1	8.4	**35.3**	**44.1**	8.6
H2	5.5	2.5	2.5	5.2

IY 2001

Age	0-4	5-17	18-64	65+
H1	7.7	**18.6**	**12.2**	0.7
H2	7.0	0.4	2.5	0.6

For both IY 2000 and IY 2001, we reject the hypothesis H1 for ages 5-17 and 18-64. We fail, however, to reject H2 for all age groups. The results suggest that assuming regression intercepts are the same between the CPS ASEC and ACS equations is not tenable for ages 5-17 and 18-64, and hence that we should reject Model B2, and by implication, the more restrictive Model B1. Given this result, suggesting some systematic difference in level between what CPS ASEC and ACS are estimating for poverty, we would be disinclined to use Model B2 for ages 0-4 and 65+ as well. On the other hand, the failure for all age groups in both years to reject H2, corresponding to the less restrictive Model C, suggests that perhaps the regression parameters other than the intercepts can be assumed to be the same in the CPS ASEC and ACS equations, so we might consider using Model C instead of Model A. The consequences of this for posterior variances are examined in the next section.

3.3 Posterior Variance Comparisons for Model C

Posterior variances for Model C were computed as discussed in Section 2.3, i.e., using equations (1) and (2) with the simulations of (ρ, s_{11}, s_{22}) obtained under Model A, and with $E(Y_{1i} \mid \mathbf{y}, \rho, s_{11}, s_{22})$ and $\mathrm{Var}(Y_{1i} \mid \mathbf{y}, \rho, s_{11}, s_{22})$ computed to account for the Model C restriction, $\beta_1 = \beta_2$. Table 3.4 presents summaries of the percent differences of the resulting posterior variances from those for the univariate model; these results can be compared to those of Table 3.2. Doing so we see that the variance reductions from Model C are, on average, substantially larger than those from Model A. The largest reductions are about 50 percent or more, and while there are some variance increases, the maximum increases from use of Model C are not as severe as those from Model A. Note also that for age 5-17 all states show variance reductions with Model C.

Table 3.4 Relative percent differences of the posterior variances from Model C and the univariate model

IY 2000

Age	Mean	Min	Max
0-4	−19.3	−49.8	15.0
5-17	−18.1	−50.3	−0.2
18-64	−17.2	−47.1	12.8
65+	−18.3	−48.4	17.4

IY 2001

Age	Mean	Min	Max
0-4	−17.2	−56.5	22.3
5-17	−17.8	−44.3	−1.6
18-64	−24.6	−55.3	23.3
65+	−28.5	−64.0	18.7

To put these results in context, Appendix B presents further tables showing the average posterior variances from Models A and C, and from the univariate model, along with average sampling error variances for the CPS ASEC and ACS direct estimates. The latter are the variances one would have from each data source if modeling was not used to improve the estimates. The tables show the substantially lower variances of the ACS direct estimates compared to those for CPS ASEC, due to the larger sample size of the ACS supplemental surveys. (This difference will increase with the production ACS.) They also show that the univariate models achieve substantial variance reductions compared to the direct CPS ASEC estimates, and that further reductions under Models A and C are not as large.

The larger variance reductions under Model C than under Model A are presumably due to increased precision in the estimation of the regression coefficients under Model C's assumption that, apart from the intercepts, the regression coefficients are common to both equations. Given the substantially lower sampling variances from ACS, under this assumption the ACS data should provide relatively more information for estimation of the common regression coefficients than does the CPS ASEC data. So there appears to be more potential for improvement from using the ACS data to improve estimation of the regression coefficients (if the assumption that they are common to both equations holds) than from using the ACS data to improve prediction of the state random effects (which is done by all the bivariate models.)

Note one important qualification to these results. The posterior variances quoted from Model C assume that the Model C restriction of $\beta_1 = \beta_2$ holds exactly (and that the other model assumptions are true, these being the assumptions for Model A). We determined that the assumption $\beta_1 = \beta_2$ was plausible by performing the chi-squared tests of Section 3.2. So to decide to use Model C, we had to actually estimate Model A and perform the tests of $\beta_1 = \beta_2$. If we then quote posterior variances calculated as if Model C were known to be true we would not be accounting for this estimation and testing and we would thus understate our uncertainty about the true poverty ratios. One way around this dilemma that we intend to explore is to use an informative prior distribution for $\beta_1 - \beta_2$ with mean zero. In the future such a prior might be developed from estimation results for previous years.

4. Summary

Our goal in this paper was to examine the potential benefits of modeling the CPS ASEC and ACS state poverty ratio estimates jointly to improve estimates of poverty ratios in the CPS ASEC equation. We examined alternative models for doing this using CPS ASEC estimates for IYs 2000 and 2001, and ACS data from the supplementary surveys of C2SS and SS01. The models included regression variables constructed from administrative records data and 2000 Census poverty ratios, along with state random effects and sampling error components. We can summarize the results from our empirical study as follows:

1. In comparing the general bivariate Model A with the current univariate model, there is, at best, a small average improvement in state posterior variances. Results are variable for individual states, with some states showing larger improvements, but a few showing substantial variance increases.

2. For IYs 2000 and 2001, chi-squared tests rejected the more restrictive bivariate Models B2 and B1 for ages 5-17and 18-64, but failed to reject model C for all four age groups. This suggests a systematic difference in level between the CPS ASEC and ACS estimates, but provides no evidence that other regression coefficients are different in the two equations.

3. Model C had substantially lower posterior variances, on average, than the general Model A. However, in order to decide to use model C, we have to estimate Model A and test whether $\beta_1 = \beta_2$. Quoting posterior variances as if Model C is known to be true thus understates uncertainty.

Further studies are needed when we have the full production ACS data.

References

Bell, William R (2000), "Models for County and State Poverty Estimates," appendix to *Small-Area Estimates of School-Age Children in Poverty – Evaluation of Current Methodology*, Panel on Estimates of Poverty for Small Geographic Areas, eds. Constance F. Citro and Graham Kalton, Committee on National Statistics, Washington D.C.:National Academy Press.

Bell, William R (1999), "Accounting for Uncertainty About Variances in Small Area Estimation," Bulletin of the International Statistical Institute, 52nd Session, Helsinki.

Fay, Robert E. and Herriot, Roger A. (1979), "Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data," *Journal of the American Statistical Association*, **74**, 269-277.

Fay, Robert E. and Train, George F. (1997). "Small Domain Methodology for Estimating Income and Poverty Characteristics for States in 1993," American Statistical Association, Proceedings of the Social Statistical Section, 183-188.

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, 160-165.

Spiegelhalter, D., Thomas, A., Best, N. and Lunn, D. (2003), "WinBUGS User Manual," Version 1.4, MRC Biostatistics Unit, Institute of Public Health, Robinson Way, Cambridge CB2 2SR, UK.

Appendix A: Frequency distributions over states of the relative percentage differences of the state posterior variances from bivariate Model A and the univariate model

Table A.1: IY 2000

Percentage Difference	age 0-4	age 5-17	age 18-64	age 65+
-20 to -15	2	0	0	0
-15 to -10	15	2	9	0
-10 to -5	19	16	21	18
-5 to 0	11	22	13	20
0 to 5	1	7	3	5
5 to 10	0	3	1	4
10 to 15	1	0	1	2
15 to 20	2	1	2	1
20 to 25	0	0	1	1

Table A.2: IY 2001

Percentage Difference	age 0-4	age 5-17	age 18-64	age 65+
-20 to -15	0	2	14	0
-15 to -10	2	11	15	0
-10 to -5	15	25	7	6
-5 to 0	21	9	8	29
0 to 5	7	3	3	10
5 to 10	3	1	0	3
10 to 15	0	0	0	1
15 to 20	1	0	1	1
20 to 25	1	0	2	1
25 to 30	0	0	1	0
30 to 35	1	0	0	0

Appendix B: Comparing average prediction error variances of direct and alternative model-based state poverty ratio estimates using CPS ASEC and ACS supplementary survey data

Age 0-4	**IY 2000**	**IY2001**
Direct estimates		
ASEC	11.794	11.633
ACS	4.083	3.479
Model estimates		
Univariate	2.973	2.265
Bivariate A	2.798	2.243
Bivariate C	2.333	1.779

Age 5-17	**IY 2000**	**IY2001**
Direct estimates		
ASEC	4.907	4.957
ACS	1.833	1.521
Model estimates		
Univariate	0.996	1.576
Bivariate A	0.974	1.472
Bivariate C	0.780	1.264

Age 18-64	**IY 2000**	**IY2001**
Direct estimates		
ASEC	1.135	1.189
ACS	0.359	0.309
Model estimates		
Univariate	0.260	0.233
Bivariate A	0.251	0.216
Bivariate C	0.208	0.167

Age 65+	**IY 2000**	**IY2001**
Direct estimates		
ASEC	3.993	4.046
ACS	1.162	0.748
Model estimates		
Univariate	0.910	0.663
Bivariate A	0.898	0.658
Bivariate C	0.716	0.439

Figure 1: Posterior Densities for Model A Parameters - Income Year 2000 - Age 5-17