UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

December 31, 2002

DSSD A.C.E. REVISION II MEMORANDUM SERIES #PP-53

MEMORANDUM FOR Donna Kostanich
Chair, A.C.E. Revision II Planning Group

From: Dawn Haines
Chair, A.C.E. Revision II Estimation Subgroup

Prepared by: Roger Shores
Decennial Statistical Studies Division

Subject: A.C.E. Revision II: Adjustment for Correlation Bias

The attached document focuses on the adjustment for correlation bias that was incorporated into the A.C.E. Revision II dual system estimates (DSEs). Specifically, we adjust for correlation bias for adult males and we assume no correlation bias for females and children. This report is based on DSSD Census 2000 Procedures and Operations Memorandum Series memorandum B-12*, written by William R. Bell.

A.C.E. Revision II: Adjustment for Correlation Bias

Roger Shores

Decennial Statistical Studies Division

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

Table of Contents

Tables

A.C.E. Revision II: Adjustment for Correlation Bias

Executive Summary

We compared A.C.E. Revision II dual system estimates (DSEs), prior to the correlation bias adjustment, aggregated to the national level for age-race-sex groups (Black versus non-Black race) against results from demographic analysis (DA) to adjust for correlation bias in the A.C.E. Revision II DSEs. We made these comparisons for sex ratios. We then used a previously developed method (Bell 1999) to adjust for correlation bias in the DSEs for adult males using the DA sex ratios, assuming no correlation bias for females or children. We made significant adjustments for correlation bias for Blacks. For non-Blacks, we made a relatively small adjustment for correlation bias in the two older age categories (30-49 and 50 and over). Because of inconsistencies between DA and the revised estimates for non-Blacks, it was not possible to reliably adjust for correlation bias for non-Blacks in the 18-29 age category.

What is correlation bias?

Correlation bias can be defined as the error that would result from comparing the estimated DSEs to DSEs based on "perfect data," that is, DSEs based on data with no sampling error and no other biases. Correlation bias results from failure of the general independence assumption underlying DSEs due to

- **causal dependence** – the act of being included in the Census makes someone more or less likely to be included in the A.C.E., or

- **heterogeneity** – Census and A.C.E. inclusion probabilities vary over persons within post-strata.

If heterogeneity exists in the sense that persons likely to be missed in the Census are also more likely to be missed in the A.C.E., then correlation bias is negative, implying underestimation by the DSEs. While causal dependence can lead to either positive or negative biases in DSEs, generally the concern about correlation bias is heterogeneity leading to underestimation.

What evidence do we have of correlation bias in DSEs ?

Historical evidence of correlation bias in DSEs comes from comparisons of results aggregated to the national level against DA estimates for age-race-sex groups (Black versus non-Black race). Comparisons against DA population totals provide a crude check because of the relatively high level of uncertainty about the effects on DA totals of errors in estimates of undocumented immigration and emigration. The DA sex ratios (number of males over number of females) are believed to be more accurate than DA totals and provide a more refined check of correlation bias for adult males (assuming negligible correlation bias for adult females). Comparisons of 1990 Post Enumeration Survey (PES) estimates (357 post-strata) against 1990 DA estimates (totals and sex ratios) gave evidence of significant correlation bias in DSEs for adult Black males and

possible correlation bias for adult non-Black males (Bell 1991). There was no evidence of correlation bias for children or adult females in the 1990 PES.

Prior to making an adjustment for correlation bias, the 2000 DA sex ratios for adult Blacks significantly exceed those for A.C.E. Revision II, strongly suggesting correlation bias in DSEs for adult Black males. DA sex ratios for non-Blacks 30-49 and 50 and over only slightly exceed those from A.C.E. Revision II, suggesting at most small amounts of correlation bias for these populations.

Why should we adjust for correlation bias?

Adjustment for correlation bias was considered but ultimately not implemented for the estimates from the 1990 Post Enumeration Survey and for the March 2001 A.C.E. estimates. A key reason, though not the only reason, was the unresolvable uncertainty about which model is most appropriate for correlation bias adjustment. This issue arises since DA provides estimates only at the national level by age, sex, and Black versus non-Black race groups and thus direct estimates of correlation bias (for adult males - see discussion of results below) are available only for these national aggregates. Various models can then be used to synthetically allocate the correlation bias estimated at the national level for adult males to post-strata within the age-race groups. The different models would produce different sub-national results. Since different models provide the same fit to the available data, there is unresolvable uncertainty about which model is most appropriate.

When concern focused on net undercount, as in 1990 and going into 2000, using DSEs without adjustment for correlation bias was expected to move from the census count towards the truth. We could do better still by adjusting the DSE for correlation bias, subject to the assumption that the female DSEs are unbiased (or at least not overestimates), although it is then necessary to consider the uncertainties about sub-national allocation of correlation bias.

Because the revised preliminary estimates of Census 2000 net undercount were close to zero, it was thought necessary to reconsider correlation bias adjustment for the A.C.E. Revision II estimates. The presence of a large number of census erroneous enumerations makes it more critical to adequately adjust for census omissions. Correlation bias results in an underestimation of omissions in the DSE. If we have a true net overcount, or a sufficiently small undercount, then if we don't adjust for correlation bias we could estimate a net overcount that may be further from the truth than the census count. For example, suppose there is a true net overcount of about 0.5%, and DSEs with correlation bias adjustment estimate about this amount, but DSEs without adjustment for correlation bias estimate an overcount of 1.5%. Then using DSEs without correlation bias adjustment produces higher aggregate absolute error (1.5% - 0.5% = 1%) than the unadjusted census count (0.5% - 0% = 0.5%), but DSEs with correlation bias adjustment are close to the truth in aggregate.

Another relevant issue that was of concern in 1990 was what would happen if dual system estimates for females were not unbiased, but instead were overestimates resulting from biases in the data other than correlation bias. If this occurred, then applying DA sex ratios to DSE aggregates for females would produce overestimates for males, making things worse. Given the overestimation by the March 2001 A.C.E. estimates caused by undetected census erroneous enumerations, this concern would have been quite relevant for the March 2001 A.C.E. Since A.C.E. Revision II attempts to adjust for other biases in the DSEs, this concern is now much less relevant.

For A.C.E. Revision II, analysts decided to adjust for correlation bias using the constant relative bias or "Two-Group" Model. Since other available models all fit the data equally well, giving the same overall result, we chose (1) the simplest model and (2) the model expected to have the lowest variance among the alternative models.

How did we adjust for correlation bias in A.C.E. Revision II DSEs (for adult males)?

We used a model known as the "Two-Group" or "Constant relative Bias" Model to adjust for correlation bias in adult males for A.C.E. Revision II. This model assumes no correlation bias for adult females and that the DA sex ratios are accurate. The defining characteristic of the model is that relative correlation bias is assumed constant across post-strata within an age-race group. Within each age-race group, the DA sex ratio is combined with the aggregated female DSEs to produce a "control total." The aggregated male DSEs times a constant are equated with this control total, and solving for the constant gives the estimated adjustment needed to correct for correlation bias within the age-race group.

This approach produced the following results for A.C.E. Revision II estimates:

- We made significant adjustments for correlation bias for adult Black males. The revised estimates for the 18-29 and 30-49 age groups were adjusted to remove correlation biases of –7.5 and –9.2 percent, respectively. Corresponding correlation bias estimates from the 1990 PES were –8.0 and –7.7 percent.

- The A.C.E. Revision II estimate for Black males 50 and over was adjusted to correct for a correlation bias of –5.2 percent. The correlation bias for this group was –8.2 percent in the 1990 PES. One explanation for the reduction is that significant revisions were made to the DA estimates for the cohort of Black males age 65-74 in 2000.

- Because of inconsistencies between DA and A.C.E. Revision II estimates for non-Blacks 18-29, we concluded that we cannot use the DA sex ratios to adjust for correlation bias for non-Black males 18-29. We made the assumption that there was no correlation bias for non-Black males 18-29 on the grounds that:

- 2000 A.C.E. Revision II estimates of relative correlation bias for non-Black males in the age categories 30-49 and 50 and over are small, and

- the 1990 PES estimate of relative correlation bias for non-Black males 18-29 was small.

Introduction

Dual system estimates (DSEs) are said to contain bias if they systematically underestimate or overestimate the true population. Biases in the sample estimates of the components of the DSE formula can lead to biases in the DSEs, as discussed in Mulry (1991) and Mulry and Spencer (1991, 1993). Even in the absence of any of these biases, DSEs can still be subject to another form of bias called *correlation bias*, resulting from failure of the general independence assumption that underlies the DSEs. This independence assumption can fail due either to *causal dependence* between the act of inclusion in the Census and the act of inclusion in the A.C.E., or to *heterogeneity* across persons with respect to the probabilities of being included in the Census and in the A.C.E. DSEs are constructed within post-strata to reduce heterogeneity in the inclusion probabilities, so heterogeneity leading to correlation bias exists only if the inclusion probabilities vary across persons *within* a post-stratum. For further general discussion of correlation bias, see Griffin (2000).

When heterogeneity exists it is generally suspected to be of the form in which persons (within a post-stratum) more likely to be missed in the Census are also more likely to be missed in the coverage survey (A.C.E.). Correlation bias resulting from this form of heterogeneity is negative, reflecting systematic underestimation of the true population by the DSEs. The direction of the effect of causal dependence, if it exists, is less certain. It could be that persons included in the Census are made more aware of the Census process, and hence are more likely to be included in the A.C.E., than are those missed by the Census. This type of dependence would lead to underestimation by the DSEs. Or, it could be that persons included in the Census feel they have already responded to the Census Bureau, and so are more resistant to being included in the A.C.E., than those missed by the Census. This type of dependence would lead to overestimation by the DSEs. It is difficult to say which of these two types of causal dependence might be more likely, so it is unclear whether correlation bias due to causal dependence would lead to underestimation or overestimation by the DSEs.

Two approaches that have been used to investigate possible correlation bias in DSEs from a given post-stratification are:

1. compare to results from demographic analysis,

2. compare to results from alternative post-stratification DSEs or more general models (e.g., logistic regression).

Both these approaches have strengths and limitations. Two other approaches that have been used that provide some evidence about possible correlation bias in DSEs are triple system estimation (Zaslavsky and Wolfgang 1993) and ethnographic observation. However, both these approaches have been used only in limited test sites, and it was therefore not feasible to use either to adjust for correlation bias in the A.C.E. Revision II estimates.

Demographic analysis (DA), discussed by Robinson et al. (1993), has the advantage that its estimates are constructed from administrative data sources, some of which (birth and death registration data) are believed to be quite accurate. Comparison of DSE results against DA estimates provides an aggregate check for correlation bias whether due to causal dependence or heterogeneity (with some qualifications regarding allowance for other biases, as noted below). However, DA estimates are adversely affected by errors in the administrative data, such as uncertainty about the level of emigration from the U.S. and uncertainty about the level of undocumented immigration. For this reason, DA population estimates (DA totals) are thought to be less accurate than DA *sex ratios* (number of males over number of females). This reflects an assumption that errors in migration estimates are not grossly different for males than they are for females.

In addition to errors in its administrative data sources, the primary limitation of DA results is a lack of detail. Difficulties in using administrative data to construct estimates of subnational migration mean that subnational DA estimates, while providing useful indicators, are significantly less accurate than DA national estimates. Also, limited racial detail in the administrative data sources, along with differences in racial classification from the Census, limits separate DA estimates by race to simply Black and non-Black. This limitation was somewhat more pronounced in 2000 than in 1990 because the allowance of multiple race responses to the 2000 Census created some uncertainty about appropriate definitions of the Black and non-Black groups for comparability of DA and A.C.E. results.

Because of the limitations of DA, to adjust for correlation bias in A.C.E. Revision II we only used DA data at the national level broken down by age, race (Black and non-Black groups), and sex. We use DA sex ratios to explore in more detail possible correlation bias for adult males while assuming no correlation bias for adult females.

Comparing 2000 A.C.E. Revision II and DA Estimates

Adjusting the DA estimates for comparability with A.C.E. Revision II estimates

For comparison to the 2000 A.C.E. Revision II results, we use DA estimates revised as of September 18, 2001 (Robinson 2001b). Definitional differences between DA and A.C.E. Revision II require adjustments to the data to make the two sets of estimates comparable. Since we wish to adjust for correlation bias in the revised estimates, we make adjustments to the DA estimates, and not the other way around. We made the following adjustments to the DA totals for Blacks and non-Blacks to make them comparable to A.C.E. Revision II results. We then computed DA sex ratios from these adjusted totals.

- We subtracted the Census count of the group quarters population from the 2000 DA totals, since the group quarters population is not part of the A.C.E. Revision II universe.

- We subtracted estimates of Black Hispanics from the DA totals for Blacks and added these estimates to the DA totals for non-Blacks. We need this adjustment because A.C.E. Revision II assigns Black Hispanics to its Hispanic race domain (domain 3), not its Black race domain (domain 4). The implied DA estimates of Black Hispanics for 2000 were obtained by inflating the Census counts of Black Hispanics by adjustment factors corresponding to the DA estimates of Black undercount, since separate DA estimates of Hispanic undercount are not available.

- Robinson (2001a) explains how alternative Census tabulations corresponding to alternative definitions of the Black and non-Black race groups can be used in comparing DA to Census results. He considers two extremes for assignment of individuals to the Black and non-Black groups. Under his Model 1, only those persons who checked only Black for the Census race question are classified as Black. Under his Model 2, persons who checked Black *and* any other race are also classified as Black. In comparing DA and A.C.E. Revision II results, however, this affects only the Census group quarters population that is subtracted from the DA totals. In particular, it affects the allocation of the group quarters population to the Black and non-Black race groups. However, the effects on the DA sex ratios were negligible, and thus so were the effects on estimates of correlation bias. Hence, we use only one set of results for sex ratio comparisons and one set for estimating correlation bias. The results are presented using Model 2.

Comparing DA and A.C.E. Revision II estimated sex ratios

Comparisons of DA and A.C.E. Revision II sex ratios, prior to the correlation bias adjustment, are given in Table 1.a. We draw the following conclusions from the results.

- 2000 DA sex ratios for adult Blacks significantly exceed those for A.C.E. Revision II, strongly suggesting correlation bias in DSEs for adult Black males.

- DA sex ratios for non-Blacks 30-49 and 50 and over only slightly exceed those from A.C.E. Revision II, suggesting at most small amounts of correlation bias for persons in these groups. (Expressed to more digits, the non-Black DA and A.C.E. Revision II sex ratios for 30-49 are 1.0060 and .9905, respectively, and those for 50 and over are .8561 and .8471.)

- The DA sex ratio for non-Blacks 18-29 (1.044) is slightly lower than that from A.C.E. Revision II (1.048). This comparison is suspect and probably reflects the underlying inconsistency of the DA and revised estimates for this group, rather than saying anything about correlation bias. Because of this anomaly, in the next section we do not attempt to adjust for correlation bias for non-Blacks 18-29.

- Except for non-Blacks 18-29 and 50 and over, the sex ratio comparisons between DA and A.C.E. Revision II are similar to the corresponding comparisons for 1990. The different treatment of the group quarters populations in A.C.E. Revision II and 1990 PES could lead to some differences. This may explain the higher sex ratios for non-Blacks 50 and over observed for both the 2000 DA and revised estimates in comparison to the corresponding 1990 sex ratios.

Table 1.a Sex Ratios from DA and A.C.E. Revision II – 2000

Age	Black[2] A.C.E. Rev II	Black[2] DA	Non-Black A.C.E. Rev II	Non-Black DA
18-29	.83	.90	**1.05**	**1.04**
30-49	.81	.89	.99	1.01
50+	.72	.76	.85	.86

Table 1.b Sex Ratios from DA and the PES – 1990

Age	Black[2] PES	Black[2] DA	Non-Black PES	Non-Black DA
18-29	.83	**.**90	1.02	1.02
30-49	.84	.91	.99	1.01
50+	.72	.78	.81	.82

Notes to Table 1:

1. The sex ratios for A.C.E. Revision II are those before any adjustments are made to correct for correlation bias.

2. For 2000, DA estimates revised as of September 18, 2001 are used. Before computing the DA sex ratios, the DA totals were adjusted for comparability with A.C.E. Revision II and PES estimates as discussed in the text.

3. Sex ratios for DA in 2000 with race assignment of the group quarters population using Model 1 are the same as those with Model 2 to the accuracy shown in the table.

Adjusting for Correlation Bias in Adult Males for A.C.E. Revision II

Alternative combining models are presented in Bell (1993) and Bell et al. (1996), and statistical refinements to these models have been developed by Elliott and Little (2000). There was insufficient time to thoroughly investigate the performance of several models, but in the appendix undercount rates and undercount estimates are provided for the model implemented and for three others. The performance of these other models, and perhaps others, may be investigated further.

Correlation bias adjustments from the "Two-Group" Model

The model that we use assumes relative correlation bias (percent error) is constant over male post-strata within adult age groups for Blacks and non-Blacks. This model is discussed in Bell (1999).

This model was implemented because its performance is believed to be competitive with that of any other model that might have been used, while offering a considerable advantage in simplicity. The model can also be called the "Constant Relative Bias" Model. It has been called the Two-Group Model because it can be derived by postulating two groups of people within each male post-stratum (say hard-to-count and easy-to-count), and assuming a parameter ϕ is constant across post-strata. To calculate an estimate of ϕ within a particular age-race group, we solve the equation $\phi \sum_i DSE\ males = r_{DA} \sum_i DSE\ females$ for ϕ, where r_{DA} is the DA sex ratio for the age-race group and the right hand side of the equation is a "control total" consisting of the DA sex ratio for the group multiplied by the summed female DSEs within the group. This provides an estimate of a multiplicative correction factor to correct each male DSE in the age/race group for correlation bias. The parameter ϕ can be defined as the ratio of the true population over the expected value of the DSE, since the control total is accepted as the true count of males in the age-race group.

Table 2 gives the parameter estimates calculated for A.C.E. Revision II under the Two-Group Model.

Table 2: Estimates of Adjustment Factor for A.C.E. Revision II

Age	Black	Non-Black
18-29	1.08	1.00
30-49	1.10	1.02
50+	1.05	1.01

Table 3.a gives estimation results for the Two-Group Model for A.C.E. Revision II, in terms of the relative correlation bias estimates expressed as percents. Table 3.b gives similar results for the 1990 PES. The results show the following:

- The two-group model estimate of relative correlation bias for Black males 18-29 for A.C.E. Revision II is similar to that in 1990, implying around eight percent underestimation by the DSEs.

- The estimate of relative correlation bias for Black males 30-49 was slightly higher for A.C.E. Revision II (–9.2 percent) compared to the estimate from 1990 (–7.7 percent).

- The estimate of relative correlation bias for Black males 50 and over in A.C.E. Revision II (–5.2 percent) is smaller than the estimate in 1990 (–8.2 percent). This may be partly due to revisions to the DA estimates for the cohort of Black males who were 65-74 in 2000.

- A.C.E. Revision II estimates of relative correlation bias for non-Black Males 30-49 and 50 and over are negative and small in magnitude, and are similar to results obtained for 1990.

- The relative correlation bias estimate for non-Blacks 18-29, 0.4 percent, is hard to explain as "correlation bias." Possible reasons for this result are that males in this group who were more likely than others to be missed in the census were more likely than others to be found in A.C.E. Revision II, or that females in this group exhibit a larger amount of conventional correlation bias than do males. Neither of these explanations makes much sense. A better interpretation of the result for this group is simply that because of the inconsistency between DA and A.C.E. Revision II results, we cannot estimate correlation bias for non-Blacks 18-29.

- Given this conclusion, we decided to proceed by assuming no correlation bias for non-Black males 18-29 on the grounds that
  - estimates of relative correlation bias for older (30-49 and 50+) non-Black males are small in 2000, and
  - the estimate of relative correlation bias for non-Black males 18-29 in 1990 was small.

Table 3.a Relative Correlation Bias Estimates for 2000 A.C.E. Revision II (results from Two-Group Model, expressed as percents)

Age	Black males	Non-Black males
18-29	–7.5	.4
30-49	–9.2	–1.5
50+	–5.2	–1.0

Table 3.b Relative Correlation Bias Estimates for 1990 PES
(results from Two-Group Model, expressed as percents)

Age	Black males	Non-Black males
18-29	–8.0	–.3
30-49	–7.7	–1.6
50+	–8.2	–1.2

Notes to Table 3:

1. These estimates use the DA sex ratios from Table 1.

2. For the Two-Group Model, the estimate of relative correlation bias, expressed as a percent, is $100\left(\frac{\text{A.C.E. Revision II sex ratio}}{\text{DA sex ratio}} - 1\right)$ for A.C.E. Revision II, and

$100\left(\frac{\text{PES sex ratio}}{\text{DA sex ratio}} - 1\right)$ for the 1990 PES.

Table 4 shows totals for the DSEs in A.C.E. Revision II for males in the different age-race categories. The first column gives the totals for the original DSEs not adjusted for correlation bias. The second column gives the totals for the DSEs after adjusting for correlation bias. The third column shows the change that adjustment produces in each category, and the final column expresses this change in terms of percentages.

Table 4 Effect of Correlation Bias Adjustment for A.C.E. Revision II

Race/Age Categories	DSE Totals Without Corr Bias Adj	DSE Totals With Corr Bias Adj	Difference Between DSE Totals	Percent Increase
Black males				
18 - 29	2,582,683	2,790,833	208,150	8.1
30 - 49	4,385,308	4,832,212	446,904	10.2
50+	2,746,088	2,895,811	149,723	5.5
Non-Black males				
18 - 29	19,047,719	19,047,719	0	0
30 - 49	36,735,919	37,312,620	576,701	1.6
50+	30,311,508	30,632,206	320,698	1.1

Note that for non-Black males 18-29 the correlation bias adjustment is forced to have the value zero. Also, the correlation bias adjustment factor for females and persons age 17 and younger is set equal to 1. Therefore, the correlation-bias adjusted DSE is equal to the original DSE for persons in these groups.

References

Bell, William, "1990 Post Enumeration Survey Evaluation Project P13: Use of Alternative Dual System Estimators to Measure Correlation Bias," 1990 Coverage Studies and Evaluation Memorandum Series, U.S. Bureau of the Census, 1991.

Bell, William, "Using Information from Demographic Analysis in Post-Enumeration Survey Estimation," Journal of the American Statistical Association, 88, pp. 1106-1118, 1993.

Bell, William, "Accuracy and Coverage Evaluation Survey: Combining Logistic Regression DSEs (Target Model) with 1990 Demographic Analysis Results," DSSD Census 2000 Procedures and Operations Memorandum Series Q-12, September 22, 1999.

Bell, William, Campbell Gibson, Prithwis Das Gupta, Gregory Spencer, J. Gregory Robinson, Mary Mulry, Elizabeth Vacca, Robert Fay, and Charlene Leggieri, "Report of the Working Group on the Use of Demographic Analysis in Census 2000," U.S. Bureau of the Census, 1996.

Elliott, Michael and Roderick Little, "A Bayesian Approach to Combining Information from a Census, a Coverage Measurement Survey, and Demographic Analysis," Journal of the American Statistical Association, 95, pp. 351-362, 2000.

Griffin, Richard, "Accuracy and Coverage Evaluation Survey: Definition and Explanation of Correlation and Related Heterogeneity Bias," DSSD Census 2000 Procedures and Operations Memorandum Series Q-35, September 29, 2000.

Mulry, Mary, "1990 Post Enumeration Survey Evaluation Project P16: Total Error in PES Estimates for Evaluation Post-strata," 1990 Coverage Studies and Evaluation Memorandum Series R-6, U.S. Bureau of the Census, July 11, 1991.

Mulry, Mary and Bruce Spencer, "Total Error in PES Estimates of Population (with discussion)," Journal of the American Statistical Association, 86, pp. 839-863, 1991.

Mulry, Mary and Bruce Spencer, "Accuracy of the 1990 Census and Undercount Adjustments," Journal of the American Statistical Association, 88, pp. 1080-1091, 1993.

Robinson, J. Gregory, Bashir Ahmed, Prithwis Das Gupta, and Karen Woodrow, "Estimation of Population Coverage in the 1990 United States Census Based on Demographic Analysis (with discussion)," Journal of the American Statistical Association, 88, pp. 1061-1079, 1993.

Robinson, J. Gregory, "Accuracy and Coverage Evaluation Survey: Demographic Analysis Results," DSSD Census 2000 Procedures and Operations Memorandum Series B-4*, February 28, 2001a.

Robinson, J. Gregory (2001b) "ESCAP II: Demographic Analysis Results," Report No. 1 to the Executive Steering Committee for A.C.E. Policy II, U.S. Bureau of the Census.

Zaslavsky, Alan and Glenn Wolfgang, "Triple System Modeling of Census, Post-Enumeration Survey, and Administrative List Data," Journal of Business and Economic Statistics, 11, pp. 279-288, 1993.

Undercount Rates and Undercount Estimates from the Two-Group Model and Selected Other Models

In this section, undercount rates and undercount estimates are presented for the Two-Group and three other models, categorized by tenure and the age/sex and major race categories used in A.C.E. Revision II.

Other Models

The other models implemented are called Modified Two-Group, Fixed Relative Risk (FRR), and Prithwis, the last named for its author, Prithwis Da Gupta of Population Division.

The modified Two-Group Model is similar to the Two-Group Model. The difference between the two is that, in the modified Two-Group Model, Hispanics are treated as a separate category, instead of being included in the non-Black category as they are in the Two-Group Model. There are therefore nine age/race groups rather than six. Correlation bias for Hispanics in the modified Two-Group Model is assumed to be equal to that of Blacks in each age category. The result is that the correlation-bias adjusted DSEs for Hispanics are significantly larger, and those for non-Black non-Hispanics smaller, under this model compared to the DSEs calculated in the Two-Group Model. Correlation-bias adjusted DSEs for Blacks do not change, but the adjustment for non-Black non-Hispanics is reduced to compensate for the larger adjustment for Hispanics.

The FRR Model, like the Two-Group Model, postulates the existence of a parameter that is assumed constant across post-strata within a particular age/race group. The parameter in this case is the ratio of the probability of inclusion in A.C.E. given capture in the census to the probability of inclusion in A.C.E. given lack of capture in the census. Through the estimation of this parameter, another method is developed by which correlation-bias adjusted DSEs can be calculated. The summed male DSEs within an age/race group are subtracted from the control total for that group. The difference is then allocated proportionately across post-strata within the age/race group according to an estimate of the total number of people not captured by the census in the group. For a particular post-stratum, the difference is multiplied by the percentage of the total number missed who are estimated to have been in that post-stratum. The correlation bias adjusted DSE consists of this product added to the original DSE. In the FRR Model, post-strata with low match rates are adjusted more than are those with high match rates. The FRR Model uses only Black and non-Black race categories.

In the Prithwis Model, there is once again a parameter that is assumed constant across post-strata within an age/race group. Here the parameter is the ratio of the census or A.C.E. coverage rate for males to the census or A.C.E. coverage rate for females. As in the other models, the Prithwis Model uses the DA sex ratios and control totals for each age/race group for calculation of parameter estimates. In this model, the correlation-bias adjusted DSE for each post-stratum in an age/race group is calculated so that 1) the parameter estimate remains constant across the post-strata in the group, and 2) the sum of these adjusted DSEs within an age/race group will be equal to the control total for the group. The Prithwis Model also uses only Black and non-Black race categories.

Tables

Results are presented in the following two tables:

Table 1: Percent Net Undercount Rates for Major Groups

This table provides the net undercount rate and standard error for 18 major demographic groups for the five models discussed above.

Note that:

The estimates for children and for females are unchanged in all models

The Two-Group Model, the Modified Two-Group Model, the Fixed Relative Risk Model, and the Prithwis Model often provide the same estimates because the totals for males are being forced to the same quantity. The only differences are between the non-Black race groups and between the two tenure groups.

Table 2: Net Undercount Estimates for Major Groups

This table shows the estimated undercounts or overcounts compared to the census data.

Table 1: Percent Net Undercount Rates for Major Groups

Appendix A

Characteristic	If No Corr. Bias Adjustment Estimate (%) (S.E.)	**A.C.E. Revision II (Two-Group)** Estimate (%) (S.E.)	If Fixed Relative Risk Model Estimate (%) (S.E.)	If Prithwis Model Estimate (%) (S.E.)	If Modified Two-Group Model Estimate (%) (S.E.)
Race/Origin Domain					
Non-Hispanic White	-1.53 (0.20)	-1.13 (0.20)	-1.17 (0.20)	-1.10 (0.20)	-1.39 (0.20)
Non-Hispanic Black	-0.53 (0.41)	1.84 (0.43)	1.84 (0.43)	1.84 (0.43)	1.84 (0.43)
Hispanic	0.42 (0.44)	0.71 (0.44)	0.89 (0.44)	0.58 (0.43)	3.17 (0.49)
Non-Hispanic Asian	-1.12 (0.68)	-0.75 (0.68)	-0.64 (0.70)	-0.72 (0.68)	-1.01 (0.68)
Hawaiian or Pacific Islander	1.81 (2.73)	2.12 (2.73)	2.47 (2.90)	0.53 (2.26)	1.90 (2.73)
AIAN on Reservation	-1.16 (1.53)	-0.88 (1.53)	-0.63 (1.57)	-0.97 (1.52)	-1.08 (1.53)
AIAN off Reservation	0.30 (1.35)	0.62 (1.35)	0.71 (1.38)	0.64 (1.37)	0.39 (1.35)
Tenure					
Owner	-1.85 (0.20)	-1.25 (0.20)	-1.40 (0.20)	-1.21 (0.20)	-1.26 (0.20)
Non-Owner	0.45 (0.36)	1.14 (0.36)	1.46 (0.37)	1.06 (0.36)	1.56 (0.37)
Age/Sex					
0 - 9	-0.46 (0.33)	-0.46 (0.33)	-0.46 (0.33)	-0.46 (0.33)	-0.46 (0.33)
10 - 17	-1.32 (0.41)	-1.32 (0.41)	-1.32 (0.41)	-1.32 (0.41)	-1.32 (0.41)
18 - 29 Male	0.17 (0.70)	1.12 (0.63)	1.12 (0.63)	1.12 (0.63)	2.69[1] (0.64)
18 - 29 Female	-1.39 (0.52)	-1.39 (0.52)	-1.39 (0.52)	-1.39 (0.52)	-1.39 (0.52)
30 - 49 Male	-0.43 (0.33)	2.01 (0.25)	2.01 (0.25)	2.01 (0.25)	2.01 (0.25)
30 - 49 Female	-0.60 (0.25)	-0.60 (0.25)	-0.60 (0.25)	-0.60 (0.25)	-0.60 (0.25)
50+ Male	-2.24 (0.30)	-0.80 (0.27)	-0.80 (0.27)	-0.80 (0.27)	-0.80 (0.27)
50+ Female	-2.53 (0.27)	-2.53 (0.27)	-2.53 (0.27)	-2.53 (0.27)	-2.53 (0.27)

All net undercounts are for the household population. A negative net undercount denotes a net overcount.

Table 2: Net Undercount Estimates for Major Groups

Appendix A

Characteristic	Census	If No Corr. Bias Adjustment Estimate (S.E.)	A.C.E. Revision II (2-Group) Estimate (S.E.)	If Fixed Relative Risk Model Estimate (S.E.)	If Prithwis Model Estimate (S.E.)	If Modified Two-Group Model Estimate (S.E.)
Total	273,587	-3,034 (543)	-1,332 (542)	-1,332 542	-1,332 (542)	-979 (549)
Race/Origin Domain						
Non-Hispanic White	192,924	-2,902 (380)	-2,151 (382)	-2,228 382	-2,098 (381)	-2,649 (383)
Non-Hispanic Black	33,470	-177 (136)	628 (146)	628 146	628 (146)	628 (146)
Hispanic	34,538	146 (153)	248 (152)	310 155	201 (150)	1,130 (174)
Non-Hispanic Asian	9,960	-110 (67)	-74 (67)	-64 70	-71 (68)	-100 (67)
Hawaiian or Pacific Islander	590	11 (16)	13 (16)	15 18	3 (13)	11 (16)
AIAN on Reservation	540	-6 (8)	-5 (8)	-3 8	-5 (8)	-6 (8)
AIAN off Reservation	1,565	5 (21)	10 (21)	11 (22)	10 (22)	6 (21)
Tenure						
Owner	187,925	-3,420 (370)	-2,320 (372)	-2,601 (375)	-2,250 (371)	-2,334 (373)
Non-Owner	85,662	386 (310)	988 (310)	1269 (319)	918 (309)	1,355 (320)
Age/Sex						
0 - 9	39,642	-180 (130)	-180 (130)	-180 (130)	-180 (130)	-180 (130)
10 - 17	32,307	-422 (129)	-422 (129)	-422 (129)	-422 (129)	-422 (129)
18 - 29 Male	21,594	37 (150)	245 (138)	245 (138)	245 (138)	598[1] (142)
18 - 29 Female	21,576	-295 (111)	-295 (111)	-295 (111)	-295 (111)	-295 (111)
30 - 49 Male	41,297	-176 (137)	848 (104)	848 (104)	848 (104)	848 (104)
30 - 49 Female	42,783	-257 (105)	-257 (105)	-257 (105)	-257 (105)	-257 (105)
50+ Male	33,798	-740 (100)	-270 (90)	-270 (90)	-270 (90)	-270 (90)
50+ Female	40,590	-1,001 (107)	-1,001 (107)	-1,001 (107)	-1,001 (107)	-1,001 (107)

All Estimates are in thousands. The Census count is for the household population. A negative net undercount denotes a net overcount.

[1] For the Modified Two-Group Model, Hispanics males in the 18 - 29 age group are assigned the same correlation bias adjustment factor calculated for Blacks in that group. In the other three models Hispanic males in the 18 - 29 age group are not adjusted for correlation bias. The result is that for the 18 - 29 age category, a large number of Hispanic males are added, but without the compensating reduction in the number of non - Black non - Hispanic males that occurs in the other age categories.

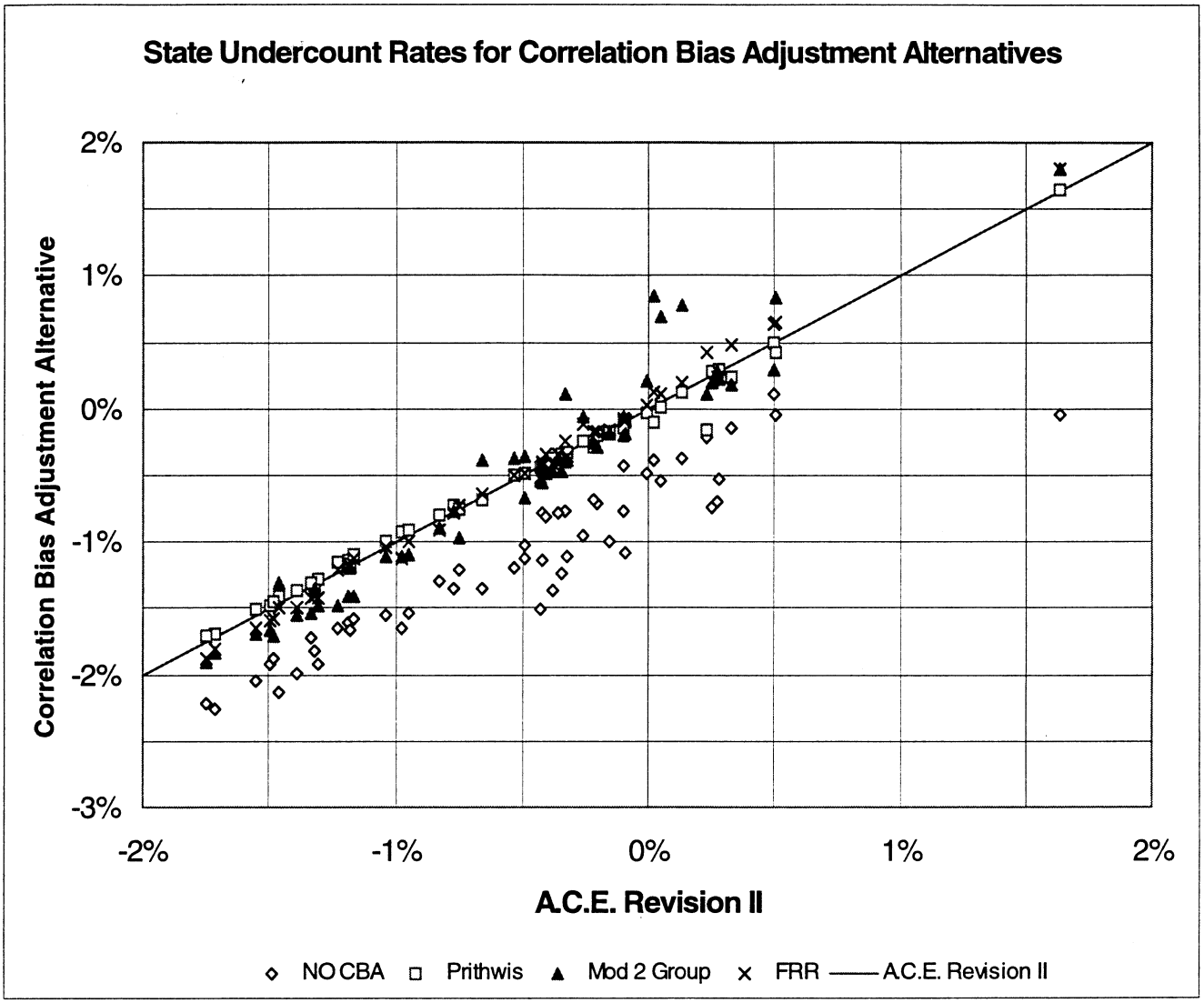
Comparison of State Level Undercount Rates Between the Two-Group Model and Selected Other Models

The following graph compares the undercount rate at the state level for several alternative approaches to correlation bias adjustment to the A.C.E. Revision II estimates, shown by the straight line. The A.C.E. Revision II approach, known as the Two-Group Model, has added males to five race/origin/age categories (Blacks 18-29, 30-49, and 50+ and non-Blacks 30-49 and 50+) proportionately to the unadjusted dual system estimates for males so as to be consistent with Demographic Analysis.

The alternative without correlation bias adjustment is always lower than the A.C.E. Revision II estimate because no males are added.

The Modified Two-Group Model, in which Hispanics are assigned the same adjustment as that for Blacks, is higher for states with large numbers of Hispanics and lower for states with few Hispanics.

Results for two additional correlation bias adjustment models are presented. One model was developed by Prithwis Das Gupta, in which additional people are allocated to post-strata based on the number of persons counted by either the census or the independent sample. Another was the Fixed Relative Risk Model, which allocates the additional people to post-strata based on the difference between the unadjusted dual system estimates and the number of correct enumerations. The results from these models are generally very close to those for the Two-Group Model used for the A.C.E. Revision II estimates.


State Undercount Rates for Correlation Bias Adjustment Alternatives
Correlation Bias Adjustment Alternative
2%
1%
0%
-1%
-2%
-3%
-2%
-1%
0%
1%
2%
A.C.E. Revision II
NO CBA
Prithwis
Mod 2 Group
FRR
A.C.E. Revision II