Sampling to Adjust the 1990 Census for Undercount

Prepared for 5 May 1998 hearing of the

United States of America
House of Representatives
Subcommittee on the Census

Philip B. Stark
Department of Statistics
University of California
Berkeley, CA 94720-3860

I thank Chairman Miller and the other members of the Subcommittee for inviting me to speak about the 1990 census adjustment.

We know from experience that, overall, the census misses some people.[1] The undercount is different in different places.[2] That leads to errors in state population shares. For many purposes, including distributing Federal funds and congressional representation, state shares matter more than the total U.S. population.[3] I will focus on the accuracy of state shares.

It would be wonderful to know how many people the census missed, and where. Then we could add them where they belong. That would adjust for the undercount, and improve state shares. But we do not know: the missing people were not counted.

Sampling is selecting part of a population to represent the whole. The Census Bureau used sampling to estimate the 1990 Decennial Census undercount, so they could adjust for it. The official 1990 census numbers were not adjusted. For the 2000 Decennial Census, there is a proposal to use sampling to adjust for undercount, and to use sampling to follow up some people who do not mail back their census forms. I will talk only about using sampling to adjust for undercount.

The 1990 and 2000 adjustments have different names and different details,[4] but they are based on the same statistical methods, so much of what I say about the 1990 adjustment applies to 2000 as well.

Would the adjustments have improved the 1990 census? Probably not, because of *statistical bias*.

Adjustment has two kinds of error: ***sampling error***, which comes from the luck of the draw --- the blocks that happen to be in the sample --- and ***systematic error*** or ***bias***, which comes from bad data, processing errors, and wrong assumptions, among other things.

Bias is a technical term: it does not mean someone is intentionally skewing the results. Sampling errors tend to average out. Bias does not.

Making estimates from a sample is like shooting a rifle. Each shot hits the target in a different place. Sampling error is the scatter in the shots. Bias is a tendency for all the shots to be off in the same direction, for example, to the left. You fix bias in a rifle by sighting it in. That is straightforward, because you can see where the shots land.

Fixing statistical bias in a census adjustment is hard. You only get one shot (because you only take one sample), and you cannot see where the shot lands (because you do not know the true undercount). The 1990 adjustment process was extremely complex,[5] so it is very hard to track down all its biases.[6]

For example, months after calculating the adjustment, the Census Bureau found that a coding error had inflated the undercount estimate by 1,000,000 people[7] --- about 20% of the adjustment. That's bias. Studies show that 40% to more than 80% of the 1990 adjustment is bias.[8] Adjustment could easily make the census worse instead of better.

New York, Pennsylvania, and Illinois lose shares in the adjustment.[9] Texas and Arizona gain shares. Arguably, it is easier to count people in Dallas and Phoenix, for example, than in the Bronx, Philadelphia, and Chicago, where the inner cities are denser.[10] Taking shares from New York, Pennsylvania and Illinois might be right --- or it might be bias from bad assumptions.

Some claim that adjustment would have made the 1990 census more accurate.[11] Their technical arguments depend on statistical models.[12] The models are false,[13] and have bizarre consequences. For example, the model for "correlation bias" says that the 1990 census missed nearly 900,000 white males, of whom only 13 --- less than 0.002% --- were between 20 and 30 years old. It also says that the 1990 census missed over three quarters of a million black males, but counted almost 30,000 too many black males under age 10.[14]

Using that incredible model, the Census Bureau estimated that about 38% of the adjustment is statistical bias. Without the model, the figure is 57%, almost 20% higher.[15] With better assumptions, the estimated bias is even higher. The study I trust most[16] puts the bias over 80%: adjustment puts in far more error than it takes out.

There is another way to estimate the total population, called Demographic Analysis.[17] The 1990 adjustment adds more people than Demographic Analysis says were missed, including about a million extra women.[18] Because of bias, the adjustment probably puts the people in the wrong place, making state shares worse.[19]

In summary, adjusting the census using sampling did not work in 1990,[20] because of statistical bias. Taking a bigger sample, as proposed for the 2000 census, could make bias even worse.

Technical Notes

[1] Edmonston, B. and Schultze, C. eds., 1995. Modernizing the U.S. Census, National Academy Press, Washington, D.C.

[2] ***Ibid***.

[3] ***Ibid***.

[4] The 1990 procedure is called the Dual-System Estimator (DSE), which uses data from the Post-Enumeration Survey (PES). The 2000 procedure is called Integrated Coverage Measurement (ICM).

Both PES/DSE and ICM take a random sample of blocks after the census is taken, and tabulate the people found in the households in those blocks who were missed by the census (omissions), as well as the people in the census who should not have been counted in those blocks (erroneous enumerations). Results are pooled for the blocks in the sample to get the fractions missed and erroneously enumerated, for various groups of people, called "post-strata." For example, black male renters age 30-44 living in the central city of a major metropolitan area in New England comprised one 1990 PES post-stratum. There were 1,392 PES post-strata in all.

The basic idea in the adjustment is that the fraction of people in a post-stratum who were in the sample blocks, but not in the census, is an estimate of the fraction of all the people in the post-stratum that the census missed. The fraction in the census in a post-stratum in the sample blocks, but not in the PES, is an estimate of the fraction of people in the post-stratum the census enumerated erroneously. The difference estimates the undercount rate for the post-stratum. Dividing the census count by (100% - undercount rate) adjusts for the undercount.

This is just a sketch: the details of determining whether or not there is a match, treating missing data, and combining numbers from different blocks to estimate fractions in post-strata are extremely complex; see **Hogan, H., 1993. The 1990 Post-Enumeration Survey: Operations and Results, *J. Amer. Statist. Assoc., 88,* 1047-1060.**

[5] **Hogan, H.,** ***loc. cit.***

[6] There is a great deal of information in **Committee on Adjustment of Postcensal Estimates, 1992.** ***Asessment of Accuracy of Adjusted Versus Unadjusted 1990 Census Base for Use in Intercensal Estimates,*** **Bureau of the Census (C.A.P.E. Report).** Here are some excerpts:

> "...additional research detected some errors and made some refinements to the levels of undercount originally reported in the spring of 1991." **C.A.P.E. Report**, p2.

The table on p3 of the C.A.P.E. Report show that uncertainty estimates (for sampling error alone) were increased by as much as 300%.

> "As a result of an error in computer processing, the estimated national undercount rate of 2.1% was overstated by 0.4%. After correcting the computer error, the national level of undercount was estimated to be about 1.7%. After making other refinements and corrections, the national undercount is now estimated to be about 1.6% [the figure is 1.58% in attachment 3, Table 2] ... The level of total bias, excluding correlation bias, on the revised estimate of undercount is negative 0.73 (-0.73%)." **C.A.P.E. Report**, p15.

Thus (2.1 - 1.58 + 0.73)/2.1 = 60% of the original estimate of 2.1% is bias. The report continues, evaluating the "revised" estimates, which correct the coding error and use different post-strata:

> "Therefore, about 45% (0.73/1.58) of the revised undercount is actually measured bias and not measured undercount. In 7 of the 10 evaluation strata, 50% or more of the estimated undercount is bias." **C.A.P.E. Report**, p15.

[7] **C.A.P.E. Report**, p15.

[8] According to the Director of the Bureau of the Census,

> "A significant amount of bias remains. The research estimates that, at the national level, removing all biases from the PES estimates would lower the estimated undercount from 1.6 to 1.3 percent. When the effect of correlation bias is not taken into account ... the estimated undercount would fall to 0.9 percent."

Bureau of the Census, 1993. ***Decision of the Director of the Bureau of the Census on Whether to Use Information From the 1990 Post-Enumeration Survey (PES) To Adjust the Base for the Intercensal Population Estimates Produced by the Bureau of the Census ACTION: Notice of final decision*****.** **Federal Register 58 FR 69.** (cited as 58 FR 69 henceforth)

That yields bias estimates of 38%-57%, depending on the treatment of "correlation bias." Correlation bias has to do with the fact that a key assumption in the capture-recapture model is false: everyone in a post-stratum in the sample blocks does not have the same chance of being found by the PES, and the same chance of being found by the census. For example, if there are people unreachable by any survey, the PES cannot detect that they are are missing from the census. That would tend to make the undercount estimate smaller than the true undercount. The correlation bias estimate uses a model to disaggregate Demographic Analysis figures to local levels. Evidence cited below (note 14) shows how unreasonable the model is.

The figure of over 80% comes from **Breiman, L., 1994. The 1991 Census Adjustment: Undercount or Bad Data?** ***Statistical Science, 9*****, 458-537.** Breiman combines information from various Bureau of the Census evaluation studies. Sources of bias include fabrications by interviewers, matching errors, census day address errors, bias in the ratio estimator, people discovered to be out-of-scope in reinterview, late census data, and the computer coding error. The following paragraphs are drawn from Breiman's work.

Small errors in the match rate can produce extremely large errors in the undercount estimates. For example, in one block cluster, an unmatched family of 5 people added 45,000 to the undercount estimate. In Census Bureau studies of matching errors, match and rematch classifications disagree by 1.8%. A June, 1991, Census Bureau memorandum states: "...approximately 75 percent of the non-matching people could have been converted to a match if the search area had been expanded." This is a huge source of bias.

The match status of about 2% of the cases could not be resolved from the records or by interview. Depending on how these cases are treated, the PES estimates range from an ***overcount*** of 1,000,000 people to an undercount of 9,000,000 people. See also **Wachter, K.W., 1991.** ***Recommendations on 1990 Census Adjustment, report to the***

***Secretary of Commerce as a Member of Special Advisory Panel*, U.S. Department of Commerce,** for "half-high" and "half-low" estimates.

The "probabilities" that unresolved cases were matches were imputed using a statistical model [**Belin, T.R., *et al*., 1993. Hierarchical Logistic Regression Models for Imputation of Unresolved Enumeration Status in Undercount Estimation, *J. Amer. Statist. Assoc., 88,* 1149-1159**] with obviously false assumptions [**Wachter, K.W., 1993. Comment: Ignoring Nonignorable Effects, *J. Amer. Statist. Assoc., 88,* 1161-1163**]. At least one explanatory variable in the model is missing for 28% of the unresolved PES cases, and 38% of the unresolved census-sample cases; those missing variables were also imputed.

[9] See Figure 1, p111, in **Wachter, K.W., 1993. The Census Adjustment Trial: An Exchange, *Jurimetrics, 34,* 107-115.** California would have gained most by adjustment; Texas, second most. Pennsylvania would have lost most; Ohio, second most.

[10] ***Ibid***.

[11] For example, National Academy of Sciences reports recommend using sampling-based adjustments for the 2000 Decennial Census; see **Edmonston, B., and Schultze, C., ed., 1995. *Modernizing the U.S. Census*, National Academy Press, Washington, D.C. 460pp.,** and **White, A.A., and Rust, K.F., ed., 1997. *Preparing for the 2000 Census*, National Academy Press, Washington, D.C. 98pp.**

I reviewed those reports. Their evidence is weak. The issue is whether the PES improves or degrades the accuracy of the census. That is very hard to determine, because the PES is subject to large biases that cannot be measured directly. Arguments in favor of the PES depend on the assumption that the errors in the PES generally go in the same direction as the true undercount, and seldom go too far.

[12] See, for example, **Mulry, M.H., and Spencer, B.D., 1993. Accuracy of the 1990 Census and Undercount Adjustments, *J. Amer. Statist. Assoc., 88,* 1080-1091.**

[13] In addition to models relating various parameters, the assumptions include:

Independence: This assumption has two parts. First, for each individual, in the sample blocks, being caught in the census is independent of being caught by the PES. Second, the probability of being caught in the census is the same for every individual in a given post-stratum within the sample blocks, as is the probability of being caught in the PES.

Synthetic Assumption (Homogeneity): In each block that was not sampled, the nonresponse rate is a weighted average of the nonresponse rates of the post-strata that intersect the block. The weights are the proportions of people in the block in the post-strata.

Violation of the independence assumption leads to "correlation bias:" see note 14. There are a number of studies of the synthetic assumption using proxy variables, for example, **Hengartner, N., and Speed, T.P., 1993. Assessing Between-Block Heterogeneity Within Post-Strata of the 1990 Post-Enumeration Survey, *J. Amer. Statist. Assoc., 88,* 1119-1129,** and **Freedman, D. and Wachter, K., 1994. Heterogeneity and Census Adjustment for the Intercensal Base, *Statistical Science, 9,* 476-485.**

Those studies find that heterogeneity within post-strata is significant. According to the Director of the Bureau of the Census,

> "...it is possible that errors due to heterogeneity in fact could be larger than all other sources of error in the adjustment." **58 FR 69**

The C.A.P.E. also studied heterogeneity:

> "The Panel cautioned that artificial population analysis ... was inconclusive about whether the homogeneity assumption held." **C.A.P.E. Report**, p30.

But their analysis had flaws:

> "A first analysis showed similar homogeneity for the 1,392 design as well as the 357 design as well as for a design with only 2 strata." **C.A.P.E. Report**, p26.

They also state:

> "The level of bias in the PES was close to the point where artificial population analysis shows that homogeneity assumption fails to hold." **C.A.P.E. Report**, p26.

[14] The model for disaggregating "correlation bias" from national DA estimates down to local levels is in **Bell, W.R., 1993. Using Information from Demographic Analysis in Post-Enumeration Survey Estimation, *J. Amer. Statist. Assoc., 88,* 1106-1118.** The consequences of that model for the cited demographic groups is on p533 of **Freedman, D., and Wachter, K., 1994. Rejoinder, *Statistical Science, 9,* 527-537.**

The C.A.P.E. also had reservations about the model of correlation bias:

> "The fourth cell in the DSE is an estimate of the number of people missed in both the PES and the census...Both the Committee and the Panel of Experts were very concerned about the negative values in the fourth cell...correlation bias should be a component of total error. However, there was concern about our method of estimating it and very serious concern about the method of allocating it." **C.A.P.E. Report**, pp22-23.
>
> "The Census Bureau ... knew of no adequate methodology to remove the bias by state, city, etc." **C.A.P.E. Report**, p30.

[15] See note 8.

[16] **Breiman, L., *loc. cit.***

[17] **Robinson, J.G., Ahmed, B., Das Gupta, P., and Woodrow, K.A., 1993. Estimation of Population Coverage in the 1990 United States Census based on Demographic Analysis, *J. Amer. Statist. Assoc., 88,* 1061-1079.**

[18] "... there was concern that the PES estimated a higher population than DA and estimated about a million more women than DA." **C.A.P.E. Report**, p27.

[19] According to the Director of the Bureau of the Census,

> "...no survey -- either the high quality, well controlled and interviewed PES of 170,000 households or a larger one -- can be used to make post-census fine tuning of an average undercount as small as 1.6 percent in all types of places, counties, and states at a level of accuracy beyond that by which surveys are usually judged....there is little or no evidence adjustment would improve the quality of substate estimates..." **58 FR 69**.

[20] "...there is no intention to adjust the 1990 census because research shows insufficient technical justification." **C.A.P.E. Report**, p33.