Making Estimates from the American Community Survey

Charles H. Alexander, Scot Dahl, Lynn Weidman

U.S. Census Bureau
Washington, DC 20233

Presented to the Annual Meeting of the American Statistical Association (ASA), Anaheim, California, August 1997.

This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. This report is released to inform interested parties of research and to encourage discussion.

INTRODUCTION AND BACKGROUND

This paper discusses the estimation methods used for the 1996 American Community Survey (ACS). In particular, the weighting strategy and weighting factors are described in detail.

The American Community Survey is a new Census Bureau initiative to update regularly the social and economic profile of the nation's communities that the U.S. census has traditionally produced once a decade. By "communities" we mean both geographic and demographic domains. The survey eventually will produce annual estimates for domains of all sizes, but these annual "profiles" will be reliable only for communities of greater than 65,000 population.[1] For smaller domains, we recommend using "profiles" based on averaging (or otherwise cumulating) 2-5 years of data, depending on the size of the domain.

The ACS is being introduced as follows:

1996-1998 Demonstration Period (four counties in 1996, nine in 1997, eleven in 1998)
1999-2001 About 40 comparison sites with 5 percent annual samples
2000-2002 National comparison sample with overall rate of 0.7 percent annually
2003-on Full introduction (three million addresses per year, including all counties).

In the demonstration period, we are evaluating and refining our survey methods, and also working closely with experts from the test sites to understand the potential uses and quality of the ACS estimates. In the 40 comparison sites, we will work with local experts to understand differences between the 1999-2001 ACS average estimates and the 2000 census long form results. The 2000-2002 national sample will be used to generate model-based estimates for small domains nationwide--down to the size of census tracts--representing "what the 2000 long-form estimate for the domain would have been had ACS data collection methods been used." These model-based estimates will be used to interpret changes between 2000 and future ACS estimates. The models will be specified based on lessons learned in the 40 sites.

This paper describes the estimation methods used in the four 1996 ACS demonstration sites: Multnomah County/Portland OR, Rockland County NY, Brevard County FL, and Fulton County PA. These areas had an especially high sampling rate in the first year, 15 percent in most areas and 30 percent in small governmental units,[2] so that detailed estimates could be studied without waiting for multiple years' data. Some initial results are presented, along with future research issues. Our ultimate interest is in implications for the best methods and likely quality of the 2003 and later annual estimates which will be based on an approximate 3 percent annual rate in most areas.

Issues of cumulating multiyear data have been discussed at length in Alexander (1996).

ACS SURVEY DESIGN

Sample Design: The ACS design is based on the rolling sample designs described in Kish (1990). Sample addresses are interviewed at most once in any five-year period. Each year's sample addresses are spread evenly across the 12 months of the year and across the entire nation. The sample is in general not clustered, although there may eventually be some exceptions in areas with unusually high travel costs and in group quarters. Each year's sample will consist of addresses not selected in the previous four years. We are still considering whether the sample should "roll" through the entire list, never reselecting an address until all have been included, or whether some reselection may be allowed after five years have passed.

Sampling Frame: The ACS sample is selected from the Census Bureau's Master Address File (MAF). The MAF is being created for the 2000 census, but an initial version is already available in many areas. It starts with the 1990 census Address Control File, linked to the TIGER geographical database. This is updated using postal Delivery Sequence Files (DSF) in areas where the DSF addresses can be geocoded based on a "city-style" house number and street name. In other areas, the MAF must be created by physically listing each block.

Special advance MAFs were created for the 1996 ACS test. The Multnomah, Rockland, and Brevard sites used MAFs based on the DSF; the Fulton site had a separate listing/drop-off method being tested for the ACS since there were no city-style addresses in most of the county. Current plans call for updating the MAF each six months after the 2000 census. However, no updating system was in place for the 1996 ACS sample. Thus, the 1996 sample does not include new addresses added during the year.

The 1996 ACS sample only includes regular housing units, not group quarters. One type of group quarters, in particular college dormitories, is included in one of the 1997 sites. Other types of group quarters will be added in 1998. Both institutional and noninstitutional group quarters will be included.

Data Collection: For each monthly sample panel, the ACS starts by mailing a questionnaire to each address about 10 days before the start of the "mailout month." There is a "pre-notice" letter, an initial mail questionnaire, and a reminder card, one week apart. After about 3 weeks, a replacement questionnaire is mailed if no response has been received.

At the beginning of the following month, nonresponding addresses are assigned to telephone nonresponse follow-up. For addresses where the telephone number can be obtained from commercial directories, a telephone interview is attempted. Mail returns continue to come in during this second month; about 19 percent of the telephone nonresponse follow-up group are removed from follow-up because of a late mail return.

The third month, any addresses still not interviewed are eligible for personal-visit follow-up. This includes addresses where no telephone number could be located and addresses where there was a number, but no interview could be obtained. One-third of these addresses are selected for follow-up by personal visit. Note that this includes most of the vacant addresses.

As an example, the March mailout panel had telephone follow-up in April, and personal-visit follow-up in May. In general, a new area introduced in a particular year starts with mailouts in November and December of the previous year, so that the normal pattern of follow-up work in January and February is in place by the time the year starts.

In some cases, more than one record is received from a household: either two mail returns, or some combination of mail return, telephone, or personal interview. There is an algorithm to select one of these records as the real record from the address, taking into account their interview dates and how completely the various questionnaires are filled out.

The telephone and personal-visit nonresponse follow-up interviews used Computer-Assisted telephone or personal interviewing, known as "CATI" or "CAPI" respectively. The computerized questionnaire has built-in automated skip patterns with response required to each question. A sample of the CATI interviews are monitored by a supervisor at the centralized facility as a check on quality. In addition, for the CAPI interviews starting late in the year, there was a telephone "recheck" of a portion of the work as a special study to probe for possible incomplete coverage of persons on the roster and also to guard against the possibility of falsification.

The 1996 ACS test was complicated by the shutdown of the Federal government during December 1995 and January 1996. There was no January mailout, and the telephone and personal visit cases for the November and December mailouts were completed on a delayed schedule. The survey data processing was done as though there was one combined December/January panel. This exceptional situation is ignored in much of the following discussion, since it would complicate the description of each step and is not enlightening about the general ACS process.

Failed Edit Follow-up: The mail returns undergo a clerical edit which includes determining

1) if the form is missing enough responses to require a callback;

2) if an initial write-in entry giving the number of persons at the address was inconsistent with the number actually included in the questionnaire;

3) if more than 5 people were listed as living at the address, since the form only collects characteristics for 5 people.

The first condition is referred to as the "content edit" and the last two as the "coverage edit." If the form fails either edit, then there is a telephone callback that attempts to fill in all missing data and straighten out any coverage problems. Any remaining missing information is handled using the imputation procedures described in a later section. The exception is that if the content edit finds that the form is completely blank, the case continues to nonresponse follow-up, as if no form had been returned. About half of the returns failed one or more edits; of these 96 percent gave a telephone number, and 92 percent of these had some further resolution from telephone followup.

Residence Rule and Reference Period: The residents of any sample address, and their characteristics, will be determined as of the time of data collection. This refers either to when the mail form is filled out or when the nonresponse follow-up interview takes place. The 1996 ACS uses a "two-month" rule to determine who is a "current resident" of an address. Anyone staying at the address more than two months is included as a current resident. People staying two months or less would also be included, unless they usually live somewhere else.

Response Rates: The weighted response rates are shown in Table 1, cumulated over the four sites. Note that some addresses are determined to be out-of-scope because the unit no longer exists or is a duplicate. The table includes vacant units, most of which are completed by CAPI.

The critical percentage in terms of nonresponse bias is the estimated 1.8 percent of the eligible housing units in the 4 sites who would not respond at all. To the extent that these units are different than the rest of the population, there will be a potential for bias.

In Table 1, we have adopted the standard survey practice of reporting the weighted response rate as an indicator of nonresponse bias. Some have suggested that the mail/telephone nonrespondents who are randomly selected to be excluded from follow-up should be included as nonrespondents in calculating an unweighted "response rate". However, that is not how the "response rate" is traditionally calculated for a "double sampling" design such as the ACS has used.[3] The population, as of a particular point in time, can be classified into three categories: i)addresses where a response would be obtained by mail or telephone, if the address is selected; ii) addresses where a response would not be obtained by mail or telephone, but would be obtained if the unit is followed up; iii) addresses where no response would be obtained even after mail, telephone, and personal followup. Our "double sampling" design ends up giving us interviews for a 15 percent representative annual sample of the first group, interviews for a 5 percent representative sample of the second group, and *no interviews at all* from the third group. The point of the nonresponse rate is to estimate what fraction of the population falls in the third category; that is what the weighted non-response rate does. The fact that the second group has a lower sampling rate does not cause any nonresponse bias, since the sampling rate is taken into account in the estimation process. Note that the results in Table 1 apply only to the four sites and cannot be extrapolated to any wider area. The sites were selected because they had close to average response rates in the 1990 census.

GENERAL ESTIMATION STRATEGY

Major Parts of the Estimation Process: The ACS estimation process consists of:

- editing the responses
- imputing missing responses
- confidentiality edit
- weighting sample households and persons
- verification

A related activity is the creation of independent population controls for counties using demographic models. The final step in the ACS weighting for persons is to bring the county-level estimates into agreement with these "control totals." Correcting the controls is not listed as part of the ACS estimation process, since these intercensal demographics estimates predate the ACS and are derived independently. In time, we expect information from the ACS and the MAF to be used in the demographic models used to produce controls for counties or smaller areas. Then the ACS and the demographic controls will be viewed as a "program of integrated estimates."

The different parts of ACS estimation are discussed in the following sections. The theoretical goal of the person weighting is to reproduce the estimate that would have been obtained if data for every person in the population had been collected for the ACS. Weighting does not correct for imperfections in the collection such as erroneous answers or failure to give all the answers. These are handled in the editing and imputation steps.

The weighting and edit/imputation steps are intermingled. The basic data items used in weighting, such as race and household relationships, are cleaned up by editing and imputation before the weighting, so the weighting is not complicated by problems with these variables. Some later stages of the edit and imputation are done after the weighting, so that weighted frequency distributions for observed values can be used to guide the process.

The "confidentiality edit" is a special step to disguise the identity of specific sample households or people with minimal effect on tabulated data or analysis of the relationships of variables for the public use microdata.

The rest of the paper starts by discussing weighting, which is the aspect covered in most detail. The details of the demographic methods for calculating population controls are beyond the scope of this paper, but are briefly summarized in the weighting section.

Use of Survey Weights: As with most household surveys, weights are applied to each sample unit to bring the characteristics of the weighted sample more into agreement with the known characteristics of the population, at least on average. These weights can be justified on any of several theoretical grounds (Little and Rubin, 1987, Chapter 4). The weights compensate for differences in selection probability for different domains, for differences between the original sample and the interviewed sample, and for differences between the sample and independent estimates of population characteristics. The weights do not compensate for the fact that some information is missing for sample units; see the later section on edit and imputation.

A weight will be assigned to each sample person and housing unit. The person weights will be used for calculating estimates of the number of persons with certain characteristics. The housing unit weights are used for estimates of housing units, either occupied or vacant. For household or family estimates, the housing unit weight is used.

When the data are tabulated, a sample person with a weight of 8 is counted as though there are 8 people in the population exactly like that person. Higher weights are given to people or housing units in groups thought to be under-represented in the sample, for example, because they were in the group subsampled for nonresponse followup or because other people in the same geographic area refused to be interviewed for the survey. The assignment of these weights is described in the next section.

The ACS data will be weighted on an annual (calendar year) basis, rather than being separated into monthly or quarterly panels that are weighted separately. The units included in a particular year's estimate will be those units whose data are collected during that calendar year, rather than those mailed out during the year. When "month" is used to define cells for use in the weighting, it will in most cases be defined by the collection month. This is a consequence of the decision to use the time of interview as the reference period for the questionnaire.

EDIT AND IMPUTATION

Editing the ACS data to identify for obviously erroneous values and imputing reasonable values when data were missing involved a complex set of procedures. The specific procedures for different sets of variables, such as marital status, education, or income, were developed by demographers and economists familiar with each specific topic. The documentation of the procedures is over 1000 pages long, so only a very general discussion will be given here.

The variables measured by the ACS were edited/imputed in a specific order, so that a complete and consistent set of data for all earlier variables would be available when the work for each of the later variables was done. The order was determined by considering the importance of each variable for guiding the edit and imputation for later variables. The edit and imputation for the main variables used in defining weighting cells had to be completed before the weighting could be done. The general order was as follows: basic population variables (sex, race, etc.), all housing variables, the remaining population variables, and variables re-edited after the weighting.

"Imputation" is the insertion of a plausible value for missing values of variables. In some cases, the missing value can be determined deterministically from reported values; for example the missing variable "sex" for a person listed as "daughter" of the reference person would be allocated as "female." In cases when the likely value is not so obvious, the value to be imputed is taken from a "donor" person or housing unit with complete data. The donor record is selected to be similar on certain other matching variables. For example in selecting a donor for missing "educational attainment," a person of the same age, race, and Hispanic origin would be used. The matching variables need to be earlier in the order than the variable being imputed, since they must have non-missing values when the search for a donor takes place.

Note that the variable to be imputed can be missing either because the respondent did not give a value, or because the edit process determined that the reported value was unreasonable and it was deleted.

The edits included "range checks" and "consistency edits", to correct implausible values or combination of values for specific housing units or people. For example, the range check for year of birth only accepted values between 1880 and 1996.

Consistency edits range from simple rules to complex algorithms. The consistency edit for fertility ("how many babies has this person ever had") deletes responses from anyone identified as Male or under age 15. In setting a cutoff like this, a decision must be made based on the data about which categories have more "false positives" than "true positives." The consistency edit for housing value involves a joint examination of value, property taxes, and other variables. When the combination of variables is improbable for a particular area, several variables may be modified to give a plausible combination with values as close as possible to the original.

The justification for the edits (and imputation) is partly statistical -- that well-chosen edits can reduce the mean squared error of the estimates by eliminating some response errors. It is also partly for the convenience of users of public use microdata, as well as cross-tabulated summary data, whose analyses would be complicated by having to deal with impossible or implausible combinations of variables. In some cases, the goals can conflict and there is debate over what edits are appropriate. For example, the 1990 census long form edits did not allow situations where the survey estimated that a tract had fewer than 50 units (but more than zero) in structures with more than 50 units. The response was changed to say that these units were actually in structures of fewer than 50 units. This edit was not used for the 1996 ACS, on the grounds that the inconsistency can arise due to sampling error as well as mis-reporting. Applying the edit when sampling error is the cause can result in a downward bias in these estimates. A question for data users is whether the occasional impossible estimate is acceptable as the price for avoiding what is probably a slight statistical bias.

FACTORS USED IN THE WEIGHTING

1. Base weight (BW). This is the inverse of the probability that the address was selected for sample for this year. In areas with a 15 percent annual sample in 1996, BW = 1/.15 = 6.6667. In small governmental units with a 30 percent sample, BW = 3.3333.

In selecting the sample for the 1996 test, we deliberately avoided selecting addresses in sample for other household surveys conducted by the Census Bureau. We applied weights to adjust for the omission of these units from the frame, but these weights were so close to 1.00 that they will be ignored in the discussion hereafter.

2. CAPI Subsampling Factor (SSF). This is equal to 3 for units selected for CAPI follow-up, zero for those not selected, and equal to 1 for units interviewed by mail or telephone. This is applied to all interviews and noninterviews that have been selected for CAPI follow-up.

3. Variations in Monthly Sample Factor (VMS). Our goal is to have a fairly even distribution of different response modes (mail, CATI, CAPI) by month of interview. However, we anticipate that monthly variations in the mail return rates will cause some variations in the number of interviews completed each month. For areas where the population size or characteristics are seasonal, this variation will cause certain months to be over- represented in the annual average and others to be under-represented. To correct for this, the total weight of units whose data are collected by CATI or CAPI in a given month is adjusted to equal the total weight (using BW) of units selected for sample for that month, minus the total weight of mail responses received during that month.

To illustrate, Table 2 gives the information used in computing the VMS factor for June interviews in Rockland County.

The total weight of units completed in June is 4099 + 874 + 2449 = 7422, which is less than the total weight of units mailed out in June. The problem is that Rockland has a relatively low mail response rate for June, while April had the highest rate for the year resulting in relatively few CAPI cases carried forward into June.

The VMS factor, applied to the CATI and CAPI cases, is calculated as

$$\frac{7888 - 4099}{874 + 2449} = 1.1402$$

Mail responses have a VMS weight of 1. Applying this factor, the total weight of cases completed in June is 1 x 4099 + 1.1402 (874 + 2449) = 7888, which equals the weight of units mailed out in June. Even after applying the VMS, the weighted proportion of units completed by mail can vary from month to month. For example, there may be more vacant units in some months than others. This variation in the "mail respondent population" from month to month is something that actually occurs in the population and should not be eliminated by the weighting. The problem the VMS is intended to correct is that the total weight of June CATI and CAPI cases may differ from the June "non-mail-respondent population", because the June CATI and CAPI cases are derived from the April and May mail respondent populations. The factor will tend to be largest in months with a low mail return rate that follow months with a high mail return rate.

4. Noninterview Factor (NIF). The NIF is a standard "weighting cell estimate" (Little and Rubin, 1987, Section 4.4, or Kalton, 1981, Chapter 3.) in which the sample is divided into classes or "cells" based on variables that are observed for both respondents and nonrespondents. Within each cell, an additional factor NIF is applied to each interviewed unit, so that the total weight of the interviews after the NIF is applied equals the total weight for interviews and noninterviews in the cell. The interviews in a cell thus "carry" the weight of their cell's noninterviews in addition to their own weight. The NIF for the noninterviews is zero; they are now dropped from the estimation process.

The variables defining the cells should be selected so that either the probability of nonresponse or the characteristics of interest to the survey (and preferably both) are homogeneous within the cells. For the 1996 weighting, the cells were based on month of interview, census tract, and whether the address was in a single-unit or multi-unit structure as determined from the Master Address File.

However, because the ACS monthly sample by tract is quite small, we adjusted to the "margins" rather than use detailed "month-by-tract" cells. First, compute the adjustment factor

$$NIF_1 = \frac{\text{Total weight of all units (using BW x SSF x VMS)}}{\text{Total weight of interviews (using BW x SSF x VMS)}}$$

for each cell, where the cells are defined by tract and single/multi-unit. Then compute

$$NIF_2 = \frac{\text{Total weight of all units (BW x SSF x VMS)}}{\text{Total weight of interviews (BW x SSF x VMS x } NIF_1}$$

for cells defined by month and single/multi-unit. Then NIF = NIF_1 x NIF_2.

This is related to "raking ratio adjustment." Note that after applying the adjustment NIF_2, the agreement by tract produced by NIF_1 is no longer perfect. A complete application of "raking" would repeat the NIF_1 calculation with the NIF_2 factor included, then redo the NIF_2 calculation with the new NIF_1 factor, and iterate these calculations until there was perfect agreement for both month and tract. Because of the low nonresponse rates, we judged that this iterative process was not worthwhile, so only "one rake" was used.

Adjusting to the margins in this way corrects for overall differences in nonresponse rates by month and overall differences by tract, but it does not take into account differences in monthly noninterview rates by tract. There are too few cases in some tracts in some months to make that adjustment.

5. The Mode Bias Factor (MBF). This factor is an attempt to correct for bias resulting from not taking the mode of interview into account in calculating the NIF. The concern is that there are known to be systematic differences between households that return census mail forms and those that do not (Love, et al, 1995). Since almost all ACS noninterviews are among the group that did not return a completed mail form, but the NIF is applied to all interviewed addresses, there may be some "mode-related" nonresponse bias that NIF does not compensate for.

To compensate for this, we calculated an alternative noninterview factor, NIFM, in which the denominator of the factor includes only cases interviewed by CAPI. In principle, the NIFM could have been used as the noninterview adjustment factor (applied only to CAPI interviews) rather than NIF (applied to all interviews). We did not do this, because the NIFM factors tend to be substantially higher, which is likely to give too much weight to a few cases and thereby increase the survey's standard errors for small areas too much. The NIFM is used only in calculating the MBF; it is not one of the factors in the final weight.

Instead the NIFM is used to compute county-level adjustments to correct for the "mode-related" nonresponse bias. For each of the cells used in the MBF adjustment, two county-level estimates for the number of housing units in the cell are calculated, once using the weight BW x SSF x VMS x NIF and once using the weight BW x SSF x VMS x NIFM. The MBF is calculated for each category as the ratio:

$$MBF = \frac{\text{estimate for cell using NIFM}}{\text{estimate for cell using NIF}}$$

With the weight BW x SSF x VMS x NIF x MBF applied to all cases, the county-level estimates for the cells used in the MBF calculation are the same as if NIFM had been used instead and applied only to the CAPI cases.

The cells used in the MBF calculation are month by marital status (married/widowed, all other) by form of tenure (renter, owner, missing). The relatively small number of units with missing tenure (mostly temporarily occupied housing units) were combined into one cell with no breakdown by month or marital status. For Rockland, the missing tenure group had MBF = 1.0271. The other MBF factors were very close to 1.00, ranging from a high of 1.0202 to a low of .9905. Future research on ACS weighting will consider other cells for the MBF.

6. Truncation Factors. To keep weighting factors from getting too large, which has an adverse affect on the variance, it is customary after looking at the data to "collapse" weighting cells that have very large factors with other cells having similar characteristics until the factor for the combined cell is below a specified limit. However, this collapsing based on the observed factors complicates the computer programming for the weighting and makes a theoretical analysis of the method complicated. For the ACS, we used a simpler alternative of "truncation", simply cutting off the factor at a maximum value for any cell. The "lost" weight, i.e., the total amount by which the non-truncated weights exceed the limit, is to be "redistributed" over the other cells by increasing all other weights by a constant factor to bring the total weight back to what it would have been without the truncation.

The limit to be used when truncating weights was set at 6 times the basic weight. This reflects a combination of the maximum ACS subsampling rate of 1 in 3 with the traditional limit of 2 used for collapsing noninterview cells (Hanson, 1978).

We had planned to truncate, if necessary, after the NIF and after the MBF. However, using the limit of 6, the planned truncation never was actually required. The low nonresponse rates in the 1996 test sites never required a NIF or MBF large enough to violate the limit. Thus, the truncation factors had no effect on the 1996 ACS weighting.

7. Furlough Adjustment Factor (FAF). A shutdown of the Federal government in late December 1995 - January 1996 interrupted the ACS data collection, essentially eliminating the January mailouts and interviews. The January mailout and CATI cases were replaced by data from the November and December 1995 mailouts. This did not require any special adjustments in the weighting, but there may be biases in estimates because much of the "January 1996" interviews actually took place in late 1995.

The greater problem is that January was to have been the first month of CATI interviewing, so there were no CAPI interviews to represent the January population. To make up for this variation from the desired seasonal patterns, February CAPI cases were counted twice. This was implemented by creating duplicate records for use in the earlier stages of the process, as far as the NIF. After the NIF calculation, the duplicate records were dropped and the February original records are given the FAF.

The FAF will be 2.0 for all February 1996 CAPI cases and 1.0 for all others. This factor adjusts the weights of February 1996 CAPI records to account for the absence of CAPI interviews in January due to the Federal government shutdown.

8. First Housing Unit Post-stratification Factor (HPF_1). The HPF_1 is a ratio adjustment factor adjusting the number of weighted sample addresses in each tract (using all prior weights) to agree with the count of addresses from an updated Master Address File (MAF). The MAF used was the latest MAF as of February 1997. This included new addresses added since the ACS sample was selected for the 1996 interviews.

The denominator of the adjustment includes the interviewed units among the occupied and vacant housing units, using all their weights (BW x SSF x VMS x NIF x MBF x FAF) defined so far. Noninterviewed housing units are represented by the NIF weights for the interviewed units. However, the denominator also includes the addresses determined to be out-of-scope, with their weight BW x SSF x VMS. These addresses are not used in making estimates--no data are collected for them--but they need to be included in the denominator since out-of-scope addresses are counted in the numerator.

We expect the main effect of the HPF_1 adjustment to be compensating for new units added to the MAF after the sample was drawn, and changes to the geographic placement of those units on the sample MAF. There also are small differences between the sample and the frame in the distribution by census tract even for units on the original MAF. These differences are not expected to be very great because both the original sampling and the nonresponse subsampling used systematic sampling from a list sorted by tract, as did the random assignment of units to month. However, some differences can occur because the systematic sample can only select a whole number of units, which means that it cannot exactly reproduce the percentage of units in the tract, and because the VMS and MBF do not take tract into account. Also, noninterviews can affect tracts unevenly, although the noninterview adjustment NIF took tract into account at the first step.

The main effect of the HPF_1 is on a few small tracts, where greater percentage variation can be expected. In one tract in Rockland County, the factor is trying to correct for what appears to be a substantial geocoding discrepancy between the 1990 census address list and the Postal Delivery Sequence Files. We have targeted this for investigation and correction. Even when the ACS is running nationwide, we expect to have a quality assurance process that investigates and corrects the worst such discrepancies.

9. The Person Post-stratification Factor (PPSF). The PPSF assigns weights to people in interviewed units so that the total weight for the county by age, race, sex, and Hispanic origin agrees with independently derived intercensal demographic estimates. Race is coded as "white", "black" and "other". Hispanic origin was a "yes-no" variable. Age was coded in five year intervals from 0-4 to 80-84, and "85+", with some prior collapsing of cells in some counties.

The numerator of the adjustment is the county-level demographic estimate for the age-race-sex-Hispanic-origin cell and the denominator is the weighted estimate for the cell using all previous weights for the sample people in the cell.

Some of the differences between the weighted estimate (prior to the PPSF) and the demographic estimates are due to random variation in which units are selected for sample. However, part of the difference is known to be due to systematic undercoverage of housing units and persons within housing units. (Hainer, et al, 1991). The adjustment partly corrects for the bias due to missing these people and units, but biases will remain to the extent that there are differences in characteristics between the missed people and other people of the same age-race-sex-origin. The effect of the PPSF on the 1996 estimates will be discussed in more detail in a separate section.

It is the independently derived demographic estimates that actually are the population values reported by the ACS. The 1996 intercensal estimates start with the 1990 census counts--adjusted for census undercount. Components of population change between 1990 and 1996 are estimated from a variety of sources including birth and death records, migration within the U.S. as reflected in data for tax filers from the Internal Revenue Service as well as other administrative sources, and immigration/emigration data. Some critical assumptions are i) that the migration patterns of non-tax-filers are similar to the patterns of filers; ii) that the racial and ethnic distribution of migrants can be predicted from the distributions in the locations to which they are going and from which they came.

The PPSF is the final factor used for constructing the person weight, which is calculated as
person weight = BW x SSF x VMS x NIF x MBF x FAF x HPF_1 x PPSF.

The person weight is applied to all interviewed persons who are residents at a sample address.

The main goal of the person post-stratification factors is to correct for undercoverage of persons, especially differential undercoverage by race or Hispanic origin. Table 3 shows the effect of these factors.

10. The Principal Person Factor (PPF). The principal person factor is used (in different versions) for the Current Population Survey (CPS) and many other surveys as a final factor to adjust household weights for under-coverage when there are no independent household estimates to use as controls. (Bureau of Labor Statistics (1986), Alexander (1990)). We have used it as an intermediate weight to adjust for differences in coverage by race, following it by an additional factor (HPF_2) to control again to total MAF counts for the county.

The principal person factor (PPF) for an occupied housing unit is equal to the person post-stratification factor (PPSF) for the 'principal person" in the household. The principal person is the wife in a husband-wife household. In a household where the reference person does not have a spouse present, the reference person is the principal person, whether male or female. The assumption is that principal persons are missed by the survey only when their entire household is missed, not because the household respondent failed to list them on the questionnaire. Thus, their PPSFs give an indication of the undercoverage of households with principal persons of their age, race, and sex.

The PPF for a vacant housing unit is equal to 1.

A variety of alternatives to the principal person factor have been developed in recent years, including modifications to the principal person idea that distinguish between male reference persons and males of the same age-race-origin who are not reference persons. Other alternatives create household weights that are consistent with the person population controls, such as generalized least squares (Zieschang, 1991) or calibration

estimation, (Valliant and Jayasuriya). These methods have not been shown to be superior to the principal person method for reducing coverage bias. How well they work for this purpose depends on the actual patterns of under-coverage (Alexander, 1990). These patterns have not been determined with sufficient specificity to determine the best method.

The effect of these newer methods on the weights is harder to predict than that of the basic principal person method; which can at most alter the weight to the extent the PPSF differs from 1. As an extreme example, some earlier versions of the generalized least squares method regularly produced a few negative weights. All these methods can give unexpectedly low weight to some units under some conditions. Therefore we did not use them for the initial ACS release where there was limited time to revise the methods if examination of the ACS results showed problems. However, these methods will be considered in future research on ways of making the household and person weights consistent.

11. The Second Household Post-stratification Factor (HPF_2). Based on the theory described in the previous sub-section, the expectation is that the PPF tends to bring the race and Hispanic origin distribution of households more in line with the population, but tends to overestimate the total number of housing units. The problem is that the assumption that the PPSF for principal persons only includes the effect of whole-household misses is probably not completely correct. The PPSF also will include some effects of within-household undercoverage, so it overcorrects when used as a housing unit weight.

Because of this concern, the final step in calculating the housing unit weight was to apply another adjustment to control to MAF counts. This factor, called HPF_2, was calculated in exactly the same way as HPF_1, except that the weights used in the denominator also include HPF_1 and PPF. This brings the total housing unit estimates by tract back to approximately where they were after HPF_1, but the race and Hispanic origin distribution of households has been improved by the PPF. This is only "approximately," because it is the number of addresses, including out-of-scope addresses that are not actually housing units or that is controlled exactly, not the number of actual housing units.

The combined effect of HPF_1, PPF, and HPF_2 is shown in Table 4.

The final weight for housing units is then:

housing unit weight = BW x SSF x VMS x NIF x MBF x FAF x $HPSF_1$ x PPF x $HPSF_2$.

12. Rounding to an Integer Value. The final housing unit weights and final person weights are then rounded to integers. For housing units the rounding is done so that the number of housing units in a county, tract, or block computed using the rounded final weights is within one housing unit of corresponding total using the unrounded weights. For persons, the rounding was performed so that estimates of total county population, population by race, population by race and sex, and population by race, sex, and hispanic origin produced using the rounded weights are within one person of the corresponding estimates using the unrounded weights.

The weighted number of households is affected by the three factors according to expectations as shown in Table 4.

USE OF SEPARATE PERSON AND HOUSING UNIT WEIGHTS

The use of separate weights for persons and housing units means that there can be inconsistencies between certain estimates produced using the ACS person weights and related estimates using the housing unit weights. These inconsistencies may be disturbing to users of census long-form data, but are familiar in estimates from the CPS and other household surveys.

Two inconsistencies will be most noticeable:

i) The total number of persons can be calculated from person weights, but can also be calculated from the table giving the estimated (using the housing unit weight) numbers of households by household size, by totaling one times the number of one-person households plus two times the number of two-person households, etc. The results of second method will not equal the estimated number of persons using the person weight. Typically the housing unit weight will give a lower estimate for the number of persons. The person weight is thought to give the better result in this case.

ii) The estimated number of households using the housing unit weight is not equal to the estimated number (using the person weight) of people who are identified as being a "householder". The housing unit weight is thought to give the better result in this case.

The inconsistencies ultimately result from the intended goals of the person and housing unit weighting processes. Weights are meant to correct for the fact that certain groups of people or housing units are under-represented (or over-represented) in the sample, with the goal of giving greater weight to groups that are actually under-represented. The weights are not intended to correct for errors in reporting the characteristics of the individual people or housing units.

The first inconsistency is expected because of within-household undercoverage. For example, some people in two-person households are left off the roster, causing these households to be mis-reported as one-person households. A goal of the person weighting is to give greater weight to other people of the same age, race/origin, sex groups and geographic location as the missed people, because this is a matter of under-representation. However, it is not a goal of the household weighting to overweight those units that are reported to have two people and underweight those units reported as having one person, because this is an error in reporting the characteristics of housing units, not an under- or over-representation of housing units. Changing the weights to adjust for this reporting error might improve the estimates of the number of one- and two-person households, but would tend to distort the distributions of household characteristics that are reported correctly.

Similarly, within-household undercoverage causes some people to be identified as householders who would not have been identified as householders had all household members been listed. A typical case is thought to be when a male who would have been called the householder is left off the roster, which creates an "erroneous" female householder. A goal of the person weighting is to give higher weights to other males (typically householders) similar to the missing male, but not to give lower weight to the female "misidentified" as a householder, since she is not over-represented in the sample, merely incorrectly reported as a householder. Thus the person weights are expected to over-estimate the number of

householders.

The inconsistencies are a symptom that there are errors in the survey process that are at best partially corrected by the weighting. It is an open research issue whether it is better to cover up the inconsistencies due to within-household undercoverage by "over-weighting" some housing units and under-weighting others, compared to their actual representation in the population (Alexander, 1990). Theoretically, the "right answer" is to identify a group of housing units that have missing people and to impute additional people to those units, and more generally to try to impute the correct characteristics to the unit for whatever characteristics that are reported in error. Weighting cannot be expected to cure these problems. The decennial census procedures provide more opportunity to assign people to an appropriate housing unit during the data collection; therefore such inconsistencies are not expected in census data.

FOOTNOTES

1. This corresponds to a 10 percent coefficient of variation (CV) for an item which is 10 percent of the population. For measuring year-to-year change, our recommended limit is 250,000 population, corresponding to a 5 percent CV for a 10 percent item. The designation "profile" means that the estimates meet the former of these reliability requirements.

2. For the 1996 ACS, "small" governmental units are local areas with 1,000 or fewer addresses. The decennial census long form samples small governmental units (defined as less than 2,500 population in 1990) at a higher rate than other areas. The ACS will also oversample such units, except for the 2000-2002 national comparison sample.

3. It is a correct observation that another measure called the "effective sample size" is reduced by the subsampling. This increases the sampling error of the estimates, but causes no nonresponse bias.

Table 1: Weighted Response Rates

	As Percent of Mailout Addresses	As Percent of Total Estimated Eligible Addresses
Completed by mail, including Late Mail Returns	60.9	62.6
Completed by CATI	8.0	8.2
Completed by CAPI*	26.7	27.4
Eligible non-respondents*	1.7	1.8
Out-of-Scope Units**	2.7	NA
	100%	100%

*Including a weight of 3 for subsampling.
**Includes subsampling weight for subsampled cases.

Table 2: Total Weight (BW) of Units Mailed and Weight (BX x SSF) of Units Completed in June for Rockland, New York

Units Mailed in June	Mail Responses Received in June	CATI Interview in June	CAPI Cases Resolved in June
7888	4099	874	2449

Table 3: Estimated Number of Persons

		Rockland County	Multnomah County	Brevard County	Fulton County
Total Persons	Before PPSF	264,031	605,353	429,104	13,515
	After PPSF	270,885	613,243	447,729	14,359
Persons with Race = "Black"	Before PPSF	31,545	32,524	33,701	38
	After PPSF	30,219	41,642	39,151	53

Persons with Race = "Other"	Before PPSF	15,483	43,423	9,206	115
	After PPSF	15,900	44,072	10,256	138
Persons of Hispanic Origin	Before PPSF	21,531	28,515	15,885	10
	After PPSF	23,147	24,540	18,714	30

Table 4: Estimated Number of In-Scope Housing Units

	Rockland County	Multnomah County	Brevard County	Fulton County
Before HPF_1	91,579	267,249	201,967	6300
After HPF_1	94,116	273,166	206,263	6329
After PPF	94,852	276,024	208,349	6584
After HPF_2	94,136	273,239	206,303	6341