THE ACCURACY AND COVERAGE EVALUATION: THEORY AND APPLICATION

Howard Hogan, Bureau of the Census
Bureau of the Census Room 2024/2 Washington, DC 20233-7600

Keywords: Dual System Estimation, Census Adjustment, Undercount.

1. The Dual System Estimation Model

The U.S. Census Bureau has made a preliminary determination to correct the initial Census 2000 population figures for measured net undercount (Prewitt, 2000). This correction will be based on the Accuracy and Coverage Evaluation (A.C.E.). The A.C.E. is a post-enumeration survey based on the dual system estimator (DSE).

This paper discusses both the general question of designing a post-enumeration survey (PES), and how these general questions are addressed in the U.S. Census Bureau's plans for the A.C.E. planned as part of Census 2000. Throughout, I will use the terms DSE and PES when a general question is discussed and A.C.E. for specific details of the U.S. 2000 Design.

The use of the dual system model is well known either for measuring the completeness of vital events registration (Sekar and Deming, 1949; Marks, et al, 1974) or for use in measuring coverage errors in census data (Marks, 1979,Wolter, 1986, U.S. Bureau of the Census, 1985.) Application of the dual system model in the context of the 1990 Census, including the issue of census adjustment, is documented in (Hogan 1992, 1993).

The standard Petersen (1896) or Sekar-Deming estimator expressed as:

$$\hat{N}_{++} = N_{+1}\left(N_{1+}/N_{11}\right) \qquad (1)$$

where

N_{11} is the number of people counted in both the census and the survey,
N_{+1} is the number of people counted in the census,
N_{1+} is the number of people counted in the survey, and
N_{++} is the total number of people.

That is, the total population is estimated by the number captured in the census multiplied by the ratio of those in the survey to those in both systems (i.e., the inverse of the coverage rate of the census, as measured by the survey).

The DSE will yield a direct estimate of the population of class *j,* as well as any sum of classes. The class *j* might be the household population of a state, of a district, of an ethnic group, or perhaps of an ethnic group within a state. Requirements for estimating small or local populations, for example, age by sex, by race, by town, often far exceed the capacity of even a very large sample. To meet this needs, the DSE is combined with a synthetic assumption to produce estimates for areas of geography smaller than that defined by the domain *j*. The synthetic estimator assumes that a proportion or ratio measured at an aggregate level applies equally to all sub-groupings. (Gongalez and Hoza, 1978) Using a synthetic assumption, we write

$$\hat{N}^{s}_{jkh} = CCF_j \, C_{jkh} \qquad (2)$$

$$CCF_j = \frac{\hat{N}_j}{C_j} \qquad (3)$$

Where,

$\hat{N}^{s}_{jkh}$ is the estimated population for domain *j*, level of geography *k* (i.e. town, tract, block) and finer demographic subclass *h*.

CCF_j is the net coverage correction factor

$\hat{N}_j$ is the DSE

C_{jkh} is the measure (usually census count) of the population available at the smaller level of geography *k* (i.e., town, tract, block) and finer demographic subclass *h*.

$$C_j = \sum_k \sum_h C_{jkh}$$

C_j need not equal N_{+1}, which is the number of people correctly included in the census. N_{+1} is estimated from sample data and is not available for all small areas. C is normally the census count, including imputations and erroneous inclusions (duplicates, etc.).

Summing over group *j* and subclass *h* yields a measured population for a given geographic area *k* (state, county, town).

$$\hat{N}^{s}_{k} = \sum_j \sum_h CCF_j \, C_{jkh} \qquad (4)$$

For example, *j* may define all 0-17 year-old Asians in owner-occupied housing units while *k* may define Orange County, California, and *h* may define 11-year-old girls.

While this produces a small-area and small-group estimate, this calculation can generate fractions. The typical user of census data prefers whole person records. The U.S. Census uses controlled rounding and person imputation to create integer numbers of person records for ease of tabulation and data acceptance.

We now turn to the detailed steps of "filling in" Equation 2. The reader interested in more of the operational and timing details of the 2000 design should refer to Childers, 2000, or Childers and Fenstermaker, 2000.

2. Measuring Correct Enumerations

The first step in operationalizing Equation 1 is to define and estimate the list or set of individuals "correctly" in the census. In this context "correctly" has four dimensions:

1. Appropriateness
2. Uniqueness
3. Completeness
4. Geographic correctness

"Appropriateness" means that the person should be included in the census. People who die before or who were born after the census reference date (April 1 in the U.S.) are not part of the population (universe) to be measured. Similarly, records that refer to fictitious "people," tourists, or animals are out-of-scope.

"Uniqueness" refers to the fact that we wish to measure the number of people included in the census, not the number of census records. If more than one record refers to a single person, the count of records must be reduced for purposes of the DSE.

"Completeness" means that the census record must be sufficient to identify a single person. If it lacks sufficient identifying information, we cannot determine whether the person was appropriately and uniquely included in the census, nor can we determine whether he or she was also included in the survey.

Although completeness is necessary for the DSE, the census count includes imputations and other incomplete enumerations. Census operations normally have a requirement for a "data-defined person." In Census 2000, the requirement is two characteristics where name counts as a characteristic. Name must have at least three characters in the first and last name together. The characteristics that are included in the counting are relationship, sex, race, Hispanic origin, and either age or year of birth. (Childers, 2000).

When a record does not meet these requirements census processing substitutes (imputes) a data-defined record. Since the census processing flags all these whole-person imputations, the quantities are known and need not be estimated. Traditionally, the number of whole person imputations is denoted by II.

Additionally, there are person records that are acceptable for census processing but insufficient for use in the DSE. This group includes records with reasonably complete data but without a person's name. Accurate matching or additional interviewing is not possible for these cases. For A.C.E. 2000, the definition for "sufficient information for matching" is complete name and two characteristics. (Childers, 2000)

"Geographic correctness" means that people are included in the census where they should be included. Enumerations outside that defined search area are counted in the census but not <u>correctly</u> included in the census. This area must be searched during the matching process as well as searched for duplicates. As the size of the search area increases, the complexity increases and the chance of false matches grows. Normally a small area is defined such as a block and enumeration area.

Two dimensions must be defined to operationalize a smaller area (1) correct location and (2) the search area around the correct location.

The "correct location" defines where, under the DSE residence rules, the person should be included. These rules may differ from the rules used in the census. The only requirement is that the location be uniquely defined and consistently applied during PES processing.

In the 1990 PES and 2000 A.C.E. the Census Bureau adopted the following rule:

> The person is correctly included in the census if he or she is included at the location where the person considers, at the time of the survey interview, to have been his or her usual residence as of April 1.

This definition generally follows the census rules. However, it makes an explicit allowance for the fact that the concept of "usual residence" is somewhat subjective. Because of this subjectivity, where the person considers his/her usual (April 1) residence may have changed by the time of the survey interview. This, by itself, does not bias the DSE. However, it does require <u>consistent</u> reporting of the "correct location."

The second dimension of geographic correctness is the area of search around the correct location, i.e., the <u>search area</u>. The concept of a search area is to accommodate errors in either the census or survey assignment of residents to a particular geography. It has the effect of lowering the variance and can, in some circumstances, lower the bias as well.

The A.C.E. uses the following definition:

> A person was correctly enumerated if the person was counted in the block cluster containing his/her usual residence; or if he/she was included by the census in the housing unit where he/she usually resides, and the housing unit was included in a block adjacent to the correct block cluster.

These concepts are used to define the number of people correctly included in the census [N_{+1}].

The definition of "correctly included" does not depend on the correctness of classification *j*. For example, if a person was really 19 years old, but was counted in the census as 17, he/she is still considered as correctly included. This is discussed in a later section.

To estimate the number of people correctly included in the census, one must take a sample of all data-defined census enumerations. This sample is called the enumeration (or E) sample. Census whole person imputations (II's) are not part of the E-sample frame.

To maximize correlation with the population sample (see below), the A.C.E. first defines a set of sample areas. These are either a single block or a group of contiguous blocks and are known as block clusters. If a block is sampled, all census records coded to that block, even incorrectly, fall into sample. If the block contains many census housing unit records it may be subsampled.

The records in the E sample will be checked for completeness. Only records that meet the minimum completeness requirement can be considered as correctly enumerated in the census. Records are then searched throughout the search area to see if the person was counted more than once within the sample block (uniqueness). Duplicate search is done using computer-assisted clerical matching. If more than one record is found, the extra records are coded as duplicates.

Appropriateness and geographic location cannot be determined from the census enumeration alone, but require additional interviewing. If interviewing locates a member of the household, or an acceptable respondent who can confirm the person's existence and that the person had his/her usual residence there on April 1, the enumeration is accepted as correct.

The E-sample follow up is required in the U.S. because experience has shown that the census contains an appreciable level of erroneous enumerations. Ideally, this source of error would be controlled and excluded in the enumeration phase but this approach has not proven practical. Some countries assume that census erroneous enumerations can be ignored.

If the respondent reports that the person did not live in the block or search area on April 1, the enumeration is excluded from the correct enumerations. This can occur when the person responded to the census but moved before April 1; the person moved in after April 1 but was enumerated by the census nonresponse follow up operation; or when a parent incorrectly reports a college student as living at home.

The interviewers may determine that the person never existed or was never associated with the block. These records are considered erroneous. It can be difficult in some cases to prove that a "person" was not real, especially in a large block.

An important source of error arises from the need to often accept proxy responses to verify many enumerations. If the proxy reports a different "correct" residence than the person himself would, an enumeration could be miscoded. The A.C.E. uses proxy interviews for households that have moved between the time of the census and the time of the A.C.E. interviews. The A.C.E. requires the interviewers to find at least three knowledgeable respondents before coding a record as fictitious. However, since the person might have lived somewhere else in the block, it can be difficult in some situations to code the record fictitious.

After missing data estimation and sample weighting, we can estimate the number of people correctly counted in the census as

$$N_{+1} = (C - II)\frac{CE}{N_e} \qquad (5)$$

Where

C = Census total records, including imputed, duplicate, fictitious, etc. (the Census count),

II = number of whole-person census imputations,

CE = weighted estimate of appropriate, unique, complete and correct enumerations,

N_e = weighted E-sample estimate of total, including duplicate, fictitious, etc.

This completes the estimation of the number of people "in the Census."

3. Measuring the Proportion of People Correctly Enumerated

Having defined the set of correctly enumerated people, the next step in DSE is to estimate the census coverage rate, N_{11}/N_{1+} .

Conceptually, estimating the rate entails (1) taking a sample of people, (2) determining whether they should be enumerated in the census, and (3) determining whether they were, indeed, correctly enumerated, using the same definitions as were used to measure N_{+1}. If an unbiased sample can be drawn of people who should have been enumerated and if we can determine whether they actually were correctly enumerated (included in the census), then the DSE will produce asymptotically unbiased estimates. If each step can be approximately correct, the results will approach an unbiased estimate.

The first step in the process is, normally, to draw a random area sample. The A.C.E. uses the same set of block clusters for this purpose that it uses to define the E sample.

Interviewers then canvass the block and prepare an independent list of people who should have been enumerated. This list constitutes the population or P sample. The (weighted) sum of the people on this list, denoted $\hat{N}_p$, estimates N_{1+}. However, it is not the number which is of interest, but the ratio of N_{11} to N_{1+}, which we approximate by the ratio of correct matches, $\hat{M}$, to $\hat{N}_p$.

Operationally, the "correctly enumerated" census records are searched to see if the P-sample people were enumerated. The (weighted) number who were matched $(\hat{M})$ estimates N_{11}.

The DSE model will work if we can approximate:

1. Operational independence
2. Consistent reporting
3. Accurate matching
4. Homogeneity within post-stratum *j*

3.1 Operational Independence

Operational independence is the easiest assumption to approximate, but still requires vigilance. In Census 2000, the A.C.E. sample is drawn and the housing units listed before the delivery of the census questionnaires. Although personal contact is minimal, some people may react differently to the census because of their inclusion in survey listing. Early telephone interviews are allowed for independently listed housing units linked to a census address with a completed census questionnaire. This operation occurs while census nonresponse follow up is still being conducted in the area. Personal visit interviewing takes place concurrently with some census "coverage improvement" interviewing. Clearly, some contamination could occur. Great care is taken to prevent the same field staff from working the same area in both Census and A.C.E. and to prevent the sharing of information. Still, some people may react differently to the survey because they were enumerated, for example, by a very polite or very surly enumerator. Others may believe that they have a duty to provide the information once, but not twice. Operational independence must also be preserved in office procedures.

Definitions of "nonresponse" or "sufficient information" are sometimes applied differently to matched and non-matched P-sample records.

The A.C.E. guards against unnecessarily introducing operational dependence by forcing the processing system to first decide whether a case is acceptable for matching and only then attempt matching. The philosophy is "Do not attempt to find a match unless you would be satisfied that, if no match is found, the person was not enumerated!"

As with the E sample before beginning the matching, P-sample records first are reviewed for:

(1) Appropriateness
(2) Uniqueness
(3) Completeness
(4) Geographic correctness

The A.C.E. will contain almost no obviously fictitious records. One important safeguard is the use of Computer Assisted Personal Interviewing (CAPI). The CAPI instrument makes falsification difficult by "time stamping" the interview and recording every key stroke. We have instituted a quality assurance process to minimize other sloppy or dishonest A.C.E. interviewing. In addition, one important exception to our "no follow up" rule are cases where A.C.E. fabrication is possible, e.g., cases where no one in the household matches, implying possible fabrication.

Out of scope records, e.g., group quarters, are screened out. Occasionally, survey duplicates occur and these are eliminated (uniqueness). Finally, if the survey interview does not meet minimal standards, the case is converted to nonresponse and is later imputed.

3.2 Consistent Reporting of Residence

To measure the number of people correctly in both systems, we must determine whether or not a P-sample person was correctly enumerated in the census. This is done by searching the correct census records in the area where the person should have been enumerated.

The same definition of geographic correctness must apply both to whether an enumeration (in the E sample) was correct and to whether the person (in the P sample) was correctly enumerated. Failure to make these concepts agree is termed "balancing error."

Specifically, we must have the same definition of "correct" location and the same search area around the correct location. Errors can result in both erroneous non-matches and erroneous matches. Difficulty comes primarily from two sources. First, both the P- and E-sample accept proxy responses. Thus, even though the person might have a clear and consistent understanding of his usual residence, the proxy respondent may not. Secondly, the way in which the question is posed in each interview could lead to different responses even from the same person. This might result in false non-match/not correctly enumerated. On the other hand, if the person was incorrectly included by the census, we could incorrectly count the person as "correctly enumerated."

The other dimension of geographic correctness is, again, the extent of search. The same area must be used to define the correct residence for determining both whether an enumeration was correct and whether a person was correctly enumerated. This is achieved by consistently applying the same search area definitions as in Subsection 3.1.

3.3 Accurate Matching

The purpose of matching is to determine whether a person interviewed in the P sample was also enumerated in the census within the defined search area. Much of the matching is now done by a computerized matching system. The system produces matches, possible matches, and non-matched cases. Repeated tests have shown that cases matched by the computer are nearly certainly correctly linked. (Belin, 1993.) Nearly all clerical matching is now computer-assisted and largely paperless. This new system makes searching easier, including duplicate search. It restricts the codes clerks can apply to only those appropriate for the situation. Since the searching and data entry is now easier, we feel it is more likely to be accurate. For example, the almost paperless system should eliminate lost and misfiled A.C.E. questionnaires.

The first-level clerks are backed up by a team of 46 technicians. These technicians have been trained since September 1999. They are supported by a team of seven permanent analysts, most of whom have been matching for many years. Each level of matching acts as quality assurance for the level before. In addition, each level can refer problem cases to the next higher level. All matching will be done in one location by one staff. The 1980 and 1990 matching operations were done in three and seven sites, respectively.

The use of the A.C.E. procedures for movers also greatly simplifies the matching. Under the procedures used in 1980 and 1990, it was necessary first to code the reported correct Census Day residence to the correct census geography before beginning matching. This procedure was difficult, especially in rural areas. Mover matching was never before automated. In A.C.E. all matching, including for movers, will be done in the E-sample block cluster or an adjacent block, using the same computer and computer-assisted clerical matching system.

3.4 The Role of After-Matching Reinterview

Some cases are sent to the field to gather further information after the initial matching is complete. This after-matching reinterview is often termed "follow up interview."

The follow up interview process, like all PES activities, must fit into the overall framework of the DSE. Specifically, it must account for:

1. Appropriate, unique and correct response
2. Independence between census and survey inclusion probabilities
3. Balancing P- and E-sample errors
4. Unique location matching rules
5. Treatment of missing data.

Follow up is only useful if it provides more accurate or consistent responses. Simply obtaining a different response is not justification. Since follow up takes place further from the census reference date than the initial interview, it is more difficult to obtain accurate responses. This is equally true for E-sample follow up and P-sample follow up. To provide better responses, follow up must use better resources, for example: (1) Better respondents (household vs. proxy), (2) A better trained, supervised or quality-controlled interviewer, or (3) Better questions or interview procedures.

The census data collection period extends from mid-March through mid-summer. Little emphasis is placed on verifying that the people were residents of the household on April 1. Quality assurance reinterview to prevent fabrication is minimal. Because of better training and supervision, and more complete questioning, the A.C.E. follow up interviewing can, in general, obtain more accurate information on residence and location than that gathered during the census process itself. Thus all non-matched E-sample cases are sent to follow up.

Follow up can, however, compromise independence. If all cases were sent to follow up, independence would not necessarily be compromised. However, cases that are matched during initial matching are seldom sent to follow up. To do so would stress the resources available for follow up. Instead, only non-matches or "possibly matched" cases are usually selected for follow up. This can introduce operational dependence.

The biases that can be introduced by follow up can occur even if the follow up interview was successfully conducted, since follow up may selectively change the defined "correct location" for non-matches but not for matches. If the follow up operation results in a non-interview, further biases can be introduced depending upon the missing data models applied to these cases. (See Section 3.7)

Choosing cases for follow up requires balancing the need for accurate and consistent information with the need for independence. As noted above, for E-sample non-matched cases, the A.C.E. contends that reporting improvements from better staff, training, and questions outweigh the losses in independence.

For the P sample, we only follow up cases when better information is likely. Cases sent to follow up include:
1. Possible matches, since with the information at hand the interviews can resolve the situation.
2. Initial proxy interviews that result in non-matches. Since we have not spoken to a household member, we have reason to doubt the accuracy.
3. Non-matched cases where, for the same housing unit, the census reports one family and the A.C.E. reports another. In order to ensure consistent reporting of Census Day address between the P sample and the E sample, these cases are sent out together.
4. Partial-household non-matches.

Cases that match and some other non-matched cases are generally not sent to follow up. For example, the A.C.E. does not follow up whole-household non-matched cases where the census missed the unit, reported it as vacant, or could not obtain an interview (last resort information only).

3.5 Homogeneity Within Post-stratum *j*

The DSE requires that the capture probabilities be independent for all individuals within estimation domains called post-strata. This is approximated by making the post-strata as homogeneous as possible with respect to the census capture probabilities and then striving for as uniform as possible inclusion probabilities for the survey.

For A.C.E. we have conducted extensive research in defining the post-strata. (Haines, 2000) In addition, we have factored in what we know about changes between 1990 and 2000 and from our experience in the Census 2000 Dress Rehearsal conducted in 1998.

In defining post-strata we must balance the need for smaller, more homogenous strata against an increase in sampling variance and in ratio bias. Ratio bias follows from the fact that the DSE is inherently a ratio estimator. This bias tends to decrease as the size of the post-stratum increases. In addition, our treatment of movers adds an additional ratio (see below). For this reason, we have designed post-strata with a minimum expected sample size of 100.

For A.C.E. we will post-stratify based on the following variables:

1. Race Hispanic Origin (7)
2. Age, sex (7)
3. Tenure (2)
4. Metropolitan area size and type of enumeration area (4)
5. Mail Return rates (2)
6. Region (4)

Where the number in parenthesis refers to the number of categories. More details on the post-strata are found in Childers and Fenstermaker (2000).

Coverage differences between racial and ethnic groups is well documented. Social, cultural, linguistic and economic differences may lead different racial and ethnic to react differently to the census procedures.

Demographic analysis and previous coverage surveys have demonstrated that people are differentially missed in different age groups and that the pattern is different for males and females. Most important in this pattern is young adults. (Robinson, 1993)

The importance of tenure was first measured following the 1980 Census and then implemented in the 1990 post-stratification. Those who live in owner-occupied houses are less mobile. They may feel that they have more of a stake in their community and thus are more influenced by the census outreach program.

Metropolitan area size obviously affects housing patterns and is correlated with the way the Census builds its address lists. The combined variable "metropolitan area size and type of enumeration area" isolates differences in housing unit coverage. It may, in addition, measure some aspects of social and economic isolation.

The mail return rate measures public cooperation with the census, an important predictor of coverage. It also measures directly the proportion of the enumeration that must be done in the census nonresponse follow up. One difficulty in this variable is that not all areas of the country use mail-back. A small proportion is done by direct interview, and obviously have no "mail return rate." We have chosen to group these areas with "high" mail response areas.

Census Region picks up, among other things, broad differences in settlement patterns and housing stock. "Brown stone walk ups" are more common in the Northeast. Mobile homes are more common in the South.

Obviously, the complete cross-classifications can lead to very small cells. The maximum set of post-strata we believe we can support is 448.

In planning these post-strata, we face the decision of how to classify people who choose more than one race category. Since multiple-race responses were not allowed in any previous census, we had only limited data upon which to decide. Haines (2000) gives our rules for the treatment of multiple-race respondents.

3.6 Treatment of Movers

People who move between the census reference date and the time of the survey interview present a challenge for designing a DSE for census application. First, people who move are more likely to be missed by the census and by the survey. Secondly, if a person has a different "usual residence" at the time of the survey than he did at the time of the census, one must decide where to sample him.

In the 1990 PES, movers were sampled where they lived at the time of the survey interview. We then searched the census records at, and only at, their April 1 usual residence. This is known as procedure B (or PES). (Marks, 1979) This approach requires both coding the address to the correct Census Day geography and then matching. These activities are complex and time consuming.

The A.C.E. uses a different procedure known as procedure C (or PES). The A.C.E. will estimate the number of movers by the number of people who moved into the sample blocks between April 1 and the time of the A.C.E. interview (in-movers). If the population was closed to international migration, deaths, movement to group quarters, etc., then the number of people who moved in must equal the number who moved out (out-movers). The true total of out-movers should equal the total of in-movers. They are the same people in the population, if not in the sample. It is normally easier to find people where they are, so the measured number of in-movers is normally a better estimate of the total number of movers than the measured number of out-movers.

The proportion of movers who are correctly enumerated is estimated by matching the out-movers to the census records for the sample block and extended search area, if appropriate. The estimated number of correctly enumerated movers is then $\hat{M}_t = \frac{\hat{M}_o}{\hat{N}_o} \hat{N}_i$

where $\hat{M}$ denotes the weighted number of correct matches; $\hat{N}$ denotes the weighted population number; and the subscript denotes total moving (t), out-movers (o) and in-movers (i).

If we denote those who do not move by the subscript n, the overall coverage rate becomes

$$\frac{N_{11}}{N_{+1}} = \frac{\hat{M}_n + \hat{M}_t}{\hat{N}_n + \hat{N}_i}$$

The effect of procedure C is to increase the effective capture probabilities in the survey for movers and thus increase homogeneity of inclusion in the survey with respect to mover status (i.e., mover/nonmover). (Griffin 2000.)

Post-stratification and reweighting movers can reduce the heterogeneity but will not totally eliminated it. This heterogeneity will tend to make the DSE underestimate the true population.

3.7 Missing Data

There will be nonresponse and incomplete response at various steps. The goal of the missing data process is to improve the estimate of the number of people correctly counted (from the E sample) or the estimate of the coverage ratio (from the P sample). In designing missing data procedures, we choose methods that support the underlying DSE assumptions.

Missing data can occur with an initial failure to get a survey interview. One possible approach is to treat these cases as not in the survey. This treatment is adequate if whole household nonresponse is not correlated with census coverage. To the extent the two are correlated, excluding these cases increases the bias due to correlation.

If household nonresponse cases are similar to survey response cases within the same cluster, we can reduce the bias due to nonresponse by reweighting the cluster. In A.C.E. we will use two sets of nonresponse adjustments in the P-sample: one applied to non-movers and out-movers (for use in estimating the *proportion* of matches), the other applied to in-movers (for use in estimating the *number* of movers). For each procedure, if there are enough households, we will form nonresponse cells within the block cluster according to the type of basic address: single-family, apartment, and other. Where it is necessary, we will collapse cells according to pre-specified rules. (Cantwell 2000)

Even if the household has been interviewed, some PES cases can be unusable. These include, for example, those lacking sufficient information for matching and those whose residence status is still unresolved. When these are screened out before matching is attempted, one again has the option of treating them as "not in the survey". Again, one can do slightly better by taking advantage of what information is known. In this case it is information about the block cluster.

An important class are cases (1) that do not initially match, (2) are then sent to the field for additional interviewing, but (3) for which the additional interview is not successful (failed follow up). An efficient matching process will find most of the matches. Thus, cases sent to follow up will be disproportionately, if not predominantly, weighted with cases that were truly not in the census. Most follow up nonresponse cases are probably, in fact, not enumerated.

In the A.C.E. for a P-sample person with unresolved residence status, we will assign a probability of being a resident on Census Day according to operational or demographic information collected on the person. The idea is to group together into imputation cells people who are similar with respect to the A.C.E. operations--matches needing follow up, nonmatches needing follow up from whole-household nonmatches, persons resolved before follow up, etc.--or demographic characteristics. Within each imputation cell, we assign to the unresolved cases the weighted average residence probability of all resolved cases. (Cantwell, 2000) Starting with the 1990 PES, the U.S. has estimated the probability a nonresponse record was correct rather than assigned a "zero/one" classification. (Schenker, 1988, Belin 1993).

4. Synthetic Estimation

4.1 The Synthetic and Dual System Model

To this point, we have been dealing with the actual DSE. However, as noted in Section 1, we use a synthetic estimator to distribute the measured net undercount to local areas and small groups as noted in Equation 2.

In A.C.E. the carrying-down is based on the same post-stratification variables as the DSE itself. The synthetic estimation is based on the assumptions that (1) the DSE estimates the true population, and (2) within post-strata, the true population is distributed proportionally to the pre-adjustment (expected) census count.

Clearly, at some level the second assumption can be only true with respect to the expected census counts. That is, even if within post-strata all people had identical probabilities of being enumerated in the census, we would observe different outcomes across blocks. The underlying DSE explicitly models the undercount as a stochastic process.

As areas get larger, two things happen. First, the stochastic effect, or the random "block effect" begins to average out. Secondly, the effect of the actual undercount from a collection of blocks becomes positively correlated with the post-stratums coverage correction factor. That is, the larger the area, the more the area's undercount determines the net correction factor.

The stochastic effect would be trivial for all but the smallest areas if Wolter's (1986) autonomous independence assumption held in practice that is, if each person was included or missed independently of all other people. In fact, it is well known that within family or block, people are often missed as a group. The whole building (or sometimes even block) might be missed by the census address listing procedure. The failure of the autonomous independence assumption does not cause a bias in the dual system model as long as the underlying probabilities are equal within post-strata. This failure can mean that observed coverage for a block is inconsistent with the estimated undercount adjustment. However, as attention is turned to larger areas the stochastic effect diminishes and is replaced with the problem of true heterogeneity of the underlying capture probabilities. (Haines, 2000 for synthetic estimation details)

A related question is how differential bias of the DSE by post-stratum might manifest itself through the synthetic estimator. For example, many people accept that, when matching is under control, the DSE will underestimate the population due to correlation bias. When each post-stratum is viewed individually, one could say that the DSE moves the measure of population closer to the truth, but not all the way.

However, if correlation bias is more pronounced in some post-strata than in others, then these post-strata would not move as close as they should to the true population count. Members of these post-strata could conceivably be made relatively worse, although their measured population would be more accurate.

We might refer to the difference between the DSE and the true population as the residual or unmeasured undercount. Obviously, we have little evidence of where these "unmeasured" people live. If correlation bias is so strong that the unmeasured people live in areas where the DSE measures the smallest undercount rates, the DSE/synthetic model may fail to produce an improvement in relative shares. However, if the unmeasured undercount exists in the same area as the measured undercount (i.e., the two are correlated), we are likely to improve relative shares.

4.2 Misclassification Error

In the discussion so far, we have accepted the post-stratum classification, *j*, as fixed. In practice, some people will be classified in different post-strata in the census and in the survey. For example, a woman may be reported as age 28 in the census and 31 in the survey, placing her in different post-strata.

Such misreporting is normally not important for matching. Name, address, month and day of birth, relation and household composition are far more important than age, race or sometimes even sex. So, assuming a match, in the above example we would have one correctly enumerated 28 year-old in the E sample and one correctly enumerated 31 year-old in the P sample. Misclassification can be seen to have two effects. To the extent the true undercount probabilities are homogeneous with respect to the true characteristics, misclassification introduces heterogeneity (and heterogeneity bias) into the observed estimation cells. Note that this is true even if reporting is consistent between the census and the survey, because it can introduce unobserved subgroups within post-strata where the probabilities of inclusion in each system are correlated.

Inconsistent reporting between the census and the survey poses a problem for the synthetic estimator as well as for the DSE. This is easily seen by ignoring census imputations and erroneous enumerations. In this case, the coverage correction factor is the inverse of the matching rate $\left(N_{11j}/N_{1+j}\right)$ where *j* represents the post-stratum. If the classification into the post-strata is inconsistent between the census and survey, we would be applying the rate, estimated from one group, to a somewhat different group. While misclassification may be ignorable at the post-stratum level, it may be important locally. The A.C.E. seeks to protect itself against the general problem by avoiding, when possible, post-stratum definitions based on variables with high reporting variability.

5. Concluding Remarks

This paper has described the theory of the DSE, and has discussed how PES in general, and A.C.E. in particular, have implemented that theory. It has described the approximations necessary in real applications and the types of errors that can occur.

Obviously, it is the role of the survey designer and survey manager to balance and minimize the errors so as to produce useful and accurate measures of the population. When this is successfully done, Census 2000 A.C.E. will produce fair and accurate population measures for use by American scholars, planners and leaders.

References

Belin, T. R., (1993), "Evaluating Sources of Variation in Record Linkage Through A Factorial Experiment." *Survey Methodology*, 19, 13-29.

Childers, D., (2000), "Accuracy and Coverage Evaluation: The Design Document" DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-DT-1

Childers, D., and Fenstermaker, D., (2000), "Accuracy and Coverage Evaluation: Overview of the Design" DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-DT-2

Cantwell P., (2000), "Accuracy and Coverage Evaluation Survey: Missing Data Procedures" DSSD Census 2000 Procedures and Operations Memorandum Series Q-19

Gonzalez, M., and Hoza, C., (1978) "Small-Area Estimation with Application to Unemployment and Housing Estimates", *Journal of the American Statistical Association*, Vol. 73, No. 361, 7-15.

Griffin, R., (2000), "Accuracy and Coverage Evaluation Survey: Dual System Estimation" DSSD Census 2000 Procedures and Operations Memorandum Series Q-20

Haines, D., (2000), "Accuracy and Coverage Evaluation Survey: Synthetic Estimation" DSSD Accuracy and C overage Evaluation Survey Memorandum Series Q-22

Hogan, H., (1992), "The 1990 Post-Enumeration Survey: An Overview", *The American Statistician*, 46, 261-269.

Hogan, H., (1993), "The Post-Enumeration Survey: Operations and Results", *Journal of American Statistical Association,* Vol. 88, No. 423.

Marks, E. S., Seltzer, W., and Krotki, K. J. (1974), *Population Growth Estimation,* New York: Population Council.

Marks, E.S., (1979), "The Role of Dual System Estimation in Census Evaluation", in K. Krotki, Recent Developments in PGE, University of Alberta press, 156-188.

Peterson, C. G. J., (1896), "The Yearly Immigration of Young Plaice into the Limfjord from the German Sea," *Report of the Danish Biological Station,* 6, 1-48.

Prewitt, K., (2000), "Statement on the Feasibility of Using Statistical Methods to Improve the Accuracy of Census 2000."

Robinson, J.G., (1993), "Estimates of Population Coverage in the 1990 United States Census Based on Demographic Analysis," *Journal of the American Statistical Association,* 88, 1061-77.

Schender, N., (1988) "Handling Missing Data in Coverage Estimation with Application to the 1986 Test of Adjustment Related Operations" *Survey Methodology* Vol. 14 No. 1, pp 87-97.

Sekar, C.C., and Deming, W.E., (1949), "On a Method of Estimating Birth and Death Rates and the Extent of Registration," *Journal of the American Statistical Association,* 44, 101-115.

U.S. Census Bureau, (1985), Evaluating Census of Population and Housing, Statistical Training Document, ISP-TR-5, Washington, D.C.

Wolter, K. M., (1986), "Some Coverage Error Models for Census Data," *Journal of the American Statistical* Association, 81, 338-346