NORC WORKING PAPER SERIES

Development of Methodology for Evaluating Model-Based Estimates of the Population Size for States

WP-2021.03 | July 30, 1998 (with minor updates in 2021)

PRESENTED BY:
NORC at the University of Chicago
55 East Monroe Street
30th Floor
Chicago, IL 60603
Ph. (312) 759-4000

AUTHORS:
Shelby J. Haberman
Wenxin Jiang
Bruce D. Spencer

AUTHOR INFORMATION*

Shelby Haberman
haberman.statistics@gmail.com

Wenxin Jiang
wjiang@northwestern.edu

Bruce Spencer
bspencer@northwestern.edu

Table of Contents

List of Tables

Abstract

Loglinear models for probabilities of census erroneous enumeration and census nonenumeration are developed for the Census Bureau's 1990 Post Enumeration Survey (PES). The PES was a stratified random sample of 5,290 blocks following the 1990 census. Census enumerations within the blocks comprise the *E* sample. A separate survey was fielded in the blocks, and the resulting enumerations are the *P* sample. In principle, all persons in the PES could be classified as in one of three mutually exclusive categories: in the *P* sample and the *E* sample, in the *P* sample but not the *E* sample, and in the *E* sample but not the *P* sample. Such classification would be useful for estimating undercount and population sizes. We demonstrate that such classifications are not possible for the 1990 PES and we discuss the reasons. Nevertheless, we are able to fit loglinear models for probabilities of census erroneous enumeration and census nonenumeration. A feature of the models is the use of latent classes to model heterogeneity in the probabilities. The models can be used to develop simulated populations, upon which the effects of alternative census designs can be modeled.

Introduction

This report describes methodological work done under contract to the Census Bureau to develop models of undercounts and overcounts in the census, with particular attention paid to heterogeneity in capture probabilities at the local level.[1] The contract report has been cited in the literature but is not widely available, and because the work remains of interest for methodological reasons it is being reissued in slightly revised form in the NORC Working Paper Series. The research described in this report may be useful for several purposes.

1. One is the modeling of probabilities of census misses and erroneous enumerations, using latent class models to account for heterogeneity in those probabilities.

[1] The work was done under contract no. 50-YABC-2-66023 by NORC for the Census Bureau. The original report was titled "Activity 7: Development of Methodology for Evaluating Model-Based Estimates of the Population Size for States, Final Report." The work was done in support of construction of a simulated population to support the Bureau's planning of the 2000 Decennial Census.

2. Another is the combination and articulation of data collected from the two arms of a Post Enumeration Survey (PES); see Section 2. A PES as conducted by the Census Bureau consists of two samples, a sample of census enumerations (E sample) and a sample of people for whom it is to be determined whether they were enumerated in the census (and if so, where) or were not enumerated (or "missed"). The 1990 PES used samples of blocks (or clusters of blocks) and attempted to use the same blocks for both P sample and E sample.
3. A third purpose is to enable construction of a simulated population that contains a record for each person on each block in the U.S. in 1990 (or, if updated, for the subsequent 2000 census). Spencer (1998) provides details. Briefly, each record contains census data and indicators for variables for the 1990 PES. If the record corresponds to a person missed in the 1990 census, the census data will be imputed. The population is based on models for undercount in the 1990 census to adjust 1990 census data for undercount and overcount at the block level, and this report describes the development of those models. The models are designed to allow explicitly for unobserved heterogeneity in capture probabilities and for consequent correlation bias. Thus, the population is to be used for simulation purposes as if it were "true." The 1990 PES variables are imputed for the vast majority of the records, based on the observed data from the 1990 PES and from the various quality control and data evaluation programs for the 1990 PES. Once the simulated population is constructed, the evaluation of alternative designs and estimators can be easily and quickly studied.2

Background on the Census, PES, and Demographic Analysis

The decennial census of the U.S. attempts to enumerate the entire population in each geographic subarea of the U.S., down to the block level. The enumerations are subject to various errors. Gross undercount occurs when people are missed entirely or not counted in their appropriate geographic area or as members of their demographic subgroups. On the other hand, gross overcount occurs when people are enumerated multiple times, are enumerated where they should not have been, or are enumerated when they should not be. The difference between gross undercount and gross overcount is net undercount.

Information for estimating the components of undercount, gross overcount, and gross undercount and for obtaining geographic breakdowns is provided by the Post Enumeration Survey (PES); see Hogan (1993) for a comprehensive overview. The PES was based on a stratified random sample of 5,290 block clusters

[2] Thus, one can imagine selecting a block from the simulated population and seeing what data would be collected from a census or PES that included the block. One can also simulate the results from a nonresponse follow-up design directly or the results of imputations based on administrative records. From this data one can calculate an estimate and directly compare that with the true number of people, for the latter is known without error in this simulated population. A significant challenge in the planning of a decennial census is the prediction of accuracy attainable under alternative census designs with alternative methods, some based on technology still under development. One approach to meet this challenge is to develop and use a simulated population that is constructed to approximate the population.

(blocks or groups of blocks). Census enumerations within these blocks (or random parts of the largest blocks) are called the *E* sample. A separate survey was fielded in the blocks, and the resulting enumerations are the *P* sample. A new listing of housing units was prepared for each *P*-sample block independently of the listing used in the census. Very large blocks were subsampled, and all other blocks had 100 percent of their listed housing units included in the *P* sample. Interviews were to be conducted in every sampled housing unit to enumerate all persons living in the housing units at the time of the interview (typically within four months of the census) and to obtain sufficient information to ascertain whether they were enumerated in the census. The *E* sample consisted of all census enumerations coded to the sample blocks (correctly or not), except for the larger blocks, which were subsampled to match the subsampling in the *P* sample. The PES sample was restricted to the noninstitutional population, and excluded people living in jails, nursing homes, and other institutions, military personnel living in barracks or on ships, and homeless people.

The *P* sample can be used to study gross undercount in the noninstitutional population. The basic idea is to see what fraction of the *P*-sample persons were not enumerated in the census, and to use this fraction to estimate the gross undercount rate. The practicalities are somewhat complex. An attempt is made to match each person in the *P* sample to a census enumeration from the *E* sample for the corresponding block. If a match is found, then the person is classified as enumerated in the census. (False matches are rare; Davis, Biemer, Mulry, and Parmer 1991.) If a match is initially not made, several alternatives are possible before a final match status is assigned.

1. Perhaps the person really was enumerated, but the census enumeration was coded to the wrong block. To account for this, a search for a matching enumeration was extended to a search area consisting of ring of single blocks (of two blocks for rural areas) surrounding the sampled block; a match within the surrounding ring is considered a match. If the enumeration was coded mistakenly to a more distant block, the match would not be made, and the person would be considered missed in the census. In an attempt to balance out the error resulting from this, an enumeration coded outside the search area is defined as an erroneous enumeration, so on average, the presence of undercount caused by a limited search area would be offset by the presence of erroneous enumerations due to limited search areas.
2. The *P*-sample person could have moved to the block since the time of the census, in which case his or her census-day address would be sought and a match would be attempted at the census-day block. If the match were made there, the person would be considered to be matched.
3. The person might have been enumerated but incorrect information was recorded in the *P* sample or on the census questionnaire (West, Mulry, Parmer, and Petrik 1991), or a clerical matching error was made (Davis, Biemer, Mulry, and Parmer 1991), leading to failure to detect a match. Mulry and Spencer (1993, 1082, Table 1) found these to be among the major sources of error in the estimate of national undercount.

4. The person's household might not have responded in the *P* sample, in which case the match would not be directly attempted. This occurred for 1.6 percent of the intended interviews (Mulry and Spencer 1993, 1089). Instead, the set of responding households was essentially viewed as a subsample of the *P*-sample households, and the subsampling probabilities were reflected by nonresponse weighting adjustments applied to interviewed households.
5. In some cases (1.9 percent of interviewed cases), insufficient information was available to be certain of match or nonmatch status. Such cases were classified as "unresolved" and statistical models were used to assign a probability of a match (Belin et al. 1993). These models incorporate field information (e.g., pre-follow-up match code and mover status) and thus are more informative than a usual adjustment for missing data would be. For an analysis of error, see Mulry and Spencer (1993, 1089), and Breiman (1994, 469–70, 526) with rejoinder by Belin and Rolph (1994, 503–04), and rebuttal by Freedman and Wachter (1994, 535–36). We will consider each person in the *P* sample to be classified as having a probability of match. If the person was definitely matched or unmatched, the probability is 1 or 0 accordingly.

The *E* sample checks all enumerations in the *E*-sample blocks (or subsampled parts of the blocks) to see whether they are erroneous or not. An attempt is made to match every *E*-sample enumeration to the *P*-sample interview data for cases within the search area. If a match is made, the person is considered to be a correct enumeration. As with the *P* sample, however, failure to find a match initially does not imply an incorrect enumeration. Some enumerations are whole-household imputations in the census, or otherwise have insufficient information for matching to be possible; these are called "II" cases. A field follow-up operation is conducted for all nonmatches except the II cases. After the follow-up, cases were classified either as correct enumeration or erroneous enumeration or, if enumeration status was uncertain, as unresolved (1.4 percent of non-II cases). Unresolved cases were assigned probabilities of correct enumeration. Reasons for nonmatch include the following.

1. The enumeration could have been correct, but the *E*-sample person could have moved out of the block or died prior to the *P*-sample interview.
2. The enumeration could have been correct, but the person could have been missed in the *P* sample.
3. The *E*-sample person could have been coded to a block outside the search area, in which case the person should be defined as an erroneous enumeration.
4. The enumeration could have been erroneous.

It is well known (e.g., Bell 1993) that different groups of people can have very different undercount rates, and this heterogeneity can cause a downward bias in estimates of population size. For insight, suppose that a group is composed of distinct subgroups i of true sizes N_i and enumeration sizes X_i, suppose that all enumerations that occur are correct, denote totals by

$$N = \sum_i N_i$$

and

$$X = \sum_i X_i,$$

and let $f_i = X_i / N_i$. The subgroups could be of size 1, in which case f_i is the probability that individual i is enumerated. The overall enumeration fraction f is a weighted average of the f_i, with weights proportional to N_i, so that

$$f = \sum_i \left(N_i / N \right) f_i.$$

If f were known, the group size could be calculated as $X / f = N$, but in applications where heterogeneity is present but not accounted for, only the X_i are observed directly. A simple but reasonable scenario is one in which the estimator of f is approximately equal to a weighted average of the f_i but with weights proportional to observed counts X_i instead of true sizes N_i, say

$$\hat{f} = \sum_i \left(X_i / X \right) f_i.$$

For example, this would arise in a PES where the subgroup enumeration fractions for the *P* sample were proportional to those in the *E* sample, for then we would have, say, $\lambda N_i f_i$ from subgroup i observed in the *P* sample, and of these $\lambda N_i f_i^2$ would also be observed in the *E* sample. The overall fraction of the *P* sample that was enumerated in the *E* sample is thus

$$\frac{\lambda \sum_i N_i f_i^2}{\lambda \sum_i N_i f_i} = \hat{f}.$$

The resulting estimator of population size, $X / \hat{f}$, will tend to be too low, because Jensen's inequality implies that

$$1/\hat{f} \leq 1/f,$$

with equality only if f_i is constant over i. The downward bias caused by heterogeneity can also be shown to obtain under more general conditions that involve positive correlation of *P*-sample and *E*-sample enumeration probabilities rather than proportionality, a phenomenon known as "correlation bias" (Sekar and Deming 1949, Wolter 1986, Alho et al. 1993, Kadane, Meyer, and Tukey 1998). The Bureau uses poststratification in attempts to achieve equal rates within poststrata, but Alho et al. (1993) demonstrated using logistic regression models for individual enumeration probabilities that enumeration probabilities vary within poststrata. Estimates of correlation bias have been developed from triple-system estimates (Zaslavsky and Wolfgang 1990, 1993; Darroch et al. 1993) for St. Louis in the 1988 Dress Rehearsal; however, those estimates are available at this time for too few areas to be generalizable.

We distinguish between "moderate heterogeneity", which means that the f_i for individuals are simply spread out, versus "extreme heterogeneity", in which f_i for some individuals are near zero. If the latter condition applies, the PES will simply be unable to provide the information needed to produce an approximately unbiased estimator of population size— there will be too few "uncountable" individuals in the PES to support accurate modeling of their enumeration probabilities, and the estimates will basically be targeting the size of the population excluding these "uncountable" people.

At least some heterogeneity error is practically inevitable for small-area estimates. In application, data from a sample of areas in the PES is applied to produce estimates for other areas. The model used by the Bureau has held that the net coverage rate is constant within poststrata, regardless of geography unspecified by the poststratification. This model is called a synthetic estimation model. There is no avoiding the use of some model, explicit or implicit, to use data from outside the area to develop estimates that apply to the area of interest, and failure of any such model to account for heterogeneity will be called "synthetic estimation error." Following Schafer (1993), consider any poststratum, suppose that the probability that a person is enumerated is a constant f, and suppose for simplicity that the number of enumerations follows a binomial distribution. (If the enumerations are not stochastically independent but are clustered within households, the variances will be even larger.) The undercount rate for an area with N persons in the poststratum will be a random variable with expected value $1-f$ and variance $f(1-f)/n.$ Since the undercount rates for the areas are random, they will tend to be slightly unequal, particularly for the smaller areas. Although the expected values of the rates will be equal, the attained rates for the actual census will not. Thus, bias from synthetic estimation error will be present. (These biases are conditional on the actual census, but that is what is relevant.)

Previous approaches to estimating synthetic estimation error have relied on surrogate variables for which direct estimates are available independent of the synthetic model; see Fay and Thompson (1993), Kim, Blodgett and Zaslavsky (1991), and Wachter and Freedman (1994). The general problem with surrogate variables is that they correlate imperfectly with net undercount rates, and conflicting results (low error for some surrogates, and high for others) are not uncommon. The approach cannot tell us whether the synthetic estimation error is appreciable for the variable of actual interest. An alternative and more fully parametric approach is to fit models of the variable of interest that include "effects" for the areas. For example, Hengartner and Speed (1993) considered block effects for the erroneous enumeration rate and gross omissions rate in the 1990 PES. A difficulty with using models with a fixed effect for each area is that the number of areas can be too large relative to the sample size. Rather than fixed effects for the areas, one may use latent-class models or random-effects models, with the effect for an area considered to have been chosen from a distribution to be estimated. (Any random-effects model may be approximated to arbitrary accuracy by a latent-class model.) The classical formulation of the dual-system estimator, as well as fitting of latent class models, requires that each person enumerated in the PES be classifiable with a "triple" of indicators— as either in the *P* sample only, the *E* sample only, or both. Unfortunately, the PES data do not allow such classifications, for reasons discussed in Section 3.

Information about the national extent of net undercount by age, race, and sex is provided by the method of demographic analysis (DA), which uses administrative and vital statistics data to construct an alternative estimate of population. Subtracting the census count from this yields an estimate of net undercount. For example, Robinson et al. (1993) calculated births - deaths + immigration - emigration by age, race, and sex using data since 1935 to estimate the population under age 55 in 1990. The vital statistics were essentially complete, although the immigration and emigration numbers are problematic. To account for incompleteness in earlier vital statistics, more complex methods were used to estimate the population aged at least 55 and under 65, and aggregate Medicare enrollments (adjusted with estimates for incompleteness of enrollment) were used to estimate the population 65 and over. An important contribution from DA is the provision of sex ratios (numbers of males over numbers of females) by age and race. The DA sex ratios are believed to be more accurate than the DA numbers of persons, and the ratios provide an important tool for assessing the accuracy of the census and alternative estimates of population. DA estimates are only available at the national level, unless very strong assumptions are invoked to account for internal migration.

Some alignment of population definitions must be done for consistent comparison of the PES, census, and DA estimates. The racial classifications in birth statistics used in DA are different than those reported in the decennial census. In the 1990 census, almost 10 million enumerations, mostly identified as Hispanic, were of "Other Race - Not Specified"; Robinson et al. (1993) allocated these to other race categories to force consistency with historical demographic statistics; for example, 250,000 of these were classified as black males and 247,000 as black females. The race classification issue does not affect the estimate of total net undercount, but it does affect the estimates of undercount by race. To make the census and DA estimates comparable with the PES, estimates of the institutional and military populations need to be subtracted from the former, although the sex ratios change only slightly as a result (Bell 1993, 1110). Sex ratios tend to be lower in the census than in DA for persons 20 and over, particularly for blacks (Bell 1993, 1110, Table 2). This indicates a disproportionately larger net undercount for males than for females. The sex ratios from the census for blacks aged 20–29, 30–44, and 45–64 were only about 90 percent as large as the ratios from DA. In addition to diagnostic information concerning undercount in the census, sex ratios provide information for checking and even adjusting the PES. One aspect is relatively prosaic, as the sampling weights in the PES can be adjusted so that the sex ratios of the *E* sample match those of the census for nonmilitary noninstitutional population. The other use of sex ratios points to a deficiency in the PES estimates of net undercount. It is not clear whether the small sex ratios for blacks in the PES are due to failure of assumptions for the models or to nonsampling errors in the data. Bell (1993) developed estimates of correlation bias for major demographic groups in the 1990 census under the assumption that enumeration probabilities were heterogeneous within poststrata for males but not females. Those methods do not directly lead to estimates of correlation bias at the local level, however, without the use of untestable assumptions.

PES Data Complexities

The PES was designed to produce two samples, the *P* sample and the *E* sample. The *P*-sample cases carry sampling weights reflecting probabilities of selection into the *P* sample (adjusted for unit nonresponse) and the *E*-sample cases carry sampling weights reflecting their selection probabilities. It is desirable to construct a single “concatenated” file including each *P*-sample and *E*-sample case with an appropriate sampling weight and the appropriate “triple” (Section 2). This would imply that the number of *P*-sample cases that matched to the *E* sample was the same as the number of *E*-sample cases that matched to the *P* sample. Unfortunately, there are several reasons why the concatenated file could not be constructed. One reason is that, for many subgroups (poststrata), the weighted number of matches in the *P* sample, that is, the weighted number of persons in both the *P* sample and the *E* sample, is larger than the weighted number of persons in the *E* sample. This will be called the “negative cell” phenomenon. It occurred in part because, although in theory the *E* sample and *P* sample included the same blocks, in practice the overlap was less than 100 percent. Gregg Diffendal suggested the following example of what could go wrong when the blocks do not overlap. Suppose the *P* sample counted ten persons in block A and they all matched. The *E*-sample block should have been the same block (A) but instead was an adjacent block (B), and it counted ten different persons and they were all correct enumerations who did not match to the *P* sample. In constructing the concatenated file, we would think there are 20 persons that should have been counted in block (A), when in fact there were only ten. In this example, the number of *P*-sample matches equals the *E*-sample total, but if the blocks had different numbers of people either number could be larger than the other, and this could occur on a random basis. Thus, sampling variability is one explanation for the negative cells.

To illustrate the problem further, we describe our experiences with two versions of the concatenated files, one based on resolved cases and the other based on resolved and unresolved together. The basic approach was to take the *P*-sample cases plus unmatched *E*-sample correct enumerations. If we take the *E*-sample file alone and look at the sum of the weights, we get around 244 or 245 million, which matches the number of census enumerations for the subpopulations (Bell 1993, 1110, Table 2). Of these, about 8.5 million were erroneous enumerations. If we look at the concatenated file based on the resolved cases only and we add up the sum of the weights for *P*-sample cases that match and *E*-sample-only correct enumerations, we get about 236 million, which is about 8.5 million short due to the excluded erroneous enumerations. So far, so good. But, when we look at the concatenated file that includes both resolved and unresolved cases, we get about 251 million for the sum of the weights for *P*-sample cases that match and *E*-sample-only correct enumerations. (These do not include the 8.5 million erroneous enumerations.) The weights for unresolved cases were adjusted so that, for example, an *E*-sample case with a weight of w and correct-enumeration probability of p contributed wp to the correct enumerations and $w(1-p)$ to the erroneous enumerations. The unresolved cases needed to be included in the analysis because their imputed match probabilities were based on additional field information. Thus, there are too many cases included, and the problem is the movers.

Let us define as an outmover an *E*-sample person who moved after census day and before the PES interview date on his or her destination block; a *P*-sample person who moved into a *P*-sample block (or subsampled part of a *P*-sample block) after census day and before the PES interview date will be called an inmover. In the *E* sample, attempts were made to see whether outmovers were correctly enumerated or not, even though they could not be matched to the *P* sample except in the very rare instances where both their origin and destination blocks were in the PES. In the *P* sample, the census-day address was sought for all identified inmovers, and a match was attempted with the enumeration record coded to the census-day address. Note that a person who moved between the time of census enumeration and the *P*-sample interview would have two chances to be in the PES, once if the person's census day address were in an *E*-sample block, and once if the block the person moved to were selected in the *P* sample. The simplest way to adjust for this would be to restrict the file so that cases had only one chance of appearing. Although *P*-sample movers are identified in the file, the movers tended to have different capture probabilities than nonmovers, and simply omitting them would lead to biased estimates of the probability of being enumerated in the census (Section 7). Outmovers have the problem that their match status to the *P* sample is almost always nonmatch because the blocks they moved to are so rarely in the *P* sample; this nonmatch status would be assigned only a small probability if 100 percent of the blocks were in the PES. Although *E*-sample movers were not identified in the files, information was collected concerning their status. Thus, at our request the Bureau reanalyzed some data from the *E* sample to identify outmovers, and a substantial fraction, though not all, were identified. However, when the *E*- sample outmovers were excluded from the concatenated file, the sum of the weights was still too large (about 269 million).

Others have coped with the problem of concatenating the files by restricting their analysis. Thus, Alho et al. (1993) treated a person as captured in the *P* sample only if the person did not move. This approach is internally consistent, but suffers the drawback that movers and nonmovers have different probabilities of being enumerated in the census; the issue is not likely that moving is causal but that it is associated with other characteristics that do cause people to be missed in the census. Also, they only used resolved cases.

As discussed earlier, the weighted number of *P*-sample matches often exceeds the weighted number of *E*-sample cases (matches and nonmatches combined), and sampling variance is a plausible explanation for that discrepancy. When one focuses on block-level counts for the sampled areas, however, sampling variance is much less of an issue for the large blocks (which were subsampled) and is not an issue at all for the smaller blocks (which were not subsampled). We made two comparisons of numbers of matches in the *P* sample and *E* sample. First, we matched blocks on the *E*-sample file (specifically, the *E*-sample "post- imputation file," wtdposte2.dat) and the *P*-sample file ("post-imputation file," fullwtdp2.dat). Table 1 shows the distribution of blocks by whether the number of *E*-sample matches was the same as, greater than, or less than the number of *P*-sample matches, for small, medium, and large blocks, when matches in the surrounding ring were included and movers were included. The last three columns show the average, the standard deviation, and the average absolute value of the discrepancies in number of matches per block. The standard deviations are about 2 for small blocks, 6 for medium, and 18 for large blocks. The latter could reflect the effect of subsampling, but for every size block it is more likely than not that the number of matches is not the same. This may be explainable by the nature of the match codes, and is not necessarily cause for alarm, but it does not augur well for construction of a concatenated file. Details on the exact criteria for declaring a match are in Appendix I.

We were concerned that some of the discrepancies in Table 1 could be caused either by matches occurring outside the block (in the surrounding ring) or by movers. Therefore, we restricted attention to matches occurring within the block and involving nonmovers. For each block we compared the number of *P*-sample nonmovers who matched within the block to the number of *E*-sample nonmovers who matched.

The data used for this analysis were the same as before, but they were augmented by identification of movers from the file prepared for us at the Bureau. The results are shown in Table 2. The discrepancies are much reduced relative to Table 1, but they are still prevalent enough to prevent construction of a completely consistent concatenated file. Details on the exact criteria for declaring a match are in Appendix II.

Table 1. E-sample Matches Minus P-sample Matches for each Block, Movers and Surrounding Ring Included

	Small	Medium	Large
Total	1601	4748	2209
$\#e = \#p$	783	741	48
$\#e > \#p$	367	2559	1700
$\#e < \#p$	451	1448	461
$\text{av}(\#e - \#p)$	-.16	.50	2.51
$\text{sd}(\#e - \#p)$	1.9	5.9	18.4
$\text{av}(\|\#e - \#p\|)$	.34	2.27	3.38

The most promising way to construct the concatenated file is to begin with (i) the full *P* sample and to include (ii) all *E*-sample non-outmovers who were not matched to the *P* sample as well as all (iii) outmovers who would not have matched if their destination block had been in the PES. To identify (ii) we took all *E*-sample cases that were coded as nonmatches and were not among the group identified as outmovers. We were not able to identify (iii), because we could not distinguish outmovers who would have matched from those who would not have matched, under the counterfactual supposition that the block they moved to was in the *P* sample (nearly all outmovers in the *E* sample were coded as nonmatch). Failure to include group (iii) is not likely significant, however, because Table 2 shows that there are relatively too many *E*-sample matches anyway among the nonmovers.

Table 2. Results on Matches at Block Level, Movers and Surrounding Ring Excluded

	Small	Medium	Large
Total	1601	4748	2209
$\#e = \#p$	1010	1866	476
$\#e > \#p$	455	2662	1697
$\#e < \#p$	136	220	36
$\text{av}(\#e - \#p)$	0.06	1.06	1.20
$\text{sd}(\#e - \#p)$	1.0	3.2	5.8
$\text{av}(\|\#e - \#p\|)$	.19	1.36	1.23

Analysis strategy

Due to the problems outlined in the previous section, it was decided to analyze the *P* and *E* samples separately and to use the results to attempt to estimate the noninstitutional population. In addition to the probabilities of correct enumeration from the *P* sample and the probabilities of erroneous enumeration from the *E* sample, basic demographic data were available for each sample concerning race and ethnicity, tenure, age, and sex and additional information concerning characteristics of the household and of the block to which the individual belongs. Information was also obtained for the *P* and *E* samples concerning whether the individual moved after the Census enumeration. The mover data are more difficult to apply because comparability of responses for the *P* and *E* samples is problematic and because the mover data are not available for the complete census.

Historical Note: The following text (through the end of Section 4) was in Section 1 of the original contract report, but it presumes details that were only discussed in Sections 2 and 3, and so it is being included here for completeness. It may be skipped without loss of continuity.

The work departed from what we had envisioned in several respects. Our Study Plan was predicated on the assumption that all persons in the PES could be classified as a "triple" with indicators of membership in one of three mutually exclusive categories: in the *P* sample and the *E* sample, in the *P* sample but not the *E* sample, and in the *E* sample but not the *P* sample. This assumption seemed quite reasonable and was used in a somewhat restricted way in prior work by Alho et al. (1993). For reasons described below, however, the assumption is not correct. Attempts to modify the data to allow the assumption to hold consumed large amounts of person-hours and calendar time but were ultimately not successful. On the positive side, the attempts did lead to increased understanding of the PES, particularly the importance of movers. These issues are described in more detail below.

The Study Plan envisioned the development of models for assigning variables for each personin the simulated population to indicate how they would (hypothetically) respond in various data collection programs. For example, we wanted to include *P*-sample variables that would be observed (whether the person was captured in the *P* sample; assigned match status with census) and *E*-sample variables that would be observed (whether the person was captured in the *E* sample; assessment of the number of times the person was classified as enumerated in the census) if the person had been selected in the PES. Moreover, to study the extent of nonsampling errors due to response error, interviewer error, etc., we wanted to include (ii) *P*-sample and *E*-sample variables that should have been observed if there were no nonsampling errors in the PES and if the person had been selected in the PES. To construct these models, we would have used the 1990 evaluation studies to generalize from the Matching Error Study and Evaluation Follow-up, among others, to the full PES, and then generalize from the PES to the full population. The models we would have developed would relate the observed variables for sampled people to the "true" variables, and inconsistencies between the two would denote nonsampling errors. We decided not to develop these models, however, for two reasons. First, and most important, the nonsampling errors for 1990 may not be useful indicators for the nonsampling errors for the nonsampling errors for 2000. Second, the evaluation studies were not infallible themselves. Thus, we decided not to build into the simulated population the assumption of similar nonsampling errors from 1990, and this part of the anticipated project was not carried out.

It was hoped that the Census Bureau's methods for creating the transparent file could be implemented for the simulated population, but this was not to be the case. To develop the models needed would involve modeling the probability of missing a household (of a given size and composition) and then modeling the conditional probability of within-household misses given the enumeration of the household. Although such joint estimation may be possible for the 2000 ICM program, it is not achievable for the 1990 PES due to the nature of the data files. The intended uses of the simulated population are such that its construction should focus on population size and not be overly concerned about the composition of housing units. Thus, Spencer (1998) concluded that adjustments for net undercoverage of housing units should not be made in this application. Accordingly, we have not developed the models for household undercoverage.

Population estimation

Given the data from the P and E samples and given national data from the Census enumeration, it is possible to estimate the number of individuals in the United States who had selected characteristics at the time of the Census enumeration. To consider the estimation procedure, consider the following model. Let S be a set with N potential census enumerations for the noninstitutional population of the United States, where N is unknown. Let J be a nominal variable defined on S with possible values EE (erroneous enumeration), C (correct enumeration) and U (unenumerated). For x equal to EE, C, or U, let $\delta_x(J)$ be the variable with value 1 if $J = x$ and value 0 if $J \neq x$, and let $\delta_x(J(s))$ be the value of $\delta_x(J)$ for s in S. Then the correct noninstitutional population total is

$$N_T = \sum_{s \in S} [1 - \delta_{\mathrm{EE}}(J(s))],$$

the number of individuals in S_T, the set of s in S such that $J(s)$ is C or U. The noninstitutional population reported by the Census enumeration is

$$N_R = \sum_{s \in S} [1 - \delta_U(J(s))].$$

Let S_E be the set of s in S such that $J(s)$ is EE or C, so that S_E consists of all Census enumerations. Observe that S_E has N_R members.

Let $J(s)$ be a random variable for s in S, where $J(s)$ is x with probability $p_x(s)$ for x equal to EE, C, or U. Assume that $p_U(s) < 1$ for each s in S. Let the adjustment factor be

$$a(s) = \frac{p_C(s) + p_U(s)}{p_C(s) + p_{\mathrm{EE}}(s)} = \frac{1 - p_{\mathrm{EE}}(s)}{1 - p_U(s)}$$

for s in S. If the $p_x(s)$ are known for each s in S, then N_T may be approximated by

$$\hat{N} = \sum_{s \in S} [1 - \delta_U(J(s))] a(s) = \sum_{s \in S_E} a(s).$$

The error of the approximation is

$$\hat{N} - N_T = \sum_{s \in S} b(s),$$

where

$$b(s) = \begin{cases} a(s), & J(s) = \mathrm{EE} \\ a(s) - 1, & J(s) = C \\ -1, & J(s) = U. \end{cases}$$

The expectation of $\hat{N} - N_T$ is 0. In the special case in which the $J(s)$ are independently distributed, the mean-squared error is given by

$$\sigma^2(\hat{N} - N_T) = \sum_{s \in S} \left[p_U(s) + \frac{p_U^2(s) + p_{\mathrm{EE}}(s) p_C(s)}{p_C(s) + p_{\mathrm{EE}}(s)} \right].$$

In practice, the $a(s)$ must be estimated by use of the P and E samples. In addition, cases arise in which the value of $J(s)$ is unclear. Let T_E be the set of individuals enumerated in the E sample, and let T_P be the set of individuals found in the *P* sample. Note that, for s in T_E, $J(s)$ is either EE or C, while, for s in T_P, $J(s)$ is either C or U. It is assumed that the probability of fabrication is negligible for the *P* sample. Let $q_E(s)$ be the probability of inclusion in the *E* sample of s in S_E, and let $q_P(s)$ be the probability of inclusion in the *P* sample of s in S_T. Let

$$W_E = \sum_{s \in T_E} [1/q_E(s)]$$

and

$$W_P = \sum_{s \in T_P} [1/q_P(s)].$$

For s in T_E, let $r_{EE}(s)$ be the probability, according to the Bureau of the Census, that $J(s) = EE$. In resolved cases, $r_{EE}(s)$ is 0 or 1. Otherwise, $r_{EE}(s)$ is strictly between 0 and 1. For s in T_P, let $r_U(s)$ be the probability assigned by the Census to the event $J(s) = U(s)$. Let X_j, $1 \le j \le k$, be real variables defined on S. Apply the logit models

$$\log \frac{p_U(s)}{p_C(s)} = \sum_{j=1}^{k} \beta_{jU} X_j(s)$$

and

$$\log \frac{p_{\mathrm{EE}}(s)}{p_C(s)} = \sum_{j=1}^{k} \beta_{j\mathrm{EE}} X_j(s)$$

to the *P* and *E* samples, respectively. Weight observations by the inverse of the sampling probability, and then apply maximum likelihood, so that one obtains the equations

$$\sum_{s \in T_E} X_j(s) \hat{f}_{\mathrm{EE}}(s) / q_E(s) = \sum_{s \in T_E} X_j(s) r_{\mathrm{EE}}(J(s)) / q_E(s),$$

$$\sum_{s \in T_P} X_j(s) \hat{f}_U(s) / q_P(s) = \sum_{s \in T_E} X_j(s) r_U(J(s)) / q_P(s),$$

$$\hat{\lambda}_{\mathrm{EE}}(s) = \sum_{j=1}^{k} \hat{\beta}_{j\,\mathrm{EE}} X_j(s),$$

$$\hat{\lambda}_U(s) = \sum_{j=1}^{k} \hat{\beta}_{jU} X_j(s),$$

$$\hat{g}_{\mathrm{EE}}(s) = \exp(\hat{\lambda}_{\mathrm{EE}}(s)),$$

$$\hat{g}_U(s) = \exp(\hat{\lambda}_U(s)),$$

$$\hat{f}_{\mathrm{EE}}(s) = \exp(\hat{g}_{\mathrm{EE}}(s)) / [1 + \exp(\hat{g}_{\mathrm{EE}}(s))],$$

and

$$\hat{f}_U(s) = \exp(\hat{g}_U(s)) / [1 + \exp(\hat{g}_U(s))].$$

Estimate $a(s)$ by

$$\hat{a}(s) = \frac{1 + \hat{g}_U(s)}{1 + \hat{g}_{\mathrm{EE}}(s)} = \frac{1 - \hat{f}_{\mathrm{EE}}(s)}{1 - \hat{f}_U(s)}.$$

Given observations $X_j(s)$ for all s in S_E, the logical estimate of N_T is

$$\hat{N}_L = \sum_{s \in S_E} \hat{a}(s).$$

If the $X_j(s)$ are unavailable for s in S_E but not T_E, then alternative estimates are

$$\hat{N}_{LE} = \sum_{s \in T_E} [\hat{q}_E(s)]^{-1} \hat{a}(s)$$

and

$$\hat{N}_{LEC} = \sum_{s \in T_E} [\hat{q}_E(s)]^{-1} [1 - r_{\mathrm{EE}}(J(s))][1 + \hat{g}_U(s)].$$

In the latter case, an attempt is made to weight observations only in the E sample that were correctly enumerated. A further refinement may be based on known population totals. Let Z_U be real variables with values 0 or 1 for u from 1 to v, and let

$$\sum_{u=1}^{v} Z_u(s) = 1.$$

Let

$$N_{Zu} = \sum_{s \in S} Z_u(s) > 0$$

be known for each u, and let

$$n_{Zu} = \sum_{s \in T_E} Z_u(s) / q_E(s) > 0$$

for each u. Then an adjusted weight may be used. Here $[q_E(s)]^{-1}$ is replaced by

$$V(s) = N_{Zu} / [n_{Zu} q_E(s)]$$

for $Z_u(s) = 1$. In this way,

$$\sum_{s \in T_E} V(s) Z_u(s) = N_{Zu}.$$

The resulting estimates are

$$\hat{N}_{LEZ} = \sum_{s \in T_E} V(s) \hat{a}(s)$$

and

$$\hat{N}_{LECZ} = \sum_{s \in T_E} V(s) \delta_C(J(s))[1 + \hat{g}_U(s)].$$

In this study, the Z_u were defined to adjust for known totals for all combinations of gender, race (black or nonblack), and age group (0–9, 10–19, 20–29, 30–44, 45–64, and 65+). Thus v was 24.

The estimates proposed have relatively complicated statistical properties for several reasons. The model is never in practice valid, so an error component due to misspecification exists. The sampling is complex due to stratification and clustering at the block and household level. Weights are used to compensate for unequal probabilities of being sampled and for nonresponse. Further biases exist due to a variety of errors introduced by nonresponse, imputation, and response error. Variability of estimates may be assessed when needed by use of jackknifing applied to block clusters. Nonrandom biases can be assessed to a limited degree by use of demographic analysis.

5.1 Estimates for subpopulations

The methods proposed for estimation of the population total may be applied with little modification to estimates of subpopulation totals. Let B be an indicator function for a subpopulation Q of S, so that $B(s) = 1$ for s in Q and $B(s) = 0$ for s not in Q. Then the true number of individuals in population Q is

$$N_{TQ} = \sum_{s \in S} [1 - \delta_{EE}(J(s))] B(s)$$

and the enumerated number of individuals in population Q is

$$N_{RQ} = \sum_{s \in S} [1 - \delta_U(J(s))] B(s).$$

Possible estimates of N_{TQ} are

$$\hat{N}_{LQ} = \sum_{s \in S_E} \hat{a}(s) B(s)$$

and

$$\hat{N}_{LECZQ} = \sum_{s \in T_E} B(s) V(s) \delta_C(J(s)) [1 + \hat{g}_U(s)].$$

Estimation of variability

To assess the variability of results due to random error, grouped jackknifing was employed. The block clusters were divided at random into ten groups, with each cluster assigned independently to a group and with the probability of membership in any particular group set to 0.1. Estimates were derived by omission of one group at a time. This procedure is approximate since stratification effects were ignored; however, such effects appear likely to be small and the required records were unavailable in any event.

Variable selection

With these problems in mind, efforts were made to select variables X_j for $1 \leq j \leq k$ that were relatively successful predictors of $J(s)$. As in Gilula and Haberman (1994, 1995), the criteria employed were based on the estimated logarithmic penalty functions

$$\hat{H}_E = -(W_E)^{-1} \sum_{s \in T_E} [q_E(s)]^{-1} [\delta_{\mathrm{EE}}(J(s)) \log \hat{f}_E(s) + \delta_C(J(s)) \log[1 - \hat{f}_E(s)]]$$

and

$$\hat{H}_P = -(W_P)^{-1} \sum_{s \in T_P} [q_P(s)]^{-1} [\delta_P(J(s)) \log \hat{f}_P(s) + \delta_C(J(s)) \log[1 - \hat{f}_P(s)]].$$

The rationale for a logarithmic penalty function can be seen by consideration of prediction of the probability that a nominal variable Z assumes specified values. Let Z have values from 1 to $v > 1,$ and let q be a v-dimensional vector such that $q(j) \geq 0$ for all integers j in $[1, v]$ and $\sum_j q(j) = 1.$ Consider prediction of the distribution of Z by the vector $q.$ Let a nonnegative penalty $f_j q(j))$ be assessed if $Z = j.$The expected penalty is then

$$\sum_{j=1}^{v} P(Z = j) f_j(q(j)).$$

Consider the requirement that for any possible distribution of $Z,$ the expected penalty is minimized and finite for $q(j) = P(Z = j)$ for $1 \leq j \leq v.$ For $v \geq 3,$ it follows that

$$f_j(x) = a \log(x)$$

for x in [0, 1] for some nonnegative $a.$ If $v = 2,$ then this formula for f_j continues to satisfy the required inequality, but other functions can also be constructed. A particularly elegant proof appears in Savage (1971).

It is desirable to select the predicting variables X_j to minimize $\hat{H}_E$ and $\hat{H}_P,$ although it is desirable as well to maintain relatively small $k.$ This choice helps reduce sampling variability and is more efficient computationally. This choice is also helpful in transfer of the methodology to smaller samples, an issue that would arise in the use of logit models within states.

In application of estimated logarithmic penalty functions, it should be noted that $\hat{H}_E$ is an estimate of the minimum expected penalty

$$H_E = -N^{-1} \sum_{s \in T_E} [p_{\mathrm{EE}}(s) \log \tilde{p}_{\mathrm{EE}}(s) + p_C(s) \log[1 - \tilde{p}_{\mathrm{EE}}(s)]]$$

for prediction of $J(s)$ given that $J(s)$ is EE or C and $\hat{H}_P$ is an estimate of the minimum expected penalty

$$H_P = -N^{-1} \sum_{s \in T_E} [p_U(s) \log \tilde{p}_U(s) + p_C(s) \log[1 - \tilde{p}_U(s)]]$$

for prediction of $J(s)$ given that $J(s)$ is C or $U,$ where

$$f_{\mathrm{EE}}(s) = p_{\mathrm{EE}}(s) / [p_C(s) + p_{\mathrm{EE}}(s)],$$

$$f_U(s) = p_U(s) / [p_C(s) + p_U(s)],$$

$$\sum_{s \in S} [p_C(s) + p_{\mathrm{EE}}(s)] X_j(s) \tilde{f}_{\mathrm{EE}}(s) = \sum_{s \in S} X_j(s) p_{\mathrm{EE}}(s),$$

$$\sum_{s \in S} [p_C(s) + p_U(s)] X_j(s) \tilde{f}_U(s) = \sum_{s \in S} X_j(s) p_U(s),$$

$$\lambda_{\mathrm{EE}}(s) = \sum_{j=1}^{k} \tilde{\beta}_{j\mathrm{EE}} X_j(s),$$

$$\lambda_U(s) = \sum_{j=1}^{k} \tilde{\beta}_{jU} X_j(s),$$

$$g_{\mathrm{EE}}(s) = \exp(\lambda_{\mathrm{EE}}(s)),$$

$$g_U(s) = \exp(\lambda_U(s)),$$

$$\tilde{f}_{\mathrm{EE}}(s) = \exp(g_{\mathrm{EE}}(s)) / [1 + \exp(g_{\mathrm{EE}}(s))],$$

and

$$\tilde{f}_U(s) = \exp(g_U(s)) / [1 + \exp(g_U(s))].$$

One has

$$H_E \geq -N^{-1} \sum_{s \in S} \{ p_{EE}(s) \log f_{\mathrm{EE}}(s) + p_C(s) \log[1 - f_{\mathrm{EE}}(s)] \},$$

with equality only if the logit model $f_{\mathrm{EE}}(s) = \tilde{f}_{\mathrm{EE}}(s)$ holds for s in S_E, and

$$H_P \geq -N^{-1} \sum_{s \in S} \{ p_U(s) \log f_U(s) + p_C(s) \log[1 - f_U(s)] \},$$

with equality only if the logit model $f_U(s) = \tilde{f}_U(s)$ holds for s in S_E. It should be noted that $\hat{N}_L$ approximates

$$\hat{N}_L = \sum_{s \in S} [p_{\mathrm{EE}}(s) + p_C(s)] \tilde{a}(s)$$

if

$$\tilde{a}(s) = \frac{1 + g_U(s)}{1 + g_{\mathrm{EE}}(s)} = \frac{1 - \tilde{f}_{\mathrm{EE}}(s)}{1 - \tilde{f}_U(s)}$$

is the adjusted factor associated with the logit models.

The estimates $\hat{H}_E$ and $\hat{H}_P$ are biased. In the application under consideration, the bias is relatively small unless k is large. Changes in H_E or H_P of 0.01 are substantial, and changes of 0.001 are rather small. The biases are roughly proportional to k/n, where n is the number of sampled subjects. As evident from Section 7.2, the constant of proportionality is generally somewhat less than 10 for the sampling procedure under study. Since n is roughly 400,000 in both the *P* and *E* samples, biases are expected to be small for k less than 40. Due to the complex sampling procedure employed, bias corrections were obtained when needed by use of jackknifing.

The bias corrections are not fully satisfactory in terms of assessment of predictions when parameter estimates are used rather than the true parameters. As in Gilula and Haberman (1994, 1995), the appropriate bias correction for this prediction assessment is about twice as large as the appropriate bias correction for estimation of H_E and H_P. This result is related to a proposal by Akaike (1970), but the present suggestion applies to complex samples and to mis-specified models.

Poststratification

A very simple example of use of logits is the case of poststratification in which each X_j is 0 or

$$1,\ \sum_{j=1}^{k} X_j(s) = 1 \text{ for } s \text{ in } S, \text{ and}$$

$$N_{jx} = \sum_{s \in S} X_j(s) p_x(s) > 0$$

for $1 \le j \le k$ and for x equal $C, U,$ or EE. In this case, if $X_j(s) = 1$ for some integer j from 1 to k, then

$$\tilde{f}_{\mathrm{EE}}(s) = \frac{N_{j\mathrm{EE}}}{N_{j\mathrm{EE}} + N_{jC}},$$

$$\tilde{f}_U(s) = \frac{N_{jU}}{N_{jU} + N_{jC}},$$

and

$$\tilde{a}(s) = \frac{N_{jU} + N_{jC}}{N_{j\mathrm{EE}} + N_{jC}}.$$

If, for $1 \le j \le k$

$$n_{jE\,\mathrm{EE}} = \sum_{s \in T_E} X_j(s) r_{\mathrm{EE}}(s) / q_E(s) > 0,$$

$$n_{jEC} = \sum_{s \in T_E} X_j(s)[1 - r_{\mathrm{EE}}(s)] / q_E(s) > 0,$$

$$n_{jPU} = \sum_{s \in T_E} X_j(s) r_U(s) / q_P(s) > 0,$$

and

$$n_{jPC} = \sum_{s \in T_E} X_j(s)[1 - r_U(s)] / q_P(s) > 0,$$

then

$$\hat{f}_{\mathrm{EE}}(s) = \frac{n_{jE\,\mathrm{EE}}}{n_{jE\,\mathrm{EE}} + n_{jEC}},$$

$$\hat{f}_U(s) = \frac{n_{jPU}}{n_{jPU} + n_{jPC}},$$

and

$$\hat{a}(s) = \frac{[n_{jPU} + n_{jPC}] n_{jEC}}{[n_{jE\,\mathrm{EE}} + n_{jEC}] n_{jPC}}.$$

One has

$$\hat{N}_{LE} = \hat{N}_{LEC} = \sum_{j=1}^{k} n_{jEC} [n_{jPU} + n_{jPC}] / n_{jPC}.$$

Results of selection

The selection process emphasized variables that have been studied by the Bureau of the Census and have received attention in previous literature. The major new feature of the approach adopted here is use of an objective criterion to compare the predictive power of different model choices. On the whole, it appeared to be the case that almost all readily achieved predictive power was achievable by relatively simple models. To obtain some perspective, the following models were compared. Table 3 summarizes findings.

5. For a baseline, a single dummy variable was used with constant value 1. Thus, all individuals are predicted to have the same associated probabilities $p_{EE}(s)$, $p_U(s)$, and $p_C(s)$.
6. Dummy variables were used for the 48 possible combinations of race, tenure, sex, and age group. In this case, $\hat{H}_E = 0.2136$ and $\hat{H}_P = 0.2578$. Here $k = 48$.
7. Additive logit models are used for the predictors race, tenure, sex, and age.

Table 3. Estimated Log Penalties

Model	No. of parameters	*E*-sample estimate	*P*-sample Estimate
1	1	.2280	.2735
2	48	.2136	.2578
3	9	.2142	.2587
4	27	.2137	.2580
5	357	.2109	.2538
6	18	.2119	.2527
7	19	.2113	.2449

8. Logit models for the predictors race, tenure, sex, and age are used that include all main effects and two-factor interactions.
9. The division of Hogan (1993) into 357 post-stratification groups was used. These groups were based on race/ethnicity, tenure, urbanization, region, age, and sex.
10. Logit models were used that were additive in race, Hispanic status, tenure, sex, age, metropolitan location, relationship to householder, marital status of household (both married man and married woman in household of age at least 15 years), region, nonresponse rate for block, and log block size. Parameter estimates can be found in Table 4.
11. Logit models were used that were additive in race, Hispanic status, tenure, sex, age, metropolitan location, relationship to householder, marital status of household, region, actual nonresponse rate for tract, log block size, and mover status. Parameter estimates can be found in Table 4.

Bias corrections are rather small, except in the poststratification case. For model 6, the bias correction for the estimate of HE is only 0.00026, so that twice the correction is only 0.00051. For model 7, bias correction for estimation of HE is only 0.00057, so that twice the correction is only 0.00115. In the poststratification case (model 5), bias corrections for HP and HE are 0.00182 and 0.00238, respectively. Doubling the corrections yields 0.00363 and 0.00475, respectively. These bias corrections somewhat reduce the relative attractiveness of the post-stratification model. The bias corrections are considerably larger than would be expected under simple random sampling. For example, in the poststratification case, bias corrections under simple random sampling would be expected to be about 0.00045.

Table 4. Parameter Estimates for Models 6 and 7

Sample	Independent variable	Model 6 estimate	Model 7 estimate
P	Constant	-3.37776	-3.48116
P	Age 0–9	.78614	.67995
P	Age 10–19	.66716	.55533
P	Age 20–29	.96487	.77624
P	Age 30–44	.63020	.55361
P	Age 45–64	.30386	.26492
P	Female	-.17306	-.17491
P	Black	.32237	.39421
P	Hispanic	.32768	.38597
P	Renter	.59617	.43340
P	Metropolitan	.07968	.13228
P	Unrelated to householder	.76149	.66357
P	North	.01815	.06302
P	South	.06262	.06917
P	Midwest	-.19773	-.19589
P	Marital status	-.35964	-.34785
P	Block nonresponse rate	.02301	.02505
P	Log block size	-.22598	-.23028
P	Nonmover		1.33445
E	Constant	-3.52883	-3.53683
E	Age 0–9	.32487	.34187
E	Age 10–19	.50184	.51755
E	Age 20–29	.78922	.81381
E	Age 30–44	.29169	.30206
E	Age 45–64	.03585	.03913
E	Female	-.13486	-.13519
E	Black	.12707	.11963
E	Hispanic	.03406	.02654
E	Renter	.27841	.30838
E	Metropolitan	.13032	.11798
E	Unrelated to householder	.18898	.22062
E	North	.28810	.27865
E	South	.14543	.14765
E	Midwest	.00219	.00185
E	Marital status of household	-.28418	-.28673
E	Block nonresponse rate	.02116	-.10981
E	Log block size	.11579	.02115
E	Nonmover		-1.05078

Conclusions on model selection

These results permit several basic conclusions. It is much more efficient to use additive models with more variables rather than more complete models with fewer variables, especially with the *P* sample. Thus, in the *P* sample, model 6, which uses 18 parameters, clearly dominates the poststratification model with 357 strata. Addition of interaction terms to models 6 and 7 yielded very minor improvements in the fit. It is possible to improve prediction substantially by use of variables not normally measured by the census, as is evident from the gain obtained by use of mover status. Thus, model error can be expected to be an issue of practical significance.

Application of model selection to population estimation

The models considered in the preceding section were applied to population estimation and to estimation of sex ratios. Estimates of population totals were derived for all combinations of race, sex, and age and for all combinations of race and sex. These estimates were used to obtain sex ratios for all combinations of age and race and for both racial classifications. Results were rather similar for all but the most primitive models. Results are summarized in Tables 5, 6, 7, and 8 for models 5, 6, and 7. Calculations use $\hat{N}_{LECZ}$. At this stage, estimates were derived from the *E* and *P* samples using $\hat{N}_{LECZ}$. The only information used from the complete census was the totals by age, race, and sex for the noninstitutional population. These estimates were compared to results from demographic analysis. The fundamental difficulty is that the estimated number of adult black males is rather low compared to demographic analysis. The sex ratios in Table 6 are particularly disturbing for blacks in this age range.

Table 5. Estimated Population Totals in Thousands by Sex, Age, and Race: Models 5, 6, and 7

Sex	Age	Race	Census total	Model 5 estimate	Model 6 estimate	Model 7 Estimate	Demog. analysis
Male	0–9	Nonblack	16,000	16,460	16,660	16,520	16,400
Female	0–9	Nonblack	15,200	15,610	15,590	15,500	15,600
Male	10–19	Nonblack	15,000	15,310	15,270	15,150	14,900
Female	10–19	Nonblack	14,300	14,530	14,380	14,300	14,300
Male	20–29	Nonblack	17,200	17,840	18,040	17,680	17,600
Female	20–29	Nonblack	17,100	17,650	17,800	17,500	17,200
Male	30–44	Nonblack	25,700	26,380	26,390	26,210	26,500
Female	30–44	Nonblack	26,100	26,550	26,680	26,530	26,100
Male	45–64	Nonblack	20,000	20,280	20,200	20,130	20,500
Female	45–64	Nonblack	21,400	21,560	21,540	21,470	21,400

Sex	Age	Race	Census total	Model 5 estimate	Model 6 estimate	Model 7 Estimate	Demog. analysis
Male	65+	Nonblack	11,100	11,180	11,080	11,060	11,300
Female	65+	Nonblack	15,900	15,930	15,870	15,830	16,000
Male	0–9	Black	2,820	3,045	3,147	3,135	3,100
Female	0–9	Black	2,760	2,956	2,949	2,922	3,020
Male	10–19	Black	2,600	2,836	2,819	2,789	2,670
Female	10–19	Black	2,620	2,821	2,758	2,744	2,680
Male	20–29	Black	2,280	2,361	2,431	2,375	2,570
Female	20–29	Black	2,740	2,943	3,048	2,985	2,870
Male	30–44	Black	3,000	3,257	3,176	3,183	3,430
Female	30–44	Black	3,690	3,861	3,906	3,887	3,780
Male	45–64	Black	2,010	2,132	2,081	2,076	2,280
Female	45–64	Black	2,540	2,660	2,672	2,662	2,550
Male	65+	Black	920	956	935	935	960
Female	65+	Black	1,470	1,496	1,500	1,500	1,450

Table 6. Estimated Sex Ratios by Age and Race for Models 5, 6, and 7

Age	Race	Census	Model 5 estimate	Model 6 estimate	Model 7 estimate	Demog. analysis
0–9	Nonblack	1.053	1.054	1.068	1.066	1.051
10–19	Nonblack	1.043	1.054	1.062	1.060	1.043
20–29	Nonblack	1.006	1.011	1.013	1.011	1.019
30–44	Nonblack	.987	.993	.989	.988	1.014
45–64	Nonblack	.936	.940	.938	.938	.957
65+	Nonblack	.699	.702	.698	.699	.707
0–9	Black	1.023	1.030	1.067	1.073	1.027
10–19	Black	.994	1.005	1.022	1.016	.994
20–29	Black	.831	.802	.797	.796	.896
30–44	Black	.807	.843	.813	.819	.907
45–64	Black	.790	.801	.779	.780	.893
65+	Black	.628	.639	.623	.623	.661

Table 7. Estimated Population Totals in Thousands by Sex and Race: Models 5, 6, and 7

Sex	Race	Census total	Model 5 estimate	Model 6 estimate	Model 7 estimate	Demog. analysis
Male	Nonblack	105,000	107,450	107,640	106,750	107,200
Female	Nonblack	110,000	111,830	111,860	111,130	110,600
Male	Black	13,630	14,589	14,589	14,493	15,010
Female	Black	15,820	16,737	16,833	16,700	16,350

Table 8. Estimated Sex Ratios by Race for Models 5, 6, and 7

Race	Census	Model 5 estimate	Model 6 estimate	Model 7 estimate	Demog. analysis
Nonblack	.961	.962	.962	.961	.969
Black	.862	.872	.867	.868	.918

An alternative analysis, Model 8, employed the mover variable from the *P* sample. In this case, a two-stage procedure was used. A weighted logit analysis was used to predict mover status from the other 11 predictors in the model with 12 predictors (Table 9). To each individual s in the *E* sample who was correctly enumerated, a prediction $\hat{a}_M(s)$ of $a(s)$ was made given that the person was a mover and a prediction $\hat{a}_{NM}(s)$ of $a(s)$ was made given that the person was not a mover. In these predictions, the parameter estimates from Model 7 for the *P* sample were used (Table 4). Given the predicted probability $\hat{p}_M(s)$ that the person was a mover, the weighted average

$$\hat{p}_M(s)\hat{a}_M(s)+[1-\hat{p}_M(s)]\hat{a}_{NM}(s)$$

was used to approximate $a(s)$. Results are summarized in Tables 10 and 11. They are rather similar to those for Model 7 in Tables 5 and 11, but population estimates from addition of the mover variable are appreciably larger. The problem of the low estimated population of adult black males remains and is not explained by the size of the standard error. For the nonblack population, estimates appear to be rather consistent with those from demographic analysis.

Table 9. Parameter Estimates for Mover Status for Model 8

Independent variable	Estimate
Constant	-3.44815
Age 0–9	0.99238
Age 10–19	1.00809
Age 20–29	1.54508
Age 30–44	0.78369
Age 45–64	0.37314
Female	-0.02554
Black	-0.33392
Hispanic	-0.24578
Renter	1.13239
Metropolitan	-0.31035
Unrelated to householder	0.58948
North	-0.01815
South	-0.05505
Midwest	-0.06767
Marital status	-0.11058
Block nonresponse rate	-0.00729
Log block size	-0.01197

Clustering effects

As noted in Section 7.2, estimates obtained in this study show quite substantial deviations from random sampling. This phenomenon appears to involve two basic issues. The weights vary substantially, and clustering is present at the block and the household level.

The weights for the *P* sample have a coefficient of variation of 1.487. If random sampling were present, weights were used, and weights were independent of all other variables, then this result suggests that variances of estimates would be multiplied by about

$$(1.487)^2 + 1 = 3.21.$$

Similarly, in the case of the *E* sample, the coefficient of variation of the weights is 1.497, and the suggested variance multiplication is 3.24.

Table 10. Estimated Population Totals in Thousands by Sex, Age, and Race for Model 8

Sex	Age	Race	Census total	Est. total	Std. error	Demog. anal.
Male	0–9	Nonblack	16,000	16,710	115	16,400
Female	0–9	Nonblack	15,200	15,630	187	15,600
Male	10–19	Nonblack	15,000	15,310	178	14,900
Female	10–19	Nonblack	14,300	14,410	85	14,300
Male	20–29	Nonblack	17,200	18,370	142	17,600
Female	20–29	Nonblack	17,100	18,050	114	17,200
Male	30–44	Nonblack	25,700	26,450	122	26,500
Female	30–44	Nonblack	26,100	26,730	112	26,100
Male	45–64+	Nonblack	20,000	20,210	39	20,500
Female	45–64	Nonblack	21,400	21,540	106	21,400
Male	65+	Nonblack	11,100	11,080	95	11,300
Female	65+	Nonblack	15,900	15,870	532	16,000
Male	0–9	Black	2,820	3,183	60	3,100
Female	0–9	Black	2,760	2,973	27	3,020
Male	10–19	Black	2,600	2,845	36	2,670
Female	10–19	Black	2,620	2,778	24	2,680
Male	20–29	Black	2,280	2,492	98	2,570
Female	20–29	Black	2,740	3,108	75	2,870
Male	30–44	Black	3,000	3,205	59	3,430
Female	30–44	Black	3,690	3,932	82	3,780
Male	45–64	Black	2,010	2,091	56	2,280
Female	45–64	Black	2,540	2,682	72	2,550
Male	65+	Black	920	939	31	960
Female	65+	Black	1,470	1,506	43	1,450

Table 11. Estimated Sex Ratios by Age and Race for Model 8

Age	Race	Census	Estimate	Std. error	Demog. analysis
0–9	Nonblack	1.053	1.069	.009	1.051
10–19	Nonblack	1.046	1.062	.013	1.043
20–29	Nonblack	1.006	1.018	.010	1.019
30–44	Nonblack	.987	.990	.003	1.014
45–64+	Nonblack	.936	.938	.005	.957
65+	Nonblack	.699	.698	.018	.707
0–9	Black	1.023	1.070	.016	1.027
10–19	Black	.994	1.024	.013	.994
20–29	Black	.831	.802	.016	.896
30–44	Black	.807	.815	.006	.907
45–64	Black	.790	.780	.004	.893
65+	Black	.628	.624	.006	.661

There is clustering at the household level and there is clustering at the block levels. These phenomena were investigated by use of generalized residuals computed for each block and each household. The basic computation can be found in the discussion of Haberman (1978, Ch. 5) concerning generalized residuals, although simplifying approximations can be made for the case under study due to the small size of the households and blocks relative to the total sample. For a given block b in the *P* sample, one computes the following quantities:

12. The number P_b of individuals s in the *P* sample who are in block b and who are correctly enumerated.
13. The estimated expected number $\hat{P}_b$ of individuals in the *P* sample who are in block b and who are correctly enumerated. This estimate is the sum of $1-\hat{f}_U(s)$ for each member s of the *P* sample who is in block b.
14. The sum $\hat{\sigma}^2_{bP}$ of the variance estimates $\hat{f}_U(s)|1-\hat{f}_U(s)|$ for individuals s in the *P* sample who are in block b.
15. The approximate generalized residual

$$z_{bP} = (P_b - \hat{P}_b) / \hat{\sigma}_{bP}.$$

The generalized residual z_{bP} has an approximate standard normal distribution under random sampling if the selected model is correct. The sum of squares of the z_{bP} was about 6 times larger than expected under simple random sample for a correct model. Given the relatively small changes in entropy measures achieved in Section 7.2 when models with many parameters were used instead of the models with 11 or

12 variables, model error cannot be expected to account for a substantial fraction of this increase in the size of the sum of squares. In addition, the distribution of the z_{bP} is very negatively skewed (skewness -6.48) and long-tailed (kurtosis 66.17). On the other hand, the interquartile range is only 1.47. Some 11.95 percent of the z_{bP} are less than −2 and 4.50 percent exceed 2.

Similar calculations were made for the *E* sample. For a given block b in the *E* sample, we computed the following quantities:

16. The number E_b of individuals s in the *E* sample who are in block b and who are correctly enumerated.
17. The estimated expected number $\hat{E}_b$ of individuals in the *E* sample who are in block b and who are correctly enumerated. This estimate is the sum of $1-\hat{f}_{\mathrm{EE}}(s)$ for each member s of the *E* sample who is in block b.
18. The sum $\hat{\sigma}^2_{bE}$ of the variance estimates $\hat{f}_{\mathrm{EE}}(s)|1-\hat{f}_{\mathrm{EE}}(s)|$ for individuals s in the *E* sample who are in block b.
19. The approximate generalized residual

$$z_{bE} = (E_b - \hat{E}_b) / \hat{\sigma}_{bE}.$$

The generalized residual z_{bE} has an approximate standard normal distribution under random sampling if the selected model is correct. The sum of squares of the z_{bE} was about 4.3 times larger than expected under simple random sample for a correct model. In addition, the distribution of the z_{bE} is very negatively skewed (skewness -4.47) and long-tailed (kurtosis 48.49). On the other hand, the interquartile range is only 1.54. Some 11.27 percent of the z_{bE} are less than 2 and 3.77 percent exceed 2.

Generalized residuals may also be computed for individual households and sums of squares may be obtained from the generalized residuals. Despite weaknesses in normal approximations due to the sizes of households, the sums of squares should be comparable to the number of households for both the *P* and *E* samples in the case of random sampling. In fact, the sum of squares for the *P* sample is 2.07 times the number of households, and the sum of squares for the *E* sample is 1.98 times the number of households. The apparent clustering of misses and of erroneous enumerations within households and within blocks has implications for estimation of undercount generally and at the small-area level in particular. First, the clustering may indicate the existence of omitted variables at the block or household level. Although block-level variables were considered, one omitted variable possibly relevant at the block level is an interviewer effect. Misses in the census occur both because a whole household is missed and because a household may be reported but some persons within the household are missed (and analogously for erroneous enumerations). It would be desirable to link the PES data from the HUCS (Housing Unit Coverage Studies) and the PES to allow two-stage modeling of misses and erroneous enumerations, along the lines of Cowan and Malec (1986). Second, the models for estimating undercount at the block level essentially assume person-to-person independence of enumeration status, which is inconsistent with the clustering observed. It is not clear how the block-level estimates of population size should account for the clustering other than, to the extent possible, including all relevant variables in the models.

Conclusions

The models for estimating undercount and erroneous enumerations fail to account for some heterogeneity, as evidenced by the improved estimation when mover status was taken into account (see Tables 3 and 5). Mover status cannot feasibly be ascertained for persons not in the PES, however, and therefore some heterogeneity may be inevitable. Improved modeling of heterogeneity might be attainable if the PES design were more integrated, so that a merged file of *P*-sample and *E*-sample persons could be constructed; however, the sample sizes needed for latent class analysis of heterogeneity are larger than occurred in the 1990 PES. Although the PES sample size for 1990 appears large at first glance, the comparison of sex ratios from the PES and demographic analysis indicate that heterogeneity may be a most significant issue for black males, and the sample size of black males is not large. Furthermore, the relevant sample number is not simply the number of black males but the number who were found in the *P* sample but not the census, and this number is only 5,368. Although the nominal size of the PES for Census 2000 is larger, it is not clear whether even that will be sufficiently large. (If separate models were to be fitted to each state's sample, the sample size would unquestionably be too small for latent-class analysis to be very successful.) Although the models we have developed are more satisfactory than the DSE in some respects, they are still deficient in the prediction of sex ratios when judged against demographic analysis. This result may reflect extreme heterogeneity (Section 2) or it may result from unknown problems with the data. For development of the simulated population, it may be appropriate to adjust the model using models similar to the FRR or FOR models in Bell (1993).

The major accomplishments of our activity were (i) to account for some heterogeneity via the mover variable, (ii) to show how logit models could be fitted to *P*-sample and *E*-sample data separately rather than jointly (which is only possible under simplified assumptions), (iii) to improve understanding of clustering of census misses and erroneous enumerations, and (iv) to show that parsimonious logit models have better predictive power than the large poststratification model used by the Bureau of Census. This implies that logit models may be useful if direct state estimates are to be produced.

Appendix I: Definitions used in Table 1

Criteria for the categories of size of blocks:

Small Original *E*-sample block has 10 records or fewer (in the bottom quartile).

Medium Original *E*-sample block has more than 10 and not more than 60 records (in the middle two quartiles).

Large Original *E*-sample block has more than 60 records (in the top quartile).

There are 8840 blocks in the *E* sample and 8826 blocks in the *P* sample. In the tables, we take only the 8558 blocks that appear in both samples.

Total Total number of blocks of a certain size category

$\#e$ Number of records in set $A1$, where $A1$ is the set of *E*-sample records in the block which are correctly enumerated. Here $\#e$ is the weighted sum of the *E*-sample records in the block, weighted by the PCE, where PCE is variable 16 in wtdpos.dat (probability of correct enumeration).

$\#p$ Number of records in set $A2$, where $A2$ is the set of *P* -sample records in a certain block who match. Here $\#p$ is calculated by summing over all *P* -sample records weighted by PMATCH, where PMATCH is variable 23 in fullwt.dat (probability of match).

$\#e = \#p$ The number of blocks of a certain size category in which $\#e$ is equal to $\#p$

$\#e > \#p$ The number of blocks of a certain size category in which $\#e$ is larger than $\#p$

$\#e < \#p$ The number of blocks of a certain size category in which $\#e$ is smaller than $\#p$

$\#e - \#p$ $\#e$ minus $\#p$ in a block

$\text{av}(\#e - \#p)$ The average of $\#e - \#p$ over all blocks of a certain size category

$\text{sd}(\#e - \#p)$ The standard deviation of $\#e - \#p$ over all blocks of a certain size category av(|#*e* − #*p*|) The average of |#*e* − #*p*| over all blocks of a certain size category

Appendix II: Definitions used in Table 2

Criteria for the categories of size of blocks

Small Original E-sample block has 10 records or fewer (in the bottom quartile).

Medium Original E-sample block has more than 10 and not more than 60 records (in the middle two quartiles).

Large Original E-sample block has more than 60 records (in the top quartile).

There are 8840 blocks in the *E* sample and 8826 blocks in the *P* sample. In the above tables, we take only the 8558 blocks that appear in both samples.

Total Total number of blocks of a certain size category

#e Number of records in set A1, where A1 is the set of E-sample nonmovers who match within the block. A record in the E sample is counted in this set A1 if and only if (1) EMOV ER is 2 or missing and (2) FINMATCH is M, where FINMATCH is variable 43 in wtdpos.dat and EMOVER is variable 10 in EMOVER.DAT.

#p Number of records in set A2, where A2 is the set of P-sample nonmovers who match within a certain block. A record in the P sample is counted in this set A2 if and only if

20. FINSTAT is not equal 2, or 3, or 6,
21. ESAMP is 1,
22. FINMATCH is M,

where FINSTAT is variable 47 in fullwt.dat, FINMATCH is variable 46 in fullwt.dat, and ESAMP is variable 7 in P EFLAG.DAT.

#e = #p The number of blocks of a certain size category in which #e is equal to #p

#e > #p The number of blocks of a certain size category in which #e is larger than #p

#e < #p The number of blocks of a certain size category in which #e is smaller than #p

#e − #p #e minus #p in a block

av(#e − #p) The average of #e − #p over all blocks of a certain size category

sd(#e − #p) The standard deviation of #e − #p over all blocks of a certain size category

av(|#e − #p|) The average of |#e − #p| over all blocks of a certain size category

References

1. Akaike, H. (1974) "A New Look at the Statistical Identification Model." *IEEE Transactions on Automatic Control* 19 (1974): 716–23.
2. Alho, J. M., M. H. Mulry, K. Wurdeman, and J. Kim. (1993) "Estimating Heterogeneity in the Probabilities of Enumeration for Dual-System Estimation." *Journal of the American Statistical Association* 88 (1993): 1130–36.
3. Belin, T. R., G. J. Diffendal, S. Mack, D. B. Rubin, J. L. Schafer, and A. M. Zaslavsky. (1993) "Hierarchical Logistic Regression Models for Imputation of Unresolved Enumeration Status in Undercount Estimation." *Journal of the American Statistical Association* 88 (1993): 1149–59.
4. Belin, T. R. and J. Rolph. (1994) "Can We Reach Consensus on Census Adjustment?" *Statistical Science* 9 (1994): 486–508, with discussion 508–37.
5. Bell, W. R. (1993) "Using Information from Demographic Analysis in Post-Enumeration Survey Estimation." *Journal of the American Statistical Association* 88 (1993): 1106–18.
6. Breiman, L. (1994) "The 1991 Census Adjustment: Undercount or Bad Data?" *Statistical Science* 9 (1994): 458–75, 505–08, with discussion 497–537.
7. Cowan, C.D., and D. J. Malec. (1986) "Capture-Recapture Models When Both Sources Have Clustered Observations." *Journal of the American Statistical Association* 81 (1986): 347–53.
8. Darroch, J. N., S. E. Fienberg, G.F.V. Glonek, and B.M. Junker. (1993) "A Three-Sample Multiple-Recapture Approach to Census Population Estimation with Heterogeneous Catchability."*Journal of the American Statistical Association* 88 (1993): 1137–48.
9. Davis, M. P. Biemer, M. Mulry, and R. Parmer (1991) "The Matching Error Study for the 1990 Post-Enumeration Survey," in *Proceedings of the Section on Survey Research Methods, American Statistical Association*: 248–53.
10. Fay, R. E., and J. Thompson. (1993) "The 1990 Post-Enumeration Survey: Statistical Lessons, in Hindsight." *Proceedings of the 1993 Annual Research Conference*, U.S. Bureau of the Census.
11. Gilula, Z., and S. J. Haberman. (1994) "Models for Analyzing Categorical Panel Data," *Journal of the American Statistical Association*, 89 (1993): 645–56.
12. Gilula, Z., and S. J. Haberman. (1995) "Prediction Functions for Categorical Panel Data," *The Annals of Statistics* 23 (1995): 1130–42.
13. Haberman, S. J. (1978) *Analysis of Qualitative Data*, vol 1. New York: Academic Press.
14. Hengartner, N., and T. P. Speed. (1993), "Assessing Between-Block Heterogeneity Within the Post-Strata of the 1990 Post-Enumeration Survey." *Journal of the American Statistical Association*, 88 (1993): 1119–25.
15. Hogan, H. (1993). "The 1990 Post-Enumeration Survey: Operations and Results," *Journal of the American Statistical Association*, 88 (1993): 1047–60.
16. Kadane, J. B., M. M. Meyer, and J. W. Tukey. (1998) "Yule's Association Paradox and Ignored Stratum Heterogeneity in Capture-Recapture Studies." Technical Report No. 682, Department of

Statistics, Carnegie-Mellon University.

17. Kim, J., R. Blodgett, and A. Zaslavsky. (1991) "Evaluation of the Synthetic Assumption, 1990 Post-Enumeration Survey," in *Proceedings of the Survey Research Methods Section, American Statistical Association*: 254–59.
18. Mulry, M. H., and B. D. Spencer. (1993) "Accuracy of the 1990 Census and Undercount Adjustments," *Journal of the American Statistical Association* 88 (1993): 1080–91.
19. Robinson, J. G., B. Ahmed, B., P. Das Gupta, and K. A. Woodrow. (1993) "Estimation of Population Coverage in the 1990 United States Census Based on Demographic Analysis." *Journal of the American Statistical Association*, 88 (1993): 1061–71.
20. Savage, L. J. (1971) "Elicitation of Personal Probabilities and Expectations," *Journal of the American Statistical Association* 66 (1971): 783–801.
21. Schafer, J. (1993) "Comment." *Journal of the American Statistical Association* 88 (1993): 1125–27.
22. Sekar, C. C., and W. E. (1949) "On a Method of Estimating Birth and Death Rates andthe Extent of Registration." *Journal of the American Statistical Association* 44 (1949): 101–15.
23. Spencer, B. D. (1998) "Activity 6: Final Report," Prepared under Bureau of the Census contract number 50-YABC-2-66023 with National Opinion Research Center, April 3, 1998. Chicago: NORC.
24. Wachter, K., and D. Freedman. (1994) "Heterogeneity and Census Adjustment for the Intercensal Base," *Statistical Science* 9 (1994): 458–75. 505-08, with discussion 496–7, 508–37.
25. West, K., M. Mulry, R. Parmer and J. Petrik. (1991) "Address Reporting Error in the 1990 Post-Enumeration Survey," in *Proceedings of the Survey Research Section, American Statistical Association* 236–41.
26. Wolter, K. (1986) "Some Coverage Error Models for Census Data." *Journal of the American Statistical Association* 81 (1986): 338–46.
27. Zaslavsky, A. M., and G.S. Wolfgang. (1990) "Triple-System Modeling of Census, Post-Enumeration Survey, and Administrative List Data." in *Proceedings of the Survey Research Section, American Statistical Association*, 668–73.
28. Zaslavsky, A. M., and G.S. Wolfgang. (1993) "Triple System Modeling of Census, Post-Enumeration Survey, and Administrative List Data." *Journal of Business and Economic Statistics* 11 (1993): 279–88.