HANDLING NONINTERVIEWS TO PROVIDE EQUITABLE COMPARISONS OF ICM ESTIMATES

Rita Petroni, Anne Kearney, Philip Gbur, U.S. Bureau of the Census[1]
Rita Petroni, U.S. Bureau of the Census, Washington, DC 20233

1. The Problem and Our Research Approaches

To develop methodology to be used in Census 2000, the Census Bureau undertook a census test program in 1995. A major goal of the test program was to develop and test a new coverage measurement methodology, Integrated Coverage Measurement (ICM). In 1995, the goal of ICM was to measure the error in coverage (overcount or undercount) of the census test in three sites: Oakland, CA; Paterson, NJ; and six parishes in Northwest Louisiana. A parallel goal was to test CensusPlus and Dual System Estimation (DSE). (Thompson, 1994.) The Census Bureau tested these two methods in Oakland and Paterson.

For CensusPlus, after the regular census enumeration, the Bureau used additional enumeration techniques to enhance the count for a sample of block clusters (ICM blocks). Households of persons in the enhanced listing are in the R-sample (essentially the census enumeration plus any additional census missed persons). Using the ICM blocks, we produced two site-level poststratum estimates--one from the R-sample and one from the census count. We then multiplied the original census site-level poststratum estimates by the ratio of the R-sample estimate to the census estimate based on the ICM blocks to obtain the final poststratum estimates. Poststrata are defined by race/ethnicity, tenure, and age/sex cross-classifications. We summed poststratum level estimates to obtain total site-level population estimates and site-level population estimates of race/ethnicity by tenure, race/ethnicity, and tenure.

For DSE, after the regular census enumeration, the Bureau independently obtained a listing of people from a sample of block clusters. Households of persons in the listing were said to be in the P-sample. We matched the independent listing to the census enumerations to determine whether each person was captured in the regular enumeration, the independent listing, or both. We then applied estimation techniques assuming independence (Wolter, 1986) to obtain site-level poststratum ratios. The ratios are based on P-sample site-level poststratum estimates and site-level poststratum correction factors from census data from ICM blocks and from census data from all census blocks. We multiplied the original census site-level poststratum estimates by the ratios to obtain final poststratum estimates and summed poststratum level estimates to obtain the other site-level estimates. (Gbur, 1996; Hogan, 1993; Schindler,1996.)

For a given household, we collected data for both procedures simultaneously using one CAPI (computer assisted personal interviewing) instrument. The instrument was expected to do equally well in obtaining interviews for both procedures. The instrument contained census rosters which were to be displayed to the interviewer after collection of an independent household roster. The independence between the roster obtained during the ICM interview and the census roster is critical for the P-sample. However, interviewers could sometimes view the census roster before obtaining the independent household roster. Thus, since the independence of the initial roster was lost due to the design of the instrument and incorrect implementation of collection instructions, we had to treat such households as noninterviews for the P-sample, but not for the R-sample. For Oakland, the P-sample noninterview rate was 15.06%. The R-sample rate was 8.54%. For Paterson, the corresponding rates were 8.49% and 2.18%. (Ikeda and Petroni, 1996.)

The noninterview rate differences for the two samples were due largely to the use of the 1995 data collection instrument for a given household for both procedures. In practice, we would use the instrument to collect data from each household for only one method at a time or we would modify the instrument to overcome design and implementation flaws. Hence, we would expect the noninterview rates to be approximately equal. The large differences in noninterview rates could introduce bias into comparisons of the two estimates if the characteristics of the noninterviewed cases differ in the two samples. For both the P- and R-samples, we account for noninterviews with a weight adjustment. This approach implicitly assumes that the average noninterviewed housing unit is similar to the average interviewed housing unit with respect to the characteristic(s) being estimated at the level we calculated the adjustment. Deviation from this assumption may affect estimates or inferences based on the data. As a result, we mounted two research efforts to lessen the differences in noninterview rates between the two samples. First, we increased the interview rates for both samples through a field follow-up. Second, we treated P-sample households that were classified as noninterviews due to the instrument problem as noninterviews in the R-sample too. We then recomputed CensusPlus and DSE estimates and compared them to the original estimates. For Oakland, we conducted both research endeavors and for Paterson, only the second. This paper focuses on comparing the two research approaches in assessing the impact of noninterview differences in the R- and P-samples for Oakland. Basically, the comparison shows that lowering versus raising noninterview levels does result in significant differences in R-sample estimates and P-sample estimates,

but has no significant effect on comparisons of CensusPlus and DSE estimates.

Section 2 briefly describes the two approaches. Sections 3 and 4, respectively, analyze the results and draw conclusions.

2. Research Approaches

2.1. Method 1: Noninterview Follow-up (NIFU) Study

The ICM NIFU study consisted of the following steps: 1) revisit a sample of ICM noninterviews (as defined for estimation); 2) obtain completed interviews; 3) match the rosters obtained to the census rosters; and 4) recalculate CensusPlus and DSE estimates. We conducted the NIFU interview using a paper and pencil instrument with no census rosters available to the interviewers. Thus, the events which produced many of the P-sample noninterviews could not occur during the NIFU interview. The NIFU interviews were considered independent of the census even though some of the selected households may have had an initial ICM interview where the census roster was disclosed.

The universe for the NIFU operation was ICM housing units in the Oakland test site which were assigned a noninterview outcome in the original estimation processing for either the P- or R-sample. These units were either considered ineligible for the sample, were taken with certainty, or were sampled at a 50 percent rate. Those housing units in another post-production evaluation sample, which were refusals in production ICM interviewing, or for which an ICM quality assurance interview was conducted during production were ineligible for the sample. Housing units in areas with a large sampling weight, more than 25 eligible noninterviews, or only one eligible noninterview were selected with certainty.

Of the 1,227 noninterviews (717 were noninterviews for both P- and R-samples, 475 for P-sample only and 35 for R-sample only), 796 were eligible for the NIFU study. We selected a total of 239 with certainty and an additional 280 by sampling for a total NIFU workload of 519 housing units.

Interviewers attempted to obtain interviews for selected ICM noninterview households. They collected information on census day residents for the sample address including demographic data and residency information.

The Bureau made several modifications to the ICM production field procedures to increase the likelihood of obtaining completed interviews for the NIFU. Specifically: 1) a simpler version of the questionnaire was used; 2) the interviewers were instructed as to which specific minimal information must be collected for the interview to be called complete; 3) no last resort interviews were accepted; 4) telephone interviews were acceptable after an initial visit to the address; 5) acceptable proxy interviews could be obtained whenever available; 6) supervisors reviewed every completed questionnaire; and 7) experienced Census Bureau current survey interviewers were used. (Gbur, 1996.)

Table 1 shows that Method 1 (NIFU) R- and P-sample noninterview rates are much lower than the original R- and P-sample noninterview rates. Although the P-sample noninterview rate is still higher than the R-sample rate, the absolute levels and the differential were sufficiently reduced to justify an examination of the effects.

2.2. Method 2: Reweighting After Noninterview Status Change

For Method 2, we identified R-sample households which matched to P-sample households which had been classified as noninterviews due to instrument problems. When the matching R-sample household had been classified as an interview, we reclassified it as a noninterview and then recomputed the CensusPlus estimate. Most of the R-sample matched households had been classified as interviewed. (Kearney, 1996.) Table 1 shows that compared to the original R-sample noninterview rates, Method 2 R-sample rates are much closer to the original P-sample rates.

3. Analysis

In this section we analyze the effect that lowering or raising the noninterview rate had on the estimates. In section 3.1 we compare differences in R-sample estimates and differences in P-sample estimates when differences in the noninterview rates are about the same for R- and P-samples to show that a comparable change in noninterview rates affects both estimates similarly. In section 3.2 we examine the reasonableness of DSE original and Method 1 adjustment factors since these serve as benchmarks to evaluate CensusPlus adjustment factors. In section 3.3 we discuss the impact of bringing the noninterview rates to similar levels for comparing DSE and CensusPlus estimates. Tests for statistical significance were performed for all comparisons.

3.1. P- and R-Sample Comparisons When Differences in Noninterview Rates Are Similar

In Method 1 we used the NIFU results to lower the noninterview rates in the P- and the R-samples. In Method 2, we raised the R-sample noninterview rate much closer to the original P-sample noninterview rate. In this section we analyze the impact that changing from a similarly high noninterview rate to a similarly low noninterview rate had on the P- and R-sample estimates. This change had a similar effect on both sets of estimates. (See Table 2 and Graphs 1 and 2.) To aid this comparison we calculated the ratio of Method 1 to Method 2 estimates for the R-sample. For the P-sample we calculated the ratio of Method 1 to the original. Overall, for the R-sample, there are approximately four percentage points fewer people estimated with Method 1 compared to Method 2. The owner estimate is short by almost eight

percentage points and the renter estimate is about even. The results are similar for the P-sample when Method 1 estimates are compared to the original estimates. There are five percentage points fewer persons estimated with Method 1 than with the original. There are nearly eight percentage points fewer owners under Method 1 in the P-sample and about two percentage points fewer renters. The lower Method 1 estimates appear to be due to the additional households included in Method 1 having smaller than the average household size for their block cluster. (Gbur, 1996.)

For most race/ethnicity categories the ratio of Method 1 R-sample estimates to Method 2 estimates and the ratio of Method 1 P-sample estimates to the original estimates are numerically similar. Ratios for race/ethnicity by tenure were more similar for owners than for renters. For both the R- and P-samples, most estimates differed under the lower and higher noninterview rates. For the R-sample, the estimates differed for all race/ethnicity by tenure categories. For the P-sample, the ratios differed for all race/ethnicity by tenure categories except Black owners and Hispanic renters.

At the race/ethnicity by tenure by age/sex level the majority of the ratios are numerically similar. Again, most of the R- and P-sample estimates differed under the lower and higher noninterview rates.

Lowering the noninterview rates created many differences between the Method 1 and Method 2 R-sample estimates, and many differences between the original and Method 1 P-sample estimates, but lowering the noninterview rates affects the R-sample and P-sample estimates in a similar way.

3.2. DSE Original and Method 1 Adjustment Factors

Based on past capture/recapture experiences(Schindler, 1996), we expected to find an undercount for minorities, renters, males, and persons age 18-29. We observed the same patterns in the original and Method 1 DSE adjustment factors. For tenure and race/ethnicity categories, the adjustment factors appear reasonable. (See Table 3.) The total renters adjustment factor is significantly greater than the total owners adjustment factor. The renter adjustment factors are greater than the owner adjustment factors for all the overall race/ethnicity categories but these differences aren't significant. Except for Asian Pacific Islanders (API's) in Method 1, the overall adjustment factors for Blacks, Hispanics, and APIs are significantly greater than the adjustment factors for All Others. Graphs 3 and 4 show that the adjustment factors for males are generally greater than the adjustment factors for females but these differences are significant less than half the time. The adjustment factors for 18-29 year olds were expected to be greater than the adjustment factors for other age groups, but this is not always the case.

3.3. DSE and CensusPlus Comparisons

In Method 1 we revisited a sample of the noninterviews in the R-sample and the P-sample and decreased the noninterview rate in both samples bringing the noninterview rates closer. In Method 2 we increased the noninterview rate in the R-sample to bring it closer to the original noninterview rate in the P-sample. Whether we bring the noninterview rates closer via Method 1 or Method 2, we found little difference in the comparison of CensusPlus and DSE adjustment factors.

3.3.1. Original CensusPlus and DSE Comparisons

As section 3.2 showed, the original DSE performed approximately as expected. However, the original CensusPlus estimator did not perform this well. (See Table 3 and Graphs 3 and 4.) Most adjustment factors for Blacks are less than 1.0 indicating an overcount. In particular, adjustment factors for Blacks, Black renters, and Black renters in the 18-29 and 50+ age categories are significantly less than 1.0. The adjustment factors suggest renters have better coverage than owners about half the time, and results suggest males are covered better than females about half the time but these differences are not always significant.

3.3.2. CensusPlus and DSE Comparisons After Decreasing Noninterview Rates

We compared Method 1 DSE estimates with Method 1 CensusPlus estimates to compare the two estimation techniques when the noninterview rate was approximately equal and low. We found that decreasing the noninterview rates has little impact on the comparisons.

Table 3 and Graphs 3 and 4 show that the Method 1 DSE adjustment factors are generally greater than the Method 1 CensusPlus adjustment factors.

Compared to the original DSE and CensusPlus adjustment factors, the overall adjustment factors are still significantly different (1.108 for DSE and 0.978 for CensusPlus). In both DSE and CensusPlus the total renter adjustment factor is significantly greater than the total owner adjustment factor. Note, however, that the Method 1 CensusPlus adjustment factor is still numerically, although not significantly, below 1.0 for total renters. In the original CensusPlus the total renter adjustment factor is less than the total owner adjustment factor although not significantly. For Blacks, the DSE adjustment factors are usually significantly greater than 1.0 as expected while all but one of the Black adjustment factors for Method 1 are numerically although usually not significantly less than 1.0 (except for total Black and total Black renters) in CensusPlus. A similar trend is apparent with the API adjustment factors.

Traditionally, coverage is worse for males than for females, but this trend is not apparent with the adjustment factors. For DSE and CensusPlus the adjustment factors for males are greater than the adjustment factors for females in about half the race/ethnicity by tenure by age/sex categories but these differences aren't significant.

The DSE adjustment factors for males are almost always greater than 1.0 (significantly so for most Black and Hispanic tenure by age/sex categories) while the CensusPlus adjustment factors for males are numerically, but not significantly, less than 1.0 more than half the time. For both DSE and CensusPlus, the adjustment factors for 18-29 year olds are generally greater than the adjustment factors for the other age groups but again most differences are not significant.

3.3.3. DSE and CensusPlus Comparisons After Increasing Noninterview Rate in the R-Sample

In Table 3 and Graphs 3 and 4, a comparison of original DSE with Method 2 CensusPlus shows the effect of high approximately equal noninterview rates in both P- and R-samples. We found that increasing the R-sample noninterview rate has little impact on comparisons of DSE and CensusPlus estimates. For the most part the DSE adjustment factors are greater than the CensusPlus adjustment factors. We expect the overall renter adjustment factor to be greater than the overall owner adjustment factor as is the case with the original DSE. The Revised Method 2 CensusPlus adjustment factors for overall tenure show a smaller adjustment factor for renters than owners although not significantly smaller. We also expect minority adjustment factors to be greater than 1.0, as is generally the case for DSE. However, adjustment factors for CensusPlus Blacks, Black renters, and many Black renter by age/sex categories are less than 1.0. Adjustment factors for males are expected to be larger than adjustment factors for females, but this is not always the case. For example, for CensusPlus, Black owners age 18-29 and age 30-49 have larger adjustment factors for females than males, as do API renters in all age groups. These differences are not significant.

4. Conclusions

Lowering versus raising noninterview rates does result in significant differences in R-sample and P-sample estimates but both approaches had little affect on comparisons of CensusPlus and DSE adjustment factors. That is, the CensusPlus and DSE comparisons were not adversely influenced by differences in R-sample and P-sample noninterview rates in the 1995 test census.

REFERENCES

Gbur, P. (1996). Integrated Coverage Measurement (ICM) Evaluation Project 3: Noninterview Followup. 1995 Census Test Results Memorandum No. 44, March 14, 1996, Bureau of the Census.

Hogan, H. (1993). The 1990 Post-Enumeration Survey: Operations and Results. Journal of the American Statistical Association, 88, 1047-1060.

Ikeda, M. and Petroni, R. (1996). Handling of Missing Data in the 1995 Integrated Coverage Measurement Sample. To be presented at the 1996 annual meeting of the American Statistical Association, Chicago, Illinois.

Kearney, A. (1996). 1995 ICM Test: Effect of Different R- and P-Sample Nonresponse Rates on the Comparison of CensusPlus and Dual System Estimation Estimates. Internal Memorandum from Singh for the Record, February 12, 1996, Bureau of the Census.

Schindler, E. (1996). Review and Evaluation of ICM Estimates for the 1995 Census Test. 1995 Census Test Memorandum Series--IS#18, February 22, 1996, Bureau of the Census.

Thompson, J. (1994). 1995 Test Census Design. Presented at the 1994 annual meeting of the American Statistical Association, Toronto, Canada.

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

ACKNOWLEDGMENTS: The authors wish to thank William Bell and Gregg Diffendal for reviewing our paper and Lisa Mundy and Alice Bell for preparing this manuscript for publication.

[1] This paper reports results of research undertaken by Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.

Key for graphs:

There are six age-sex categories numbered 1 through 6 on the graphs. The numbers represent the following age-sex categories:

1. 18-29 Male
2. 18-29 Female
3. 30-49 Male
4. 30-49 Female
5. 50+ Male
6. 50+ Female.

There are four race/ethnicity categories differentiated by the four sets of age-sex categories. The first group of six corresponds to Black, the second group to Non-Black/Non-API Hispanic, the third group to API, and the fourth group to All Other.

Table 1. Oakland Noninterview Rates

Estimate	R-Sample	P-Sample
Original	8.54	15.06
Method 1	2.38	5.58
Method 2	13.70	--

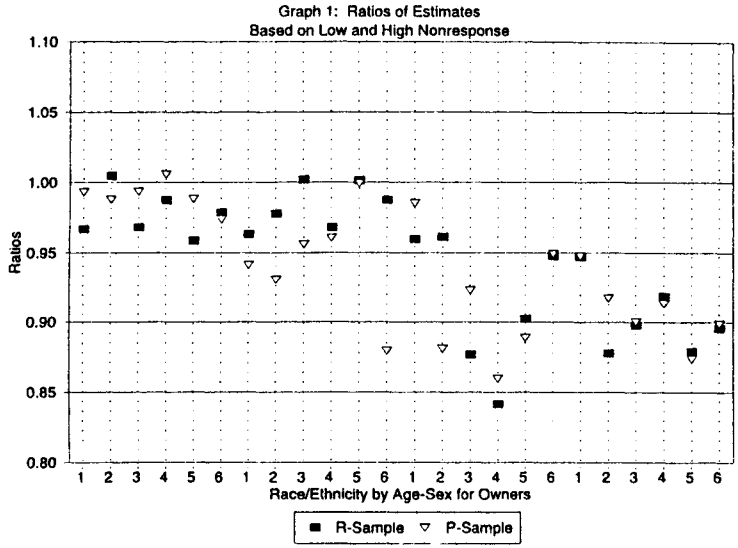
Table 2. R-and P-Sample Original and Revised Estimates for Oakland

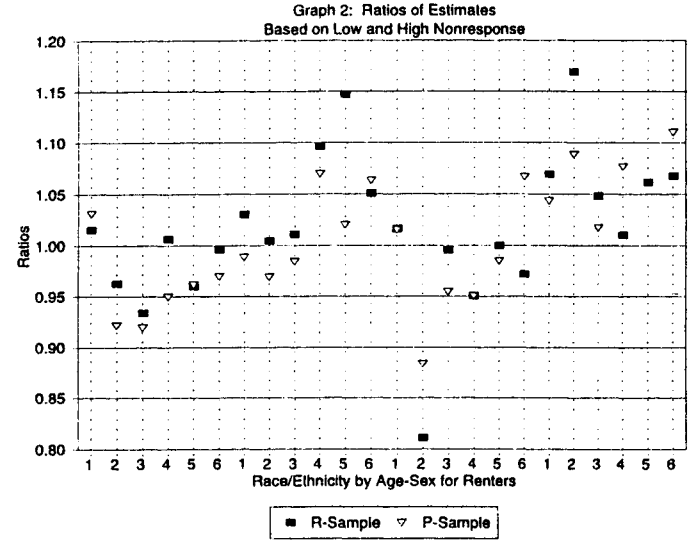
Characteristics	R-Sample				P-Sample		
	Original (1)	Method 1 (2)	Method 2 (3)	(2) / (3) (4)	Original (5)	Method 1 (6)	(6) / (5) (7)
Total	345192	333493	347284	0.960	262010	248918	0.950
Owner	173930	162134	175324	0.925	146730	135472	0.923
Renter	171262	171359	171960	0.997	115280	113446	0.984
Black	127019	124907	128091	0.975	91291	87820	0.962
Owner	51975	50938	51882	0.982	39856	39566	0.993
Renter	75044	73969	76209	0.971	51435	48254	0.938
Hispanic	54992	56559	55724	1.015	36940	36445	0.987
Owner	19543	19672	20079	0.980	15059	14486	0.962
Renter	35449	36887	35645	1.035	21880	21959	1.004
API	53066	48244	52663	0.916	39628	36758	0.928
Owner	26900	23980	26842	0.893	22379	20147	0.900
Renter	26166	24264	25821	0.940	17249	16611	0.963
All Others	110114	103783	110805	0.937	94151	87895	0.934
Owner	75512	67544	76521	0.883	69436	61273	0.882
Renter	34602	36239	34284	1.057	24715	26621	1.077

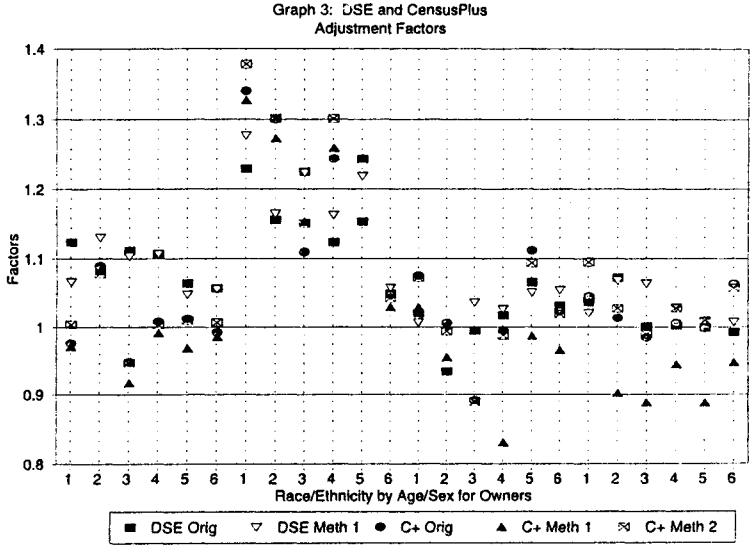
Table 3. DSE and CensusPlus Adjustment Factors* for Oakland

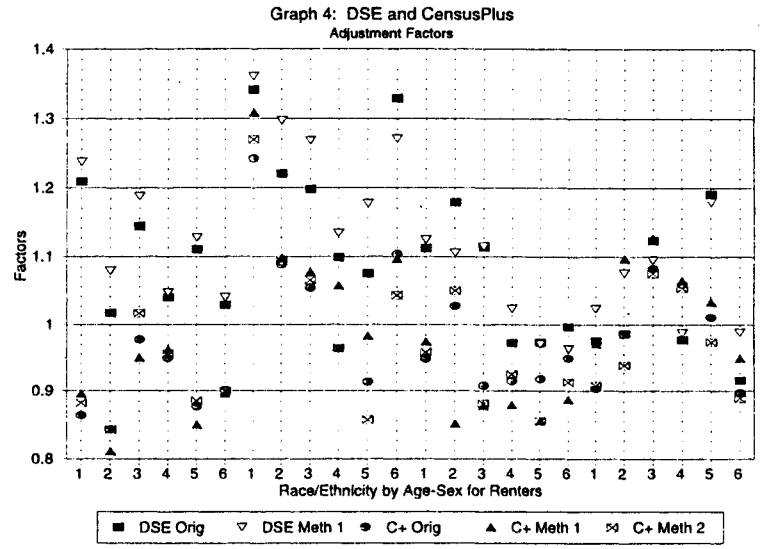
Characteristics	DSE		Census Plus		
	Original	Method 1	Original	Method 1	Method 2
Total	1.087	1.108	1.005	0.978	1.010
Owner	1.060	1.074	1.026	0.964	1.034
Renter	1.107	1.135	0.989	0.988	0.992
Black	1.105	1.121	0.950	0.934	0.958
Owner	1.097	1.091	0.987	0.968	0.985
Renter	1.109	1.138	0.928	0.915	0.942
Hispanic	1.203	1.245	1.132	1.163	1.148
Owner	1.176	1.203	1.212	1.220	1.246
Renter	1.219	1.269	1.086	1.130	1.091
API	1.075	1.083	0.976	0.891	0.967
Owner	1.034	1.046	0.981	0.877	0.978
Renter	1.104	1.109	0.972	0.900	0.958
All Others	1.007	1.035	1.039	0.998	1.042
Owner	0.999	1.027	1.021	0.915	1.035
Renter	1.019	1.046	1.064	1.114	1.053

* We use the term adjustment factor loosely. Technically the term implies a factor that would be applied to a census count to produce an estimate. In this sense, the only adjustments are at the poststratum level. That is, technically the above are not adjustment factors, but these factors minus one represent coverage rates.









*See key for graphs at end of text.