

DISCUSSION

Ansley J. Coale, Office of Population Research, Princeton University

The basic problem of evaluating the coverage of a population census is one of finding or generating reliable independent data against which the census can be checked. There are two categories of comparison which may be applied:

1. Aggregate comparisons
2. Individual or name-by-name comparisons

In an aggregate comparison an independent source is found that purports to cover the whole of the population or the whole of some portion that is delineated within the census. If the independent source can be accepted as correct the aggregate comparison indicates how fully the census has covered the whole population or the delineated sub-group. If the independent source cannot be accepted as standard the comparison shows the relative completeness of the two.

Three examples of aggregate comparison are:

1. The population under 25 can be compared with estimated survivors of the registered births from 1935 to 1960, adjusted for under-registration of births and net immigration. Such a comparison has been carried out by Akers, of the Census Bureau. The results show an estimated undercount of persons under 25 of 2.8 per cent (white males 2.6 per cent, white females 1.6 per cent, non-white males 8.3 per cent, non-white females 6.2 per cent)
2. In 1940 selective service registration of males of military age could be compared with the census enumeration of this group from six months earlier with due allowance for mortality and differences in coverage. The results showed a 4.5 per cent undercount of white males and an 18 per cent undercount of non-white males at ages 21-35.
3. The population in a given census can be compared with an aggregate obtained by updating adjusted earlier census records. Dr. Melvin Zelnik and I have recently completed such a comparison, which is fully reported in a book now in the hands of the publishers. The results show an undercount for white males of 2.6 per cent and for white females of 1.6 per cent.

The independent data utilized in these examples, especially the first, are of fairly convincing reliability, but it remains less than certain that they provide a wholly trustworthy independent determination of the number of persons.

In an individual or name-by-name comparison the independent source again covers the whole or some delineated portion of the population. In this case the comparison is not one of total figures but is an individual match of per-

sons in the two sources. The advantage of such an individual comparison is that even an incomplete source can indicate the completeness of coverage of the census provided that the chance of omission from the source is independent from the chance of omission in the census. Thus if one had a truly independent list of 10,000 persons who should have been covered by and census and found that 9,700 individuals had been so covered he could reasonably conclude that the census was 97 per cent complete, even though the 10,000 names were not a complete coverage of any well defined sub-group. Examples of individual comparisons are:

1. A comparison of registered births in the three months prior to the census with the enumeration of children no more than three months of age in 1950. This comparison indicated that about 4 per cent of these babies had been omitted.
2. Re-enumerative sample surveys. The result in 1960, reported to us today, indicates a 1.6 per cent net omission.
3. Reverse record checks where the independent source is a list of persons constructed from birth certificates, persons enumerated in earlier censuses, immigration records and the like. Before such a list can be utilized, however, it is necessary, outside of the census records, to locate the 1960 address of each person on the list.

Individual comparisons provide a valid estimate of census coverage only in the absence of correlation between the chance of omission from the census and the independent source. Unfortunately in re-enumerative surveys the correlation is probably very high rather than nearly zero. The chances are very strong that a person omitted from a more intensive re-enumerative survey, for example, persons with no usual place of residence or persons whom the respondent is omitting because of possible difficulties with the law, will be missed in both the census and the survey.

In summary, it is very difficult to achieve complete coverage in a census or to improve the 97-98 per cent completeness already attained. It is also very difficult to determine precisely the extent of coverage when it is incomplete. As the chairman has remarked, the Bureau of the Census is constantly experimenting to improve coverage and, as the papers under discussion demonstrate, is pioneering in the evaluation of its own success. Nonetheless, I would say that as of the moment we do not yet know how complete the 1960 census count was. If we do find out when the evaluation program is completed, my private guess is an undercount of about 3 per cent, compared to some 3.5 per cent in 1950.