RESEARCH REPORT SERIES
(Survey Methodology #1951-01)

Measuring the Accuracy of Enumerative Surveys

A. Ross Eckler
Leon Pritzker

Center for Behavioral Science Methods
Research and Methodology Directorate
U.S. Census Bureau
Washington, D.C. 20233

Report issued: February, 2021

Disclaimer: This report is released to inform interested parties of research and to encourage discussion of work in progress. Any views expressed are those of the authors and not those of the U.S. Census Bureau.

Abstract

One of the most important statistical developments in recent years has been the great increase in the amount of attention paid to measuring the amount of error present in data collected in censuses and sample surveys. One reason for this increased emphasis is undoubtedly the great expansion in the number and types of sample surveys, which are usually designed with reference to the degree of precision required. In planning such surveys, the problem is to measure the total error of a statistical figure including both sampling and non-sampling errors. Important pioneering work in this field has been done by Mahalanobis and, more recently, by workers in France, Great Britain and the United States. As a result, more and more users of statistics have come to realize that certain sources of error exist, whether the survey is based upon a small sample or covers the entire universe. We have learned indeed that in certain cases the most accurate measure of a particular item may be obtained through a relatively small sample survey where it is economically feasible to put great emphasis upon reduction of the non-sampling error through careful training and follow-up. The decrease in non-sampling error thus achieved will often more than counterbalance the increase in sampling error. This paper, then, has the objective of calling attention to the techniques of measuring non-sampling error employed by the Bureau of the Census and to some of the results obtained by the use of these techniques. We shall deal with two general lines of approach: 1) the study of statistical totals and 2) the study of response errors.

Keywords: non-sampling error, precision, surveys, total survey error

Suggested Citation: A. Ross Eckler**,** Leon Pritzker**. (1951). Measuring the Accuracy of Enumerative Surveys.** *Research and Methodology Directorate, Center for Behavioral Science Methods Research Report Series (Survey Methodology #1951-01).* U.S. Census Bureau. Available online at <https://www2.census.gov/adrm/CBSM/sm1951-01.pdf>

DEPARTMENT OF COMMERCE
Bureau of the Census
Washington 25

MEASURING THE ACCURACY OF ENUMERATIVE SURVEYS

by

A. Ross Eckler and Leon Pritzker, Bureau of the Census
Department of Commerce, Washington, D.C., U.S.A.

(As presented by A. Ross Eckler, Deputy Director, Bureau of the Census, at the 27th Session of the International Statistical Institute, held at New Delhi, India, December 5, 1951.)

MEASURING THE ACCURACY OF ENUMERATIVE SURVEYS

One of the most important statistical developments in recent years has been the great increase in the amount of attention paid to measuring the amount of error present in data collected in censuses and sample surveys. One reason for this increased emphasis is undoubtedly the great expansion in the number and types of sample surveys, which are usually designed with reference to the degree of precision required. In planning such surveys, the problem is to measure the total error of a statistical figure including both sampling and non-sampling errors. Important pioneering work in this field has been done by Mahalanobis and, more recently, by workers in France, Great Britain and the United States. As a result, more and more users of statistics have come to realize that certain sources of error exist, whether the survey is based upon a small sample or covers the entire universe. We have learned indeed that in certain cases the most accurate measure of a particular item may be obtained through a relatively small sample survey where it is economically feasible to put great emphasis upon reduction of the non-sampling error through careful training and follow-up. The decrease in non-sampling error thus achieved will often more than counterbalance the increase in sampling error.

A second reason for the increased emphasis upon measuring the accuracy of statistics is to be found in the fact that new and powerful tools are provided for this purpose through recent developments in the sampling field. Except for these developments, the cost of any adequate check of the quality of a survey would still be prohibitively great.

The present policy of the United States Bureau of the Census is to provide as consistently as possible a measure of the accuracy of all censuses and surveys which it conducts. It is our belief that we have a continuing responsibility to inform the public concerning the quality of statistics available from the expenditure of appropriated funds. It is perhaps well to note briefly the types of errors which we hope to measure in addition to those arising from sampling; 1) errors of coverage, which may involve either the failure to identify and count units which should have been included or the erroneous inclusion of units not properly a part of the census or survey; 2) content errors, such as those arising in the reporting of income, sales of farm products or work histories, and 3) processing errors.

We believe that a reasonably satisfactory measure of the total amount of non-sampling error from the elements just noted, serves the following purposes: 1) It helps to make our statistics more effective tools for necessary decisions on the part of government, business and the public in general. 2) It contributes to educating the public to be more discriminating in the selection and use of statistical information. 3) Public confidence in, and support for, our large scale statistical programs will increase as a result of a straight forward approach to the limitations of the data.

This paper, then, has the objective of calling attention to the techniques of measuring nonsampling error employed by the Bureau of the Census and to some of the results obtained by the use of these techniques. We shall deal with two general lines of approach: 1) the study of statistical totals and 2) the study of response errors.

(more)

Methods of studying the accuracy of a statistical total. Let us first consider some specific techniques -- and their applications to census statistics - for studying the accuracy of a statistical total directly. First, the method of comparison with another figure purporting to represent the same thing is illustrated by comparisons of our 1940 Census statistics on number of males with Selective Service Registration statistics 1/. Since compulsory registration for Selective Service was required in the United States in October, 1940, approximately six months after the 1940 Census of Population, Selective Service figures were used to check on the 1940 Census.

In order to make the comparison, minor adjustments were necessary to allow for "aging" and mortality, and to account for persons already in service at the time of Registration and for inmates of institutions. After making these adjustments, we found the Census figures to be approximately 5.5 percent below Selective Service registrations for all males 21 to 35 and approximately 15 percent below Selective Service for Negro males aged 21 to 35.

The 5.5 percent difference between Census and Selective Service on total males age 21 to 35, while not conclusive, in view of certain points still to be brought out, does strongly suggest that there was probably some under-enumeration in this age group. Young, adult males are, in general, the most mobile group in our population and are, in consequence, one of the groups most difficult to enumerate.

The presumed 15 percent under-reporting of Negro males is a much more striking phenomenon. If this discrepancy is real, it would seriously impair the validity of many important statistics, e.g., the published death rates for Negro males in the United States. There are, however, certain factors which partially invalidate the comparison of Census and Selective Service figures. The most important of these factors is the strong evidence of over reporting in the Selective Service figures. Over-reporting in Selective Service figures stems largely from their origin with autonomous local boards which were not necessarily uniform in their application of statistical definitions and procedures. As noted by one analyst (Myers), "Each local board was very desirous of obtaining as many registrants as possible 'to show what a good job it was doing' so that, rather than accepting the changes of address new registrations may have been accepted in many instances." Myers also cites Selective Service figures showing gross inaccuracy in reporting of race, for example, 50,000 Negro registrants in Hawaii as against no Negro registrants in Puerto Rico, whereas Census data showed a negligible Negro population in Hawaii and a very large proportion of Negroes in Puerto Rico.

We may be sure, however, that errors exist in both sets of records. It is precisely this fact that complicates the interpretation of gross total comparisons. Without case-by-case examination and reconciliation, it is difficult to determine the source of error or even its real magnitude.

1/ a. Price, Daniel O. A check on under-enumeration in the 1940 Census. American Sociological Review 12 (i): 44-49. February, 1947.
b. Myers, Robert J. Under-enumeration in the census as indicated by Selective Service Data. OP. cit. 13 (3): 320-325. June, 1948.
c. U. S. Bureau of the Census. Factors in the under-enumeration of males of military age. Unpublished manuscript.

(more)

A second technique for direct measurement of the accuracy of statistical totals is the inflow-outflow technique developed by the demographers. One application of this technique has been in the assessment of the under-enumeration of children under five in the 1940 Census. The technique has been described as follows:

> "a) Record the number of births registered during the five years ending on the Census date, b) record the number of deaths of these children which occurred prior to the census date, c) correct the data in "a" and "b" for incomplete reporting, d) deduct the adjusted deaths from the adjusted births, and e) compare the remainder with the number of children under five enumerated in the Census." 1/

The use of this technique led to the following estimates of under-enumeration of children under five for the United States as a whole in 1940:

TOTAL	7.6 percent
White	6.4
Nonwhite	15.2
Urban	5.0
Rural Nonfarm	5.8
Rural Farm	13.2

The use of the inflow-outflow technique is at best subject to considerable risk. Where use of this technique gives results consistent with the Census being investigated, one may feel somewhat more confident of the results of that census. Where they cannot be reconciled, one is faced with the problem of deciding whether the error lies in the Census or in the interim statistics of births, deaths and migration.

A third technique for direct measurement of the accuracy of data is illustrated by the use of expected age distributions. Biases in age reporting have been detected by comparing the distributions obtained in a Census with expected distributions based on survivors from a previous census. A discussion of age errors in the 1930 Census 2/ indicates the following types of error:

> a) under reporting of children under five, b) over-reporting of age 21 by men (but not by women), and c) over-reporting of age by the older age groups, particularly nonwhites.

1/ Whelpton, P.K. Forecasts of the population of the United States 1945-1975 U. S. Bureau of the Census, Washington G.P.O. - 1947 p. 34.

2/ Myers, R.J. Errors and bias in the reporting of ages in Census data. Transactions of the Actuarial Society of America, Vol. XLI, Part 2, No. 104.

In making population forecasts based on the 1940 Census it was decided, on the basis of an examination of the age distributions, 1/ that the nonwhite population 55 years old and over as enumerated had to be redistributed. We increased the number aged 55--64 by 8 percent, decreased the number 65--74 by 14 percent and increased the number 75 and older by 7 percent.

This kind of technique, although it may indicate that all may not be well with the Census statistics is subject to risks similar to those outlined above. Also, in view of the fact that errors of reporting (content errors) cannot be separated from errors of enumeration (coverage errors), precise measurement of selective age biases becomes difficult.

Measuring accuracy by the study of response errors. Now, I should like to direct your attention to another type of approach to measuring accuracy through a case-by-case comparison of figures pertaining to individual persons or establishments. This approach is considerably more costly than the one we have been discussing. The two techniques that have been developed for this approach involve an examination of the "responses" obtained in enumerative surveys. They deal with communication errors in the sequence of steps from interviewing to processing and tabulation. Since they have been developed primarily to deal with the field phase of an enumerative survey, they have come to be known as techniques for the study of "response error" and, indeed, the general field of study of the accuracy of enumerative statistics has also come to be known as the "response error" field.

The first technique for measuring response errors is that of checks against the records available from some administrative public program. Usually the responses of a subsample of the subjects of a survey are compared on a case-by-case basis with records containing the same information. As part of the program for the measurement of accuracy of selected statistics to be published from our 1950 Censuses, checks against a number of records are being made for a small subsample of the United States population. (Of course, these record checks are being made for statistical purposes only.) Checks are being made against the following records, for the purposes indicated:

1. Birth certificates on file in our State Offices of Vital Statistics -- to check age reporting.
2. Records of military service on file in our Veterans Administration -- to check reporting on veteran status.
3. Tax returns on file in our Bureau of Internal Revenue offices -- to check income reporting.
4. Reports by employers on f ile in our Social Security Administration -- to check reporting on wage and salary income and reporting on industry.
5. Naturalization and alien records on file in our Immigration and Naturalization Service -- to check reporting on citizenship status.
6. Returns from the 1920 Census -- to check age and birthplace reporting for persons who were thirty years old or over in 1950. This check is being made to supplement the birth-certificate check because it is known that our birth certificate records for many States are decidedly incomplete for 1920 and earlier years.

No results are yet available from these record checks.

1/ Whelpton, P. K. Op. cit. P. 35

(more)

For the 1940 and 1950 Censuses we have conducted a kind of record check in reverse. This has taken the form of a "birth registration test." During the census enumerations, the enumerators were instructed to obtain additional data for infants born in the four-month period, December 1, 1939 to March 31, 1940 and in the first three months of 1950, respectively. The information obtained was used to match the census lists of infants with birth certificates on file with the State vital statistics offices.

The results of the match in 1940 indicated an over-all figure of 92.5 percent for completeness of registration. There were marked differentials by color and residence. Ninety-four percent of the White infants were registered, and 82 percent of the Nonwhite. In cities of 10,000 or more, about 97 percent were registered, while the percentage was 88 elsewhere. Except for a preliminary figure on total birth registration, no results are yet available from the 1950 test.

The test now under way will also apply to the other side of the coin -- a record check to study the completeness of reporting infants in the Census. The results from this test are not yet available.

Except for the birth registration test, our experience with record checks has been rather limited. The results of one rather small scale study may be cited as indicative of the problems encountered. A sample of 249 returns from an income survey covering the year 1946 was matched against wage and salary records of the Social Security Administration. Of these, 23.7 percent were ruled out of the study because they represented persons presumed not to be in the labor force. Another 25.3 percent were ruled out because they could not be found in the files, because no income was reported to both the Census Bureau and the Social Security Administration, or because the persons were employed in industries not covered by Social Security records. Of the 51.0 percent of the cases that ultimately were compared with Social Security Administration records, about 60 percent showed unreconcilable disagreements of $100 or more in wage reporting for the calendar year. Thus, we end with 60 percent of the matched cases showing discrepancies in income reporting. These cases represent, however, only 30 percent of the original sample.

This experience on matching income data, as well as what we are now learning from the record checks in the 1950 Censuses, indicates that the technique, while valuable, has the following limitations:

a. Incompleteness of the file of records. It is known, for example, that despite marked gains over the years, even today some births are not reported. In some other files, however, similar improvement has not taken place.

b. The file of records may cover only a certain class of the population. Thus, the Veterans Administration files do not contain the names of non-veterans. Hence, when a person cannot be located in the Veterans Administration files, there are three possible explanations: he may be a non-veteran; the file of veterans may be incomplete; or there may be an error in the matching process.

(more)

c. The matching process itself is often quite inconclusive. In addition to errors of filing, spelling and other clerical errors there are risks associated with making a decision as to what constitutes a "match" in a particular case. There are hundreds of John Smiths in New York City for example, and a considerable number born in the same year.

Thus, while the technique is useful in producing direct measurements on the accuracy of survey results for those cases where matching can be completed, the significant gaps left by unmatched cases seriously limit the inferences that can be drawn. The method may prove to have its greatest value in providing validation for the quality check approach, and it is for this purpose that we expect it to be of the greatest use.

The second technique for the measurement of response error, consisting of a "quality check" involving the recanvass of a sample of persons or establishments, is the technique which the U. S. Bureau of the Census has been attempting to develop in its programs for measuring the accuracy of its published statistics. As developed by the Bureau, this technique has the following characteristics:

a. It consists of a re-enumeration of a sample of the population and of the geographic area included in the original enumerative survey or census. Hence, we have come to call our quality checks Post-Enumeration Surveys.

b. The re-enumeration generally has two phases - a "coverage check" designed to measure the extent of under-counting and of over-counting of the units of enumeration (people, dwelling units, farms, business establishments) and a "content check" designed to measure the extent of erroneous reporting on such inquiries as age, number of rooms in the dwelling unit, cropland harvested, sales by a business establishment.

c. A number of steps are taken to try to make the "quality check" a valid measuring device of the error in the statistics obtained in the original enumeration:

(i) Generally the sampling design is a multi-stage one drawn with probability proportionate to the number of elements of the population being investigated - persons, farms, business establishments, etc. The ultimate units of investigation for coverage purposes are small areas or segments - ranging in size from part of an apartment building for a check on our Population Census to an entire small town for a check on our Business Census. In the sample design, the attempt has generally been made to draw the segments so that they contain an equal number of elements.

(ii) Special mapping materials are provided for the segments. A great deal of use is made of detailed maps showing each building in each block of a large number of cities and towns in the United States. In rural areas, we use county highway maps which show roads, streams, other ground features, and buildings in existence at the time the maps were prepared.

(more)

(iii) To conduct our quality checks, we attempt to select the best qualified persons in our field organization. For the quality check of our 1950 Census, for example, we used a special selection test as an aid in choosing 250 out of the 140,000 Census enumerators to do the job.

(iv) Special questionnaires are designed. A limitation is placed on the number of inquiries to be checked with the objective of providing a number of more analytical and probing questions for each inquiry that is checked.

(v) Special training materials are provided and the training period for enumerators is considerably longer than that for the original survey or census.

(vi) The re-enumeration is so designed as to provide for a considerably more intensive canvass of the area, a longer period of interviewing per case, and more elaborate supervision and control. Also, we have placed sharp restrictions on the respondents. In our censuses we cannot afford the call-back and other costs that would accrue from insisting on the best possible respondent for each case. In our quality checks we have put the following kinds of restrictions in an attempt to get better information:

(a) To obtain information on the characteristics of a person, we have insisted that the person himself, or, if a child or a person not competent to furnish information for some other reason, his parent or guardian be interviewed.

(b) To obtain information about a farm, we have required the operator be interviewed.

(c) To obtain sales information about a business establishment, we have required that the proprietor, manager or accountant be interviewed.

(vii) In some of our quality checks, the original census information is furnished to the interviewer. We have tried to furnish this information in such a way that the interviewer may obtain independent answers to his questions, after which he compares his answers with the original census data and tries to determine which answer is better in each case and to write an explanation of any discrepancy. As a check on possible biases of this technique, the interviewer is assigned some cases for which the original information is not furnished to him.

Thus, the postulate underlying our use of the quality check technique is that by a more elaborate design and a more painstaking execution of a post-enumeration survey than is possible from a cost standpoint for the original enumeration, we can obtain a measure of the accuracy of the statistics compiled from the original enumeration.

(more)

As was pointed out above, this postulate and the techniques that grew out of it derive from the sampling field. From this field, have come the idea of the measurement of error, the complex sampling designs that are needed for this kind of measurement and the beginnings of some mathematical models to provide a rational basis for measurement.

Work with this quality check technique has not been limited to the United States. We have been much interested in the work of Gabriel Chevry and his associates at the French National Institute of Statistics. The post-census coverage check which they employed after the 1946 French Census of industrial and commercial enterprises led them to take the courageous step of not publishing the results.

Review of Major Quality Checks undertaken by the U. S. Census Bureau. For the U. S. Bureau of the Census there has been a fairly uninterrupted line of development of this quality check technique, starting with the 1945 Census of Agriculture and continuing with the 1947 Census of Manufactures, the 1948 Census of Business, the Current Population Survey (from 1948 on), t he pretests for the 1950 Censuses of Population, Housing and Agriculture and the 1950 Censuses themselves.

For each quality check in turn then, I should like to outline briefly the Census or Survey which preceded it, what was done in the quality check, and some of its principal findings.

1) The check of the 1945 Census of Agriculture.

a. The Census. The Census itself was conducted early in 1945. Each enumerator was instructed to list every farm in the Enumeration District assigned to him, to account for all the land in the farms, to obtain production and inventory data for crops and livestock, to count the number of dwelling units on the farms and to obtain other information about the socio-economic status of the farms such as the extent of work off the farm by the operator and the tenure relationship. A "farm" was defined as a place having: either three acres or more with agricultural operations or $250 or more of agriculture products.

b. The check of the Census. The quality check began in September 1945 and continued through the next several months. A cross-section sample of 600 area segments, containing about 6,300 farms, throughout the United States was re-enumerated. These segments were completely recanvassed by picked enumerators of the Bureau's permanent field staff, who were given special training and supervision. The principal tasks of the check enumerators were to account for every acre of land in the segments and for every farm lying wholly or partly in a check segment. Each enumerator was required to construct a map of each of his segments in order to show this accounting. The enumerators were also instructed to check the count of farm dwelling units and to obtain a second report on non-farm work by farm operators. Since this was one of our earliest experiences with the quality check technique and since it

(more)

was not always possible to obtain an explanation for an apparent error in the original Census enumeration, the editing of the schedules obtained in the check was such as to give the original enumeration the benefit of any doubt.

The results of this check point up a finding common to most of our work in the study of Census errors: The errors tend to be selective; they appear more frequently among the marginal, or border-line classes of the populations (of factories, business establishments, farms, etc.) covered by our Censuses. Thus, in this quality check, we found border-line farms and non-productive land to account for most of the statistical error:

a) There was a net under-enumeration of 11 percent of the farms. 1/ This was the net resultant of 14 percent of the farms which were not counted, partly balanced by three percent consisting of places which should not have been counted (because they did not meet the Census definition, or because they were actually parts of farms or because they were enumerated twice.) However the farms omitted had little agricultural activity.

b) There was a net under-enumeration of 8-1/2 percent of the land in farms. This consisted of a gross under-reporting of nine percent and a gross over reporting of one-half of one percent. The under-enumeration of items of agricultural production, livestock and value of products although no measures were obtained, is believed to be much smaller than that of land in farms since the farm land not enumerated was primarily in small and marginal farms and included considerable areas of land not used for crops or pasture.

c) There was a 12 percent under-enumeration of occupied farm dwelling units, a situation most frequently encountered on the marginal places.

d) Where the farms were properly enumerated in the Census, 29 percent of the operators that should have been reported as working off the farm were not so reported. The reporting was poorest where the ~~number of days worked was~~ small.

2) The check of the 1947 Census of Manufactures.

a) The Census. The Census itself consisted primarily of an enumeration by mail. Mailing lists were made up from the 1939 Census and from listings of other Government agencies and of trade associations. The unit of enumeration was the plant or establishment with a definite single geographic location rather than the company or firm.

1/ For an excellent discussion of the problem of defining a farm, see L. E. Truesdell, "The Problem of Quality in Census Data," I. A. S. I., Estadistica, June 1951.

(more)

b) The check of the Census. This was primarily a coverage check designed to study the effect of the almost total reliance on a mail-list enumeration on the completeness of coverage. The check was not designed to study the industry classification or the reporting of sales, payrolls, etc. of establishments properly included in the Census. It was designed only to find omitted establishments and to obtain some measure of their economic significance. In order to find omitted establishments in a universe of such a small size relative to those of our other censuses, a sample of some 10,000 small areas in the United States was drawn. The sample was stratified so as to bring in a fairly large proportion of blocks having factories and smaller proportions of commercial and residential blocks and was supplemented in order to bring in rural areas.

These 10,000 small areas were canvassed by specially trained enumerators who were instructed in all cases to list and usually to conduct interviews not only in all factories, but in all commercial establishments. The basic question which was asked in these establishments was: "Is anything made, processed, fabricated or assembled at this location?" If a "yes" answer was obtained, further information was obtained relating to the employment, sales and classification of the establishment. When the returns from the field check were received in Washington, an extremely careful review was made to see if the establishment should be classified as manufacturing. If there was insufficient information, additional data were obtained by consulting the establishment again or consulting the records of other government agencies. In addition to the classification review, a very thorough check was made against the Census tabulations to determine completeness of coverage in the original canvass.

c) Results. The check showed the following:

i. Establishments omitted from the Census accounted for 1.8 percent of the manufacturing employment and 1.3 percent of the total wages and salaries included in the Census tables. These figures apply only to t he over-all totals for all manufacturing in the United States as a whole.

ii. There was an approximate under-coverage of 10 percent of the manufacturing establishments. This is a less important finding since classification becomes quite uncertain for the marginal cases.

iii. The extent to which the under-reporting is limited to the small and marginal plants is shown by the following:

(a) The omitted establishments averaged approximately 12 employees each compared with the census average of 60 employees.

(more)

(b) The incorporated establishments accounted for 90 percent of census employment but for only 60 percent of omitted employment.

(c) Industries on the border-line between manufacturing and non-manufacturing accounted for only 15 percent of Census employment, but for 35 percent of the omitted employment.

iv. Nearly half of the under-coverage in the Census consisted of establishments not included on the mailing lists. The remainder consisted of those for which reports were not received, or were received and misclassified as non-manufacturing, etc.

3) The check of the 1948 Census of Business.

a. The Census. The field listing and enumeration for the Census itself did not start on a major scale until the late Spring of 1949 and was not completed until the Winter of 1949. For this Census, the main reliance was placed on a field listing of establishments by enumerators. The enumerators were aided in completing their listings by the establishment records of the Social Security Administration. The enumerators had the following major responsibilities: i) to list practically every type of establishment of an industrial or commercial character (factories, retail, stores, banks, etc.) ii) to classify the listed establishments into four groups: retail, service, wholesale and not in the scope of the Census. iii) to leave the appropriate kind of report form with each "in-scope" establishment for completion and mailing. (Where it could be done rapidly, the enumerator completed the report).

b. The check of the Census. This quality check, which took place from November 1949 to January 1950, used a multi-stage area sample consisting of about 2,500 segments in about 500 counties throughout the United States. The sample was drawn with probability proportionate to number of establishments and was designed to include one in two hundred of the establishments within the scope of this census. To conduct the check, an interviewing staff was drawn from among the best Census enumerators and supervisors. To preserve the independence of the check, the location of the sample areas was withheld from the field staff until after the completion of the Census enumeration. Also, the rule was established that no check interviewer could work in an area for which he had been responsible in the Census. The check enumerators were supplied with the original Census listings for each Enumeration District in which a sample segment was located. Their job was to canvass each segment thoroughly, fill out a classification form for each establishment in the segment, whether listed in the Census or not, make a new, independent classification of each establishment and obtain the equivalent of a Census report form for each establishment regarded as within the scope of this census.

The enumerators were restricted as to whom they might interview in an attempt to get the best possible information regarding the 1948 activities of the establishments.

As with the 1947 Census of Manufactures check, there was a thorough review of the classification of omitted establishments and a careful check against the Census tabulations to make sure that these establishments were actually omitted. In addition, work is now underway on the "content check" side, the analysis of sales, employment and classification of establishments properly included in the Census.

An attempt is also being made to estimate the sales of establishments which were in business during only part of 1948 and which had no successors in business on December 31, 1948. Primarily because of the practical difficulties involved in obtaining data from "part-year operators" with no successors, the 1948 Census excluded them by design. It is hoped that the quality check will give an estimate of the effect of this decision on the reports of aggregate sales for 1948. The quality check was designed to do this by the provision that the check enumerators obtain what was in effect a short business history of each segment. They were required to list each establishment that was in business at a segment address at any time from January 1, 1948 to the date of the check. The type of investigation required for this kind of information would be manifestly impossible from a cost standpoint for the Census as a whole.

c. Results. Only preliminary results are available on the coverage aspect of the quality check. Because the results are generally similar for the three classes of establishments included in the Census of Business, figures are quoted only for retail trade. The following are estimates of under-coverage in the 1948 Census caused by the failure of enumerators to list retail establishments and by the time lag between the "Census date" (December 31, 1948) and the listing date: There was an under-coverage of 1.3 percent of the 1948 sales and 0.9 percent of the November, 1948 employment reported in the Census for all kinds of retail trade in the United States as a whole. This under-coverage resulted from the omission of 3.6 percent of the retail establishments in the United States. As in the case of the 1947 Census of Manufactures, the omitted establishments were considerably smaller than the average establishment tabulated in the Census. The average 1948 sales of omitted establishments were $26,000 as against $74,000 for the enumerated establishments; the average November, 1948 employment of the omitted units was 1.0 as against 3.9 for the enumerated establishments.

(more)

4. The checks of the Monthly Current Population Survey.

a. The Current Population Survey (CPS). This monthly survey conducted by the Bureau of the Census provides a current measure of population characteristics, with particular emphasis upon labor-force changes. As some of you know, the survey is based upon an area sample in 68 primary sampling units, usually consisting of one or two counties. Within each primary sampling unit, a sample of small areas (segments) is continually being drawn so that a particular segment is used in the survey for about a six month period. Prior to the use of a segment for monthly interviewing, the structures and dwelling units in it are listed. The survey interviewer uses the listing for conducting an initial interview in each household in the segment which determines the size and composition of the household as well as the labor force characteristics. In subsequent months, the interviewer tries to determine if there are any changes in the household in addition to obtaining the labor-force information and other facts relating to the population.

b. The check of the CPS. Results of a systematic coverage check of a subsample of the segments conducted from July 1948 to December 1950, are available. The checks have been conducted by the field supervisory staff and have been concerned solely with finding omissions and erroneous inclusions of persons and dwelling units as a consequence of error in the initial listing, or in the initial interview, or in that part of the succeeding interviews concerned with finding changes in household composition.

c. Results. The following figures summarize cumulative experience over thirty months.

i. There has been a net under-counting of 1.4 percent of persons, resulting from the omission of 2.9 percent and the erroneous inclusion of 1.5 percent of the total population.

ii. There has been a net under-counting of 1.8 percent of dwelling units, resulting from the omission of 3.9 percent and the erroneous inclusion of 2.1 percent of the total number of dwelling units.

iii. Errors once again appear to be concentrated in certain types of situations. The under-enumerated dwelling units have a smaller number of persons, on the average, than those included in the survey. One of the principal causes is that a substantial number of the missed dwelling units consist of new construction or conversions since the date of listing.

(more)

5. The check of the October, 1948 pretest of the 1950 Censuses.

a. The Pretest. The primary objective of the pretest was to study alternative techniques of enumeration: self enumeration as against direct enumeration; the use of a separate schedule for each household as against the use of a line type schedule covering many households. The pretest was conducted in four counties and in the city of Minneapolis.

b. The check of the Pretest. For this check, members of the Bureau's technical staff were assigned to do the field work. They were required to recanvass entire Enumeration Districts in a search for missed dwelling units and farms and to conduct interviews in a sample of enumerated dwelling units and farms in a study of content error. Over-enumeration was also studied by means of this latter sample.

c. Results. The results of the check are cited in order to indicate the orders of magnitude of errors in some population statistics. Much more precise indications will come from the check of the 1950 Censuses. Some of the findings of the check of the pretest are:

i. There was a net under-enumeration of 1.3 percent of the population in the sample areas.

ii. There was a net under-enumeration of 0.9 percent of the dwelling units. (The result seems to run counter to some of our experience with the monthly survey, which indicates that the proportion of under-enumerated dwelling units is higher than the proportion of missed persons, because missed dwelling units tend to be small.)

iii. The following estimates of percentages of cases with a discrepancy between original and check answers to the same inquiry may be of interest:

Previous year's residence	3.9 percent
School attendance	1.6
Highest grade of school completed	17.5

(The rather high discrepancy on highest grade of school completed pointed up weaknesses in our question on education and led to a revision of the wording of this inquiry on the final Census schedule)

(more)

6. The check of the 1950 Censuses. The field work for this, the most detailed and extensive of our quality check programs, took place in the Summer and Fall of 1950. The program had the objectives of measuring coverage error in the Population, Housing and Agriculture Censuses and of measuring content error for many of the important inquiries of these Censuses. For this purpose a sample of 25,000 households and 7,000 farms was drawn in about 500 counties in the United States. Fairly detailed questionnaires were drawn up and an intensive program of training and supervising the specially selected field staff was devised. Results and analyses of this Post-Enumeration Survey will be published in the next year or two.

In addition to the Post-Enumeration Survey, the program of assessing the accuracy of the 1950 Censuses included a match of the information for April 1950 as collected for identical households by the decennial census enumerators and by the Current Population Survey enumerators. The match has two aspects, a coverage comparison and a content comparison for identical persons. The coverage comparisons have not yet been completed, and only preliminary results are available from the content comparisons. These results, which may be subject to considerable re-interpretation in the light of the coverage comparison, for a sample of about 21,000 households indicate the following:

a. About 2.7 million more persons (4.3 percent) were reported in the labor force by the Current Population Survey than by the Census. This number was made up of 650,000 added to the unemployed total by the Current Population Survey; 930,000 added to agricultural employment and 1.1 million added to non-agricultural employment. (Some of these differences are accounted for by the 1.5 million persons for whom employment status was not reported in the Census.)

b. The differences between the Current Population Survey and the Census varied considerably among age-sex groups. Persons in the 14-17 year group (the entrance group for the labor force) showed the highest percentage discrepancy in the labor force membership category. Thirty percent more of these persons were classified in the labor force by the Current Population Survey than by the Census. The only other age group that showed a comparable discrepancy rate was the females, 65 years old and over: 26 percent more of persons in this group were classified in the labor force by the Current Population Survey than by the Census. The other age-sex group discrepancy rates range from about two percent to about 10 percent. Of course, these net differences were all considerably smaller than the gross differences, which were in part offsetting.

In addition to content comparisons involving the 21,000 matched households, information is available which compares the Current Population Survey statistics for April 1950 with a preliminary sample tabulation from the Census. These comparisons indicate,

(more)

generally, larger differences between the Census and the Current Population Survey. The direction of some of the differences is opposite to that reported for the match of identical households. Some explanation for this is found in the following:

a. The Census coverage of persons was probably somewhat better than the Current Population Survey coverage. (Although this is probably a general phenomenon, a sample can still be used to check the adequacy of coverage of a complete enumeration. The reason for this lies in the method of estimation. An independent estimate of the total population from a quality check sample might turn out to be poorer than the Census estimate. It is when the quality check is directed toward the estimate of errors in the original enumeration of the omissions and erroneous inclusions, that it becomes most useful.)

b. There is considerable sampling variation.

After this review of our work with the quality check techniques, it is appropriate to raise a question regarding the present stage of development of this approach. We have some evidence that indicates that the methods we have developed thus far can stand considerable improvement. The ideal of absolute accuracy for even a quality check is unattainable. We believe, however, that the technique as now developed is capable of furnishing data of high accuracy and can, therefore, be used to measure with satisfactory precision the errors in our surveys and censuses.

There are many developmental and policy problems yet to be solved. Some of them are:

a. How much money should be expended on a quality check program? How large do the samples have to be? How much emphasis should be placed on selecting and training personnel, in canvassing, etc.?

b. The quality checks we have conducted are post-enumeration surveys. How serious are the time lags between the original enumeration and re-enumeration? Would it be possible in some cases to use a pre-enumeration survey?

c. Can the record check technique be used to validate the quality check technique? For the 1950 Census studies, the method proposed is this: Examine the cases where the quality check indicates the census to have been in error and then determine if the record check, in a significant proportion of the cases, supports the quality check finding. Also, if the record check does not disagree where both census and quality check agree, there is some additional basis for validation. However, the problems we raised earlier about using the record check to validate the census are still with us here, although to a lesser degree.

(more)

d. What is the possibility of developing differential techniques for measuring the relative contribution of marginal categories to survey error? In our demographic and our economic surveys -- to a greater extent in the latter than in the former perhaps -- the marginal classes such as the extremely small factory, the borderline farm, the newly constructed or the dilapidated dwelling unit, the unpaid family worker in the labor force contribute a disproportionate share of the number of errors in our statistics, but they affect most totals very little. With the development of more precise criteria for distinguishing marginal units, it may be possible to conduct more efficient surveys and quality checks.

S U M M A R Y

In summary I should like again to point out that it is our judgment that considerations of scientific presentation and of public policy appear to support very strongly the present program in various countries for attempting to estimate the accuracy of some of the statistical data being published. Accordingly, it appears to be highly appropriate that so much emphasis is being given at the present conference to the problem of measuring accuracy.

We conclude that much work remains to be done in the development and application of techniques for measuring the accuracy of data from enumerative surveys and censuses. Several methods which have been most effectively used for appraising the accuracy of statistical aggregates often shed much light upon data, but leave the analyst with questions of interpretation.

The methods based upon the analysis of facts for individual units offer promise of giving more precise knowledge of the quality of statistical data than do global comparisons. Checks with administrative records and intensive re-interviewing are the methods which we believe to be particularly worthy of serious experimental investigation. Our present view is that the most promising approach appears to be the quality check, supplemented by the record check, in which the present techniques of sampling make possible reasonably accurate measures of accuracy with only a relatively small number of cases covered. We believe, however, that much work needs to be done in this field and we look forward to many opportunities in this conference and in later meetings to exchange experience with representatives of other nations.