Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

On the Interpretation of Censuses as Samples

W. Edwards Deming & Frederick F. Stephan

To cite this article: W. Edwards Deming & Frederick F. Stephan (1941) On the Interpretation of Censuses as Samples, Journal of the American Statistical Association, 36:213, 45-49

To link to this article: https://doi.org/10.1080/01621459.1941.10502070

ON THE INTERPRETATION OF CENSUSES AS SAMPLES

By W. Edwards Deming and Frederick F. Stephan
Bureau of the Census

For inventory purposes *an enquiry must be complete.* Both by historical precedence and by law, one of the primary functions of the census is to provide a count or inventory of the population as it actually existed on a specified census date (April 1st for the censuses of 1930 and 1940). Certain actions, involving the apportionment of taxes and Congressional representation, are based on the census figures, purely as an inventory as of a certain date. States and cities likewise utilize the census returns for reapportionment of the state legislatures and the city councils, for revising election districts, and for many other administrative purposes. These inventory functions are limited by the fact that the census information concerning individuals is held confidential; consequently, states and cities often maintain other lists for inventory purposes, examples being tax rolls, city directories, the registration for selective service, and social security registration.

For such purposes, the census must by definition, be complete. It differs not only quantitatively but qualitatively from any sample less than 100 per cent that might be taken. This distinction is independent of any defects that may occur in carrying out either the census or the sample.

As a basis for scientific generalizations and decisions for action, a census is only a sample. In addition to serving the function of an inventory as of a certain date, the census tabulations serve also another important objective, namely, as bases for prediction. Any social or economic generalization, and any recommendation for a course of action, involves a prediction. For such purposes, the census takes on the character of a sample.

Any census gives data of the past, but the generalizations and courses of action that are based on it concern the population as it will exist at some time in the future. A census describes a population that is subject to the variations of chance, because it is only one of the many possible populations that might have resulted from the same underlying system of social and economic causes.

The births, deaths, vocations, migrations, and educational attainments of a population are changed and directed by a myriad of chance causes, superimposed on certain underlying social and economic cause systems. A census shows what resulted from this combination at a certain time in the past, but any generalizations that are not restricted to

a particular date and place must recognize the fact that some other population might have resulted, and must in fact be expected to arise in the future from the same underlying causes. Because of these statistical fluctuations, it follows that as a basis for scientific generalizations and decisions for action, the distinction between complete and sample coverages is often only a matter of degree.

Then, too, the state of a population today depends not only on the actions of chance today, but also on the actions of chance in the past. In other words, the population of any area today depends partly if not largely on that population as it existed a year ago, or two years ago, etc. A common technical expression is to say that there exists a serial correlation between the states of a population on two different dates. It follows that for predicting the state of a population next year, the census of that population as it actually existed on the last census date is more useful than a census of this same population as it might have existed by possible different actions of chance. Fortunately, the only census that we can take is the one that is the most useful. Moreover, because of serial correlation, it is desirable that the census or sample of the population on the census date be as precise as possible.[1]

These ideas are not new. The sense of them permeates a considerable amount of statistical work and writing, yet they are far from being well recognized, and there is a dearth of explicit statements of them. Because of their extreme importance it seems worth while to put them into print, even if it has been done before.

The two excerpts given below are quoted for the clarity with which they summarize the arguments of this section.

> A so-called 100 per cent sample from the viewpoint of scientific method is, as soon as taken, a sample of the past. The usefulness of such a sample is only as a basis for drawing an inference about the future and in this case the sample (even a 100 per cent sample) is but a finite sample of a potentially infinite one that might result from the cause system existing at the time the sample was taken.[2]

> In all scientific investigations the object is to find, not the situation that prevails at a particular moment, but the underlying tendencies and relations which, with chance modifications determine these situations approximately. The chief object in sampling the population of Seattle is to find out something about the underlying relations that make that popula-

[1] In substance, this paragraph was expressed by Dr. Leon E. Truesdell in a conversation with the authors.

[2] Extracted from a letter written by Walter A. Shewhart to W. Edwards Deming, dated 9 May 1940.

tion what it is. The situation in the city fluctuates continually, and what is found today will not be exactly true tomorrow.[3]

Comparative advantages of a census and other kinds of samples in scientific investigations. When a given sum of money is to be expended on a scientific study of some subject, the object is invariably to arrive at generalizations and conclusions on which to base predictions. A well designed sample will then elicit the greatest possible amount of information; in fact, for such purposes there is no procedure of enquiry that is not a sampling procedure—some are only more efficient than others. A number of small samples, taken in different parts of the country, or at different times, may be much more valuable than a sample consisting of a complete coverage of one small area. Likewise, sample enquiries spaced a year or two apart may be preferable to a complete coverage of the same area taken every ten years.[4]

As the proportion of any attribute of the population diminishes, a larger and larger sample is required for estimating this proportion with a given degree of reliability. Hence small samples are not adequate to provide data that are to be tabulated in fine classifications, or by small geographic areas, such as small cities or townships. The reason is that in the resulting tables those cells that are to be used for making predictions must not be reduced to a point where statistical fluctuations may disturb the reliability of the predictions. Small samples are adequate, however, for many questions on which data are needed principally for the larger population groupings, cities, states, regions, or for the entire United States, and which in tabulation are to be divided into relatively few sub-classes; for these a properly designed sample will yield results that are as satisfactory for generalizations and courses of action as the results obtainable from a complete canvass, quicker and at much reduced cost. The small loss of precision and detail in a sample is usually much more than offset by the accompanying savings in expense, and such savings open up the possibility of covering a wider territory or conducting several enquiries spaced in time, or carrying a greater number of questions than would otherwise be included on the schedule, or of publishing more tabulations on whatever data are taken.

A striking result of what has just been said is that, except when a census is by definition merely an inventory, small numbers in the cells of a table derived from a complete canvass do not have much significance—no more than if they were derived from a sample. There is a limit, therefore, to the amount of detail that will be found useful in the

[3] Extracted from a letter written by Harold Hotelling to Lester Frankel, dated 30 July 1938.

[4] These ideas are further developed by Samuel A. Stouffer in his chapter "Sociology and Sampling," Ch. XVI in *The Fields and Methods of Sociology*, edited by L. L. Bernard (Long and Smith, 1934); see in particular pp. 479 and 481.

cross-tabulation, not only of sample data, but of a complete census. Of course, the limit is reached quicker with the sample data; and the smaller the sample, the quicker the limit is reached.

Some notes on criteria for appraising the adequacy of sampling. In deciding whether the frequency of a certain attribute of the population can be estimated satisfactorily on a sample basis, consideration needs to be given not only to the detail that will be required in tabulation, and the relation of this detail to the proportion p with which the attribute appears in the population (i.e., whether $p=.5, .1, .01$), but also to the question of how much information is already available concerning this attribute. If little is known about it, neither a sample nor a single complete coverage may suffice; a number of samples or complete coverages, separated either in time or space, may be required. On the other hand, when a considerable amount of information has already accumulated, a small sample taken every five years or every decade may suffice for continuity. Thus, questions on parentage have been a part of the census for several decades, and this is one reason why a sample will provide all the new information that is required on this subject.

A sampling method is said to be satisfactory for the questions under consideration if it can be depended on to yield samples (less than 100 per cent) that will lead to the same action as would have been taken on the basis of a complete count. In the usual statistical terminology, the population sampled is the population as it existed at the time the sample enquiry was made. A number of samples taken on this date would be expected to exhibit statistical stability, in the sense that they would all be drawn from the same population. From the point of view being expressed here, however, even a complete census, for scientific generalizations, describes a population that is but one of the infinity of populations that will result by chance from the same underlying social and economic cause systems. This infinity of populations may itself be thought of as a population, and might possibly be called a super-population. A sample enquiry is then a sample of a sample, and a so-called 100 per cent sample is simply a larger sample, but is still only a sample. In order to study the underlying cause systems, it is necessary to study several members of this infinity of populations; i.e., it is necessary to make sample or census enquiries on a number of different dates, preferably far enough apart to be independent, or nearly so.

From what has been said, however, it follows that a sample may be at once satisfactory for the questions under consideration (in the sense of leading to the same action that would be taken on the basis of a com-

plete count), yet unsatisfactory for being representative of an unrepresentative population. Thus, for a severely fluctuating population, even a 100 per cent sample on a particular date may lead to action not at all fitting to the future. For example, at the time the census of South Bend was taken in August 1939, the Studebaker plant was just in the process of preparing for the production of new models, with the result that the number of men at work that week was unusually low, and was in fact very different from the number of men at work the next week. Any action that did not take account of the fact that this was an unusual situation would not have been fitting to the population being studied.

As a basis for action, a sample may be preferable because of shorter processing time. Census and sample are both subject to the common error that arises from the fact that the social order is not static, but dynamic. It takes time to process and tabulate a small sample, but more for a complete canvass. For this reason, in some kinds of enquiries, the complete canvass, by the time it is processed and tabulated, may not be as good a basis for social action as the earlier returns of a sample coverage would have been. To the extent that sampling methods produce quicker results, they also produce more accurate results for purposes of prediction. The error that lurks in predicting the future from the past may thus be greater when the prediction is based on a complete canvass than when it is based on a sample.

The authors have had the advantage of numerous discussions with Drs. Walter A. Shewhart, Harold Hotelling, Leon E. Truesdell, and Philip M. Hauser concerning the interpretation of census and sample data, and express with pleasure their indebtedness to these friends, in particular for the letters quoted in the text.