

I. SOME PROBLEMS CONCERNING THE USE OF OFFICIAL STATISTICS IN SOCIOLOGICAL RESEARCH

EDWARD G. STOCKWELL

The purpose of this paper is to indicate briefly some of the problems and needs that are associated with the use of *official statistics* in sociological research.¹ Although specifically concerned with problems that graduate students may encounter in using official statistics while preparing their theses, the problems cited are not necessarily limited to students, but are ones that must be faced by virtually all persons who are engaged in research that entails the use of official statistics.

First, it is reasonably safe to assume that the majority of graduate students preparing theses today are primarily involved in intensive rather than extensive analyses. This, of course, brings them up against one of the major limitations of official statistics as far as their sociological relevance is concerned: *They tend to be extensive rather than intensive.* With regard to this problem, three points can be cited that would greatly enhance the value of official statistics for sociological research. The first two will not come as a surprise to anyone who has used official statistics in his research; they relate to the needs for: *greater detail in the cross-tabulations of the data*, and *a greater number and variety of small-area statistics.* These needs are particularly serious for graduate students. Time as well as financial limitations more or less require that thesis topics be relatively

narrow; yet when a topic is narrowed down to a "researchable" problem, it is often found that the official data are tabulated in too broad a fashion to permit the really intensive analysis that is necessary for a good thesis. In defense of the agencies responsible for the collection, compilation, and publication of official statistics, it must be said that the amount and variety of detailed cross-tabulations and small-area data have been increasing in response to these needs; nevertheless, there is still room for a great deal of improvement.

The third need, which ties in quite closely with the preceding two, relates to the familiar problem of financial support for research. The needs for more detailed cross-tabulations and local-area data can, in many cases, be met by having the agencies concerned prepare special tabulations. This costs money, however, and money is generally a problem for most graduate students. Therefore, it would be helpful if there were increased sources of funds quickly available for deserving students and projects.

There is also the *problem of gaining access to official statistics that are not readily available in published form*. This problem is found primarily at the local level. When graduate students and others engaged in sociological research approach city or even state officials with a research problem that requires access to official records (birth certificates, sales tax returns, criminal statistics, public health and welfare records, etc.), it is not at all unusual for them to encounter a hostile reception. Although this is fortunately not universally true, the fact remains that many local administrative officials tend to be suspicious of sociologists and their motives; and it is often difficult, if not impossible, to break through this barrier of distrust. There is a need, then, to *educate the public to the fact that the value of official statistics goes far beyond their purely administrative value*. In particular, *government officials should be taught a greater appreciation of the need for more sociologically relevant data*. They should be encouraged to enhance access to unpublished data, and also to improve (or at least to maintain) the relevance of existing data. Points that come to mind in this respect are the efforts that have been made recently (some of them successfully) to eliminate the item on race or color from official records; the possibility of adding

an item on education to birth and death certificates; and the perennial arguments over the research needs for data on religion.

There is, in this writer's opinion, a serious need for *better channels of communication among members of the profession* with respect to what is being done with the available data and what special tabulations have been made for what purposes. A specific instance concerns the public use samples that can be purchased from the Bureau of the Census. Who has bought them? What problems have they encountered in using them? And, particularly, what is being done with these data? This last question can be especially serious for graduate students, since it is possible to get quite far along on a thesis research project only to find that someone else has been working on the same problem. In this regard, some sort of central agency for the exchange of information would be useful. Some of the professional journals do, of course, carry occasional lists of research in progress; nevertheless, this, too, is an area in which there is room for a great deal of improvement. It would also be helpful if some of the persons who have purchased the public use samples would publish an article or research note indicating what they are doing with the data—and particularly what some of the problems are that they have encountered.

A number of other problems could be cited—such as the small size of the Current Population Survey which limits the detail in which the data can be tabulated; the heavy reliance on sampling in the 1960 Census of Population which seriously reduced the reliability of social and economic data for small areas; changes in the definition of basic concepts from one census to another; changes in the manner in which the data are tabulated. These are familiar problems to those who have done research with official statistics. However, we might reverse the point of reference and discuss a need that concerns members of the sociological profession rather than the official statistics themselves. Although the value of these statistics for research is admittedly limited, the changes that have taken place in recent years with regard to the amount and variety of available data mean their value is much less limited today than a generation ago. Many, however—students and faculty members as well—do not seem to be fully

aware of this fact. For people trained in sociology 15 to 20 years ago, many of the basic concepts, theories, problems, and general methods of social research are still applicable; but the nature of the empirical data that are available for research has changed considerably. While most sociologists have kept pace with the more general theoretical and methodological developments, *many have not kept pace with the developments regarding official statistics*. For instance, this writer has encountered sociologists who were supposedly research-oriented but who had never heard of the Current Population Survey and were completely ignorant of the nature of the statistics that are collected and published periodically between the decennial censuses.

There is also the possibility that many who may be aware of the existence of a variety of official statistics *do not fully appreciate their value* and are often reluctant to use them for one reason or another. There is a general misguided belief that anything collected by a government agency for administrative purposes cannot have much relevance for sociological research. This belief, of course, is simply unfounded. There are three basic uses to which official statistics can be put in sociological research: 1. they can be used as general *background* materials against which to view more specific data obtained from other sources; 2. they can be used to support or *supplement* data obtained from other sources; 3. they can be used as the sole basis for *testing hypotheses*. The last-mentioned use, in particular, is one that many sociologists do not fully appreciate. Many sociologists who may use official statistics as background or supporting data frequently balk at using them as the sole basis for testing hypotheses. They regard them as too statistical and as not being “real” sociology. Actually, of course, much good sociological research can be, *and is*, based entirely on the use of official statistics.²

The lack of familiarity with (or appreciation of) official statistics among sociologists today is doubly serious in the sense that *it is being passed on to the students*—at both the graduate and the undergraduate levels. Conventional courses in research methodology devote a lot of time to a discussion of the scientific method, problems of questionnaire and schedule design, sampling techniques, multiple and partial analysis, and so forth; but they give little—if any—considera-

tion to the research potentials of official statistics. This is not to imply that the material covered is useless; on the contrary, it is very useful and very essential. It is not really enough, however, and it has created a situation in which many students complete their training, write their theses, and enter the profession with the idea that empirical research has to be limited to the data they can collect for themselves. In the course of teaching graduate seminars in demography, the writer constantly encounters students who have only a vague notion of what the decennial census is about, and who think a birth record is something to prove that you are old enough to buy a glass of beer. One might argue that a seminar in demography provides the means whereby students learn the value and limitations of official statistics for sociological research. However, not all graduate students take such courses, and those who do not are at a real disadvantage when it comes to writing a thesis and when they later attempt to do research in sociology. If for no other reason, they are at a disadvantage in the sense that much less research data are available to them than to students whose research training has included exposure to, and experience in, the use of official statistics. A serious need, in this writer's opinion, is *the re-education of sociologists in this country so that they will have a better awareness and appreciation of the value of official statistics in sociological research.*

In making the accusation that many sociologists today are either not fully aware of the vast amounts of official statistics that are available, or are not sufficiently appreciative of their potential value for social research, it is only fair to add that part of the blame must be placed on the communication problem mentioned earlier. The writer feels that the various agencies concerned with collecting, compiling, and publishing official statistics do not make a sufficient effort to publicize the data that are available. Those of us with a demographic background may not readily appreciate this, having become reasonably well aware of what is available during the course of our training and research experience. Furthermore, most of us receive the official publications, or announcements thereof, as soon as they come off the press. The majority of American sociologists, however, are not in this position. In short, there is need for *improved com-*

munication not only among members of the profession, but also *between the profession and the various government agencies responsible for the compilation of official statistics*. To offer one suggestion in this regard, it would be of immense value if such agencies as the Bureau of the Census and the National Center for Health Statistics would produce a monograph describing in some detail the nature of the statistics published by their respective agencies, their availability (either in published or unpublished form), the major concepts and definitions employed and the potential uses and limitations of the statistics for social research. Perhaps a monograph of this type might even suggest some specific research studies for which there is a need.

There are, of course, several publications somewhat similar to the one described above. For the most part, however, these tend to be limited to a particular subject, to reach a limited audience, and to be written generally in a style that is not easily understood by persons untrained in demography. What is really needed is something in a semipopular or lay style, geared specifically to the beginning student who has no background in population analysis. To make one further suggestion, the various professional journals might be put to greater use if they provided more adequate communication between the members of the profession and the agencies concerned with official statistics. The *American Sociological Review*, for example, could perhaps devote a page or two of each issue to informing members of the profession about new tabulations that have been or are being made and new reports that have been published.

SUMMARY

Although the increasing amount and variety of official statistics that are available has greatly enhanced their potential research value, there are still a number of problems that limit their usefulness for sociological analyses. Not the least of these problems concerns the fairly widespread lack of familiarity with, or appreciation of, the potential usefulness and value of official statistics for sociologically relevant research.

REFERENCES

¹ By "official statistics" is meant statistics pertaining to various characteristics of the population (social, economic, demographic) that are published periodically by official government agencies such as the United States Bureau of the Census, the National Center for Health Statistics, and other administrative bodies, at the national, state, and local levels.

² With regard to the use of official statistics in sociological research, this is one place where American sociologists seem to be lagging behind their European contemporaries. The postwar period in Europe has witnessed a notable decline in the traditional armchair theoretical approach to sociology, and a growing emphasis on empirical research. This does not mean that there are no "thinking sociologists" left, or that everyone is busily engaged in empirical research; it means that empirical research is much more respectable today than it was a generation ago, and that more of it is being done. Moreover, European sociologists who have been doing this type of research have made greater use of official statistics than have sociologists in the United States—in spite of the fact that our approach to sociology has traditionally been more empirical than theoretical. This fact was brought quite forcefully to the writer's attention recently when one of our European colleagues asked him, as a demographer, to explain why American sociologists do not take greater advantage of the wealth of data available to them in the various official government publications. The answer to this question, as indicated above, is that many American sociologists are not aware of the amount and variety of official statistics that are available; and many others simply do not appreciate their potential value for sociological research.