Frank H. Godley, National Center for Health Statistics

Data collection procedures, formats, and content will be revised in 1968 by the two major arms providing population statistics for the United States. The U.S. Bureau of the Census will hold a dress rehearsal of the 1970 Decennial Census of Population in April 1968, while the National Division of Vital Statistics will have instituted its new Standard Certificates of Birth, Death, Marriage, and Divorce by January 1 of that year.

This paper is intended to gain an historical perspective on the collection of population statistics in the United States. Specifically, investigation is of the contributions of Lemuel Shattuck, a co-founder of the American Statistical Association in 1939.

In his publicizing <u>Circular</u> of April 1840, Shattuck listed nineteen subjects in which the newly formed American Statistical Association was interested. Number two on this list was "Population.--The Census of different periods; the Births, Marriages, and Deaths, specifying the diseases, sex, age, and months of the year, when each death took place; Boards of Health, prevalence of Epidemics, and other diseases" (1). Shattuck himself was to become the most significant contributor, as we shall see, to population statistics in America.

As one of the chief planners of the 1850 Census of the United States, Lemuel Shattuck introduced not only new questions, but also new analytical concepts. The first of these, and the most revolutionary, was that of making the individual, rather than the family, the unit of analysis. This innovation he first applied to his 1845 Census of Boston.

The consequences of shifting from the analytically cumbersome family unit to the individual unit were profound. One consequence is, of course, that much more information (age, sex, education, etc.) can be collected for the individual than for the household. The census schedule needs only to provide a single line for each member of the family, so that records for families or households may be kept together (2).

The purposes of such a census of population fall in three classes, according to Shattuck: political, public health, and social-scientific (3). In addition to the political purpose of allocating legislative Representatives and raising revenue (as stated in the United States Constitution), census enumerations, Shattuck recognized, are indispensable in assessing the morbidity and mortality of a community. This is the public health purpose, which introduced the concept of relating the number of cases of disease or death to the size of the population at risk. A third purpose of a census was socialscientific. Census data on the age composition, for example, can be used to forecast the population of school age and thereby to plan school building.

In order to achieve such social-scientific and public health purposes, Shattuck listed ten classes of facts that should be included in every census. They are age, sex, race, nativity, residence in cities, occupation, marital status, education, economic status, and housing facilities. Age, believed Shattuck, was the most crucial in public health studies. He stressed the necessity for exactness in the reporting of age. To him, the principal value of a population by age was its comparability with annual decedents by age. That is, age-specific mortality was to be a powerful tool for comparing the sanitary conditions of different communities or of the same community with its past. While most of the facts recommended by Shattuck were intended for public health purposes, they have gradually acquired utility to the social scientist.

One of Shattuck's recommendations was to tabulate the occupations of males aged 15 years or older. Occupation in the 19th Century revealed significant differentials in mortality, but these have largely disappeared in modern times. More important is the current use of occupation as an index or component of socio-economic status.

Education, in those embryonic days of public health, was measured by the crude standard of literacy. The proportion literate, thought Shattuck, should influence the sanitation of a community. His proposed question on the census was "Can you read or write?" In modern times, literacy is often taken as the minimum level of educational attainment and is used in crosscultural comparisons. Within cultures or countries, however, the number of years of school completed has become the preferred index of educational attainment. In conjunction with occupation or by itself, years of school is used as an indicator of social class and in studies of social mobility.

Suspecting that property owners enjoyed better health conditions than tenants, Shattuck recommended a census question on the ownership of real estate. This indicator of economic status was phrased "means of subsistence and comfort."

Housing facilities were the last of Shattuck's ten facts essential to a census. The recommended questions were the number of persons per family and per house. It is probable that the overcrowding from rapid urbanization and centralization of the population in the Nineteenth Century impressed upon Shattuck the sordid health conditions arising therefrom.

A principle can be drawn from these early census recommendations. It is that census questions must have a basis either in some current, unanswered questions of a practical nature, or in some body of knowledge of a theoretical nature. In the Nineteenth Century, the most serious and pressing questions were posed by the high and wide-spread levels of acute contagious disease. Medicine provided the theoretical context for posing these questions.

Today, in the mid-Twentieth Century, the most pressing problems are perhaps somewhat different: they raise questions of the functioning of the society and man, the social animal, rather than man, the biological animal. Behavioral sciences, more often, provide the theoretical framework for answering these questions. Just as the census officials have responsibility for recognizing this, we as social scientists have an obligation of posing the practical and theoretical problems of the day in terms of data needs. This is what Lemuel Shattuck did one century ago.

The significance of Lemuel Shattuck's census proposals were also largely to aid the vital statistics branch of population statistics. By first removing vital statistics from an unreliable collection mechanism, the census, and then by substituting an effective registration system, Shattuck provided for a continuous file of potentially accurate birth, death, and marriage records.

Secondly, recognizing the need to relate vital statistics to the exposed population, those at risk of the event, he revolutionized the census by substituting the individual for the family as the unit of analysis. This created a base of detailed, cross-classified populations for comparing the incidence of vital events as well as that of disease in different communities or time periods. As a result, detailed studies of demographic processes were possible. Vital events could be properly related to the individuals exposed to specified risks of birth and death. It became possible to calculate person-years of exposure, as Shattuck demonstrated in his reports on the 1845 Census of Boston.

The greatest problem of early vital registration statistics was their uneven coverage of different areas of the United States. The problem of regional variation in the quality and quantity of vital statistics is a common one faced by all growing registration systems. A few highly developed registration centers, often in the large cities, report vital events relatively accurately and completely, while in the less developed areas, bias and nonreporting of vital events is high.

These considerations force a choice between two alternative registration systems: in one, geographic coverage is complete, but error and nonreporting are frequent; the alternative provides high-quality vital statistics, but entails partial geographic coverage.

In the Nineteenth Century United States, the latter alternative was embodied by the registration systems of a few major cities such as New York, Boston, and Baltimore. The alternative to this, however, was carried out by the Census of Population, using a survey method to count vital events and measure vital rates for the Nation as a whole. Shattuck disapproved of the Census method, arguing that it was highly unreliable and yielded virtually useless vital data.

It was unreliable to tally births and deaths from a census, he argued, because the time lag between the occurrence of the vital event and its recording was simply too long. In the interim, people forget. By census time they tend to misstate not only the date of the occurrence of the birth or death, but also the characteristics of the persons concerned. Moreover, the family units and their associated households may be lost by migration, marriage, divorce, and death.

These disadvantages also apply to the sample survey methods used by some of the newer nations to estimate vital events. Again there is a time lag between the date of a vital event and that of the interview. Moreover, a survey at one point or period in time may not necessarily represent the calendar year. There is a seasonal variation of births, deaths, and marriages. Unless overlapping or continuous surveys are undertaken, the respondents are forced to tax their often unreliable memories about the dates and particulars of events occurring in other seasons of the same year and sometimes in previous years. Some additional reasons have been advanced in modern times. A. Mitra, the Registrar-General of India, contends that the sample survey method can never replace universal vital registration because sample data for local areas contain too large a sampling error, especially when cross-classified, and that these data are of no use for legal identification purposes (4).

But notwithstanding the vehement protestations of Lemuel Shattuck, the United States Census served as the principal means of collecting National statistics on births, deaths, and marriages from 1850 (when, ironically, Shattuck himself designed the questions and tables) until 1910. Consequently, reliable vital statistics were not collected in the United States until the Twentieth Century, or nearly 50 years after Shattuck's proposals.

Despite all of Shattuck's highly desirable and advanced plans, considerable organized opposition was voiced. In his <u>Report to the Sanitary Com-</u> <u>mission of Massachusetts, 1850</u>, Shattuck listed some of the more persistent arguments against asking detailed questions of every person in the census:

1. The invasion of personal and family privacy would be intolerable.

2. It constitutes an invasion of "Divine

Providence," who alone should decide when and of what humans should die; public health shouldn't intervene.

3. Local communities are potentially threatened by government possession of such ancillary information as is proposed for the census.

4. The census would be "too statistical; you can prove anything by figures."

Alongside each of these objections and the others, Shattuck published rebuttals in his <u>Report</u> (5). Since most of his proposals were incorporated into the 1850 Census of the United States - excepting that of no birth and death questions - and in view of a complete nationwide system of vital statistics today, the efforts of this pioneer in American demography appear to have been successful, the obstacles overcome.

REFERENCES

(1) <u>American Statistician</u>, "125th Anniversary Issue," Volume 19, Number 2 (April 1965), p. 31.

(2) National Office of Vital Statistics,
<u>Vital Statistics of the United States</u>, <u>1950</u>,
Volume I (Washington, 1954), pp. 5-6.

(3) Lemuel Shattuck, et. al., <u>Report of the</u> <u>Sanitary Commission of Massachusetts</u>, <u>1850</u>, (reprinted, Cambridge, <u>1948</u>), passim.

(4) India Ministry of Home Affairs, The Registrar-General, <u>Vital Statistics of India</u> for <u>1961</u> (New Delhi, 1964), p. 1.

(5) Lemuel Shattuck, et. al., <u>op</u>. <u>cit</u>., pp. 283-292.