

Philip M. Hauser, University of Chicago

How will social statistics in 2000 in the United States compare with social statistics as we know them in 1976? In the attempt to answer this question three factors must be considered: first, in the present state of fields of statistics and trends in respect thereto; second, in the anticipated change in the social economic and political milieus in the approximately one human generation that remains in this century; the third, in the continuing advance of the state of the statistical arts--in the design of conceptual frameworks, in the measurement of social, economic and political phenomena, in compilation and tabulation procedures, and in the invention of new techniques of analysis.

Development of Fields of Statistics. At present in the United States economic statistics in contrast with social statistics are better developed, more systematically integrated and more subjected to policy and programmatic use. Economic statistics are more often analyzed in relation to explicit economic goals than are social statistics in relation to explicit social goals. That is, economic statistics serve as economic indicators more than social statistics serve as social indicators. A statistic is an indicator when it is interpreted as measuring progress or retrogression towards or away from an accepted goal.

The history of statistical developments in the United States reveals that over time new inquiries of census schedules and new current series of data, whether based on administrative records or sample surveys, reflect the transition of the country from a rural, agrarian, folk, to an urban, industrial mass society.¹ The emergent, more complex, interdependent, and vulnerable economic and social orders have required increased government intervention in the exercise of planning and regulatory and evaluative functions. The increase in such functions called for more and more data as a basis for policy formation and administrative action. Since I have documented this assertion elsewhere,² I shall use only a single illustration for purposes of clarity here.

In the Census of the United States there was no systematic effort to measure unemployment and the total work force as of a given time period until after the unprecedented level and duration of unemployment of the deep depression of the 1930's. The efforts of the government to alleviate the distress of the unemployed and to cope with other aspects of that depression led to a variety of survey experiments--local, state, and national--which resulted in the abandonment of the "gainful worker" approach and the adoption of the "labor force" or "active population" approach to the measurement of the labor force and its employment status. The adoption of the labor force approach in the 1940 Decennial Census was followed by the monthly series of labor force data based on a sample survey as reported in the Current Population Reports or more specifically in the Monthly Report on the Labor Force.³

In similar fashion in the transition from an agrarian to an urban society, emergent problems which required national attention led to new census inquiries, new sample surveys, or new administrative statistics, as new agencies were created to deal with the new problems. Among the

statistics which proliferated were data relating to internal migration, housing, income, fertility, education, and place of work.

In respect of economic statistics the development of the National Statistical Accounts, not only in the United States but in the world as a whole, reflects the changing character of the economic order and the growing interest and intervention of government in the operations of the economy. In the United States the Employment Act of 1946, which created the Council of Economic Advisers and required the annual Economic Report, certainly greatly stimulated the development and use for policy and programmatic purposes of economic statistics.

Increasing concern in the U.S. with social problems such as intergroup relations; the "urban crisis;" delinquency, crime, and the administration of criminal justice; welfare; health and medical care; education and recreation has given rise to proposals for a Council of Social Advisers to parallel the Council of Economic Advisers.⁴ Should such an agency be established there can be no doubt that social statistics would be subjected to stimulation similar to that of economic statistics for further development and to increased policy and programmatic use.

The increase in the scope and use of statistics has, of course, not been confined to government. Similar developments have taken place in the private sector in business, labor, welfare, civic, and educational organizations. Especially notable, by reason of their impact on public policy, are the public opinion polls such as those conducted by Gallup and Harris.

Anticipated Social Change. Next let me state explicitly my assumptions in respect of changes in the social, economic, and political milieu of the United States during the remainder of this century:

1. The transition from an agrarian, folk, to an urban, mass society will continue.
2. The frictions of this transition will be exacerbated and require increasing government surveillance and intervention. The increasing role of government will probably be more resisted in the United States with its greater addiction to its inherited frontier psychology and old economic religion than has been the case in other advanced countries.
3. Government interest in the realm of the "social" as distinguished from the "economic" will increase to match government interest in the "economic."
4. In consequence, government will adopt more explicit social goals to complement present economic goals.
5. The need to measure progress or retrogression in respect of social goals will result in a proliferation of social statistics and in efforts to integrate them.
6. Social statistics will, therefore, tend in use increasingly to become social indicators.
7. Governments in the United States will be engaged in more central (as well as regional) state, and local planning and action to supplement and complement the play of market forces in

dealing with collective, as distinguished from private, problems and needs.

8. By 2000, it is assumed that the priority of social over personal rights will have been so established and that the use of the computer in the preservation of privacy and confidentiality will be so advanced that comprehensive data banks will have been created. These data banks, described further below, will make it possible to collate significant information for individuals, households, and institutions on a local, state, and federal level for statistical purposes, while safeguarding privacy and preserving confidentiality. Regulatory and administrative agencies will have access to the data bank but will not be able to obtain more information about individuals, households, or institutions than they now possess or should possess in light of the future developments. The information collected for administrative and regulatory purposes, however, would be available for statistical purposes.

State of the Statistical Arts. What statistics are available at any point in time depend in part on existent conceptual frameworks and what phenomena can be measured. Gross National Product statistics, for example, did not exist until the necessary conceptual apparatus was developed and the necessary components could be measured. Similarly, although the need for statistics on underemployment has been discussed for many years only scanty data on this subject are available because of conceptual and measurement problems. In efforts to measure levels of living, areas of concern have been mentioned from time to time for which statistics were desirable but little, if any, data exist in respect of them because of difficulties of measurement. Examples of such items are "human freedoms," "health," "security," "opportunity," and "happiness." Also resistant to measurement has been the synthesis or integration of social statistics into a single index comparable to GNP in the realm of economic statistics.⁵

It may be anticipated that as measurement techniques improve new statistics will be developed to include the types of items considered above. Improved measurements may be anticipated on aspects of personality, on attitudes and values, and on other psychological, social psychological, and sociological phenomena which will be of public concern.

The advent of the computer has, of course, greatly and positively affected the scope, timeliness, and quality of data. New more powerful and more efficient generations of computers yet to come will continue to have similar effects.

Finally, new techniques of analysis may be expected to influence statistics of the future. Such innovations as log-linear models for multivariate analysis;⁶ the proposed "demographic accounting" procedures;⁷ a "health accounts" system;⁸ and analytical models of various types will undoubtedly influence future statistical developments.

Central Data Banks by 2000. In view of the above considerations, the most important single development in statistics by 2000 will be the emergence of comprehensive data banks for individuals,

households and institutions. By that time in the United States a file will be initiated for each individual beginning with a birth certificate, including a record of every significant event in realms in which government has an interest or program, and closing with a death certificate. Among the files which would be continuously maintained would be a medical file--immunizations, disease episodes, contacts with physicians and paramedical personnel; health facilities used and outcomes. The file would also contain information on marriage, divorce and re-marriage; schools attended with fields of concentration and certificates, diplomas or degrees; employment, unemployment and underemployment; information on internal and international migration; income received, including transfer payments; taxes paid; housing; arrests, indictments, convictions, sentences and institutionalizations. Personal characteristics would, of course, also be included such as age, sex, ethnicity or race, etc.

A similar record would be maintained for households with appropriate entries indicating the person's departures from the household including the new household created with new household formation. Comparable files would also be maintained for institutional households.

I am aware that this prophecy is bound to evoke reactions of consternation and visions of Orwell, if not by 1984, then by 2000. I should state that I am in complete sympathy with the right to privacy and the obligation of governments to maintain the confidentiality of information collected for statistical purposes. But I am convinced of the following:

1. That the existence of discrete record files for diverse purposes, such as vital registration, social security, medicare, internal revenue, census tabulations, voting registration, etc., will prove increasingly costly and inefficient; and impose frustrating constraints in the production of data increasingly required for planning and administrative purposes;

2. It is possibly easier to preserve privacy and maintain the confidentiality of information with the computer, even with the present, let alone future generations of computers, than it was in the pre-computer era;

3. It is conceivable, and it will come to pass, that various safeguards would make it impossible for a Richard Nixon, or a J. Edgar Hoover, let alone lesser men, to have access to any data other than that specifically required and permitted by statute about any person, family or institution that could be used in a way inimicable to their interests. That is, the same data bank could be used by regulatory agencies to obtain the individualized data they need for authorized administrative purposes without their having access to other information in the data bank inserted only for statistical purposes. Simultaneously, the data obtained for regulatory and administrative purposes would be available for statistical use under provisions that would not violate privacy and guarantee confidentiality.

I am convinced that even now, let alone by 2000, the necessary combination of ingenuity and technology can be marshalled to achieve these goals. Such a record system maintained on a

decentralized basis, by city or metropolitan area and by state, with provision for national aggregation would make possible the implementation of Richard Stone's demographic accounting proposal; and such proposals as made by the Committee to Evaluate the National Center of Health Statistics in the U. S. for a national health accounts system. It would make possible statistics with agreed upon periodicities of the stocks and flows of human beings in significant categories and functional units.

Needless to say the establishment of such a record system requires much in the way of advance planning--agreement on standard definitions and practices; and the development of classification systems and taxonomies yet to be devised. The types of problems and considerations involved are discussed at some length by Stone in his OECD volume.⁹ Such a data bank could be an economical and efficient way of producing many statistics now being produced in a discrete and overlapping manner which defy integration and synthesis. It also could considerably reduce the items now obtained through censuses and sample surveys and duplicate and overlapping surveys and files; and it would certainly, to a considerable extent, offset the costs of the data bank and the derivation of statistics therefrom.

Finally, it should be observed that similar record systems may be established by 2000 for business and industrial enterprises so designed as to serve government, administrative and regulatory as well as statistical needs. Again it is emphasized that adequate safeguards will have been established to maintain the confidentiality of data for individual enterprises.

Specific Examples of Anticipated New Statistics. The new statistics which will emerge will reflect national priorities as the governments turn successively to deal with various problem areas as they become acute and engage national attention.

Before turning to the purely domestic scene let me first focus on development of statistics generated by international interest in helping developing nations to cope with widespread poverty and to accelerate the advancement of levels of living. By 2000, it is likely that much progress will have been made in the measurement of social and economic development and in the synthesis of an index of development for all nations. Progress in this direction has already been made by the United Nations Research Institute for Social Development.¹⁰ It is conceivable that standardized data collections, a world data bank and common tabulation and analytical procedures will enable each nation to see how it compares with other nations at the same and different levels of development. The United States will be an element in this international statistical system and will have a synthetic index of development as well as component indicators.

Next let us turn to specific domestic developments. On the assumption stated above that the increasingly complex urban, industrial, mass society which will characterize the United States will require increased government surveillance and intervention, major innovations in statistics may be anticipated in areas such as the following:

health and medical care; poverty and income maintenance; underemployment as well as employment and unemployment; welfare; minority group status; social mobility; housing; education; crime and the administration of criminal justice and recreation.

Health and Medical Care. By 2000, it may be anticipated that greatly strengthened and broadened health and medical programs will have developed in the U. S. as in other advanced countries, including provisions for comprehensive health insurance or equivalent and improved delivery of medical services. In the development of such national health care system, cooperative health statistical systems will be established linking and integrating present discrete data in the public and private sectors relating to the health industry. The data banks described above would enable a health accounts system to be developed which could relate inputs to outcomes with attention to intermediate processes and flows. Linked and integrated would be such components as vital statistics, health survey data, health manpower and facilities data, medical intervention information, health programmatic data, and medical costs information. Also to be anticipated is the development of data which could make possible the evaluation of medical intervention by relating procedures to outcomes. Needless to say the latter will not be achieved without great controversy. That is, with increasing government interest, surveillance and participation in health and medical care programs, the relation between physician and patient will become a matter of public concern and evaluation of medical practice a consequence.

Poverty and Income Maintenance. By the century's end it will be recognized, even in the United States, that poverty has its origins in deficiencies and frictions of the economic and social orders as well as in deficiencies of the family and the individual. Furthermore, it will be recognized that whatever the cause the government will have the obligation to provide an adequate income flow to all persons and families--as far as possible through payments for services performed. It will have long been recognized that "welfare state" is not a pejorative term; that the government must be "the employer of last resort;" and that the problem is not whether the nation is to be or not to be a welfare state, but rather how equitable welfare provisions can be made.

With such orientation it may be anticipated that "poverty" statistics will be greatly strengthened, comparability over time and space much improved, and the many vexing technical problems by reason of changing consumer baskets of goods and inflation reasonably resolved. On the assumption that the United States will have made the elimination of poverty a national goal, poverty statistics will have become poverty indicators and data on income distribution, consumer expenditures, savings and wealth will have become greatly strengthened, routinized and increasingly monitored and used for social as well as economic policy and programmatic purposes.

Labor Force. It has become increasingly evident that the "labor force" or "active population"

approach is not meeting the needs of developing countries for manpower policy and program purposes. One of the reasons for this deficiency lies in the failure of the standard approach to measure underemployment in addition to unemployment. By 2000, it may be anticipated that the measurement of underemployment will have become a standard practice throughout the world in accordance with the recommendation of the Eleventh International Conference of Labor Statisticians in October 1966, held under the aegis of the ILO.¹¹

Some indication of the type of data which will become available is afforded by the "labor utilization framework" made operational by the writer¹² for which experimental data in various degrees of completeness are now available for nine nations.¹³

Needless to say, the data bank discussed above will make possible longitudinal as well as cross-section statistics to provide a much better understanding of patterns and changes in labor force participation in relation to other social and economic variables.

Other. Space does not permit even the sketchy information presented above for other statistical areas. In quick summary the following observations are in order:

Housing. Housing statistics will be strengthened and elaborated so that continuous data on stocks and flows will be available; and on quality of housing in relation to occupancy and characteristics of occupants.

Education. Education statistics will be strengthened and planned in significant social and economic context by reason of the central data bank. Measurements will become available on the quality and content of education and on the efficiency and success of schools and educational procedures.

Crime. Data on crime and delinquency and the administration of criminal justice will be improved and integrated and outcomes evaluated. Population surveys and data bank files will provide much more accurate information about the level of criminal and delinquent behavior than obtainable through reports based on police or court records.

Minority Group. The increasing insistence of minority and underprivileged groups, including women, for full equality of opportunity will make much more data on the socio-economic status of such groups available. These statistics will be used as social indicators to measure progress in the elimination of discriminatory practices.

Social Mobility. On the assumptions stated above and the basic assumption that the United States will still be a democracy, it is anticipated that social mobility will be a matter of increasing national concern.¹⁴ In consequence it may be anticipated that periodic statistics will become available on increase or decrease in the social mobility of the population as a whole and on sub-groupings of the population.¹⁴

Concluding Observations. It is clear that the need for a brief presentation precluded comprehensive coverage of the statistical firmament and permitted only sketchy considerations of the specific areas covered. I have tried to present

a framework within which the direction of change in statistics during the rest of this century can be visualized. In the specific area to which I have made some reference it should not be surprising that I have, in the main, concentrated on aspects of social rather than economic statistics. I have done this not only because it is the area with which I am most familiar but, also, because it is the relatively underdeveloped statistical area. The major thrust of my remarks is that with the anticipated changes in the social, economic and political milieus, social statistics will not remain relatively undeveloped.

Without question the most controversial of my prophecies will be that relating to the central data banks. It may be useful to point out that emotional reactions against such a development may be attributable to the fact that the reaction flows from 1976 attitudes and realities not from the attitudes and realities of the year 2000.

It is fitting to close with the thought that should central data banks of the type discussed come to pass, many of the frustrations that face statisticians today will have disappeared; and that this new and greatly enriched source of information will tax the ingenuity of the statistician in producing more and better data in the public interest.

Footnotes

1. Philip M. Hauser, Social Statistics in Use, New York: Russell Sage Foundation 1975, pp. 5-13; also "Social Accounting" in Paul F. Lazarsfeld et. al. The Uses of Sociology, New York: Basic Books, 1967, pp. 839-846.
2. *ibid.*
3. Bureau of the Census and Bureau of Labor Statistics reports. For example, see U.S. Department of Labor, Bureau of Labor Statistics, Employment and Earnings, vol. 24, No. 7, July 1977. Washington, D.C.: U.S. Government Printing Office, 1977.
4. U.S. Congress, Senate Committee on Government Operations, Subcommittee on Government Research. Hearing on the Full Opportunity and Social Accounting Act (S.843). 90th Congress 1st Sess. Washington, D.C.: Government Printing Office, 1967.
5. Philip M. Hauser, *op. cit.*, Chapter 16.
6. For example, see Leo A. Goodman, "Guided and Unguided Methods for Selection of Models for a Set of T Multidimensional Contingency Tables," Journal of The American Statistical Association, 68 (1973) pp. 165-175. Leo A. Goodman, "A Note on Cohort Analysis Using Multiplicative Models and The Modified Multiple Regression Approach," Unpublished Manuscript, Department of Statistics, University of Chicago, 1975.
7. Richard Stone, Demographic Accounting and Model-Building, OECD Education and Development, Technical Reports, Organization for Economic Cooperation and Development.
8. U.S. Department of Health, Education, and Welfare. Public Health Service Health Resources Administration. Health Statistics Today and Tomorrow: A Report of One Committee to Evaluate The National Center for Health Statistics, Vital and Health Statistics - Series 4, No. 15, Washington, D.C.:

Government Printing Office, 1973.

9. Richard Stone, op. cit.
10. D. V. McGranahan, et. al. Contents and Measurement of Socio-economic Development, New York: Praeger, 1972.
11. ILO, International Recommendations on Labor Statistics, Geneva: International Labor Office, 1976.
12. Philip M. Hauser, "The Measurement of Labor Utilization" Malayan Economic Review, Vol. 19 No. 1, April 1974, pp. 1-15. Philip M. Hauser "The Measurement of Labour Utilization-- More Empirical Results." Mimeograph. International Statistical Institute, 1977.
13. Hong Kong, Indonesia, Malaysia, Philippines, Singapore, South Korea, Taiwan, Thailand, United States.
14. For example of study based on such data, see Peter Blau and O. D. Duncan, The American Occupational Structure, New York: John Wiley & Sons, Inc., 1964.