By: Frederick F. Stephan, Princeton University

Our purpose in meeting here this evening is two-fold. First, we wish to take a broad look at "social statistics" and consider the directions in which this realm of statistical activity may develop advantageously in the future. Second, but not secondarily, we seek to recall the many contributions to social statistics made during his career by Samuel A. Stouffer, who, were he alive today, would be up here fulfilling our first purpose more skillfully than I.

Perhaps no one can match Sam Stouffer as a pioneer in social statistics and social research. Few men in his generation have had his unique combination of far-sightedness and practicality, ingenuity joined to persistent and concentrated effort in pursuit of elusive truth, high standards of research craftsmanship, and restless intellectual curiosity.

These remarkable qualities of his, as well as his great contributions to the progress of social statistics and social research, are vivid memories for many of you who are here tonight and who knew him as an inspiring teacher, helpful adviser, stimulating colleague, and warm friend. For others who were not fortunate enough to have known him this well a brief review of his professional career may serve to show the variety of his interests and achievements. It will also remind us of his connection with many of the trends in social statistics which we plan to discuss together.

Samuel A. Stouffer's Professional Relations to Social Statistics and Social Research

Sam Stouffer became interested in a career as a social scientist after he had taken his bachelor's degree and, indeed, a M.A. in English. He was already well started on a career in journalism, following the footsteps of his father who was publisher of a newspaper in a small city in Iowa. Quite by accident, while on vacation, he read one of the classic books on sociology and, with a newsman's knack for spotting what is significant, he decided to become a sociologist. He entered the University of Chicago as a graduate student. It was there that he became interested in statistics and in related aspects of scientific method. Perhaps the greatest influence in this direction came from L. L. Thurstone who taught him the elements of statistical method and a good deal of psychophysics. W. F. Ogburn, who had just come to Chicago from Columbia to introduce courses in statistics in the Department of Sociology, also had a hand in arousing Stouffer's interest in statistics and shaping his outlook.

Stouffer took his Ph.D. in 1930 with a dissertation on "An Experimental Comparison of Statistical and Case Study Methods of Attitude Research."(1) This was before the day of modern attitude surveys; it was, in fact, a validation of a newly developed Thurstone scale on attitudes toward Prohibition. Characteristically Stouffer sought to do genuinely scientific work on a controversial subject. His report was a thorough presentation and scrutiny of his experimental data, restrained in its claims and conclusions, and carefully tested in many ingenious ways to make sure the mutual confirmation of the two approaches to the determination of attitudes was not spurious. It stands as a masterpiece in attitude research and a clear indication of the kind of research Stouffer was to do during his subsequent career.

Stouffer taught statistics briefly at the University of Chicago, in place of Ogburn who was busy with a gigantic study of social trends. He became convinced that the social research of the future would need men with better training in mathematics and mathematical statistics. Here again he pioneered. He applied for and was awarded a Social Science Research Council fellowship to study for a year with Karl Pearson and R. A. Fisher. After he returned from London to teach statistics in the Sociology Department at the University of Wisconsin he prepared a number of papers on the application of tests of statistical significance to specific studies in sociological research including one with Clark Tibbitts on a problem in criminal statistics involving the analysis of experience in predicting the risk of violation of parole. Thus he immediately applied what he had learned of mathematical statistics to current problems in sociology. He also encouraged his students to obtain a good grounding in mathematics and mathematical statistics. This was well in advance of the great interest that sociologists now show in such preparation.

Stouffer also had a deep concern with the improvement of the data of social statistics collected by state and federal agencies and he sought to extend their usefulness by supplementary projects for data collection. In the early 30's, he worked with the Registrar of Vital Statistics in Wisconsin on a study of trends in birth rates in that state based on data obtained by a well executed mail survey coupled with the birth registration records.(2) He also collected data by mail on marriage and divorce to supplement the incomplete reports then obtained by the Bureau of the Census(3).

It was in this period that Stouffer contributed two judicious discussions of current sampling practice.(4) He continued to have an active interest in sampling and in the modern methods which were just then gaining attention and being developed.(5)

It is not surprising that Stouffer was appointed a member of the Committee on Government Statistics and Information Services which was formed by the American Statistical Association and the Social Science Research Council at the beginning of the Roosevelt Administration to advise in the improvement and development of the statistical work of the Federal Government. He also served for 5 months as a member of the technical staff of the Committee and subsequently became a member of the staff of the newly formed Central Statistical Board.(6) Among the subjects to which he gave special attention were population, marriage and divorce, and vital statistics.

Stouffer became Editor of the Journal of the American Statistical Association toward the end of 1934. He almost became full-time Secretary of the Association as well. Had he done so the improvement of social statistics collected by government agencies might have remained his principal interest. However, he decided to return to teaching and academic research. He accepted a professorship at the University of Chicago but in his return to academic life he did not escape frequent calls to serve elsewhere as an adviser and researcher. He orgamized a series of studies for the SSRC on research that could be conducted to determine the social consequences of the great depression of the 30's. He was chairman of a Committee on Prediction of Personal Adjustment which studied logical and technical problems of prediction in such fields as vocational selection, school success, marriage and rehabilitation of criminals.(7)

He also made a study of the Population Census for the National Resources Committee, (8) prepared estimates of the effect on population growth of the reduction in the marriage rate during the depression of the 30's, participated in the International Population Congress in Paris in 1937, and served as a principal staff member in Gunnar Myrdal's comprehensive study of The American Negro. In the midst of all this, he devoted himself as fully as he could to an abstract problem in migration, namely how to explain the distribution of migrants from a common place of origin as they spread over the pattern of places to which they move. What excited him most was the theory he developed in terms of the pattern of the final location of migrants from all sources (intervening opportunities) rather than geographic distances (9) He returned to this problem with a refinement of his theory only a few years ago.

Had it not been for the Second World War, Stouffer might well have become absorbed in the development of mathematical theories in sociology, pioneering further and inspiring others to join him in exploration. Instead he became director of the professional staff of the Research Branch, Information and Education Division of the War Department, in 1941 and spent 5 years directing highly original and significant studies of the attitudes of soldiers and officers on a great many matters pertinent to programs and policies in the Army and Air Force. Immediately after the War he and a group of his colleagues prepared the well known and monumental reports on The American Soldier and Measurement and Prediction. His work in this great research enterprise coupled with his previous achievements led Princeton University to confer on him in 1948 an honorary

degree of Doctor of Science.

In 1946, Stouffer went to Harvard as Professor of Sociology and Director of the Laboratory of Social Relations. Here as elsewhere he led a busy and fruitful life, working with colleagues and students on many studies including social mobility and education, role conflicts, motivation, attitudes toward communism and civil liberties, and finally, with the Population Council, on motivation for the control of fertility in economically underdeveloped areas. He served as President of the American Association for Public Opinion Research and of the American Sociological Association. He also was chairman of the Committee on Measurement of Opinion, Attitudes and Consumer Wants appointed jointly by the National Research Council and Social Science Research Council.

It may appear that Stouffer's early interest in the improvement of the data of social statistics was displaced by his later emphasis on theory and his involvement in studies of controversial subjects and important public problems. This is not quite so; again he was pioneering. He pleaded for a shuttling back and forth between applied research and pure research as the best way social scientists could make progress since thereby they might more readily bring ideas into effective relations with data. He held out controlled experimentation as the ideal research method but also pleaded for rigorous statistical analysis, with more than two or three appropriate variables held constant or controlled, as a good approximation to the ideal when experimentation is not feasible. This calls for greater detail in tabulation, especially more cross tabulation, and places more exacting requirements on the collection of social statistics.

Sam Stouffer was humble in his view of his work. He stated in the opening paragraphs of <u>The American Soldier</u> that "Science...is cumulative, in the sense that a scientific achievement is most successful when it stimulates others to make the concepts and techniques it has used look crude and become obsolete as rapidly as possible." This will be very difficult for us to do but it is worth our greatest effort; he has given us a great challenge as well as a great example. We will remember him not only this evening but again and again as we continue the quest for better social statistics and work our way ahead in the many directions in which he pioneered.

Past Progress and Current Accomplishments

Sam Stouffer's unique contributions to social statistics were part of a larger development that started a long time ago with the first sporadic assessments of human and economic resources and the first quantitative determinations of risks and probabilities in human affairs. These two grand traditions followed different courses and at times were in conflict but repeatedly they were joined in ways that served individual and national needs and advanced the development of social science. Stouffer was particularly successful in blending data and abstract analysis.

The statistical tradition concerned with systematic collection of data on population, birth, deaths, marriages, migration, education, and many related aspects of family life and individual behavior has made remarkable progress in recent years considering the great difficulties that impede it and the substantial costs of this kind of data collection. About one third of the thousand pages of the Statistical Abstract of the United States are devoted to social statistics and these pages present only a very condensed summary in small type of the information which is published by the numerous collecting agencies of the Federal and State Governments and private organizations. We are all greatly indebted to the people who serve in such agencies to produce these data. They work day after day and year after year performing tasks that are often routine, uninspiring, and tiresome. From the top administrators to the clerks, interviewers, and machine operators their work is made more difficult by public indifference and inertia, by the natural human errors made by other people, and often by the lack of adequate facilities and appropriations. While some of them are lax in their performance, most of them are conscientious, highly competent, and motivated by the highest objectives of public service.

We have no time to review the progress made in each specific field of social statistics but any of us can find for himself what has been accomplished in fields of special interest to him by comparing the data available now with data available say 30 years ago. This will produce some shocks as well as some pleasant surprises; the progress has not been uniform and some fields have made little advance or even fallen back.

We can more readily take a general view of the progress of the last three decades. There are a number of prominent landmarks to guide us. At the time that Sam Stouffer was getting started in statistics our Association had a Committee on Social Statistics which arranged three joint sessions with the American Sociological Society and published the papers in 1930 in a volume in Statistics in Social Studies, edited by Stuart A. Rice, the chairman of the Committee.(10) A few years later Rice completed a comprehensive survey of social statistics for the President's Committee on Social Trends. Unfortunately his reports are not readily accessible though they were reproduced and circulated at the time. The various reports of the Committee on Government Statistics and Information Services provide many more landmarks, as do the pertinent parts of the Mills and Long report on the Statistical Agencies of the Federal Government prepared for the Hoover Commission and published in 1949.(11) These reports are dominated by problems of economic statistics but they attempt to deal comprehensively with the objectives, organization, and performance of statistical data collection. There are many other landmarks for specific fields within social statistics, two examples of which may be cited:

1) the mammoth collection of papers on demography edited by Hauser and Duncan in 1959 (12) and 2) the reports on criminal statistics by Beattie in 1959 and 1960.(13) Finally there is that masterly appraisal of the present situation and future prospects of social statistics presented to us by Bowman, Gall, and Rubin at Stanford last year.(14) I purposely avoid repeating what they have said so well and commend it to you for reading and rereading.

If I may be so bold as to attempt to summarize what we may expect to learn from a general survey of social statistics guided by these landmarks I will do so in the following terms:

1. <u>Maintain past achievements and</u> move ahead.

By dint of repeated efforts we have developed a great body of data in the various fields of social statistics, particularly in the field of population. Yet there are serious gaps and deficiencies still to be eliminated and there are many unrealized opportunities for improvement.

> 2. Get things done by cooperation within the existing organization of statistical activity.

Our decentralized systems of statistics in the Federal and State Governments enable us to collect many statistics as by-products of administrative procedures which otherwise would be too expensive or difficult to obtain. Yet this involves serious problems of coordination and comparability as well as great variations in the quality and performance of statistical work. Consequently it is very important to have the leadership and coordinating functions of such agencies as the Office of Statistical Standards in the Budget Bureau and similar offices in the State capitols plus a general willingness everywhere to work out problems cooperatively in the absence of central control of statistical operations. We can not readily change this system and we should do all we can to make it work well inspite of its weaknesses and inconsistencies.

3. Determine needs more specifically.

We need greatly a clearer formulation of the needs to be served by various sets of statistical data and more precise specification of the accuracy, detail, and other characteristics of the data that are essential to their effective use. Resources should not be utilized for the collection of unimportant data when they are needed for the collection of other data of critical importance. Unfortunately, the demand for data is expressed too much in terms of continuing what has always been done or in terms of pressure groups rather than clear demonstration of value and need. While the widespread use of social statistics for many kinds of individual decisions and actions in addition to their general use in the legislative, judicial and executive functions of national, state and local governments, makes it difficult to assess the need for particular kinds of statistical information, the inevitable choices made in the regular operations of the data collecting agencies deserve better guidance and justify a major effort to determine what is needed enough to warrant the cost of obtaining it.(15)

> 4. <u>Continue technical progress and</u> <u>extend its applications further</u>.

Great progress has been made in the technical aspects of data collection through standardization of classification schemes, improvement of the techniques of collecting data at the source, advances in data processing equipment, and the application of modern mathematical statistics in sampling, estimation, quality control, and analysis. The benefits of this progress are still to be fully realized in many realms of social statistics and there are still some technical problems to be solved in every field. Further advances may be expected but there are difficulties connected with maintaining comparability with similar data collected in the past, shortages of adequately qualified personnel, and practical details of operation to be worked out.

5. <u>Gain public understanding and</u> support.

Social statistics have an increasingly important part to play in a free-enterprise, democratic society such as ours as well as elsewhere around the world. While there are differences in opinion about the appropriate role of the Federal, State and local governments in some programs such as education and public welfare, it is clear that they are of concern to individual citizens and that citizens need dependable statistical information about them. It is true that statistics play a central part in other countries where state planning in socialistic and totalitarian governments require it for regulating their economies and the actions of their citizens. In a free society individual citizens need ample statistical information to regulate

themselves. Their associations, companies, and representative governments need statistical information on many subjects for which legislation, policy determinations, and administrative decisions must be provided to keep the free society free and healthy and to make it prosper. We should not resent the common jokes about statistics but we should concern ourselves about prevalent misunderstanding and unwarranted attacks as well as ignorance of the important role statistics play in a free democratic society.

> 6. Cultivate the interplay of data and ideas that advances science and enhances the value of applied statistics.

Finally, as Stouffer recognized so well, not only good statistical data but valid analysis and theorizing are necessary for the advancement of our understanding of human life in all its ramifications and, indeed, for the effective application to statistical information to important, even controversial, public problems. This means that our collection of data will be guided by the development of our ideas, our social sciences, just as in turn our ideas will be tested and developed by their fruitful association with data. I argued two years ago for greater articulation of theoretical concepts and statistical definitions.(16) Now I would only plead that social statisticians work more closely with other social scientists and that we again arrange joint sessions with them as the Committee on Social Statistics did 32 years ago.

A Program for the Future

As social statisticians we have many common interests and similar problems. Although we tend to be absorbed in our particular subdivisions and specialties we have profited from each other's experience in the past and borrowed in a neighborly way many useful ideas and procedures. We can and should do this increasingly in the future. In my opinion there are several rather definite ways in which we can do this within the Social Statistics Section and in the other opportunities we have to work together. 1) <u>Annual review</u>. We should continue the practice of making a comprehensive survey of recent developments in social statistics. This review should not be patterned too closely after Ray Bowman's paper last year or this paper of mine but should be adapted freely to the possibilities of broadening and strengthening our comprehension of our common purposes and how well they are being accomplished. Each year there may be a special emphasis that is appropriate for its time.

2) Planning and specifications. We should study together by individual research projects and special committee inquiries the complex problems of determining what is needed, what is feasible, and what should be chosen among the almost infinite variety of data that conceivably can be collected. The problems of deciding when data should be taken. how often, from what sources, in what detail, in what relation to other data, and with what degree of accuracy are familiar to all who have had some responsibility for initiating or for revising a system of data collection. These problems get more tightly interconnected as we elaborate the system and tune it up to meet more exacting requirements. Hence we can well justify devoting more time to preliminary planning and systematic analysis of data collecting operations than we could when everything was easier and simpler. Moreover our accumulating experience should crystallize and definite principles should emerge as we work on this problem of design and decision. The discussions at meetings of our Section should contribute to this part of the program at least in examining well chosen cases if not in effecting the final synthesis of experience.

3) <u>Development of new concepts and</u> reformulation of old ones.

We should explore actively the possibility of enriching the vocabulary of basic concepts we use in dassification and measurement. Examples are common in every field. The evolution of the concept of unemployment is familiar to most of us. Just now it is being examined again; it may be improved or clarified to meet better the current need for statistical information about this important public problem. In the 1960 Census of Population the concepts of dwelling unit and urban population were revised. We need better concepts of social class or socio-economic status, occupational mobility, social cohesion, education, health, and many other matters of importance both to social scientists and to the public. Our Section can well discuss this phase of the program and stimulate work on these problems elsewhere.

> 4) Development of new and improved means for measuring, classifying, and combining variables.

Most of our measurements have been made by simple counting such as size of family, years of schooling, or weeks of unemployment. The Committee on Social Statistics was interested in the possibility of constructing an index of social welfare. Stouffer was greatly interested in Thurstone's and Guttman's systems of scaling and in Lazarsfeld's latent structure analysis. These are only beginnings in the creation of measuring procedures that are needed for the analysis of social statistics and their more effective application to practical problems. The long-range program of the Section may well allot a good fraction of our time and attention to progress in measurement.

5) Appraisal and testing. As we become more specific about the needs to be met by our data, we become more concerned about their quality. We can no longer afford to be blissfully ignorant of their shortcomings but must test and validate to make sure they meet reasonable specifications and serve effectively the purposes for which they are collected. Too often in the past we have assumed that numbers, and the words that give them meaning, when printed in black and white or reported by cooperative respondents are correct as given.

We have only recently come to a clear realization that virtually all data are subject to some degree of inaccuracy and most collections of data are to some degree incomplete. Instead of assuming that a set of statistics is correct until proven guilty of error we should consider every set on probation and possibly inaccurate, incomplete, and imperfectly defined until sufficient evidence is produced for assuming otherwise. Many agencies still shrink from public discussion of the deficiencies of their data. Others, surprisingly, are not aware of these deficiencies. The progress made by the Census Bureau in its Post-Enumeration Survey and other tests of its data is an example of what can be done. (17)

The trend in this direction is evident in the papers on validation and response errors presented at recent meetings of the International Statistical Institute. (18) At times an independent examination of important statistical results is needed to give the public assurance, if they are trustworthy, or reveal their deficiencies more clearly, if they are not. The Section may well discuss this increasingly important aspect of all important statistical reports.

Regular appraisal practices are the mark of professional maturity. Those agencies are most advanced which risk embarrassment in determining the accuracy of their statistical data and reporting their determinations frankly and fully to the users. They are to be commended in leading the way that others should follow. It might be appropriate for our Section or our Association to give public recognition from time to time to accomplishment in assessing data and informing the public of the result.

Conclusion

Five major parts of a long range program have been suggested. Clearly they are all interrelated; certainly others should be added. In the spirit of Sam Stouffer's own endeavors and the grasp he had of what still needed doing, as well as in continuation of the efforts of all the many others who contributed to the progress of social statistics in the past, we should strengthen our present efforts and, where necessary, start new work so that our statistical work can better fulfill the expectations we have for it and better serve our free society.

FOOTNOTES

- (1) <u>Abstracts of Theses, Humanistic</u> <u>Series, University of Chicago</u> <u>Press, 1932, Vol.VIII, pp. 263-270.</u>
- (2) "Trends in the Fertility of Catholics and non-Catholics," <u>American</u> <u>Journal of Sociology</u>, Vol. 41, 1935, pp. 143-166.
- (3) Samuel A. Stouffer and Lyle M. Spencer, "Recent Increases in Marriage and Divorce," <u>American</u> <u>Journal of Sociology</u>, Vol. 44, 1938-39, pp. 551-554.
- (4) "Sociology and Sampling," in L. L. Bernard, Ed., <u>Fields and Methods of</u> <u>Sociology</u>, New York, Ray Long and R. R. Smith, 1934.

"Statistical Induction in Rural Social Research," <u>Social Forces</u>, Vol. 13, 1935, pp. 505-515.

- (5) Frederick F. Stephan, "History of the Uses of Modern Sampling Procedures," Journal of the American <u>Statistical Association</u>, Vol. 43, 1948, pp. 12-39.
- (6) <u>Government Statistics</u>, New York, Social Science Research Council, Bulletin 26, 1937, 174 pp.
- (7) <u>The Prediction of Personal Adjust-</u> <u>ment</u>, Social Science Research Council, Bulletin 48, 1941, 455 pp.
- (8) "Problems of the Bureau of the Census in Their Relation to Social Science," Section 7 of <u>Research</u>
 <u>A National Resource, Report of the Science Committee to the National Resources Committee, Vol. 1,</u>
 <u>Washington, U. S. Government Printing Office, 1938, pp. 195-232.</u>
- (9) "Intervening Opportunities: A Theory Relating Mobility and Distance," <u>American Sociological Review</u>, Vol. 5, 1940, pp. 845-867.
- (10) Philadelphia, University of Pennsylvania Press, 1930, 222 pp.

- (11) Frederick C. Mills and Clarence D. Long, <u>The Statistical Agencies of</u> <u>the Federal Government</u>, New York, National Bureau of Economic Research, 1949, 201 pp.
- (12) Philip M. Hauser and Otis Dudley Duncan, Editors, <u>The Study of Popu-</u> lation, An Inventory and Appraisal, Chicago, University of Chicago Press, 1959.
- (13) Ronald H. Beattie, "Sources of Statistics on Crime and Correction," Journal of the American Statistical Association, Vol. 54, 1959, pp. 582-592. Ronald H. Beattie, "Criminal Statistics in the United States -1960," Journal of Criminal Law, Criminology and Police Science, Vol. 51, 1960, p. 49.
- (14) R. T. Bowman, Alexander Gall, and Israel Rubin, "Social Statistics: Present Conditions, Future Needs and Prospects," in <u>Proceedings of</u> <u>the Social Statistics Section</u>, 1960, Washington, American Statistical Association, 1961, pp. 74-81.
- (15) An example of such effort is the Appraisal of Census Programs, Report of the Intensive Review Committee to the Secretary of Commerce, Washington, U.S. Government Printing Office, 1954, 119 pp.

- (16) "Relations of Some Social Science Concepts to Statistical Data," in <u>Proceedings of the Social Statis-</u> <u>tics Section, 1959</u>, Washington, <u>American Statistical Association,</u> 1960, pp. 170-171.
- (17) <u>The Post-Enumeration Survey: 1950</u>, Technical Paper No. 4, Washington, U.S. Government Printing Office, 1960, 93 pp.
- (18) P. Depoid, "Rapport sur le degre de précision des statistiques démographiques," <u>Bulletin de l'Institut</u> <u>International de Statistique</u>, tome XXXV, 3ème livraison, 1957, pp. 119-230. Henri Bunle, "Sur les erreurs entachant les statistiques et des recensements de la population et de l'état civil," <u>loc. cit.</u>, pp. 283-288.

See also the many papers on nonsampling error, validation, and error assessment in the volumes of the <u>Bulletin</u> which contain papers presented at the Stockholm, Brussels, and Tokyo meetings.