Robert Parke and Eleanor Bernert Sheldon Social Science Research Council

In its discussion of the various types of users of the products of the federal statistical system, the President's Commission on Federal Statistics gave first place to policymakers. Consideration of the use of statistics by policymakers was followed, in order, by the uses by program managers, evaluators of government programs, exploratory research, industry and trade associations, state and local governments and, finally, the public. This sequence of presentation accords fairly well with the priorities assigned to audiences for statistics in the budget justifications of federal statistical agencies--priorities which in turn reflect the expectations of appropriations committees.

In the same report, the Statistics Commission quoted, with evident approval, a government official, as follows: ". . . the government is simply not good at defining what it wants to do in terms of needed social science research . . . the government, in general, can only articulate the area in which it needs information."² The juxtaposition of these comments--priority to the data needs of policymakers, and the limited ability of government to articulate its data needs--is the basis for our topic, which has to do with the developing relationships between social statistics and public policy.

The comment of the anonymous government official must, of course, be gualified by recognition of a number of outstanding statistical programs designed to directly serve policy purposes. The unemployment statistics, the surveys of the aged population which provided much of the empirical basis for Medicare, and the statistics on poverty are a few examples of the many that come to mind, which serve directly to measure needs and to provide the factual basis for government programs to meet those needs. These are a few illustrations of social statistics in the service of policy. At their best, they represent well-planned surveys and analyses in response to an expectation of policymakers that a major program shift is in order. These types of statistical programs require, for their policy usefulness, the active participation of policymakers in determining their substantive content, since such programs are visualized as contributing to the solution of social problems. In these circumstances, the policymaker's perceptions of need are properly accorded priority in decisions to undertake the programs.

Our comments are addressed to the primacy accorded the expressed information needs of federal policymakers in shaping the social statistics enterprise generally. In our view, this set of priorities distorts the enterprise in two key ways. First, by emphasizing the expressed needs of policymakers it undervalues the potential contribution of social statistics to the <u>definition</u> of policy problems as well as to their solution. Problems that can be anticipated often go unrecognized until it is too late, and statistics capable of assisting in the resolution of problems often are based on an excessively narrow definition of what the problem is. Second, by emphasizing the needs of federal policymakers, the prevalent priorities undervalue the role of other constituencies for social statistics, which play important roles in the policymaking process. Such constituencies include state and local governments and substantial segments of the public at large.

These considerations lead us to suggest that the federal activity in social statistics should display more initiative in communicating the meaning of the data, through development of analytical time series and of measures which derive their significance from models of social processes, through projections of social data, especially for areas below the national level, and through social statistics chartbooks which draw upon all sources of competent data, and other means of communication. We also believe that federal statisticians should do more to cultivate their role as one key element in a complex of interrelated activities which includes the best of privately conducted social research and which serves all levels of government, private associations and businesses, and the public at large. By such efforts, statisticians can contribute to policy by assisting in the perception of problems, the anticipation of problems, and outlining the need for "adaptational" as well as "manipulative" policy strategies.

Policy Knowledge

Of policymakers, the Statistics Commission said: $^{\rm 3}$

"Policy-making, as the term is used by this Commission, is a set of activities including definition of problems, identification of solutions, and choice among them. The definition of a problem is a political-technical process involving translation of public <u>perception</u> of a situation into remedial legislation."

The Statistics Commission differentiated policymakers from program managers, whose data needs it "broadly characterized as performance data, describing resource uses and outputs of the programs administered." A related distinction is made by Biderman in his differentiation of the several levels of social information.⁴ The summary is by Peter Henriot:⁵

". . Biderman notes that there are three distinct uses of data which should not be confused mentally or organizationally. The lowest or most specific level of data is 'information'--data intended for use at the operational level. The next level of data is that designed for overall administration and management purposes, and is termed 'intelligence.' The third and highest level of data is termed 'enlightenment,' and is designed for contributing to public understanding and formation of general policy . .

"An example of the various types of data outlined by Biderman would be helpful. Take the example of police records:

"--'Information,' operational data, might be the number of complaints, of routine pickups, of nightly investigations, etc.

"--'Intelligence,' administrative data, might be the number of squad cars available, of business and residences within a given area, of insurance rates on homes, of persons on probation, etc.

"--'Enlightenment,' public policy data, might be the number of criminal acts as related to characteristics of the neighborhood, the percentages of expenses spent on prevention, apprehension, and correction, etc."

Biderman's formulation views "public understanding" as a key function of statistics for policy use. For, in his words, "In a democratic social order . . . policy knowledge and public enlightenment are closely related." The same view was expressed by Otis Dudley Duncan at the 1972 meetings of this association. Our views concur with theirs. A society that relies as much as ours does on the activity of private associations and businesses, and on a high degree of participation by all levels of government and the public at large in the development of public policy, cannot afford to let its statistical agenda be limited to the perceived information needs of federal public policymakers.

Broadening Perception: Health

If the policy uses of social statistics are seen primarily as a matter of meeting the perceived information needs of administrators, we limit policy-relevant social statistics to questions raised by those whose responsibilities may cause them to overlook some important questions, and to those matters covered by the authority, competence, and traditions of administrators. No one will deny the importance of such matters. But this is no justification for letting the statistician's agenda be governed by the perceived information requirements imposed by a social engineering approach or letting it be limited to "policy-manipulable" variables--that is, those subject to the control of the agencies responsible.

Let us illustrate the distinction we are making. In 1973, a book appeared entitled <u>Health</u> <u>Status Indexes</u>. The book presents the edited results of a conference held in October 1972.⁶

It appears from the papers and the published discussion that the intended audience for the book consists of decisionmakers whose chief stock in trade is the delivery of health services, and that the purpose of the work represented in the book is to devise summary measures of population health status which will assist decisionmakers in the implementation of a policy whose main outlines have been decided upon. For example, there are papers on estimating needs for health services from household survey data; papers discussing how to incorporate the prognosis characteristically associated with a condition into a measure of current health status; and studies empirically assessing the value that patients place on an impaired life, relative to the value they place on a surgical procedure with a specified probability of killing them or providing a full cure.

The fruition of these efforts should contribute substantially to rationalizing the mix of health services and their deployment. But it will contribute very little to the definition of health problems. That has been and is being accomplished in part by statistical studies of smoking and lung cancer, drinking and drug abuse, tension, and auto accidents. To a health establishment committed to dealing with those conditions that can be dealt with by providing health services, such studies as these may appear to be of limited use for decisionmaking. But some health agencies support such studies, appreciating that, if they provide little direct assistance to decision, they make an immense contribution to cognition, that is, the intellectual mapping operation that plays such a critical role in the definition of the problem. The role of statistics in assisting the redefinition of health problems is probably most spectacular in the case of smoking, partly because it led to new departures in health policy.

Defining the Problem: Manpower Policy

The 1973 Manpower Report of the President contains a chapter which represents a new departure in the attempt to define the problems with which manpower policy must cope. Entitled "Population Changes: A Challenge to Manpower Policy," the chapter describes recent national population changes stemming from the wide fluctuations in national birth rates over the past 30 years, and outlines their implications. Drawing in part on materials developed by the Population Commission, this report reviews the probable impact of these changes on the female labor force, low income and minority groups, the geographic mobility of workers, unemployment, GNP, consumption patterns, schools, health services, and other matters of policy concern. The intent is to anticipate the effects of population changes and to sensitize manpower policy to the requirements these changes impose. The editors of the Report acknowledge the major contribution to the chapter made by the Bureau of Labor Statistics.

The contribution here is to invite the attention of policymakers to developments not under their control that may alter the way they wish to conduct affairs that are under their control. Such guidance serves policymakers by assisting in the development of what Biderman calls "adaptational strategies." Biderman constrasts adaptational strategies (that is, coping with the consequences of developments in realms the policymaker has no control over) with "manipulative strategies," which deal with matters the policymaker can control or influence. For example, our interest in weather reports is not eliminated by our inability to manipulate the weather; we need them because they "facilitate our adaptations to the phenomenon."

Social intelligence of this sort is also illustrated by statistics showing pre-Watergate declines in public expressions of confidence in government, as well as declines in the proportion of the electorate voting--a shift which reputable analysts have attributed to an increasing public sense that their vote makes little difference. Whether or not a policymaker can develop shortterm responses to such developments, they are (as Otis Dudley Duncan noted at these meetings last year) surely something he ought to know about.

More on Defining the Problem: Modeling Social Processes

In speaking of the policy relevance of social statistics, we wish to emphasize its contribution to the definition of problems as distinguished from their solution. This contribution is pervasive, but is particularly clear in the case of statistics incorporated in models of social processes. For example, Featherman and Hauser have presented a cogent statement about the policy relevance of their work on pathanalytic models of occupational mobility. This work, using the survey resources of the Census Bureau, will provide comprehensive trend data on the degree of inequality of opportunity in the United States and the changing importance of ethnicity, family background, education, and other factors in occupational advancement. Featherman and Hauser say:

". . . there are five areas in which we think our study can contribute to policy formation.

"(a) In the assessment of widespread beliefs about equality of opportunity and factors affecting it. We think that the 'debunking' function of OCG findings should not be underrated. For example, findings from the 1962 study cast a great deal of doubt on the utility of concepts of a 'cycle' or 'culture' of poverty, and more specifically on the suggestion that family instability was the major source of white-black achievement differentials . . .

"(b) In locating and defining the problems of specific population subgroups . . .

"(c) In providing an overall model of the process of social and economic achievement which can serve as a frame of reference for discussions about specific aspects of that process . . .

"(d) In providing a set of current trend estimates on major features of the process of social achievement . . .

"(e) In improving the measurement of processes of social and economic achievement. We think that our investment in improved measurement techniques--and also in new techniques of data reduction and analysis--will be quite as important for the policy applications of our findings as for their purely academic uses. We think our proposed innovations can contribute both to the quality and legitimacy of the information we can supply and to the development of methods for future replications and other related studies."

Halliman Winsborough has developed and presented a method for preparing annual estimates of age, period, cohort, and education effects on earnings by race.⁸ It is intended to illustrate the potential uses of the Current Population Survey basic data files for social trend analysis, through models of social change. This is a use for which the CPS is uniquely adaptable given its design and size, the period it covers (since 1947), and the care with which over-time consistency of statistical practice has been ensured. Mason and Hodge have proposed a related analysis, with additional features which recognize that period effects are not uniform throughout the country, and which seek to relate the CPS data to It data on political and institutional changes.³ would be a significant loss to social statistics if these proposals were to founder on the public unavailability of such files for the period before 1968, and the question of whether the earlier data will ever be usable.

Projecting Consequences: The Supply of Teachers

The social statistician knows something about how the society works that is capable of informing and directing the deliberations of policymakers. However, with all their competence, industry, and integrity, statisticians remain fairly passive when it comes to exploiting what they know to assist in the anticipation of policy problems. Problems that can be anticipated too often go unrecognized until it is too late.

Let us illustrate with an implausible case. Take a closed population for which high-quality current estimates of the population by age are available, and for which good current statistics on school enrollment and teaching manpower are regularly published. Should not such a population be expected to avoid major imbalances in teacher supply relative to demand? One might expect so, but that is not what has happened. Instead, we have, over the past 15 years, seen a drastic shift from a situation of undersupply to oversupply of teaching personnel. We do not refer here to the debate over whether graduate schools have generated an oversupply of Ph.D.'s, but rather to the situation with respect to the supply of persons qualified to teach elementary and secondary school. Briefly, estimates and projections published by the National Education Association in 1971 showed that in the late 1950's the nation's teachers colleges were producing half the annual number of additional teachers needed for primary and secondary schools, in 1968 they were producing just about the number needed, and by 1972 they would be producing double the number needed.¹⁰

This major dislocation in the labor market deserves to be recognized as one of the most anticipatable dislocations on record. The baby boom of the 1950's and the decline in the absolute number of births starting in 1961, in combination with the rise in teachers college enrollments, made a severe oversupply of teachers in the early 1970's next to inevitable. One needed no birth projections to anticipate it; recorded births would have shown it.

We have not studied the process by which this dislocation was permitted to develop. Hence we are not in a position to state at what juncture the breakdown in foresight and planning occurred. We suspect that a significant role was played by the fact that the statistics which provided a basis for anticipating the problem were known at the national level but not at the state level at which primary and secondary school policy is made. The decline in the population under five years old was obvious in national population estimates by the mid-1960's but the state population estimates and projections of the Census Bureau contained no age detail, and no middecade census was available to show it. The Office of Education's projections of school enrollments and teaching graduates were published for the U.S. only, while unpublished projections for states were available on request. Thus the states were left to their own devices. Local school districts took school censuses, but they must always hedge against the substantial effects of net migration on their enrollments, and in any case they have nothing to say about teacher supply. In sum, it seems as though state policy with respect to graduating teachers generally operated on the assumption that each state could export any excess, although it would have been obvious, had anyone looked, that the sum of such implied exchanges netted out to a substantial national oversupply.

More on Projecting the Consequences: Declining Metropolitan Areas

Projecting national trends onto local areas is, of course, a highly uncertain business. No national agency, or any other, is in a position to forecast with precision which states or localities will suffer what degree of dislocation. But that <u>some</u>, perhaps most, areas will experience it is obvious, and it should be within the resources of a national agency to give statistical description to a few plausible projections, any one of which will yield odds that a state or locality risks substantial dislocation.

A development of this sort is in the offing with respect to the population of metropolitan areas. In the 1960's only one large metropolitan area experienced an absolute decline in population. Yet it is certain that if the national birth rate maintains its current low level or continues to drop, large numbers of metropolitan areas will experience substantial absolute de-clines in population in the near future.¹¹ This is the inevitable result of a national convergence toward zero population growth in combination with the pattern of intermetropolitan migration flows. With the drying up of nonmetropolitan sources of migration to cities, intermetropolitan migration dominates. The pattern of intermetropolitan migration is that migration is heavily focused on a few rapidly growing metropolitan areas. Most areas have maintained positive growth rates despite net outmigration, thanks to the excess of births over deaths. As that excess declines, we may expect absolute population declines in many areas.

The consequences and the policy implications of such a development are not at all clear. But they will never be clear until the prospect is addressed, and it will not be addressed until it is stated. Our own view is that, on balance, such a development is probably to be welcomed. But others may not welcome it if it takes them by surprise, and there may be real, as distinct from symbolic adjustments, which are required.

The fact is we have no experience of this particular development. We have experienced population decline in railroad towns, in agricultural marketing centers, in half the counties in the nation (mostly rural) and in several metropolitan central cities. But (the trivial decline in Pittsburgh aside), we have never experienced it in an entire large metropolitan area. What will be the consequences for land values? What will be the consequences for the amortization of fixed costs, such as roads and schools? What adjustments to decline, if any, are required? No one knows, because no one has asked. And no one will ask until the prospect is stated in numbers. Such a statement does not require that an agency estimate rates of projected decline for individual areas. But, distributions of areas by growth and decline, given plausible assumptions as to national population growth and patterns of intermetropolitan migration? Surely that would be an easy and a legitimate extension of recognized projection techniques.

Critiquing the Policy: "Balanced Growth"

We have spoken of the contribution of social statistics to the definition and the anticipation of problems addressed by policymakers. Such statistics can also play a role in critiquing positions set forth in policy discussions. Much of the discussion of "balanced growth" in recent years has been founded on the notion that the United States has too many of its people in cities, and that excessive urban growth could be countered if migrants found opportunities elsewhere, in new towns or in "growth centers," that is, communities removed from the major urban centers, which show promise of rapid growth if deliberately stimulated by the government.

The professional skills of statisticians give them no more qualification than anyone else to comment on the judgment that we have too many people in cities. Nonetheless, statisticians are acquainted with a range of data that can be used to put this judgment in perspective. Furthermore the plausibility of a growth centers policy as a means of attentuating population growth in major urban centers can be, and was, subjected to specific examination by demographers, economists, and statisticians on the staff of the Population Commission, a body whose responsibilities included the preparation of recommendations on population policy for the United States.¹²

With respect to the judgment that we have too many people in cities, statistics were brought to bear in support of the following points:

1. The United States is a metropolitan nation. Most families live in metropolitan areas; most births, deaths, and migrations take place within or between them. Today 70 percent of the American people live in SMSA's.

2. Metropolitan population growth is a basic feature of the social and economic transformation of the United States, that is, the transition from an agrarian, to an industrial, and now to a service-oriented economy. Metropolitan growth is the form that national and regional population growth have taken.

3. The most rapid growth in the decade of the 1960's occurred not in the largest areas, but in areas with populations of between one and two million. At the same time, as metropolitan populations have grown, many central cities have been losing population.

4. The growth of metropolitan population is now mainly due to the excess of births over deaths in metropolitan areas and to immigration from abroad; not to rural-to-urban migration. A projection prepared by the Census Bureau, at the Population Commission's request, showed that if current trends continue, other parts of the United States will contribute four million migrants to the metropolitan population between now and the year 2000, while <u>immigrants</u> will add about 10 million.

The foregoing considerations suggested that the trend toward bigness, if undesirable, could not be substantially checked except as national growth is slowed or stopped. This was specifically examined by construction of a hypothetical growth-centers policy to learn how much the growth of the large metropolitan areas might be reduced if the growth of smaller, less congested places were stimulated. The results showed that if these places were stimulated to grow by 30 percent each decade from 1970 to 2000, their population might absorb about 10 million of the growth which is otherwise expected to occur in areas of one million or more, assuming the 2child national projection. However, these large areas would still increase by 70 million persons under the same national projection.

These and related findings speak directly, and unfavorably, to the policy choice formulated by the President's National Goals Research Staff in its 1970 report:¹³

". . . we need to decide on whether or not we will adopt a deliberate strategy to encourage internal migration to negate the forecasts of ever-growing urban congestion in a few megalopoli. A viable option for such an alternate strategy is a policy of encouraging growth in alternate growth centers away from the large urban masses, coupled with a complementary effort of the use of new towns."

The Population Commission's findings say, in effect, that at the present stage of the nation's demographic development, there are no redistributional substitutes for lower national population growth rates. Whatever may be said for a "growth centers policy," the reduction of metropolitan population growth is not one of them. The policy contribution of social statistics in this instance is to point out that some policies under consideration are live options and some are dead ones. The statistical analysis does not tell us which of a variety of real-world options is preferable. What it can do, if heeded, is to narrow the range of consideration to those options that make sense in terms of a scientific understanding of how, in relevant respects, the world works.

Communicating the Meaning of the Data

The federal statistical establishment annually disgorges immense amounts of social data, only a fraction of which gets distilled as knowledge about social structure or social change. Salutary exceptions to this rule include the Census Bureau's Current Population Reports on changes in the social and economic conditions of blacks, Mexican-Americans, youth, and city populations. In addition, recent publications on "Age Patterns in Medical Care, Illness, and Disability," "Health Characteristics of Low-Income Persons," and other topics, have shown a recognition by the National Center for Health Statistics that its data, traditionally focused on specific health conditions and on types of health care, can be reorganized to describe important social trends.

In general, however, federal statistical reports devote a great deal of attention to how the data were put together (as they should) and very little attention to what the figures mean. This is not, of course, to suggest that statistical reports ought to contain policy comment; they should not. But, if the careful documentation of methodology is the primary responsibility of the statistician, surely a close runner-up is the responsibility to utilize accepted analytical techniques and methods of data presentation which will enable the data to tell the story that will

not be told in the absence of analysis.

An example is the information on income shares which has been published by the Census Bureau now for a quarter of a century. These are the figures showing what fraction of total personal income is received by the one-fifth of families receiving the smallest incomes. The analysis and presentation of data by income shares has a number of features of a good indicator. It is established over a long period of time with annual readings. It has been worked over by statisticians to a point where they have a pretty good idea how it acts, and why. It packs much data into few figures, and it answers a pointed question about an important aspect of our collective life.

Expectation of life is another figure with many of the same properties. An interesting feature of this figure is its seeming simplicity and its actual complexity. It is a hypothetical figure based upon the mathematical manipulation of population and mortality data. The simple statistics in this instance are the numbers of persons in the population classified by age, sex, and color, and annual numbers of deaths classified the same way. Only when the demographer or actuary processes these numbers through his mathematical model do the figures stand up and tell a story. Moreover, the message they convey is relatively unambiguous compared with such ostensibly simpler measures as the crude death rate. Expectation of life has the added virtue that it is manipulable as a component in mathematical models of the population, which renders it easily used in analyses of population growth and in population forecasting.

Possibilities for development of analytical measures may be illustrated by the prospect for an advance in knowledge about family dynamics, using available statistical resources. In a compendium of statistics on changes in the American family, Abbott Ferriss juxtaposed official statistics on divorce, classified by year of marriage, against the original number of marriages, so that one could track the divorce experience of people married in a given year.¹⁴ Kenneth Land used Ferriss' data as the basis for development of a mathematical model describing what Land calls the "divorce trajectory" of marriage cohorts.¹⁵ Land's work accomplishes an admirable summarization of many observations in terms of two numbers: The proportion of a marriage cohort ultimately divorcing, and the velocity at which divorces occur over the cohort's lifetime. Furthermore, it expresses the results in terms amenable to manipulation in a model of the family life cycle. Land's work on this topic addresses an area which has thus far been relatively innocent of the sophisticated analytical treatments that have been developed for mortality. It may be taken as a model for the development of a system of measures which can do much to enhance knowledge and understanding of family processes using data that are now available in the federal government.

The resources and techniques exist also for

annual series of indexes of the level of residential segregation by race, and other topics. What is missing in these and other areas is a mandate and support for the development of analytical measures of major social trends comparable to the emphasis given economic trends.

Short of new analytical measures, much can be done to communicate the meaning of the masses of data already produced. Five years ago, Beverly Duncan showed how much more can be learned from educational attainment data by retabulating them in a cohort format.¹⁶ There now exist powerful techniques of multivariate analysis which render obsolete much of the traditional inspection and oversimplified summarization (means, percentages, standardization on one variable out of many) which has characterized the analysis of complex cross-tabulations of social data.

In economic time series, our official agencies think nothing of adjusting the data to eliminate seasonal influences; they have done it for so long that their readers expect it, and the press relays reports of seasonally adjusted data routinely. The analogous operation for complex cross-tabulations of social data is readily available through multivariate analysis. We have recently received a Norwegian analysis of health data by the new techniques.¹⁷ There is no reason why U.S. agencies should not follow suit. Failure to do so results, often enough, not merely in failure of communication, but in distortion of meaning.

Summary and Implications

We believe it is a mistake to identify the usefulness of social statistics for policy with the question of whether statistics meet the perceived information needs of federal policymakers. Such an identification probably assumes too great an ability by policymakers to identify their data needs. It clearly undervalues the contribution of statistics to the definition and anticipation of problems as distinguished from their solution. And it undervalues the role that other constituencies for statistics play in the policymaking process. We have urged, in particular, that statistical agencies apply accepted statistical techniques for the purpose of communicating far better than they now do, the meaning of the data they collect. And we have urged that the data be developed to a point where they more readily illuminate the problems of states and localities as well as the federal government.

These observations have several implications for the conduct of statistical work both inside and outside the federal government.

1. <u>Analytical Function</u> -- The communication function involves a good deal besides adequate staffing of public information offices, valuable as that may be. It is in part, and irreducibly, a statistical function, involving the preparation of analytical measures, the exploitation of models of social processes, and the compilation and analysis of time series from all competent sources to describe a topic. Such work requires a continuity of attention that only an organization can provide, but it requires an organization which is committed to measurement and analysis as distinguished from information collection.

Both of us have served on the U.S. National Committee on Vital and Health Statistics. We recall the discussions from time to time in that committee, of the need for a "two-tiered" statistical agency, one part devoted to information collection, the other part to analysis and to communication in the sense in which we have used that term. The separation of function recognizes the fact that in an information-collection agency, the imperatives of information collection and data production invariably take priority over analysis whenever there is a choice, as there always is. The separation of function also recognizes what the Statistics Commission recognized--the disinclination of data-producing agencies to devote time and attention to exploiting the potential of data other than their own. The three compendia by Abbott Ferriss, on changes in education, the family, and the status of women, show what can be accomplished by statistical publications organized by topic rather than by agency.¹⁸ The Statistical Abstract, OMB's forthcoming compendium on social indicators, and topical publications by the Census Bureau's Population Division cover the output of several agencies. But these publications do little or nothing beyond reporting data as reported by the originating agencies, they do little standardization, and they undertake no analytical measures. In general, Ferriss' lead has not been followed. The rule seems to be that the communications mission of data-producing agencies is defined by the datacollection mechanisms they operate. In the analytical tier of a two-tiered agency, the mission is not defined by a mechanism but by a set of topics or problems.

2. Involvement of University Research -- The analytical function need not necessarily be housed in a government agency, although there is much to be said for it, including access to the staff who designed the surveys, supervised processing, and designed record layouts and tabulations. Beyond this, government agencies have longevity, and the accumulation of statistical expertise, knowledge, lore, and professional tradition that goes with it. That kind of longevity is essential, but is rarely found outside government. The National Bureau of Economic Research has it, and economic statistics was the beneficiary of the years of tender loving care the Bureau devoted to the analysis of economic time series. With the uncertainties of project funding, few university-based research centers are in a position to provide the continuity of attention necessary to a similar endeavor in social statistics.

If that continues to be the case, there are nonetheless critical ways in which the involvement of university researchers in the social statistics enterprise needs to be cultivated. The analytical effort we have suggested presupposes that government agencies will actively draw upon the capabilities of university-based researchers to develop concepts, measures, and models of social processes. We have given a few illustrations of opportunities in this regard, and there are many more.

Closely related to this is the need to actively provide researchers with the materials they need for the analytical work from which developments in conceptualization, measurement, and modeling emerge. Some of this, of course, is provided by regular publications of the statistical agencies. But increasingly, scientific development presupposes access to basic data files, as illustrated by the work of Featherman and Hauser, of Allen, and of Winsborough, exploiting recent files of the Current Population Survey. The utilization of such files, while scientifically rewarding, can be extremely arduous, as users of the CPS files well know. Except for the public-use samples of the decennial censuses, it is generally the case that the micro-data files from government surveys are formatted and documented to the extent necessary to enable an in-house programmer who is experienced with the survey to use them, and no farther. This places tremendous burdens on the user, greatly restricts the number of users, and effectively impedes much of the analytical work which is needed to enhance the meaningfulness and communicability of the agencies' own data outputs.

3. Exploitation of Data Resources -- Much that we have said falls under the general rubric of exploiting existing data resources to improve knowledge about social change. In closing, we take this opportunity to return to a theme developed earlier. That is the utilization of data resources to assist in the statement and analysis of problems affecting areas below the national level. We have elaborate models of economic processes at the national level. Rudimentary beginnings have been made on a few models of social processes at the national level; that work needs to be encouraged and continued, and the validated results of it ought to be sought out and utilized in government social statistics programs.

At the state and local level, where a great deal of social policy is made, we know of few examples of subnational model building which lends significance to local data.¹⁹ Model building for any one local area is beset by the difficulty that any area is related to others in a multiplicity of systems. Hence it is necessary to study systems of areas. A suitable place to begin would be with data capable of describing, for a set of communities, changes in a variety of facets of community life--employment, industrial composition, population and vital statistics, education, crime, and so forth, and to begin to seek structure in the way the data array themselves and in the relationships of change on one characteristic to change on others.

Over a period of time, such work might yield relationships which would lend far greater meaning to local observations than is possible at present, for it would permit the explanation of local phenomena by reference to the behavior of the systems of which a community is a part. Figures on the migration patterns of a community, its educational level, or conceivably its crime rate, could be approached nct only by asking "What was it last year" or "What is the measure for similar communities" but by asking "What did you expect?" in terms of the measure's relationships in a system of measures.

Why should not the tape files of the County-City Data Book be exploited for such a purpose? In these files we have quinquennial readings since 1952 for counties, cities, and metropolitan areas, on a wide range of key characteristics, which ought to be applicable to the problem we have posed. To be sure, a good deal of work would be required to ensure consistency of items from one year to the next, to document inconsistencies, and for other preparatory tasks. But the fruit of such efforts should be substantial. If this approach is judged not optimal, others have been proposed for consideration. In any case, it is the task which interests us at the moment, rather than the means for its accomplishment. The task is the development of measures and of models that lend them meaning, and for this purpose to exploit the resources at hand. Here, the role of the statistical agencies is to facilitate the exploitation of the data, and to incorporate in their own programs the validated results of the analysis such an effort makes possible. The policy contribution is to assist in the understanding of social processes and the definition and anticipation of social problems.

Footnotes

1. President's Commission on Federal Statistics, Federal Statistics, Vol. I, Washington, D.C., 1971, Chapter 3.

2. <u>Ibid</u>, p. 89, quoted in Otis Dudley Duncan, "Federal Statistics --Nonfederal Statisticians," American Statistical Association, <u>Proceedings of</u> the Social Statistics Section, 1972.

3. Ibid, p. 82.

4. Albert D. Biderman, "Information, Intelligence, Enlightened Public Policy: Functions and Organization of Societal Feedback," <u>Policy</u> <u>Sciences 1</u>, 1970, p. 217-230.

5. Peter J. Henriot, "Public Policy Implications of Social Indicators Research and Development," unpublished paper presented at <u>Symposium on Social</u> <u>Indicator Research in an Urban Context</u> (Ottawa: Ministry of State for Urban Affairs, 1973).

6. Robert L. Berg, ed., <u>Health Status Indexes</u>, Chicago, Hospital Research and Education Trust, 1973.

7. David L. Featherman and Robert M. Hauser, "Design for a Replicated Study of Social Mobility in the United States" in <u>Social Indicator Models</u>, edited by Kenneth Land and Seymour Spilerman, Russell Sage Foundation, forthcoming.

8. Halliman H. Winsborough, "Age, Period, Cohort,

and Education Effects on Earnings by Race -- An Experiment with a Sequence of Cross-Sectional Surveys" in <u>Social Indicator Models</u>, edited by Kenneth Land and Seymour Spilerman, Russell Sage Foundation, forthcoming.

9. W. M. Mason and R. W. Hodge, <u>A Proposal to</u> <u>Use the Tapes from the March Current Population</u> <u>Surveys to Study the Changing Socioeconomic</u> <u>Location of Blacks and Women in the United States</u> (draft, 1973).

10. National Education Association, <u>NEA Research</u> <u>Bulletin</u>, 1971, Vol. 49, No. 3.

11. William Alonzo, "The System of Intermetropolitan Migration Flows," in U.S. Commission on Population Growth, Research Papers, Vol. 5, <u>Popula-</u> tion Distribution and Policy, Sara Mills Mazie, ed., Washington, D.C: G.P.O., 1973.

12. U.S. Commission on Population Growth and the American Future, Final Report, <u>Population and the American Future</u>, Chapter 3, Washington, D.C.: G.P.O., 1972

13. President's National Goals Research Staff, Toward Balanced Growth: Quantity with Quality, Washington, D.C.: G.P.O., 1970, p. 60.

14. Abbott Ferriss, <u>Indicators of Change in the</u> <u>American Family</u>, New York: Russell Sage Foundation, 1970.

15. Kenneth C. Land, "Some Exhaustible Poisson Process Models of Divorce by Marriage Cohort," Journal of Mathematical Sociology, Vol. 7, 1971, p. 213.

16. Beverly Duncan, "Trends in Output and Distribution of Schooling" in Eleanor Bernert Sheldon and Wilbert E. Moore, eds., <u>Indicators</u> of Social Change: Concepts and Measurements, New York: Russell Sage Foundation, 1968, Chapter 8, pp. 601-672.

17. Asmund Langsether, "Indicatorer for Helsetilstanden I Befolkingen: En videre analyse av Statistisk Sentralbyras helsundersøkelse. INAS Report No. 71-3 of the Institute of Applied Social Research. Oslo, February, 1972.

18. Abbott L. Ferriss, <u>Indicators of Trends in</u> American Education, 1969; <u>Indicators of Change in</u> the American Family, 1970; <u>Indicators of Trends</u> in the Status of American Women, 1971. New York: Russell Sage Foundation.

19. James G. Anderson, "Causal Models and Social Indicators: Toward the Development of Social Systems Models," <u>American Sociological Review</u>, 33 (1973), p. 285.