I

PLANNING OF SAMPLE SURVEYS FOR SOCIAL RESEARCH IN UNDERDEVELOPED COUNTRIES

Chairman, Leslie Kish, University of Michigan

Survey Sampling and Implementation for Development Programs — Roe Goodman, Bureau of the Census

Obtaining Data on Consumption and Saving in Underdeveloped Countries — E. Scott Maynes, University of Minnesota

SURVEY SAMPLING AND IMPLEMENTATION FOR DEVELOPMENT PROGRAMS

Roe Goodman, Bureau of the Census

The main considerations in this paper are those relating to the implementation of sample survey in underdeveloped countries, hence this term is to be regarded as the key word in the title. Since certain factors are important in relation to all phases of sample surveys, the discussion on the whole is broader than that of survey sampling alone. However the illustrations are taken from the field of sampling.

It appears that the use of sample surveys for development is currently enjoying a wholesome growth. Various factors are tending to stimulate the use of this statistical tool in underdeveloped countries. Outstanding examples of current sample surveys for fact-gathering purposes are agricultural censuses based on samples which are now in progress in several Middle and Far Eastern countries, as a part of the FAO-sponsored World Census of Agriculture. Expansion of survey research designed specifically for policy making or analytical purposes can now be anticipated.

Considering the intrinsic difficulties in carrying through sample surveys in underdeveloped countries there is no doubt that the motivation of responsible survey officials is crucial for the success of this activity. More than this it is my observation based on experience in several countries that adequate motivation of survey officials cannot be taken for granted. To quite an extent the duties and responsibilities required from survey statisticians and field directors has little precedence in these countries. And only those who can see sufficient reasons for high standards of execution can really be expected to implement surveys well.

A major contention of this paper then is that in order to develop and maintain adequate motivation for survey work in underdeveloped countries great emphasis must be placed on the scientific character of sample surveys. I am stressing this point here in order to facilitate an orderly presentation by first showing the utility of this approach and second showing its necessity. The persons whose motivation could be affected by emphasis on scientific work are in general those who are college graduates, a substantial part of the staff of sample survey organizations being in this category.

To amplify a little on the notion of science and surveys, the idea is to emphasize that work on sampling and sample surveys is scientific work and that those participating in it are a part of the large body of persons engaged in scientific undertakings in the world today. This emphasis does not imply that the results from a survey are necessarily going to be used scientifically (or even used at all). Rather it is that the scientific approach permits those actively involved to feel that they are capable of uncovering new information of unquestioned validity. They can thus identify themselves with something that has prestige now and is likely to be even more promising for the future.

An emphasis upon science fits well with the idea expressed at the beginning of this paper — that proper implementation is a key consideration in survey work in underdeveloped countries. Obviously, if survey work is to be called scientific the sampling and other phases as well must be carried out strictly according to plan. Since in underdeveloped countries there has been little tradition to see to it that instructions are followed such an orientation becomes particularly appropriate.

When it comes to sampling, the textbooks usually state that the practicability of the design of a probability sample is a cardinal consideration. This matter, however, is often dismissed without specific amplification. In underdeveloped countries especially it must be borne in mind that practicability means that the plan must produce samples which can be identified correctly and easily by the field worker.

What has been said is merely that, in order to be considered scientific, the sample as implemented in the field must be a true probability sample. The very extensive work that has been done on sampling efficiency is also highly scientific. However, an efficient sample which makes great demands on the field worker for its proper implementation may be carried through so badly that the undertaking can no longer be called scientific. In my work in underdeveloped countries I have found that an insistence upon practicability often requires a sacrifice of sampling efficiency. The emphasis on science highlights the need for adhering to a thoroughly practicable, even if inefficient, plan at all costs.

A few examples may serve to clarify what has just been said. The first example, though beginning with an inefficient sample, is one in which the efficiency of the sample can be improved over time. The case is one of repeated surveys, twice a year, in order to estimate the acreages under various crops in East Pakistan. For this purpose a sample of fields is selected by a multi-stage sampling. Data on land use of sample fields are then obtained and recorded by Agriculture Department employees on the basis of personal observation. For the beginning years the idea was to select fields in clusters of about ten contiguous

fields within the villages selected at the preceding stage. Good maps, though not up to date (often forty years old), were available for the use of the field workers and it had been found repeatedly that randomly selected fields could be located on the spot with the aid of the maps alone provided sufficient care were used. In order to more or less force the field worker to stay in the proper vicinity for enough time to verify the field identification the use of the clusters of fields was decided upon. Of course, based on analyses carried out in many parts of the world there could be no doubt that the selection of fields in smaller clusters or scattered individually throughout the village would have been much more efficient statistically. But sober reflection led to the use of the relatively large clusters anyway in order to provide greater guarantee that the field work would be done strictly in accord with the plan. Naturally it is intended in cases of this kind that steps would be taken to increase the efficiency of the sample by gradually reducing the size of clusters. This would be done year by year as evidence accumulates that the field workers are performing well and that fields would be located and reported upon properly in smaller clusters. Although these surveys generally use the same sample from one survey to the next, there is no reason why a transition cannot be safely accomplished from clusters to individual fields within sample villages over time.

The second example is of an inefficiency of a different type in that it involved extra work at one step in the sampling process but finally led to the selection of a sample of normal efficiency. In this case (a family expenditure survey in Santiago, Chile) blocks had been selected with probabilities proportionate to the number of segments, the segments having been allocated to the blocks on the basis of estimated numbers of households in each block. The particular step involved the dividing of the selected blocks into the pre-assigned numbers of segments. For this purpose the field workers were asked to proceed around the blocks, going into and out of 'cites' and 'pasages' as they came to them and listing the addresses of all dwelling places one by one. Subsequently the listings were utilized in the office to establish readily identifiable points of division for the segments and one of the segments thus defined was selected at random within each block. An alternative would have been to have sent persons out to determine and note down suitable points of division within blocks without the preparation of any listing. The Director-General of the National Statistical Service, however, believed that the second plan would have led to the using of many unsatisfactory division points between segments. In this connection it is worthy of note that the original listing did not eliminate the necessity for a subsequent visitation to the families within selected segments, which was an additional preliminary to the selection of individual families for the family expenditure survey. The detailed listing of 100 or more addresses per sample block had the sole purpose of making sure that the segments as defined could always be accurately identified in the field and was an extra task carried out for this purpose.

As a final example, there is the case of the sample of villages which was selected in Pakistan for enumeration in the country's 1960 (first) Census of Agriculture. As selected and used this sample was one of the largest samples ever taken anywhere in the world but it could have been much smaller with a more efficient design. In planning the sample it was concluded that the only possible means of doing sub-sampling within selected villages would be to place complete reliance upon officials within the different districts (1) to see that accurate lists of households were obtained, and (2) to make random selections from these lists. At the same time it was well known that to the district officials the idea of sampling was relatively new and the necessity for really rigorous work in such matters was entirely unrecognized. The inescapable conclusion, therefore, was that if subsampling were to be attempted it could not be expected that the households finally to be enumerated would be a true probability sample of the rural households of Pakistan. Accordingly, the design that was decided upon was a stratified random sample of whole villages. The entire work of sample selection was done at a central point for each of the two Provinces, a total of about 12,000 villages being selected in all. Instructions were then issued that within each selected village all cultivators (farmers) were to be enumerated, which meant that although the Census was based on a sample the work of an enumerator within a village was in all major respects exactly the same as it would have been if a complete Census were being taken. The sample was, of course, a grossly inefficient one (however, a much more 'efficient' operation than a complete Census would have been) but the rigorous adherance to a scientifically selected sample seemed a goal worthy of the price which was paid.

To put these examples of inefficiencies in their proper light it should be mentioned that obtaining the necessary funds to do good survey work, however inefficient the sampling design, does not constitute a major problem. In fact, the examples given are instances in which nationals in the countries either suggested or at least agreed wholeheartedly with the procedures used.

If the problem is not one of money where then does the problem lie? Fundamentally

the problem is that the type of motivation among survey workers which is generally effective in the U.S. is not present in sufficient measure in underdeveloped countries. It is probably pointless to go into extensive discussion of these matters especially since a number of the factors involved are rather negative in character. However, it may be useful to mention the points which I believe have the greater relevance. The enumeration of these problems does not imply that persons within the countries are not aware of them or that no progress is being made in resolving them.

Some of the critical points are:

(a) Those at the highest levels who have authorized a survey may not really expect to see useful results emerge (but have approved the undertaking for a variety of other reasons) and consequently they do not go to any particular pains to see that personnel, equipment and so forth are available at exactly the right time. Moreover, they often underestimate the difficulties and more or less assume that progress will be made in due time once approval has been given.

(b) Officials at a high level may be keenly interested but due to lack of professional training they are not especially impressed with the need for rigorous methods and consequently they do not support the decisions of their well-trained subordinates and do not make adequate provision for necessary personnel or other resources.

(c) An enthusiasm for the work which might develop on the part of technical staff is stifled by knowledge that no matter how much they may be interested their superiors will not give them any particular encouragement or credit for a job well done.

(d) Professional workers, and non-professional too, have good reason to believe that no matter how hard they apply themselves to the tasks at hand they will receive no advantage in the way of promotion over those who take their responsibilities lightly. It does not matter that a new day may be dawning and they may in fact receive faster promotion; what is controlling is that they do not think they will.

(e) Officials at higher levels may have ordered the inclusion of too many questions or too difficult questions and the professional worker may not dare to advise his chief of the impossible character of the assignment given to him.

(f) Rigidity in personnel policies often means that relatively unsuited persons are placed in responsible positions in survey organization (perhaps due to seniority) and their better suited colleagues are therefore aware of the inability of the organization as a whole to function properly under the existing arrangements.

(g) Administrative ineptness either in the way of delayed approval of expenditures or inopportune transfers of personnel may place serious pitfalls in the way of the normal progress of a sample survey and require that the well-trained official do his work under particular difficulties.

Isolated instances can be found of even more extreme obstacles which rear themselves in the way of successful surveys. From one point of view all these obstacles may be seen as inefficiencies which will be overcome in due time just as sampling efficiencies can and will be overcome also. The effects of the administrative deficiencies taken all together may, however, be sufficiently extensive to create a feeling of futility and discouragement among professional workers.

What is therefore needed is a survey design which will enable those who want to make scientific progress to achieve limited objectives which will meet the test of scientific scrutiny. Even if a large part of an early survey is inevitably a failure, for some of the reasons mentioned, key professional workers can with sufficient imagination salvage information on a few items or gain methodological insights which point the way to future progress. It is necessary primarily that the desire to follow scientific procedures be firmly established in the minds of the professional staff at all stages of the work. Required is the discovery of procedures which are sufficiently simple to permit scientific standards to be upheld.

The manner in which the relatively inexperienced but well-trained workers in many countries around the world are striving to do good scientific work in what to them are pioneering days is worthy of whole-hearted support. There is an ever-growing number of people in these countries who have received good training at home or abroad, which is a portent of constructive development for the future in the area of sample surveys.

OBTAINING DATA ON CONSUMPTION AND SAVING IN UNDERDEVELOPED COUNTRIES*

E. Scott Maynes, University of Minnesota

* As will be evident, this paper rests heavily on my 1958-59 experience as Economic Consultant to the National Council of Applied Economic Research (NCAER), New Delhi, India. While with NCAER I directed--in collaboration with Mr. I.R. K. Sarma--the Delhi Saving Survey. My debts are great to Mr. Sarma, my chief collaborator; Dr. P. S. Lokanathan, Director-General of NCAER who wholeheartedly supported the project; and Dr. H. F. Lydall of Oxford University Institute of Statistics who induced me to go to India in the first place.

Any criticisms of persons or organizations are made in an impersonal, constructive spirit.

Preparation of this paper was made possible through generous grants of released time and funds by the School of Business Administration and the Graduate School of the University of Minnesota.

This paper has been improved as a result of criticisms by Professors Marcel K. Richter and Norman J. Simler and Mr. J. S. Uppal.

Introduction

My comparative advantage with respect to this subject arises from a year spent in India, where I directed a saving survey. The particular survey, called the Delhi Saving Survey,[1] was a pilot survey designed to ascertain by personal interviews the net worth, income, saving and other relevant financial data of households in Old and New Delhi. Whether my comparative advantage is substantive or only apparent, you will be able to judge only after exposure to this paper. Therefore I use this lever to obtain your attention.

One can only talk about what one knows. I know the problems of surveys in Delhi well, those of all urban India less well, and those of other underdeveloped countries almost not at all. You should be forewarned as to the relatively narrow base of experience on which this paper is based. However, knowledge is relative and this is my defense. Furthermore, in discussions with economists, survey research specialists, sociologists, and anthropologists, I gain the impression that the problems faced in the conduct of sample surveys are similar in many underdeveloped countries.

This paper has two parts. Part I analyzes potential sources of response and nonresponse errors in Indian economic surveys and discusses techniques developed to minimize such errors.

Part II focuses specifically on the Delhi Saving Survey and presents evidence on the degree of success achieved in minimizing response and nonresponse errors in this survey.

PART I

Response Errors: Source and Techniques to Minimize Them

Any discrepancy between the answer written down in the questionnaire and the true value for that item constitutes a response error. Given a perfect question, perfect cooperation on the part of the respondent, accessibility to the required information on the part of the respondent, and perfect communication between the respondent and the interviewer, we would have zero response errors: the reported value would exactly equal the true value. Nonresponse, by which we mean failure to provide any information at all on a particular item, may be viewed as a particular, extreme case of response error.

Classification precedes analysis and therapy. In the case of response errors, a set of classifications developed by John B. Lansing has proved particularly useful.[1] Lansing distinguishes three sources of response errors:

1. Motivational Factors. The respondent may not wish to give to the interviewer the correct answer to the question. He may be indifferent as to whether he conveys the correct information. He may even wish to conceal or to distort the facts.

2. Failure of Communication. The respondent may not understand what information is required of him. Or, even if the respondent understands and tries to communicate the correct information, the interviewer may not understand and record correctly what the respondent is trying to tell him.

1/ The first results of the Delhi Saving Survey, as well as estimates of aggregate saving for India, were presented by Dr. P.S. Lokanathan at the December, 1959 Meetings of the American Statistical Association in a paper entitled, "A Study of Saving in India."

A comprehensive report on the Delhi Saving Survey is reported to be in press.

1/ Classifications and their description are quoted directly from a draft memorandum by Lansing. Cf. John B. Lansing, "An Analysis of Response Errors in Economic Surveys," (Ann Arbor Survey Research Center, 1959, mimeographed), pp. 12-13.

3. Inaccessibility of the Information to the Respondent. The information may be available to him only with varying degrees of difficulty. The respondent may recall the information easily and accurately, but not all the desired information is likely to be so accessible to him. There will be psychological forces at work which will influence the rate at which different items are forgotten and may lead to distortions in what is remembered. Resort to financial records may be difficult: they may be lost, they may be poorly organized or incomplete, and they may be physically remote.

Analyzing the record of American economic surveys, Lansing--in a paper being given at these Meetings--concludes that response errors in American surveys are attributable primarily to motivational factors. 1/ The analysis of Indian economic surveys which follows suggest that response errors in Indian economic surveys are attributable more frequently to failures in communication. If true, this is a relatively happy situation because these are often the easiest failures to correct.

The discussion which follows records the experience of the Delhi Saving Survey (DSS). In the planning of this survey, our strategy was to employ the best of American survey techniques. 2/ If these proved unsuccessful, an alternative approach would have been attempted, stressing qualitative rather than quantitative data. This second step proved unnecessary. Thus, the techniques actually employed follow rather closely those developed in the American Surveys of Consumer Finances.

Inevitably one asks: How widely applicable are "lessons" gleaned from a survey conducted in Delhi? Certainly they have no applicability at all to village India, a feudal, partially monetized economic society with built-in defenses against outsiders: most observers agree that new and different techniques will be required to obtain good-quality data from villages. Though villages account for 80 percent of the population, urban areas account for 55 percent of aggregate saving (if the Reserve Bank estimates are accepted). In the large urban areas - Calcutta, Bombay, Madras, and Hyderabad-Segunderabad - the lessons of Delhi should have utility. Their applicability to smaller cities will be tested when the NCAER completes its National Urban Survey of Saving, now in the planning stage.

1/ John B. Lansing, "Studies of Validity in Reporting Financial Data."

2/ The results of methodological explorations in the collectinn of saving data sponsored by the Inter-University Committee on Consumer Behavior were not yet available. A report of one of these was published last year: Robert Ferber, Collecting Financial Data by Consumer Panel Techniques. A Pilot Study. (Urbana: Bureau of Economic and Business Research, 1959).

Motivational Factors.

Essentially, the objective of any interview motivationally is to get the respondent to accept as his own the objectives of the interviewer (and thus, indirectly, the objectives of the directors of the study). When successful, the respondent optimally will want to give full and accurate information concerning his income, assets, purchases, savings, financial plans, and attitudes. The theory of interviewing has been set forth fully by Kahn and Cannell in their recent volume. 1/ Here we shall sketch only the basic ideas.

In selecting and in training interviewers and in designing the questionnaire and other aspects of the study, our objective was to create and support certain positive psychological forces and to allay or eliminate certain negative psychological forces. Enumerated below are these positive and negative psychological forces. In each case simple statements describe first the relevant theory and then its operational implementation:

1. Theory: Respondents will accept the goals of a study on the basis of a "temporary friendship" with the interviewer. Several actions were taken to facilitate the establishment of such friendships. Implementation: (a) Attractive, intelligent, "friendly" interviewers were hired. Reflecting the relative underemployment of university graduates, all interviewers held M.A.'s (equivalent to an American Bachelor's degree); (b) These interviewers represented all major "communities" (Hindus, Sikhs, all regions of the country). Insofar as possible, likes were sent to interview likes--that is, Sikhs to interview Sikhs, etc.; (c) The initial questions in the questionnaire dwelt with the financial problems of the sample family and were designed to indicate the interviewer's interest in the particular family he or she was interviewing.

2. Theory: People like to feel important; by making people feel important, they will be willing to disclose the required information. Implementation: (a) Interviewers emphasized the "honor" of being selected in the sample; (b) Numerous opinion questions were inserted in the questionnaire. The interviewer stressed the

1/ Robert L. Kahn and Charles F. Cannell, The Dynamics of Interviewing (New York: Wiley, 1957).

importance of these questions.

3. Theory: Respondents are concerned with their own economic self-interest; they may feel that an awareness of the economic situation of their particular group resulting from a survey report will affect government economic policy and hence their personal economic well-being. Implementation: Interviewers told respondents that--though this survey was not government-sponsored--the results would be reported to the government and would probably be taken into account in government planning.

4. Theory: On an altruistic basis, respondents will wish to take actions to help their government, party, or community. Implementation: Interviewers made various appeals to respondents such as: "This survey will help the government in drawing up the Five-Year Plans." "This is the first saving survey to be conducted in an Asiatic country." 1/

5. Theory: People like to be entertained; if the entertainment value of an interview is sufficient, they will be willing to provide information not directly entertaining to themselves. Implementation: (a) Questinns of opinion (thought to be more "entertaining") and fact were alternated in the question sequence; (b) intrinsic interest in question content themselves afforded considerable entertainment value.

Negative psychological forces were treated in an analogous manner:

1. Theory: Respondents may fear that in disclosing financial information to interviewers they place themselves in possible financial jeopardy through (1) theft, (2) "adverse" use of data by tax or other governmental authorities, or (3) by swindlers. Operational Counteraction: (a) All interviewers carried evidence of the authenticity and reputability of the sponsoring organization and themselves. For example, each interviewer carried (1) an identity card with his picture, and a statement both in English and Hindi telling of the organization and its purposes; (2) a newspaper clipping telling of the Delhi Saving Survey; (3) a "brochure" of the National Council of Applied Economic Research, impressively printed on coated paper and containing full-page pictures of the nationally prominent persons constituting the governing board of NCAER; (4) a letter from the Director General of the National Council (a national figure in his own right) asking the cooperation of the respondent and assuring him of confidentiality; (5) in addition to attesting to the integrity of the organization and its representative, these documents were designed to convince the respondent of the importance of this particular survey.

2. Theory: The greater the number of activities in which the respondent is involved, and the greater the degree of involvement, the less willing he will be to devote time to a survey interview. Operational Counteraction: (a) As just mentioned, we hoped to convince the respondent that the survey was important; (b) fortunately, the busiest respondents were those who best understood the usefulness of economic research in general and particularly the appropriateness of a saving survey.

Obviously, the strength of these and other psychological forces--not discussed here--varied from one respondent to another. Our better interviewers--like good interviewers anywhere--successfully gauged the strength of the relevant forces and adopted his (her) appeal to them.

Favorable Motivational Factors in the Delhi Environment. Happily for the conduct of economic surveys there are a number of favorable motivational factors in the urban Indian environment just waiting to be exploited appropriately. In the first place, there is a strong Indian tradition of hospitality. Very few of our interviewers were ever denied immediate access to a household. Thus, to anyone who has ever attempted to take survey interviews in a large American city, New York for example, Delhi would be a strange but pleasant survey environment. In the second place, much financial data which is regarded as private information in Western society is regarded in India as "public" information. A question frequently asked of me early in an acquaintance was, "How much do you make?" Similarly, in an interview with a high official of the city of Delhi, his reaction after a very brief explanation of the purpose of the interview was: "Go ahead, I have nothing to hide."

The relative lack of competing entertainment in the Indian environment confers on a sample interview an entertainment value of considerable degree. Movies are relatively expensive; there is no television; radio ownership is considerably less frequent than television ownership in this country.

Indians tend to have a high regard for authority figures. This regard, coupled with the fact that there are fewer authority figures in India, gives a greater leverage value to

1/This is not quite true. A Ceylon Survey of Consumer Finances" was undertaken in 1952-5; however, the data on saving thus obtained were judged "unreliable." Cf. Margaret G. Reid, "Survey of Ceylon's Consumer Finances." American Economic Review, December, 1956, pp. 956-964.

particular authority figures, such as those on the Governing Board of the National Council of Applied Economic Research. Thus, among high-income respondents, particularly, it was not uncommon to have the respondent assent to an interview after learning who the members of the National Council's governing board were.

The spirit of nationalism still runs strong. Thus, a suggestion that this was the first saving survey to be undertaken in an Asiatic country evoked a positive response and positive cooperatinn on the part of many respondents.

Failure of Communication.

There are two issues here: (1) whether to use a questionnaire or schedule; (2) how to handle the problem of multiple languages.

The use of schedules rather than questionnaires has been traditional in Indian survey work.1/ Typically, the study director lists in a "schedule" the particular items he wishes to measure. Then the "investigator" records the required information on the schedule. The exact meaning of concepts and the treatment of problem cases is exhaustively specified in a manual which the investigator is to read and to master. These manuals often run to a thickness of an inch. They never specify the phrasing of the question to be asked.

Under two conditions factual information can be validly collected by use of the schedule: (1) when the subject matter under investigation is extremely simple; (2) when the interviewer is extremely articulate and has attained complete mastery of the concepts involved in the investigation. Certainly, economic surveys seeking to ascertain saving, income, and consumption can scarcely be classified as simple. Further, for the operational conditions likely to be met in underdeveloped countries it is extremely doubtful that the second two requirements--verbal facility and mastery of concept--would be fully met on the part of many interviewers.

These considerations led us to adopt the use of a questionnaire in the Delhi Saving Survey. The use of questions specified in advance places the responsibility for the verbal specification of concepts squarely on the shoulders of the study director. Further, it insures that this verbal specification is uniform from one respondent to another: all respondents react to the same question (assuming that interviewers have been well drilled in the routine of reading question wording exactly as they appear in the questionnaire.) Pre-specification of questions releases the interviewer from the difficult problems of specifying questions on the spot and permits him to concentrate on the important problem of rapport-building, data-recording, and evaluation of consistency and completeness of information provided. These tasks are sufficient to keep him fully occupied.

The Indian people speak 14 major languages. The cost of translating and printing questionnaires in 14 different languages (as would be required in an all-India survey) is undoubtedly very great. Nonetheless, in an extension of the principles just developed above, it is surely the way of wisdom to make central translations and to require that interviewers utilize these standard translations. Our objective, of course, is to measure constant concepts, regardless of language or sub-culture. By utilizing one standard translation for each language, translating errors will be kept constant and minimal, if the job is carefully done and reviewed by a number of people.

In the Delhi Saving Survey the respondent could choose to reply to an English or Hindi questionnaire, whichever was easier. These two questionnaires provided access to the majority of the Delhi population inasmuch as Hindi and Punjabi are orally identical, differing only in written script. Errors of translation were surely more frequent and less consistent in the few cases where it was necessary to take interviews with respondents who spoke neither Hindi, English, or Punjabi.

Accessibility of Information to the Respondent.

In order for the respondent to be able to provide the interviewer with the desired information, this information must be accessible to the respondent himself. Fortunately, several factors work to increase accessibility in India as compared to Western countries. In the first place, the relatively lesser affluence of Indian families means that Indians are likely to have more perfect knowledge of what they possess. The bane of American economic surveys--that families have so many possessions that they cannot describe them all with accuracy--is met much less frequently in India.

1/ The Rural Credit Survey, conducted by the Reserve Bank of India, represents a partial exception. Both schedules and pre-worded questions were employed. However, in the places where pre-worded questions were used, their phrasing tended to reflect more the requirements of the survey directors rather than the understanding of the village cultivators to whom they were addressed. Two examples:

1. "Have you in your possession as creditor under usufructuary mortgage any land?"
2. "If you have no insurance policies, is it because:
   (1) of superstition?
   (2) there are no local facilities?
   (3) it is too complicated?
   (4) there are difficulties in paying premia regularly?
   (5) the money cannot be readily realized?

In many places, these questions were translated into local (perhaps more colloquial) languages, though in some places (Mysore, Kerala, Madras, Assam, parts of Andhra) this English version was used. Cf. Committee of Direction, All-India Rural Credit Survey, Volume III Bombay: Reserve Bank of India, 1956), Appendix, ff. 963.

In the second place, the structure of the Hindu undivided family facilitates accessibility. Operationally the identity of the head of the undivided family is hardly ever in doubt. 1/ Further, the head is almost invariably conversant with the finances of each individual member of the household, even when that member maintains his finances separately. Thus financial interviews can be taken with one person whose knowledge of family finances is great.

As is well known, gold and silver jewelry constitute an important secondary financial reserve for Indian families, particularly women. To an outsider it was interesting that the monetary value of such assets, possessed by 75% of Delhi families, was known rather exactly, being given either in rupees (value) or in tolas (weight). That their value is known so well reflects their importance as financial reserves.

Memory errors are important in India, as in the U. S. As just noted, the relative non-affluence of Indians minimizes this problem with respect to financial assets. Nonetheless, to further reduce errors from this source, interviewers were instructed to ask respondents to consult records wherever available. On the basis of a rough tabulation, perhaps 20% of respondents consulted one record or another; in addition, 30% of respondents were "very sure" that the figures given were accurate.

The estimation of saving may be affected greatly by memory errors, depending on the particular approach employed. Though conceptually saving may be measured by either (1) calculating the difference between income and consumption expenditures, or (2) by obtaining the algebraic sum of changes in assets and liabilities, the income-minus-consumption approach gives rise to serious memory errors: income tends to be underestimated and consumption to be over-estimated. This was conclusively shown in the experience of the Bureau of Labor Statistics' Survey of Consumer Expenditures, 2/ and in the family budget studies conducted by the Cambridge Department of Applied Economics in Cambridgeshire, England. 3/

The income-minus-consumption approach has been employed in several, small scale financial surveys in India with unhappy results.

1/ In planning the DSS, much effort was spent in developing criteria and questions which would identify the "head" of the household. The time was misspent: operationally, it was sufficient to ask household members who the head was. Disagreements on this point were rarely encountered.

2/U.S. Bureau of Labor Statistics, "Survey of Consumers' Expenditure in 1950: Interpretation and Use of the Results," Monthly Labor Review, 1952, pp. 425-428.

3/ Dorothy Cole and J. E. G. Utting, "Estimating Expenditure, Saving and Income from Household Budgets," University of Cambridge, Department of Applied Economics, Reprint Series No. 127, 1957.

In the Delhi Saving Survey, saving was obtained by measuring the net change in assets and liabilities. The memory bias problem dominated the decision to employ this approach, but its adoption was also influenced by the desire to gain better cooperation by minimizing the length of the interview.

PART II

Evidence of Success

Having analyzed sources of response errors in the urban Indian environment, it behooves us to provide evidence as to the correctness of the diagnosis and the efficacy of the prescribed cure. This is the task of this section.

Validation. How well did the Delhi Saving Survey succeed in its stated objective, the measurement of saving? Ideally, to pass on validation of the saving data obtained, we should require for each household in the sample access to written records from which we could independently estimate the saving of that household. The comparison between this estimate of saving based on written records and the estimate secured through the Delhi Saving Survey questionnaire would give us an accurate measure of response errors arising from all of the major sources--motivational factors, failures of communication, inaccessibility of information to the respondent. Unhappily, this technique of validation cannot be used: The required written records are non-existent and/or non-available to us.

Unfortunately, even the cruder test of comparing aggregate saving as measured by the survey with aggregate saving as measured by other methods is not practicable: No other estimates exist of aggregate saving of households in metropolitan Delhi. In this connection it is worth noting that, because of the extreme skewness of the distribution of saving, surveys constitute feeble and inappropriate means of obtaining aggregates. 1/

1/ An arithmetic example will show why this is so. Assume: (1) It is wished to estimate aggregate income from a survey of 3,000 households, drawn from a population of 57,000,000 households; (2) That, of the 1,000 "income millionaires" in the population, one falls in our sample; (3) That the mean income of non-millionaire households in our sample is $5,000.

The inclusion of the millionaire will increase our mean from $5,000 as follows:

$$\bar{X} = \frac{1\ (\$1{,}000{,}000) + 2{,}999\ (\$5{,}000)}{3{,}000} = \$5{,}332$$

This would increase the aggregate from:
57,000,000 ($5,000) = $285 billion
to 57,000,000 ($5,332) = $304 billion
(an increase of $19 billion)

The chances of a millionaire falling in the sample are 1/19. Similar considerations apply to saving. A memorandum by Benjamin Tepping in 1953 first alerted a number of us to these possibilities.

Since direct validation is impossible, we must fall back upon other and less satisfactory indications of success. I have organized these under three headings: (1) Partial Measures of Validity, (2) Evidence of Willingness to Cooperate; (3) Evidence of Success in Motivating Respondents.

Partial Measures of Validity. For an accounting period, one can measure the sources and uses of funds of a household. Sources--where the money came from--include income earned, withdrawals from bank accounts, proceeds from the sale of assets, increases in debt, etc. Uses--where the money went--include rent paid, increases in bank accounts, decreases in debt, money loaned out, etc. In the Delhi Saving Survey sources were completely covered while many, but not all uses of funds were covered. Logically, in this situation, sources of funds must exceed uses of funds by a "reasonable" amount, this depending on such things as family size, income level, etc. For each interview in the survey this sources-uses check calculation was made; the number of negative or "unreasonable" residuals was very small. In these few cases interviewers revisited the respondent, ascertained what was wrong, and corrected the interview. This test is asymmetrical; it screens over-estimates of saving. The Delhi Saving Survey passed this crude test with flying colors.

Interviewers' ratings of the accuracy of saving estimates (very good, good, middle position, poor, very poor) provided a second measure of partial validity. Subjectively, they presumably take account of such factors as record consultation, perceptions of such things as the accuracy of recall, deliberate efforts to conceal or distort, the internal consistency of all information given in the interview.

Since interviewers in this survey participated in the subsequent summarization of the data and since they were highly trained and closely supervised more than the usual confidence, in my judgment, can be placed in these ratings.

As Table I indicates, three-quarters of households gave saving estimates rated "very good" or "good" by interviewers. By contrast, only 7% were rated "poor" and less than 1% were rated "very poor." As among occupations, the self-employed businessmen contributed the largest proportion of middle position and poor ratings. This is not surprising in view of the intrinsic difficulty of estimating the "business saving" of these households. Table 2 classifies households by amount of saving and the interviewers' estimate of accuracy. The large, positive savers contributed the largest part of the middle position or "poor" ratings. Whether this is attributable to (1) large savers giving inaccurate estimates, or (2) large savers being mainly self-employed businessmen and hence giving less accurate reports cannot be determined.

In all saving surveys until now aggregate change in liquid assets had tended to be underestimated, a negative number. This phenomenon, which has its explanation in the theory of recall, is not fully understood. Unlike these earlier surveys, the Delhi Saving Survey's estimate of this magnitude is positive. Though favorable to the Delhi Saving Survey, this constitutes weak evidence of validity for this component of saving.

Evidence of Willingness to Cooperate. As argued before, willingness to cooperate is a prerequisite to accurate saving reports. A survey's nonresponse rate is one index of cooperation. Judged by this index, the Delhi Saving Survey was remarkably successful. The overall weighted nonresponse rate was 7%, which compares favorably with comparable figures for other countries:

United States	15%
United Kingdom	25%
Israel	15%
Puerto Rico	4% 1/

Of the sixty-three nonrespondent households in the DSS, twenty-nine nonresponses occurred by reason of non-cooperation.

Nonresponse rates by occupation - see Table 3 - varied unexpectedly, the rate being highest among managers and officials followed by self-employed businessmen and self-employed artisans and hawkers (my candidates for the number one position). For other groups the non-response rate ranged from 2% to 4%.

To investigate actual and perceived cooperation, the following question was asked at the end of the interview of all respondents:

> "Do you think that most people will give us accurate answers to questions like how much they have in their savings accounts or not? What do you have in mind?"

Our hopes for the question were two: (1) To ascertain what stereotype existed relative to "financial candor"; (2) To elicit from some relatively naive people responses which would presumably give us knowledge of "honest" answers on their part (e.g., "I don't know about others, but I gave accurate information.").

This gambit was not a striking success. The proportion of respondents who believed disclosure of financial information likely (49%) was greater than those expressing the contrary belief (33%). Ten percent felt that "the poor, the middle-class, and government employees" would give correct reports; another 10% thought that "businessmen" and "high-income people" would not give accurate information. These proportions were invariant with respect to respondent's occupations, education, and rating as to accuracy of saving data.

1/ The Puerto Rican survey was confined to the uppermost 10% of income receivers, thus making it an even more remarkable performance.

"Can't say" and "don't know" responses to this question did vary, however, with education, being most numerous among the less educated. Can an opinion question of this sort yield a valid distribution of responses when it is put to both educated and illiterate respondents? Further investigation is surely warranted.

The "naive" response was given by 14% of respondents. Doubts as to its validity arise when we note that its frequency does not vary significantly among groups whose estimates of saving were rated "very good", "middle-position" or "poor." Alternatively, it may be that the ratings are not valid.

Evidence of Success in Motivating Respondents. Lying behind this evidence of respondents' willingness to cooperate is further evidence relating to respondents' subjective reactions to (1) the interview and (2) the study itself. Ceteris paribus, if the respondent "enjoys" the interview and expresses the view that the study is"worthwhile", then an impartial observer should be entitled to infer that the motivation situation existing during the interview was conducive to a high degree of cooperation on the part of the respondent.

1. "Do you think that most people will enjoy the interview or not?"
2. "Do you think that most people will think that this study is important and worthwhile, or a waste of time?"

"Projective" interpretations were given to answers to these questions: that is, positive responses about how "most people" would react were interpreted as statements of the respondents' own feelings.

Most respondents (77%)--See Table 4--found the interview "enjoyable." To the extent that there were occupational differences in degree of enjoyment, it is interesting to note that the managerial-official group--the group with the largest nonresponse rate--was least enthusiastic about enjoying the interview.

As to worthwhileness--see Table 5--most respondents (58%) took an affirmative view; only a few (13%) thought it a "waste of time." Oddly enough, the professional group, as well as the self-employed artisans and hawkers, was most skeptical about the worth of the study. In general, opinoons of worthwhileness were more frequently held by the better educated though this may be a spurious correlation, resting on the greater verbal skills of educated persons.

With respect to the relationship between "enjoyment" and "worthwhileness" on one hand and the accuracy of saving data on the other, one might argue syllogistically as follows: (1) Positive responses to the "enjoyment" and "worthwhileness" questions provide indices of a favorable motivational situation; (2) a favorable motivational situation yields a willingness to cooperate; (3) cooperation on the part of the respondent yields more accurate information concerning saving. Direct evidence of the relationship between the factors under (1) and the accuracy of saving data is available from the survey if one is willing to accept the interviewers' ratings as valid.

This evidence tends to confirm the theory just set out. As Tables 6 and 7 show, those who enjoyed the interview most and those who were most convinced of the worthwhileness of the study tended to provide the most accurate information on their saving. Both, with respect to enjoyment and worthwhileness, those furnishing the poorest quality saving data tended perhaps to disguise or conceal their negative reactions by giving vaguer and more equivocal responses to the two questions.

The evidence of Part II is encouraging: financial surveys in urban India should yield, if they are carefully conducted, better quality saving data than similar surveys in the United States and the United Kingdom. This is not strong praise. Though analyses of America saving surveys have yielded useful descriptive data and useful analytical insights, 1/ discrepancies between survey results and aggregate estimates have led to suggestions for drastic changes in procedures. As a prudent offset to any optimism induced by this paper, I would strongly recommend a careful reading of the reports in which these discrepancies are documented and evaluated. 2/

1/ Report of the Federal Reserve Consultant Committee on Consumer Survey Statistics. (Smithies Committee), Washington, 1955, pp. 275-290.

2/ Cf. ibid.; also Report on the Federal Reserve Consultant Committee on Statistics of Saving (Goldsmith Committee) Washington, 1955; Irwin Friend and Stanley Schor, "Who Saves?" Review of Economics and Statistics, Supplement to May, 1959 issue.

Conclusions

From this paper these conclusions emerge:

1. Techniques utilized in American and British financial surveys, when suitably adapted, will "work" in the urban Indian environment.

2. Several cultural factors in the Indian environment facilitate the conduct of financial surveys and lead to the conclusion that better-quality financial data might be obtained in India than in Western countries.

3. In Indian survey work it appears that response errors have been chiefly attributable to failures of communication rather than of motivating the respondent (the chief problem in the United States). If this is true, it is fortunate since these sources of error are more susceptible to correction.

4. In India success in making the interview "enjoyable" and in convincing the respondent of the worthwhileness of the study appear to result in better-quality financial data.

The above conclusions were based on detailed analysis of the Delhi Saving Survey and less systematic observation of other Indian survey work.

TABLE I

Percentage Distribution of Household According to Interviewers' Estimates of the Accuracy of Saving Data within Occupation Groups 1/

Interviewers' Estimate of Accuracy	All Occupations	Professional and Semi-professional	Self-Employed Businessmen	Managers, Supervisors	Clerical and Sales	Skilled Operatives	Unskilled & Service	Self-Employed Artisans and Hawkers
Very good or good.	75	68	59	73	89	77	76	82
Very good.	21	19	10	47	32	23	21	6
Good.	54	49	49	26	57	54	55	76
Middle position.	16	9	32	13	6	10	12	16
Poor	7	12	7	13	3	13	12	1
Very poor	-	1	-	-	-	-	-	-
Not ascertained.	2	10	2	1	2	-	-	1
	100	100	100	100	100	100	100	100
Percent of households	98 2/	6	25	8	24	12	10	12
Number of cases	513 2/	49	96	88	134	37	39	39

1/ Interviewers were asked to rate the accuracy of saving data in each interview on a scale from "very good" to "very poor." Only 3 cases of "very poor" are retained in the sample.

2/ Households with incomes of $25,000 or more were excluded from this and the following tables. This total includes the 3% of households headed by retired or unemployed persons.

TABLE 2

Percentage Distribution of Households According to Interviewers' Estimate of the Accuracy of Saving Data and Amount Saved. 1/

Interviewers' Estimate of Accuracy.	Amount saved or Dissaved (Rupees)							
	All Households	Plus 2000 or More	Plus 1000 - 1,999.	Plus 500 - 999	Plus 200 - 499	Plus Below 200	Minus Below 500	Minus 500 or more.
Very good or Good.	75	54	54	73	69	77	89	71
Very good	21	23	32	19	40	13	18	31
Good	54	31	22	54	29	64	71	40
Middle position	16	39	32	10	27	10	7	18
Poor	7	6	10	9	-	13	3	10
Not ascertained.	2	1	4	8	4	-	1	1
	100	100	100	100	100	100	100	100
Percent of households.	98	8	7	10	9	27	28	9
Number of cases	513	77	48	58	54	99	105	72

1/ See notes to Table I

TABLE 3

Weighted Nonresponse Rates by Occupation 1/

Occupation	Nonresponse Rate
All Occupations:	7%
Professional and Semi-professional	4%
Clerical and Sales	3%
Skilled and semi-skilled operatives	4%
Unskilled operatives	2%
Service workers	3%
Retired	2%
Self-employed businessmen	8%
Self-employed artisans	7%
Managers and officials	11%

1/ Each interview was assigned a weight equal to the inverse of the probability of inclusion in the sample. In general, high-income households had smaller weights, greater frequencies of nonresponse.

TABLE 4

Percentage Distribution of Households by Respondent's "Enjoyment" of the Interview, Within Occupation Groups.

Respondent's "Enjoyment" of the Interview.	All Occupations 1/	Occupation of Head: Professional Semi-Professional.	Self-Employed Businessman	Managers, Executives,	Clerical and Sales	Skilled & Semi-Skilled	Service and Unskilled	Self-Employed Artisans and Hawkers
Favourable	77	65	69	68	88	100	76	63
"I enjoyed it"	23	21	19	28	34	14	27	14
Most will enjoy it	33	12	32	19	26	58	40	42
Educated people will enjoy it.	13	30	13	9	20	13	1	-
They should enjoy it.	8	2	5	12	8	15	8	7
Unfavourable	17	22	21	31	13	15	-	17
I did not enjoy it	1	1	-	-	-	-	-	11
Most will not enjoy it	9	11	8	24	11	·2	-	6
Uneducated people will not enjoy it.	5	2	11	1	2	13	-	-
I cooperated to be helpful	2	8	2	6	-	-	-	-
Equivocal answers	15	28	23	15	10	6	21	12
"Depends," "Can't say," "Don't know"	9	17	14	12	4	-	14	11
Some will enjoy it, some won't.	6	11	9	3	6	6	7	1
Attitude not ascertained	4	1	4	1	3	-	3	12
TOTAL 1/	113	116	117	115	114	121	100	104
Percent of households	100 2/	6	25	8	24	12	10	12
Number of cases.	543	51	113	97	135	37	39	40

1/ The total adds to more than 100% since some respondents gave more than one answer.

2/ Includes unemployed (1% of all households) and retired (2% of all households.)

TABLE 5

Percentage Distribution of Households by Respondent's Views of the "Worthwhileness" of the Study Within Occupation Groups.

Respondent's View of Worthwhileness of the Study	All Occupation	Occupation: Professional Semi-Professional	Self-Employed Businessman.	Managers Executives	Clerical and Sales	Skilled & Semi-Skilled	Service & Unskilled	Self-Employed Artisans & Hawkers
Worthwhile	58	38	54	53	78	69	47	37
I think it is useful	36	22	31	38	50	43	27	28
Most will think it worthwhile	13	18	15	6	14	17	14	7
Educated people " "	6	7	6	5	11	7	-	1
They should " "	3	1	2	4	3	2	6	1
A Waste of Time	13	23	8	16	18	20	1	5
I think it is a waste of time	3	6	4	5	5	-	1	-
Most people " " "	4	11	1	7	5	-	-	5
Uneducated people " "	6	6	3	4	8	20	-	-
Equivocal answers	30	46	35	32	12	17	44	48
Some will say worthwhile some a waste /	3	-	6	9	1	-	2	-
Can't say, Don't know	17	14	25	12	4	12	22	38
Worthwhile if it improves conditions	10	32	4	11	7	5	20	10
Attitude not ascertained	5	1	5	5	3	-	9	12
TOTAL 1/	106	108	102	106	111	106	101	102
Percent of Households	100**	6	25	8	24	12	10	12
Number of cases	543	51	113	97	135	37	39	40

1/ The total may exceed 100% since some respondents gave more than one answer.

** Includes unemployed persons who constitute one percent of the population

TABLE 6

Percentage Distribution of Households by Respondent's "Enjoyment" of the Interview and by Interviewer's Estimate of Quality of Saving Data.

Respondent's Enjoyment of the Interview	Interviewer's Estimate of Reliability				
	All groups.	Very Good	Good	Middle Position	Poor
Favourable.	77	91	83	52	52
"I enjoyed it"	23	40	23	11	7
Most will enjoy it	33	34	36	37	7
Educated people will enjoy it	13	5	16	2	36
They should enjoy it.	8	12	8	2	2
Unfavourable	17	14	15	21	29
I did not enjoy it	1	-	3	-	-
Most will not enjoy it	9	11	8	8	-
Uneducated people will not enjoy it	5	2	3	8	26
I cooperated to be helpful	2	1	1	5	3
Equivocal Answers	15	8	9	34	43
Depends, can't say, Don't know.	9	3	6	14	39
Some will enjoy, some won't.	6	5	3	20	4
Not Ascertained	4	2	4	5	3
TOTAL 1/	113	115	111	117	127
Percent of population	100	21	54	16	7
Number of cases	543	135	261	100	34

1/ The total exceeds 100% since some respondents gave more than one answer.

TABLE 7

Percentage Distribution of Households by Respondent's Views of the "Worthwhileness" of the Study and by Interviewer's Estimate of the Accuracy Saving Data.

Respondent's View of Worthwhileness of the Study.	Interviewer's Estimate of Accuracy of Saving Data				
	All Groups	Very Good	Good	Middle position	Poor
Worthwhile.	58	65	66	44	17
I think it is useful	36	40	41	28	12
Most will think it worthwhile	13	11	17	5	5
Educated people will "	6	8	6	9	-
They should "	3	6	2	2	-
A waste of time	13	11	11	13	33
I think it is a waste of time	3	5	2	4	5
Most people think waste of time	4	2	4	4	5
Uneducated people think waste of time	6	4	5	5	23
Equivocal Answers.	30	23	27	40	48
Some will say worthwhile, some a waste	3	3	3	2	2
Can't say, don't know, etc.	17	13	11	33	34
Worthwhile if it improves	10	7	13	5	12
Attitude not ascertained	5	5	5	7	2
TOTAL 1/	106	104	109	104	100
Percent of households	100	21	54	16	7
Number of cases	543	135	261	100	34

1/ The total may exceed 100% since some respondents gave more than one answer.

II

STATISTICS OF CRIME AND CORRECTION — I

Chairman, Clarence C. Schrage, University of Washington

Current Problems in Police Statistics — John I. Griffin, The City College of New York

Court and Probation Statistics — George F. Davis, California Bureau of Criminal Statistics

Prisoner Statistics - National and State — James A. McCafferty, U. S. Bureau of Prisons

Discussion — Edward V. Comber, San Francisco Police Department

Discussion — Marie Vida Ryan, California Department of Corrections

CURRENT PROBLEMS IN POLICE STATISTICS

John I. Griffin, The City College of New York

The index to the Journal of the American Statistical Association covering the period 1888-1939 contains one entry which reads "police records, inadequacy and non-uniformity." Under the entry "crime" there are several references showing that the incidence of crime and crime statistics were the topics of several papers early in the history of the association. The index to the Journal covering the years 1940-1955 does not even contain the subject entries "police statistics," or "crime." May we conclude from this that the problems which concern our session today became so unimportant in the last 30 years as not to merit professional attention, until rediscovered by Dr. Beattie in his illuminating paper in the September 1959 issue of the Journal.

Thirty-two years ago the Annual Meeting of the Association heard a paper by Louis N. Robinson in which he said "There is nothing very encouraging in the history of criminal statistics during the last twenty-three years.....If progress in the future is to be at the same rate of speed as in the past, we who are gathered here today and our children's children will all be dead and gone, perhaps swept away by recurrent crime waves, before we have the kind of criminal statistics which plain common sense now dictates we should have." At the same meeting in 1928, Lent D. Upson said "neither the public nor the authorities responsible for curbing offenses have any usable knowledge of the number and character of crime committed or the cost of these depredations. The facts - when records exist at all - lie buried in the 'squeal books.' Police can not ask for adequate statistics from others until they have shown an inclination to remedy the defects existing in their own procedure." It would appear that only modest progress has been made and that the current problems in police statistics might well be examined in the light of the new movement toward professionalization in the police service.

Since law enforcement is primarily a local matter in the United States, the raising of the standards of statistical work becomes a particularly difficult task. Improvement can not effectively be legislated or directed from above, certainly not from the national level. Since the laws governing crime and police procedures are generally state wide, in certain states, where effective state data reporting procedures now exist, there is some hope that the state statistical agencies might work effectively with their local law enforcement agencies. However, any useful statistical training program would imply a considerable expansion of the duties of these state agencies. Since only a few states now haveeffective state-wide crime reporting, other means will be necessary in most of the United States. It should be noted that the growing acceptance of state-mandated minimum police training standards may eventually result in the development of well organized police training centers serving entire states. While such state laws presently deal with the assurance of a minimum number of hours of training for individual police officers, the pattern could be extended to specialized unit training, such as training for officers in charge of statistical units. Of course, at the present time, the minimum state standards are well below the training standards of the best local police departments. It is suggested, however, that mandated minimum training standards may prepare local police agencies to accept minimum standards in the technical services, such as statistics and records.

Fundamental to any improvement in police statistics is a change in the attitude of police commanders and the rank and file of police officers toward statistics. There can be no real enthusiasm for improved statistical procedures unless the police officers themselves can be made to realize that statistical methods are a useful tool for them in the more effective performance of their duties. All of us who have worked with police officers of high and low rank have, no doubt, been impressed with two fundamental attitudes toward statistics. First of these is an impatience with statistical reporting, regarding this as a routine chore which is added to their many, already irksome, reporting responsibilities. In large measure this attitude is explained by the lack of "feedback". Complaints are constantly heard about statistical reports flowing from the field commanders to headquarters, from headquarters to state agencies and to Washington, without anything coming back. Unless the originating unit can see a return to it from the efforts expended in preparing statistical reports, these reporting operations will become strictly routinized. On higher levels, if periodic reporting under the regulations of state agencies or in the course of cooperations with the Uniform Crime Reporting program is regarded as simply an addition to an already over burdened work schedule, the main positive gains in such a broader based reporting system will be lost. Certain fairly obvious administrative steps can be, and are, taken to involve as many members of the police department as possible in the statistical reporting cycle. Summaries of departmental annual reports may be made available to each police officer, in the same way that many business corporations make available their annual stockholders reports to their employees. In addition superior officers, down to and including the rank of patrol sergeant can be supplied with daily or weekly statistical summaries. Despite the statistical pitfalls in the comparison of uniform crime reports between cities, these dangers should not preclude vigorous discussions of the reports submitted by communities of generally comparable character, in order to help identify factors bearing upon efficient police service. Such discussions would be appropriate on the level of district or precinct commanders and chiefs. It remains a source of amazement how many senior police officers have never made any use of the Uniform Crime Reports except to assert that comparisons can not be made.

The second and far more damaging attitude is a wide-spread cynicism in respect to the validity of police statistical reports. This skepticism in regard to their own reports reflects a long

standing conviction in many police agencies that the "truth" can not be reported, or at least should not be reported in such matters as the number and types of crime complaints received. It will be recalled that in the autobiography of Lincoln Steffens he tells how he and Jacob Riis created "crime waves" in the New York City newspapers by publishing crime complaint information which was buried in a pigeonhole of a police officials desk. Many ingenious procedures have been instituted in subsequent years to assure that no crime complaint "gets lost." Experience suggests, however, that human ingenuity will defeat the most soundly conceived system. Fundamental to this problem is the fact that the reported statistics on crime complaints are used to determine the efficiency rating of the very police commanders who report the data. There is, inevitably pressure to "close the squeal book" and therefore to improve the clearance rates. Since experienced police commanders know that, with the exception of certain types of crimes against the person, the clearance rates even with vigorous police action are likely to be low, the situation from their point of view can be improved by minimizing crime complaints reported. On the level of a police chief it may be difficult for him to present to the political leaders of his community and to the public, a picture which is too black, because the police department may be blamed for these conditions which are, of course, not of their making. Occasionally a fearless new administration will shock the public with crime complaint figures which are realistic. Such shock treatment often results in substantial increases in police department budgets and manpower. However, matters soon seem to return "to normal." There would appear to be need for drastic action to separate the function of data reporting and analysis from that of efficiency rating of police commanders. Until the commander of a precinct is convinced that he will not be penalized by reporting an actual increase in crime complaints,there will be a tendency for him to take the same attitude toward his periodic reports as a salesman toward his daily activity statement. Certainly in a period when statistical standards in respect to politically sensitive series of data like the number of unemployed, industrial production and the cost of living are now of the highest, the same professional approach to police statistics is needed.

There would appear to be four major areas in police statistics which promise the greatest rewards from intensive professional work. These are: effective presentation of data, development of sample survey procedures, intensive use of small area data, and the introduction of electronic data processing. The first of these selected areas concerns the comparatively simple problem of effective tabular and graphic presentation. Unfortunately many police departments now fail to communicate their message to the community. This failure makes more difficult the accomplishment of the police mission in the community. In a period when the competition for public attention is more intensive than ever before, a poorly prepared and presented report reflects adversely upon the public opinion of the members of a police department and the manner in which they are performing their duties. A small investment of time and effort will produce effective tables and charts which will give the public an impression of professional competence. An examination of the current standards of presentation in the annual reports of police departments will quickly reveal the lost opportunity in this regard. The use of simple and effective charts in newspapers and on television programs has been found in a number of cities to be of great assistance in telling the police story to the public. From the technical point of view, instruction in the methods of statistical presentation can be easily absorbed by designated police or clerical personnel assigned to these duties. Much of the good opinion that a community may have of its police department as a progressive organization, may be lost by a poorly prepared report. The dramatic and human story of the work of a police agency should be revealed in the show window that is the annual report. The statistician can make a real contribution to police work in this simple area.

The second area of promising development is the adaption of sample survey procedures to police problems. Just as sample surveys provide useful estimates in many important areas of economic and social data, so police commanders could use scientific sampling in order to obtain information needed for planning and operations. The reluctance of many police officials to accept sampling on the ground that they do not wish "estimates" is paralleled by similar attitudes in industrial and accounting applications. The massive evidence now available of successful applications of sampling procedures should be brought to the attention of the police. Perhaps one of the most interesting and potentially significant applications of sampling is in the making of attitude surveys in sensitive areas. Such surveys might be addressed to such questions as police-community relations, attitudes of minority groups and the like. Inspector G. Douglas Gourley's pioneering study published in 1953 should be read in this connection. Because of the confidential and sensitive nature of the work of a police sampling unit, the personnel should probably be police officers trained in statistical procedures. The organization and operation of such a sampling unit will require professional guidance which could well be provided in most parts of the country by sample survey specialists.

The growing abundance of small area statistical data has led a number of police departments to code their crime complaints and activity reports by city block, census tract or other small area unit. With the results of the 1960 Census of Population and Housing soon to be available, the opportunity will present itself for many more police departments to relate police activity to small area economic and social data. Fundamental in such a procedure is the adoption of a coding system by blocks or at least by census tracts. Coding police reports by beat and by precinct or district, while important from the point of view of administrative analysis, does not permit meaningful statistical comparisons. It is best, of course, to lay out police unit boundaries so as to agree with census tract boundaries. In some cities police representatives serve as members of the local Census Tract Committees in cooperation with business and other groups. Police representatives can gain much by working together with

local public utilities, school officials, health department representatives and academic students interested in the making of current estimates of the population of small areas. In some communities the professional isolation of the police has caused them to duplicate analyses which have already been made by other agencies. One of the more interesting by-products of the correlation of small area data and police crime complaints and activity reports should be the development of defensible schemes for patrol force allocation. The pioneering work of Dean O. W. Wilson needs to be followed up by intensive correlation analysis using data from the larger cities. Out of this, uniformities should emerge which will permit the development of rational manpower allocation plans.

The last of the suggested areas for major work is the application of electronic data processing to police statistical problems. While punched card equipment has long been used in major cities, the potentials of E.D.P. have not been explored. It would appear that file computers with large rapid access capacities would be most promising since police records require rapid search as well as rapid print-out. The manufacturers of E.D.P. equipment should be encouraged to develop programs which would be applicable to the larger departments, and groups of smaller agencies using centralized files. By all these means police statistics will advance.

In the accomplishment of these objectives the professional statisticians of the United States have a responsibility to assist the police profession. Comparatively few professionally trained statisticians are now employed in law enforcement agencies and, except for the largest cities, the employment of such personnel would not be justified on a fulltime basis. It seems necessary therefore to involve teachers of statistics and statistical practitioners in this task. Perhaps, the local chapters of the American Statistical Association might seek to assist their police departments with professional advice. Colleges and universities, which now have police training programs, are, in many cases, as at The City College of New York, offering special classes in statistical methods which are attended by selected law enforcement personnel. Such courses could be offered more widely, but should be oriented toward police applications. Following the lead of many professional groups, a series of institutes could be organized in the larger cities to which police departments could send selected officers for an intensive, short period training. Such a program might well be supported by a foundation interested in improving public administrative standards.

References

Beattie, Ronald H., "Sources of Statistics on Crime and Correction", Journal of the American Statistical Association, 54(1959) 582-592.

Foley, Donald L., "Census Tracts and Urban Research," Journal of the American Statistical Association, 48(1953) 733-742.

Gourley, G. Douglas, Public Relations and the Police, Springfield, Thomas, 1953

Griffin, John I., Statistics Essential for Police Efficiency, Springfield, Thomas, 1958.

Leonard, A.E., Crime Records in Police Management, New York, Institute of Public Administration, 1952.

COURT AND PROBATION STATISTICS

George F. Davis, California Bureau of Criminal Statistics

History of Criminal Court Statistics

During the past century, the development of criminal court statistics in the United States has proceeded in a somewhat uneven fashion. For all practical purposes it may be said that there was very little interest in this field prior to World War I. Some statistics on federal court operations were collected by the United States Attorney General as far back as 1875, and a few states also collected court data at this time, but only on a very limited scale. The impetus for the development of criminal court statistics can be attributed to a series of independent surveys that were made during the Twenties. The best known of these were the Cleveland Survey, the Missouri Survey, the Reports of the New York Crime Commission, Sub-committee on Statistics, the Illinois Survey, and the Oregon Survey. All of these surveys dealt to some extent with judicial criminal statistics--some more than others. The Cleveland Survey was the first to present a statistical analysis of case mortality between arrest and final disposition, while the Oregon Survey concentrated on judicial criminal statistics and demonstrated one of the earliest examples of individual case reporting.

In the early Thirties, two events occurred which gave criminal court reporting the necessary stimulus to move ahead. The Wickersham Commission in their Report on Criminal Statistics in 1931 considered the problem of acquiring comprehensive criminal statistics on a nationwide basis. As a result, they recommended the establishment of a central national bureau of criminal statistics and the drafting of a Uniform Criminal Statistics Act which would establish a central bureau in each state. At about this same time, the Johns Hopkins Institute of Law sponsored a series of investigations that explored the reporting possibilities of judicial statistics on both the civil and criminal levels. As a result of this study, forms and procedures for court reporting were developed and standardized terms were promoted to define and classify judicial operations.

One other pioneer study that should be mentioned occurred in California in the middle Thirties. The purpose of this study was to evaluate the work of the courts in the administration of criminal justice, and further to devise and test a method for the reporting of court work on an individual defendant basis through a central collecting agency. The three most populous California counties reported criminal court statistics on the individual defendant system for one year, and these data were then coded and placed on machine records cards for tabulation and analysis. The data were published in 1936 in a report entitled, A System of Criminal Judicial Statistics for California. This report set the stage for the future development of court statistics in California.

As a result of the report of the Wickersham Commission and of the other studies that demonstrated the feasibility of criminal court reporting, the United States Bureau of Census in 1932 set up a system of court reporting from the individual states. At the inception, 16 states agreed to report criminal judicial statistics on a summary basis to the Census Bureau where they would tabulate, analyze, and publish the data. This project lasted from 1932 to 1945 and had a maximum of 30 states participating at any one time. Unfortunately, due to a multitude of difficulties, the data collection was discontinued with the last report being issued in 1946. The accomplishments of this collection though were enough to justify the energy expended on its behalf. The superiority of the individual card system was unequivocably recognized, and states such as Ohio and Minnesota that had been reporting on an individual defendant basis continued this system even after the reporting to the Bureau of the Census was discontinued. The states of Michigan and California began such a system shortly thereafter and certain other states continued their own collection of criminal judicial statistics on a summary basis.

History of Probation Statistics

In reviewing the literature on probation statistics in the United States, one soon concludes that, except in isolated instances, there has been very little development in this field. The Wickersham Commission in 1931 noted that statistics on adult probation were woefully inadequate due to ".... the absence of an accepted guide or standardized procedure for compiling probation material". The Commission went on to comment that ".... statistics concerning probation, like statistics concerning nearly every other aspect of work with offenders are distressingly inadequate in the United States". However, unlike the subsequent development of judicial criminal statistics, probation statistics did not move ahead. In the United States Attorney General's Survey of Release Procedures, Volume II, published in 1939 it was stated that "the two principal sources of information regarding the extent of probation are the courts vested with the power to grant probation and the organizations charged with the supervision of persons so released. However, machinery for collecting this information from either source in a uniform and comprehensive manner for the country as a whole has not been created".

There was some improvement--although slight--in probation statistics in the United States during the Forties. Some states began rudimentary reporting programs that culminated in published reports of one type or another. The Federal

Government strengthened its position by transferring probation statistics relating to United States courts to the newly created Administrative Office of the United States Courts rather than let it remain with the Bureau of Prisons. The state of Michigan entered into the collection of probation data as did Wisconsin and Massachusetts. Other states such as New York, Minnesota, New Jersey, Ohio, Washington, Louisiana, Texas, and Pennsylvania developed centralized agencies to collect and report data on some phase of criminal statistics. Unfortunately, it is difficult to summarize just what type of criminal statistics are being collected in the states just mentioned, as there is no uniformity in the type of information requested, the degree to which it is gathered, or the agency to whom it is reported. It would not be inaccurate to say, however, that the development of probation statistics in most stages lags considerably behind the development of crime, court, and penal statistics.

Criminal Court Statistics in California

This then brings us up to a relatively current period of time and to the development of court and probation statistics in California. First though, a word about the formation of the California Bureau of Criminal Statistics. The history of the Bureau in its present form dates back to 1944 when the statistical operation that had been a part of the Bureau of Criminal Identification and Investigation was removed from that Bureau and placed in the Division of Administration as a Bureau of Statistics. From the beginning, the Bureau has operated as a central bureau of criminal statistics as described in the Uniform Criminal Statistics Act that was originally recommended by the Wickersham Commission, and later adopted by the Commissioners on Uniform State Laws. In 1955, the California legislature formally adopted the Uniform Criminal Statistics Act--California being the only state to do so. The primary units within the Bureau are crime and arrest statistics, court and probation statistics, and juvenile delinquency statistics. The following discussion will be concerned solely with the development of court and probation statistics.

In 1947, a pilot study was instituted in the San Francisco District Attorney's Office whereby they would supply the Bureau of Criminal Statistics with an individual card on each defendant prosecuted in the superior court of that county. A card was designed calling for such information as: the defendant's name and sex; the charge filed; the type of filing; the plea or pleas entered; the type of trial and verdict; and the sentence imposed. Except for name and sex of the defendant, no information was requested as to the characteristics of the defendant, his social history or background, or his prior criminal record. The reason for these omissions was that, for the most part, the data could not be obtained from the records of the prosecuting attorney, or at best, obtained only to a limited degree. The completed cards were submitted monthly to the Bureau of Criminal Statistics where the information was transferred to machine records cards. The results of this study were first reported in mimeographed form in a pamphlet entitled, A Study of Criminal Cases Closed in San Francisco County, Calendar Years 1948-1949, which was released in August of 1950. In 1951, the idea of expanding this system, which had now reached a high degree of refinement, was presented to the district attorneys of the remaining 57 counties. The response was overwhelming--so much so that by July 1, 1952 all 58 counties were in the system and reporting on a monthly basis. The first annual report of the Bureau--Crime in California, 1952--covered crime and arrest statistics from police and sheriffs' departments and superior court prosecutions from district attorneys. A similar report has been published each year since, with the most recent publication being for the year 1959.

There is one other aspect of judicial reporting that should be mentioned in order to complete the picture of criminal court work on the felony level. What appears in the superior courts does not represent all of the defendants on whom felony charges have been filed. This is due to a certain amount of drop-out or reduction at the preliminary hearing level. The preliminary hearing is conducted by a magistrate of either a municipal or justice court and is essentially an examination to determine if a felony offense has, in fact, been committed and if there are reasonable grounds to believe that the defendant committed the offense. As a result of this hearing, the defendant may be held-to-answer for superior court prosecution or the felony complaint may be dismissed or reduced to a misdemeanor. Until recently, the mortality in felony complaints could only be approximated because the disposition of felony complaints not reaching superior court prosecution were not reported. In order to provide information at this level, a pilot project was inaugurated in 1958 with the district attorneys once again providing the source material. On this occasion, the district attorneys were requested to complete and forward a card to the Bureau of Criminal Statistics on each felony complaint dismissed or reduced in the lower courts. This card contained spaces for name, sex, county, offense charged, date of dismissal, and reason for dismissal. The cooperation received from the district attorneys on this level of reporting was very gratifying and by January 1, 1959 all 58 counties were participating.

Probation Statistics in California

Probation statistics in California were reported for a number of years by means of a simple summary report of caseload movement. This was completely unsatisfactory for detailed analysis and the data were often unreliable. In 1953, a project was begun whereby each county probation department (probation being administrated locally in California) was asked to supply the Bureau of Criminal Statistics with a card on each superior court defendant referred for probation, and then, for those granted probation, a card on each change of status that the defendant

experienced while on probation. The initial probation card that was devised, and that is presently in use, requests more identifying information on the defendant than was possible to obtain by way of the district attorney card. This is because the face sheets of most probation reports contain these basic data and they can be located and transcribed with relative ease. In addition to the personal characteristics of name, sex, age, and race of the defendant, data are also requested on the method of conviction, the dates of referral and judgment, the type of recommendation, the judge's name, the offense of which convicted, and the judgment of the court. The cards are submitted at the point at which probation is either granted or denied, and the data are then transferred to machine records cards and held for tabulation. The change of status card, on the other hand, identifies the probationer by name and number and then reports the type of change of status that has occurred. Generally, three different types of action may be reported; (1) probation may be modified to some extent, (2) probation may be revoked, or (3) probation may expire or be terminated early. As these data arrive, the information is added to the defendant's master card, and in turn, punched onto the machine records card. This information is available in the Bureau's annual publications of Delinquency and Probation in California, 1954 - 1959.

Applications of the Statistical Data on Courts

In addition to the very practical value of court data for the purposes of accounting for the workload of an agency, and for explaining the processes involved in the administration of criminal justice, there is the added value of assessing the data in light of the goals that are set forth in the administration of criminal justice. Some of these goals have been stated as; certainty of apprehension and conviction, equality of justice, and swiftness of procedure. The studies that the Bureau has done with the information collected on courts may point the way for further research in these vital areas. For instance, in judicial criminal processes it has been noted that there is a tremendous variability among counties in the proportion of felony complaints dismissed. As reported in Crime in California, 1959, some district attorneys dismiss or reduce as many, or more, felony complaints in the lower courts as they file in the superior courts. Other district attorneys seldom agree to reductions or dismissals in the lower courts. The range of felony complaints dismissed during 1959 was from under 10 percent in some counties to over 50 percent in others. This would imply that similar defendants committing similar offenses are being handled differently depending on the counties where the offenses are committed. This is particularly noticeable in certain offense groups--forgery and checks is one such example. In one county, out of 120 felony complaints disposed of on felony insufficient funds check violations, 67 were reduced or dismissed and 53 went on to superior court prosecution. In another county, out of 29 insufficient funds complaints, only 2 were reduced or dismissed and 27 were prosecuted in the superior court. District attorneys obviously use different standards in deciding on the merits of prosecution in criminal cases. They are entrusted with a considerable amount of discretion in deciding whether a felony complaint should be issued in the first place, and also whether the prosecution should be carried through to the superior court or dismissed or reduced at the preliminary hearing. These decisions may be entirely consistent within counties and yet inconsistent when two or more counties are compared.

The judicial decisions, of what type of sentence to impose on those convicted, also evidence considerable variability from county to county. As may be seen in Crime in California, 1959, the percent sentenced to prison varied from a high of 59 percent in one county to a low of 21 percent in another county. Probation judgments were as high as 57 percent of the total convicted in one county and as low as 15 percent of the total convicted in another county. Jail sentences were imposed in 38 percent of the convictions in one county and in only one percent of the convictions in another county.

The time involved in disposing of superior court cases is another area of variability among counties. In some of our larger counties during 1959, the median time from filing to disposition for superior court cases ranged from a low of 16 days to a high of 66 days. Some of the variability here can be attributed to the makeup of the counties--the urban, more metropolitan, centers having generally more crowded court calendars. However, even between the smaller, less urban counties there were differences in the disposition time of criminal cases that must be attributed to differences in procedure which allow one county to employ its resources more effectively than another.

Applications of the Statistical Data on Probation

Perhaps the most interesting of all developments in adult probation statistics is the recent study published in Delinquency and Probation in California, 1959, concerning a longitudinal or cohort evaluation of probation violations. This type of study requires a number of years of data collection, and only within the last year has the Bureau accumulated enough data to meet the requirements for such a study. The procedure involves selecting a representative group of defendants who have been granted probation and following them throughout their probation periods. In this instance, it was assumed that the defendants granted probation during a year's time would be representative of the general probation caseload. Therefore, the defendants granted probation from 56 out of 58 California counties during 1955 were selected from the files of the Bureau of Criminal Statistics, and a tabulation was made of all violations that these probationers experienced up through December 31, 1959. The results were grouped in such a way as to show the number of defendants with no violation record as of

December 31, 1959; the number of defendants with one or more violations as of that date but who were not revoked; and the number of defendants who were revoked. Of the total of 3,142 defendants in the 1955 cohort group, approximately 67 percent had no violations, 8 percent had one or more violations but were not revoked, and 25 percent were revoked. This latter percentage may be used as an indication of the probation failure rate in California, while the first percentage may be used as an indication of the probation success rate. What the middle group represents will depend upon the strictness of the definition of success or failure.

The validity of the above rates depends, to a large extent, upon the completeness of the reporting and the amount of uniformity in judicial decisions that exists among counties. There is no assurance that the defendants with no violations on record did not in fact have additional violations that were not reported. Also, the action that one judge takes as a result of a violation of probation may be entirely different from the action that another judge might take in a similar situation. The problem then is twofold; to be able to obtain an adequate and complete description of all violations that occur, and to have a standard for evaluating these violations uniformly from county to county. In respect to the first problem, the counties must be encouraged to report all violations of probation routinely. The Bureau is now getting reports on only the most serious violations, the ones that demand some sort of action, with only a few reports on the less serious violations, or those of purely a technical nature. The second problem is how to assess the violations that occur, not in terms of the action taken by the courts--which may be very inconsistent--but in terms of whether the violations represent a failure on probation. If all violations were reported, the Bureau could then set up a scale based on the seriousness of the violations and compare each violation against an agreed upon definition of failure. This definition of failure would be independent of the action taken by the courts. This would eliminate the variations from county to county in judicial and probation department policies and give a truer picture at that level.

Summary

The administration of criminal justice in the United States is a matter of vital concern to everyone, but unfortunately, it is also a subject that we often know very little about. The only way to comprehend such a vast area is to collect statistical data that will accurately and reliably give some indication of the processes involved. The ultimate goal in this regard is to work within the framework of the Uniform Criminal Statistics Act with each state, through a central bureau or agency, being held responsible for the criminal statistics of that state. The material collected should cover law enforcement statistics, court statistics, probation statistics (both juvenile and adult), and penal statistics. These data should then be brought together and analyzed by the individual states in light of their own procedures and laws. When all these data are collected, tabulated, and analyzed, the next step would be to supply the data to some agency that is operating on a national level so that they, in turn, could combine the information from all states, analyze, and interpret the variations that will inevitably occur, and release it for national distribution. The agency doing this final collecting, tabulating and analyzing could be within the Federal Government if this were feasible, or a quasi-governmental, or non-governmental agency providing such an agency could take on a nationwide level of operation.

The outlook for the future in court and probation statistics depends to a large extent on how quickly we can implement the proposals contained in the Uniform Criminal Statistics Act. Court statistics nationwide will probably progress more rapidly than probation statistics because the data on court work are more readily obtainable and do not require so many individual interpretations. Many states, as noted previously, are now routinely collecting these data and publishing them in annual report form. Probation statistics, on the other hand, are in their infancy, but even here ideas for reporting on a nationwide level are being formulated, and the National Council on Crime and Delinquency is doing much to encourage the development of some centralized repository for data that may some day be collected by the individual states.

PRISONER STATISTICS - NATIONAL AND STATE

James A. McCafferty, U. S. Bureau of Prisons

INTRODUCTION

The growth of prisoner population in State and Federal institutions is one of the critical areas of public administration. The total number of adult prisoners confined in these institutions at the close of 1939 was 179,818. By 1959 the number rose to 207,513, an increase of 15.4 percent. Should the current trend continue, conservative estimates based on projections of the U.S. population place the prisoner population at 263,140 in 1970 and 318,632 in 1980. (Table 1)

These estimates show that in the next ten years we can anticipate a 26.8 percent increase in prisoners confined. In twenty years, the number of prisoners confined will increase by over fifty percent (53.6). With this potential on the horizon there is a growing awareness among correctional administrators that there is a continuing need for adequate statistical data on prisoners.

Prisons are now coming under the scrutiny of management conscious leadership. Though some administrators continue to "push buttons" on the basis of tradition and the particular urgency of the situation, greater numbers are setting into motion new or revised policies for the custody and treatment of prisoners based on meaningful statistical and research resources.

We recognize that correctional programs are one of the last welfare frontiers to accept the sophistication of statistical inventories, and operational and experimental research. Outdistanced by the mental health field, especially with respect to institutional programs, correctional departments are fast becoming cognizant of this lag and their current pace should close the gap in another decade.

This paper is aimed at describing recent developments in the National and State prisoner statistics operations. After this descriptive statement we will turn our attention to five major considerations. Finally, we will briefly discuss research avenues in the prisoner field.

National Prisoner Statistics

Ten years ago, the Bureau of Prisons undertook the responsibility for the National Prisoner Statistics program. Begun in 1926 by the Bureau of the Census, this series represents the longest continuous national collection of criminal statistics in the United States.

To some extent my appearance here represents in part a "progress report" on the National Prisoner Statistics program since it was December 1951 when a preliminary statement was presented at the Boston ASA meeting. (1)

The last ten years can be summarized by showing how the Bureau of Prisons has achieved four major objectives. (2)

The first of these was to obtain the cooperation of all State institutions for adult offenders in reporting detailed data on court commitments and discharges. Since 1952, all the States and the District of Columbia have fully cooperated in the program. Beginning with the current year, 1960, the program will cover Hawaii and possibly Alaska.

The second objective was to streamline the reporting system. Improved consistency and reliability have been realized by revising the forms and reducing the number of cases to be reported. Whereas the Bureau of the Census collected data on each admission and each discharge which represented as much as a quarter of a million cases, today the Bureau of Prisons limits the collection to court commitments and prisoners released for the first time on their sentence. This change reduced the processing workload to about 140,000 cases, a drop of 44.0 percent.

A third objective was to speed up the processing of data so that it would be available for analysis on a current basis. Part of this was achieved by reducing the workload noted above. However, the immediate problem was to absorb the coding and punching of admissions and discharges.

The solution was the establishment of a Coding and Punching Unit at our Federal Reformatory for Women, Alderson, West Virginia. Because of our excellent experience in training young men in a modern data processing unit at the National Training School for Boys, we felt certain that the Alderson Unit could, under civilian supervision, process the National Prisoner Statistics data. The Alderson Unit has met our expectations and currently handles all punching requirements of the Bureau's Research and Statistics Branch.

Concurrent with the Alderson Unit's development there has been a steady growth in the utilization of punch cards furnished by State correctional statistics offices. This unique State-Federal cooperative system is based on the premise that where the State and Federal government are collecting essentially the same information on prisoners, it is better to depend upon the central correctional statistical office for all information pertaining to prisoners than on the widely scattered State institutions. Also, because of statistical controls provided by a central office, information provided by one agency tends to be more consistent and reliable.

The roll call of States now participating in this unique program is formidable. At present Alabama, California, Michigan, Ohio, New York, Washington and Wisconsin furnish the Bureau of Prisons with punch cards for the National Prisoner Statistics program. In the coming year Florida, North Carolina and Pennsylvania plan to furnish cards. Other States considering similar proposals are Iowa, Indiana, Minnesota and Nebraska. In addition to these so called "punch card" States, effort is made to handle statistical matters with the Department of Correction where central record keeping systems exist. With the continued centralization in such Departments of statistical data on prisoners, we now have 90 "reporters", compared to some 150 ten years ago. These 90 "reporters" furnish data on some 230 separate State and Federal institutions for adult offenders.

Our fourth objective has been to bring up-to-date the summary and detailed NPS reports. Partially realized through the tabulation by mechanical means of prisoner population movement and prison

TABLE 1

SENTENCED PRISONERS CONFINED IN STATE AND FEDERAL INSTITUTIONS FOR ADULT OFFENDERS, BY SEX, DECEMBER 31, 1939 TO 1980
(Data subject to revision)

Year[a]	All institutions[b]			Federal institutions			State institutions[b]		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
Projected:									
1980..	318,632	307,800	10,832	34,893	33,797	1,096	283,739	274,003	9,736
1975..	293,527	283,455	10,072	32,086	31,057	1,029	261,441	252,398	9,043
1970..	263,140	254,050	9,090	28,591	27,659	932	234,549	226,391	8,158
1965..	232,566	224,446	8,120	25,038	24,207	831	207,528	200,239	7,289
1960..	214,569	207,086	7,483	22,982	22,220	762	191,587	184,866	6,721
Actual:									
1959..	207,513	199,889	7,624	22,492	21,610	882	185,021	178,279	6,742
1958..	205,643	198,208	7,435	21,549	20,774	775	184,094	177,434	6,660
1957..	195,414	188,113	7,301	20,420	19,678	742	174,994	168,435	6,559
1956..	189,565	182,190	7,375	20,134	19,375	759	169,431	162,815	6,616
1955..	185,915	178,790	7,125	20,088	19,367	721	165,827	159,423	6,404
1954..	182,901	175,907	6,994	20,003	19,305	698	162,898	156,602	6,296
1953..	173,579	166,909	6,670	19,363	18,743	620	154,216	148,166	6,050
1952..	168,233	161,994	6,239	18,014	17,457	557	150,219	144,537	5,682
1951..	165,680	159,610	6,070	17,395	16,897	498	148,285	142,713	5,572
1950..	166,165	160,357	5,808	17,134	16,672	462	149,031	143,685	5,346
1949..	163,749	157,663	6,086	16,868	16,410	458	146,881	141,253	5,628
1948..	155,977	149,739	6,238	16,328	15,886	442	139,649	133,853	5,796
1947..	151,304	144,961	6,343	17,146	16,648	498	134,158	128,313	5,845
1946..	140,079	134,075	6,004	17,622	17,150	472	122,457	116,925	5,532
1945..	133,649	127,609	6,040	18,638	18,112	526	115,011	109,497	5,514
1944..	132,356	126,261	6,095	18,139	17,502	637	114,217	108,759	5,458
1943..	137,220	131,054	6,166	16,113	15,546	567	121,107	115,508	5,599
1942..	150,384	144,167	6,217	16,623	16,053	570	133,761	128,114	5,647
1941..	165,439	159,228	6,211	18,465	17,947	518	146,974	141,281	5,693
1940..	173,706	167,345	6,361	19,260	18,631	629	154,446	148,714	5,732
1939..	179,818	173,143	6,675	19,730	19,121	609	160,088	154,022	6,066

(a) Number of prisoners confined in State and Federal institutions for adult offenders by sex for years 1939 to 1959 are actual except for small estimates indicated in footnote (b).

Projections for the years 1960 through 1980 were obtained by two steps. First, it was assumed that the proportion of prisoners aged 15 to 64 appearing in the 1950 population of State and Federal institutions for adult offenders would continue in subsequent years. These proportions were applied to the years 1955, 1956 and 1957 and rates of prisoners per 100,000 of the civilian population were computed for the same years. These were averaged and became the basis for the second step. In the second step the 1955-56-57 rates of prisoners per 100,000 of the civilian population for specific age intervals were multiplied by the projected United States population including Armed Forces overseas (Series IV) for the years 1960, 1965, 1970, 1975 and 1980. (Series IV implies the following assumptions as to fertility: The 1955-57 level of births will decline to the 1942-44 level by 1965-70 and continue at that level to 1980.)

No other factors were taken into account when computing these estimates.

Material based on Bureau of the Census publications: Special Reports, Institutional Population, 1950 and Current Population Reports, P-25, Nos. 146, table 3; 187, tables 1, 2 and 3; 193, table 3. Prisoner data from National Prisoner Statistics bulletin No. 24, July, 1960 and from unpublished data on file at the Bureau of Prisons.

Basic computations developed by Miss Karen Dunkin under direction of James A. McCafferty.

(b) Includes estimates for State prisoners confined in Georgia for years 1939 to 1946 and 1948. Estimates on females were made for Georgia for years 1939 to 1948, for Mississippi 1939 to 1944 and for Alabama in 1941.

Source: Research and Statistics Branch, U. S. Bureau of Prisons, Washington 25, D. C., August 9, 1960.

personnel data, summary bulletins covering these areas are published on an annual basis. In addition to the prison population and personnel bulletins there is the executions report which in recent years has served as the major source for data on the trend of capital punishment in the United States.

The progress in publishing detailed data on court commitments and first releases has been at a slower pace. Since 1950, three reports have been issued. One of these covered court commitments for the year 1950 with summary trend tables for the years 1942 to 1950. The other two reports were on first releases, one of these for the year 1951 and the other covered the years 1952-1953.

Two reports are in process. One covers court commitments for the years 1956-57 with trend data for the period 1942-1957. The other presents first releases data for the years 1954-55-56. These should be available in early fall.

There are two new proposals on the horizon of the National Prisoner Statistics program. The first of these is the plan to carry out a Prisoner Population Survey at the close of 1960. Originally this Survey was proposed for June 30, 1960; however, it was delayed when it was determined that this would conflict with the 1960 decennial census of the U.S. population. The Survey when completed will provide urgently needed information on the characteristics of sentenced prisoners confined in State institutions for adult offenders.(3)

The second proposal is aimed toward the collection of minimal data on prior commitments beginning with calendar year 1961. Remembering the experience of the Bureau of the Census as well as several State prisoner statistics programs which have collected such data for several years, it was determined to keep the request on recidivism as simple as possible. The proposal requires only a YES or NO answer to the question -- "Has this court commitment ever served a sentence in an adult correctional institution?" If so, the answer would be YES, if not, the answer would be NO. (4)

Based on a sampling of States which operate central correctional statistics offices, we are certain that this simple request can be answered by all participants in the NPS program.

The last ten year period might be best described as a "shoring up of a sagging program." In the next ten years the NPS program should be consolidated and the efforts of the many persons who cooperate in it will be reflected in the publication at frequent intervals of detailed reports on court admissions and first releases.

State Prisoner Statistics

The historical development of present-day prisoner statistic systems presents no single pattern. Generally, their origin can be traced to the centralizing of welfare functions at the State level which oftentimes incorporated correctional institutions into Boards of Welfare or Departments of Institutions. Because of the many specialized State welfare programs, information resources had to be developed in order to determine the number and characteristics of individuals who became the responsibility of public welfare programs. Some of these statistical services have become highly sophisticated due in part to the Federal Government's role in the "grant in aid" programs since adequate statistical data had to be provided as part of the conditions of financial aid. In recent years these State statistical offices have expanded their programs to include highly significant research efforts.

In instances where Departments of Corrections were incorporated into Boards of Welfare or similar State agencies, the research and statistics offices of such Boards rather timidly, and sometimes with little forehand knowledge, entered the correctional statistics field. In almost every instance the State statistical office was faced with little cooperation and in some cases complete antagonism by one or more of the prisons operated by the State. Prisons, generally speaking, have been one of the last public agencies to accept administrative direction from the executive departments of the State. However, with the growing public concern about prisons, the opening of the prison gates to the press and interested groups and continued progress in prison management, there has been a swelling need for reliable facts about prisoners, personnel, programs and plants. (5)

In many States, correctional administrators saw the need for reliable statistics on prisoners and eventually invited the central statistical office to conduct feasibility studies prior to establishing a statistical program. Such studies brought about widespread recommendations, many of which entered into matters outside the collection of statistics. For example, administrators learned that unnecessary and duplicate forms could be eliminated and that other forms needed modification. In order to take advantage of such changes it was necessary to centralize records collections in the institutions so that an integrated and uniform set of administrative-statistics forms would be suited to mechanical handling with modern data processing facilities.

In the prison field the establishment of acceptable forms and instructions for filling them out paved the way for the collective understanding of prisoner records and their ultimate purpose within and among the prisons. In a sense, the administrative statistical forms provided an entering wedge for more overall control of prisons and certainly furnished an important link in the chain of good prison administration. (6)

In 1950 there were about 10 centralized State correctional statistics offices. Today there are 25 with half of these utilizing modern data processing equipment. In addition to prisoner statistics, several of these offices have developed comprehensive collections of probation and parole and adult and juvenile court statistics. California, Ohio, Pennsylvania, Michigan, Minnesota, Wisconsin and Washington have been leaders in this comprehensive approach. (See list of State correctional statistics offices at end of paper.)

Basic Operational Patterns of Criminal Statistics Systems

In any discussion of prisoner statistics there is need to fit it into the general scheme of criminal statistics operations. Among the fifty States and the District of Columbia there are three basic operational patterns in handling criminal statistics. These can be best described as (1) concentric; (2) fragmented; and

(3) a compromise between the concentric and fragmented operations.

Probably the two foremost proponents of the concentric approach are Dr. Thorsten Sellin who authored the famous Uniform Criminal Statistics Act, and Mr. Ronald Beattie, Chief, Bureau of Criminal Statistics for the State of California who has put many of the Act's proposals into practice. Though recent developments in California include the establishment of special Research Divisions in the Youth and Adult Authorities which illustrates a fragmentation approach, nevertheless, the California Bureau of Criminal Statistics continues to provide a good example of the concentric concept.

In the concentric operation the central statistical agency collects either in the field form (raw data) or in predetermined classifications such information which is needed for a single report on crime and its treatment. The publication, Crime in California, is an example of providing an overall report on criminal matters for an entire State. Until similar reports are available in other jurisdictions, it will be impossible to determine the amount of crime and the effectiveness of control and prevention.

The concentric pattern provides obvious advantages besides a single documentation of criminal statistics. It can provide certain guidelines for the collection of data. This would include establishing a unit of count which would be acceptable to all correctional agencies collecting statistics; preventing duplication of requests for data and providing one single agency for resolving criminal statistics matters. But the greatest advantage would be the centralization of all matters relating to criminal statistics. This would give executives, legislators and the public access to a single reservoir of criminal data.

Among the disadvantages are the inability to modify reporting systems too rapidly; the likelihood that the contributing agencies' statistical programs would operate at the same level, and the danger of unclear communication patterns between the central statistical collector and the operating agencies furnishing data.

In the fragmented approach, which is simply each operating correctional agency developing its own criminal statistics services, there are several good points. Among these are the possibility of experimentation in the collection and analysis of data. Because such programs are not hidden under a morass of statistical hiearchy, which can develop in a concentric pattern, they can be examined on their own merits. Also there is the advantage of a closer relationship of the statistical-information processes to the persons charged with the guidance and performance of agencies operating in the law enforcement and judicial correctional fields. In these instances, leadership has a more positive attachment to its own statistical program than if there were a single agency in the State responsible for all criminal statistical matters.

The disadvantages of the fragmented approach are evident. In the main these would be the lack of coordination and possible duplication in statistical collections. The opportunity for uneven development of the various fragmentary statistical resources. Another possibility ever present, especially where statistical services depend upon political leadership, is the constant fear of reprisal for statistical indices which are unfavorable and with the danger of complete discontinuance of statistical services.

The concentric-fragmented approach combines the major characteristics of the separate concepts, however, there is a duality in functions. The central agency is responsible for collecting and analyzing all criminal statistical matters while separate operational agencies also build up sizeable staffs and equipment to carry out approximately the same objectives. This type of development will occur especially when the central statistical service fails to provide the amount and type of data required by the operational facility.

It would be worthwhile to explore the possibility of placing in each operational agency professional staff who can serve as the resource and liaison person to the administrator on statistical-information matters. Certainly, such individuals could be called on to assist in programming statistical and research problems in the agency through cooperative effort with the central statistical office. If there were several correctional agencies, these individuals might collectively represent an inter-departmental committee on correctional statistics matters and might serve as an advisory group for the central criminal statistics agency.

Prisoner statistics can be developed along any of these three operational lines. With the ever increasing complexity and interdependence among agencies seeking to control and prevent crime, the combined concentric-fragmented approach with the suggestion noted above, appears to offer a hopeful solution.

Major Overall Considerations

(a) Utilizing comparisons in criminal statistics

We live in an age of comparison. Criminal statistics, because they are the only indices for determining the effectiveness of police departments, courts, probation and parole agencies and correctional departments, are continuously compared. Prisoner statistics, which represent one small part of the criminal statistics picture, receive their share of study.

Because prisoner statistics are somewhat easier to collect than other criminal statistics, there has been a tendency to overuse prisoner data without relating the data to the story behind the figures. Take the matter of State-to State comparison of prison population. The National Prisoner Statistics program provides a complete inventory of all prisoners received, confined and released by State and Federal institutions for adult offenders. However, because of the varying policies among the States with respect to the type of institution where sentences can be served, not all adult prisoners are enumerated by the NPS series. On the other hand, a few States use adult institutions for confining a substantial number of misdemeanants and youthful offenders.

To illustrate the first situation, we turn to the Commonwealth of Pennsylvania. At the close of 1959, according to the most recent National Prisoner Statistics population bulletin, Pennsylvania shows 7,924 prisoners confined in the

Bureau of Correction institutions. (7) On the same day according to a separate Bureau of Correction publication, Pennsylvania's county jails reported there were 1,247 prisoners serving maximum sentences of two years or over, including one prisoner serving a life term. Another 2,033 county prisoners had maximum terms under 2 years and of this number it is estimated that one-third or 677 prisoners had sentences of over one-year and under two years in maximum length. (8)

In the National Prisoner Statistics bulletin the rate of prisoners confined per 100,000 of the civilian population in Pennsylvania was 70.1. However, if all of the county prisoners with sentences of over 1 year had instead been confined in the Bureau of Correction institutions the rate would have been 87.1. (See table 2)

TABLE 2-PRISONERS CONFINED IN PENNSYLVANIA BUREAU OF CORRECTION AND COUNTY PRISONS: DECEMBER 31, 1959

Institutional facility	Prisoners serving sentences	Rate per 100,000 of the civilian population
Total prisoners.....	9,848	87.1
Bureau of Correction institutions.....	7,924	70.1
County Prisons:		
*With sentences --		
Maximum over 1 year, under 2 years..	677	6.0
Maximum over 2 years	1,247	11.0

*Excludes 928 prisoners serving minor judiciary sentences.

SOURCE: U.S. Bureau of Prisons, National Prisoner Statistics, No. 24 and Pennsylvania Bureau of Correction, Directorate of Research and Statistics Census of Pennsylvania Prisoners by County, December 31, 1959, table 14 and correspondence with Mr. John Yeager, Director of Research and Statistics, Pennsylvania Bureau of Correction. Rates based on Population figures appearing in U.S. Bureau of the Census, Current Population Reports, P-25, #210, Provisional.

To illustrate the second situation, we find in the State of Maryland approximately one-fourth of the prisoners received from court into four State institutions have sentences of one year or under. In 1959 a total of 3,896 prisoners were received from court in Maryland adult institutions with a rate per 100,000 of the civilian population of 131.1. By dropping one quarter of the prisoners who had sentences of one year or less, the 2,922 court commitments provide a substantially lower rate of 98.4 per 100,000 of the civilian population. (9)

These two illustrations demonstrate the difficulty of comparing State prisoner figures without knowing the policies of the States with respect to sentencing. To further complicate the matter we need to know the extent to which probation is used. For those who are sentenced to prison we need to know the regulations and legislation surrounding release procedures. It is obvious when the median time served for first releases in 1956 ranged from a low of 9 months in Vermont to 31 months in the District of Columbia and Illinois that the reasons for these figures must be determined through the analysis of commitment and release procedures, sentence lengths, and the philosophy of the prison officials who reflect the temper of the people in their concept of the meaning and purpose of imprisonment. (10)

Prisoner statistics are a useful tool if used with caution. However, without the related statistical data on crime and corrections, such as those collected by the police, courts and probation and parole agencies, prisoner data provide only a portion of the picture.

(b) Criminal Career Statistics

Prisoner statistics have their part to play in the proposed criminal career statistics programs suggested by Dr. Daniel Glaser of the University of Illinois. (11) One of the major obstacles in such a program is the difficulty of maintaining a continuous follow-up record on individuals who have violated the law. The basic problem appears to be a universal acceptance of a unit of count, that is the individual or the case, and adequate identification for statistical purposes from jurisdiction to jurisdiction. At present each time an individual violates the law numerous identifying numbers are applied as he proceeds through the halls of justice. Some States are attempting to have a single identification number adopted for prisoners who enter the adult correctional system. Eventually such States could expand this one number concept to cover all law violators. But what about persons from other States with previous records who are arrested and convicted for the first time in a State which maintains a career record system? Or the recidivist who goes elsewhere and is lost to the State attempting to follow up those individuals who are involved in subsequent criminal activity?

One answer would be for each State to adopt a criminal career statistical number. However, this would furnish the possibility of fifty-one separate numbers plus that of the FBI fingerprint identification number. perhaps the answer is the adoption of a single national criminal career statistical number such as the FBI fingerprint identification number. At the speed in which electronic data processing equipment is being developed we may be forced to accept the single number concept if we are indeed serious about developing career criminal statistics programs. (12)

(c) A Comparative Inventory of Law Enforcement-Judicial-Correctional Processes

It has been 29 years since this country has had a comprehensive report on law observance and enforcement. The famous Wickersham Reports issued by the National Commission on Law Observance and Enforcement in 1931 needs to be re-written and updated.

Twenty-one years ago the Attorney General's Survey of Release Procedures, a six volume report, was issued. This needs to be carried out again.

To some extent the Journal of the National Council of Crime and Delinquency, and Law and Contemporary Problems series issued by the School of Law, Duke University, and the Journal of Criminal Law, Criminology and Police Science of the Northwestern University together with recent criminological texts help greatly to give us a contemporary picture, but fall short of the overall comprehensive view provided by the Wickersham Reports and the Attorney General's Survey of Release Procedures.

No small task, such studies would demonstrate the meaningfulness of various criminal statistics systems and thereby place prisoner statistics in the proper perspective as an indicator of the State's ability to control and prevent crime.

Perhaps in the 1960's we will see a task force assigned to the problem of determining contemporary law enforcement, judicial and correctional practices in the United States.

(d) Handbook on Criminal and Delinquency Statistics

With the general interest in crime and delinquency and the various efforts to measure this phenomenon, various public and private organizations are developing at a rapid pace statistical resources, many of which provide data which are non-comparable to other jurisdictions. These organizations need assistance, but unfortunately have no where to turn but to persons and literature which provide the standard statistical processes. Criminal statistics are a subtle field and persons who have no acquaintance with the undercurrents of such statistics find themselves swept under in a whirlpool of contradiction.

Serious attempts have been made to provide the correctional field with useful systems for developing criminal statistics programs. For example, A System of Criminal Judicial Statistics for California, by Mr. Ronald Beattie, provided for the judicial statistics field what is needed in the prisoner statistics area, (13) as well as for probation and parole services. As the need increases it will be met in such a way that those who are interested in following through the crime phenomenon will be able to make useful comparisons.

Perhaps as a companion volume for the inventory of law enforcement-judicial-correctional processes we need a Handbook on Crime and Delinquency Statistics. This Handbook should be divided into two parts. The first part would provide a complete study of the history and resources, as well as their recommended future direction, of criminal and delinquency services.

The second part of this compendium would provide basic codes and definitions for criminal statistics data. This proposal would assist various jurisdictions in utilizing uniform definitions and codes. Those jurisdictions which adopted similar procedures would be able to make regular comparisons by the exchange of punch cards or magnetic tapes. With the use of common codes, readers of documents containing criminal statistics would be able to make more meaningful comparisons.

(e) Human and Material Resources

Today there are more criminal statistical programs operated by technically qualified individuals than there were a decade ago. However, we are still unable to meet the demand for well trained correctional statisticians and analysts. The correctional field which is now making its impact on college training programs necessarily must require that academic work include courses in statistical areas which relate to corrections. The subject matter of corrections is too unique to permit the average statistician to gain any proficiency except by a trial and error method.

Though correctional statistics systems can train statisticians in the subject matter field, this is a long tedious process with the rewards sometimes tenuous. Eventually academic programs will include training in statistics which qualify the individual at an undergraduate level to handle criminal statistics programs. And there is the possibility that the speciality may be further divided into police court, probation, parole and prisoner statistics.

Graduate programs should include major study in criminal research areas. Persons exposed to this systematic study will greatly enhance our efforts.

Until this more methodical program is developed for training criminal statistics personnel, we must continue to take our risks with individuals who feel they like statistics and the subject matter, crime and delinquency. For research personnel we will have to depend upon foundation grants and publicly supported correctional research programs to provide the testing ground for new researchers. Indeed, not a little of the research going on today is providing the training of future competent researchers.

With adequate personnel, there is the need for adequate equipment to process data quickly, efficiently and economically. In the prisoner statistics systems, the punch card is the basis for maintaining control. Some of these systems can be likened to a vast accounting system where the unit of count is the person rather than the dollar. Today, with vast prison populations, the only sensible way to handle this inventory is through a machine processing system.

Some of these systems are rudimentary and contain only the basic elements needed to carry out assignments. Others are remarkably sophisticated. Of this last group, Ohio is probably outstanding.

A little over three years ago it was determined that for the most economical operation of data processing one central computing unit should be established. This was done in the Ohio Department of Finance. Here under the able leadership of Mr. Donald Smeltzer, some 100 employees work three shifts keeping Ohio's statistical house in order. Using the latest in electronic equipment, the Data Processing Center programmers working with knowledgeable individuals in the other State Departments provide all required information. This centralization of data processing operations illustrates one way to obtain maximum use of expensive computers.

In other States, especially in the prisoner statistics area we find the Boards of Welfare or even the Corrections Departments maintaining equipment to good advantage. But many times such Boards of Departments cannot afford the cost of advanced equipment and therefore must depend upon basic machine components which provide answers at a slower pace and sometimes too late to be of any value.

Personnel and equipment requirements will for a long time concern criminal statistics operations. Prisoner statistics as a part of this greater field may have to provide the leadership toward obtaining well trained personnel and adequate equipment.

Research in the Prisoner Area

Prisoner population furnishes the possibility for three major areas of research: operational or functional, experimental and applied.

Using Dr. Elmer Johnson's definition of operations research as "a scientific method of providing executive departments with a quantitative basis for decisions regarding the operations under their

control", (14) it can be seen that for this purpose we have substantial foundation for research in the statistical collections now available in several States. These potential reservoirs have been overlooked to some extent; however, where they have been utilized the results have been most encouraging. For example, Wisconsin's study of sex offenders, parole violation and the famous Huber Law could not have been accomplished without the heritage of a consistent and reliable data collection program.

Johnson further states that operational research provides an opportunity for generating "a constellation of facts", requiring "a multidisciplinary study", pooling together numerous ideas on a single problem and provides opportunity for demonstrating scientific methods for problem solving. (15)

Turning to experimental research, we have had an extensive experience in this area. Though some of the efforts could be termed operational in purpose, such research for the most part has tried to determine if one method of handling prisoners is more effective than another. The current study of the Effectiveness of the Federal Correctional System is an example of experimental research. In the literature there are many more illustrations of experimental research among prisoners. (16)

In this trilogy of research we turn to applied research which is the translation into regulations and law the findings of operational and experimental research. The implication of applied research is that once a particular principle is applied, there will be continuous evaluation of its effectiveness. To change a procedure without continuous study of its effects is foolhardy.

We will see more emphasis on research in the coming years. Every effort should be made to maintain a constant flow of communication between those involved in operational, experimental and applied research programs.

SUMMARY

It is anticipated that by 1980 the United States will have in its adult prisons 318,670 individuals, one half more than the 207,513 now confined. Prisoner statistics potentially is one of the most critical areas facing public administrators. Through management conscious leadership, vast prisoner statistics operations have been initiated and will continue to flourish. In 1950 there were 10 State central correctional statistics systems. Today there are 25 over half of whom have data processing equipment.

During the last decade the National Prisoner Statistics program successfully accomplished four objectives: (1) it has obtained cooperation of all States in reporting all prisoners serving sentences in State institutions for adult offenders; (2) streamlined the reporting system; (3) speeded up the processing system which included the unique provision for utilizing punch cards furnished by the States and (4) made some inroads in publishing current NPS data.

State prisoner statistics systems have moved in three directions: (1) concentric where all criminal statistics collections of which prisoner statistics is a part is handled by a single State agency; (2) fragmented where each correctional agency develops its own criminal statistics services and (3) concentric-fragmented which combines the major aspects of both systems; however, there is a duality of functions.

A major consideration facing the criminal statistics field, of which prisoner statistics is a part, is the need to understand the difficulties of utilizing available criminological data for State to State comparisons. A simple rate of prisoners confined per 100,000 of the population needs an evaluation of the resources available for handling convicted law violators, the policies for admission to various institutions and for those sentenced to prison, the regulations and legislation surrounding the time to be served and method of release.

A second consideration is the proposal that career criminal statistics programs be established. However, two matters effecting such programs are the need for accepting a universal unit of count and a numbering system which would make possible a continuous and positive identification of individuals who violate the law.

A third consideration is the need for a comparative inventory of law enforcement-judicial-correctional processes similar to the famous Wickersham Reports issued by the National Commission on Law Observance and Enforcement in 1931. Also there is need for another study on prisoners similar to the Attorney General's Survey of Release Procedures of 1939.

A fourth consideration would be the compilation of a Handbook on Criminal and Delinquency Statistics. The growth of criminal statistics systems are at such a pace that such a Handbook is needed to provide some basis of comparability among the systems. Prisoner statistics has followed fairly similar programs; however, in the fields of police, probation, parole and court statistics there are wide diversities. A Handbook would establish standard codes and collection patterns sorely needed if we plan to make State to State comparisons.

A fifth consideration is the need for developing personnel and for obtaining sufficient equipment to carry on criminal statistics activities. Criminal statistics agencies must turn to colleges for personnel requirements.

The expense of data processing equipment calls for an evaluation of current practices. Perhaps, smaller agencies should maintain simple machine units and give to an advanced data processing unit the larger tasks. One State, Ohio, has practiced this for three years.

Research in the prisoner population area is now occurring in the operational and experimental areas. Application of the principles learned from these areas merits the need for research on the effectiveness of such principles.

In conclusion, prisoner statistics which are a part of the vast universe of criminal statistics is becoming the intelligence tool for purposeful administration of correctional agencies. No longer can administrators depend upon hunches or subjective evaluations, but must turn to the empirical data collections made available through valid prisoner and related criminal statistics programs.

FOOTNOTES

(1) Henry Coe Lanpher and James A. McCafferty, "The National Prisoner Statistics Program," presented at the Annual Meeting, American Statistical Association, Boston, Massachusetts December 27, 1951.

(2) These four objectives were discussed at length before the Harrisburg Chapter of the American Statistical Association, June 23, 1960.

(3) James A. McCafferty, "1960 Prison Population Survey", Proceedings, American Congress of Corrections, 1960, pp. 67-81.

(4) See NPPA Journal, July 1958, which is devoted to a complete study of recidivism. Also, James A. McCafferty, "Can We Find a Standard Statistical Definition for Recidivism?" Proceedings, American Correctional Association 1958, pp. 190-206.

(5) James A. McCafferty, "Federal Criminal Statistics and the National Prisoner Statistics Program", p. 18, (unpublished).

(6) Ibid., p. 19 and 20.

(7) U.S. Bureau of Prisons, National Prisoner Statistics, "Prisoners in State and Federal Institutions, 1959", No. 24, July 1960.

(8) Pennsylvania Bureau of Correction, "Census of Pennsylvania Prisoners by County, December 31, 1959", Report CII, March 1960.

(9) U.S. Bureau of Prisons, National Prisoner Statistics, No. 24 and estimates based on data appearing in Maryland Department of Corrections, 32nd Report, 1958, p. 60.

(10) Data on median time served appears in National Prisoner Statistics, Prisoners Released From State and Federal Institutions 1954, 1955 and 1956, in process.

(11) Daniel Glaser, "Institution Statistics", Proceedings American Congress of Correction, 1956, pp. 279-283.

(12) For more discussion about the unit of count see James A. McCafferty, "The Unit of Count", Proceedings of the American Correctional Association, 1957, and Herbert Bryan, The One Number Concept in Crime Statistics, Proceedings of American Correctional Association, 1958, pp. 155-162.

(13) Ronald H.Beattie, A System of Criminal Judicial Statistics for California, Berkeley: University of California Press, 1936. Also same author, Manual of Criminal Statistics, New York: American Correctional Association, April, 1950.

(14) Elmer H. Johnson, "Latent Functions of An Administrative Statistical System in Corrections", to be read at the American Congress of Corrections, Denver, Colorado, August 30, 1960.

(15) Ibid.

(16) For example, in the Proceedings of the American Correctional Association, 1959 see "Use of Research in Determining Administrative Policy" by Sanger Powers; "Social Role, Social Position, and Prison Structure" by Clarence Schrag, and "Social Organization and Inmate Values in Correctional Communities" by Stanton Wheeler.

Persons Responsible for Preparation of Correctional Statistics on Adults in State Institutions August, 1960

*DPM Alabama: Mr. Henry Garrett, Supervisor, Machine Tabulating Unit, State Board of Correction, Montgomery.
Delaware: Mr. David W. Dean, Delaware State Board of Corrections, New Castle Correctional Institution, Wilmington 99.

*DPM California: Mr. Ronald Beattie, Chief, Bureau of Criminal Statistics, 505 State Office Building, Sacramento 14 and Miss Vida Ryan, Statistician, Research Division, California Department of Corrections, State Office Building, Sacramento 14.

*DPM Florida: Mr. R. B. Gramling, Deputy Director for Inmate Treatment, Florida Division of Correction, Doyle Carlton Building, Tallahassee.
Georgia: Mr. Robert J. Carter, Chief Clerk, State Board of Correction, State Capitol, Room 415, Atlanta.
Hawaii: Warden Joe C. Harper, P.O. Box 3289, Honolulu 17.

DPM Illinois: Mr. Fred C. Fieker, Statistician, Department of Public Safety, Room 601, Armory Building, Springfield.
Indiana: Mr. Robert O. Conklin, Supervisor of Field Services, Department of Corrections, 141 South Meridan Street, Indianapolis 4.
Iowa: Mrs. Hazel C. Garner, Director of Statistical Services, Board of Control of State Institutions, Des Moines.
Kentucky: Mr. Harold E. Black, Director, Division of Correction, Department of Welfare, Frankfort.

DPM Louisiana: Mr. F. E. Stockwell, Research Statistician, Department of Institutions, State Capitol Building, Baton Rouge.

*DPM Michigan:Mr. Robert Glass, Administrative Analyst, Department of Correction Box 212-A Station A., Lansing 26.

DPM Minnesota: Mr. James H. Alexander, Director, Division of Administrative Services, Department of Corrections, State Office Building, St. Paul 1.

*EDPM New Jersey: Mr. Douglas MacNeil, Director, Division of Statistics and Research, Department of Institutions and Agencies, Trenton 7.

*DPM New York: Mr. Herbert Bryan, Chief, Research and Statistics, State Department of Correction Albany.

DPM Nebraska: Mr. John Wenstrand, Statistician, Board of Control, State Capitol, Lincoln 9.

DPM North Carolina: Mr. Martin Peterson, Assistant Director, North Carolina Prison Department, Raleigh.

*EDPM Ohio: Mr. Grover Chamberlain, Chief, Research and Statistics, Department of Mental Hygiene and Correction, State Office Building, Columbus 16.

*DPM Pennsylvania: Mr. John G. Yeager, Director, Research and Statistics, Bureau of Corrections, Box 200, Camp Hill.
Rhode Island: Warden Harold V. Langlois, Adult Correctional Institutions, Pontiac Avenue, Howard.

DPM Texas: Mr. J. C. Roberts, Chief, Bureau of Records and Idenfification, Texas Prison System, Huntsville.

DPM Virginia: Mr. J. O. Guinn, Chief, Department of Welfare and Institutions, Bureau of Research and Statistics, 429 South Belvidere Street, Richmond.

*DPM Washington: Mr. Larry Shull, Research Analyst, Research and Statistics Section, Department of Institutions, P.O. Box 867, Olympia.

*DPM Wisconsin: Mr. John W. Mannering, Chief Statistician, Bureau of Research and Statistics, State Department of Public Welfare, Madison 2.

Federal Programs involving State and local correctional statistics:

NPS: Mr. James A. McCafferty, U.S. Bureau of Prisons, Washington 25, D.C.

Juvenile Courts and Training Schools: Mr. Richard Perlman, Children's Bureau, Department of Health, Education and Welfare, Washington 25, D.C.

Police Departments: Mr. Jerry Daunt, Chief, Uniform Criminal Reporting Section, FBI, Washington 25, D.C.

*These States furnish the National Prisoner Statistics Program with punch cards.

EDPM - Electronic Data Processing machines - intermediate.

DPM - Data Processing machines.

DISCUSSION

by

Edward V. Comber, - Director Criminal Information
Police Department, San Francisco, Calif.

MR. CHAIRMAN, DOCTOR JOHN GRIFFIN, LADIES AND GENTLEMEN:

I come before you as a representative of the field of law enforcement to serve as a discussant of the paper that has just been presented by the previous speaker, Doctor John Griffin. The time schedule that has been established for these sessions does not permit any detailed discussion of "Current Problems in Police Statistics" hence I will only be able to touch upon a few of the many items in Doctor Griffin's paper.

I found the presentation of considerable interest and must admit that it demonstrates that the author has an appreciable knowledge of many basic problems that have tended to limit the past development of meaningful crime statistics. I am in agreement with Doctor Griffin on a number of his observations however I find that I must voice some exception to several comments and at the same time make specific reference to one or two of his statements that might be misunderstood by a lay audience.

COMMENTS

1. Dr. Griffin related that if we were to review the index to the Journal of the American Statistical Association we would find that data pertaining to crime statistics are practically non-existant.

If this statement is true, and I accept the speakers statement on face value, then I would like to emphasize that THE FAILURE TO FIND MANY REFERENCES TO CRIME STATISTICS IN THE JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION SHOULD NOT IN AND OF ITSELF BE ACCEPTED AS EVIDENCE OF ANY OR ALL OF THE FOLLOWING CONCLUSIONS:

A. Crime statistics do not merit serious consideration, or
B. No problem exists with respect to the complete and satisfactory treatment of crime data, or
C. No serious attempts have been made toward the development of sound crime statistics, or
D. The law enforcement services are not interested in the analysis of crime data, or
E. The police service is incapable of applying acceptable statistical analysis in the solution of police problems.

The fact that crime statistics have not been the subject of more attention by the American Statistical Association must be attributed to causes other than disinterest on the part of law enforcement. On the contrary, the principal source of available crime data is the result of individual and cooperative effort of the many law enforcement agencies. The Uniform Crime Report Procedure followed by most law enforcement organizations is currently supervised by the Federal Bureau of Investigation. The program was in fact originally developed by the International Association Chiefs of Police. The Uniform Crime Reporting System is itself subject to continuous study and review. Every police association that is worthy of the name devotes some measure of attention to the up-grading of police records and statistical processes.

The charge that police administrators generally lack an adequate understanding of the true nature and limitations of statistical standards is well taken. Some of the factors that account for this deficiency have been touched upon by Doctor Griffin. I will also comment briefly on this problem in a few minutes.

2. Dr. Griffin mentioned in his paper that one of the major impediments blocking the development of a sound and acceptable system of crime statistics is the deep seated concern for "home rule" that is exhibited by law enforcement agencies.

This element of "home rule" (local independence of action) does exist in all jurisdictions to a varying degree. Blind adherence to this socio-political philosophy promotes the same inadequacy and shortsightedness that can be identified with other areas of sterile in-breeding that is frequently encountered in human affairs. From the practical standpoint the local police administrator faces the constant task of making one of two choices. Either he will decide to confine his policy and operations to such limits as will satisfy local attitudes - thereby keeping the local body politic temporarily happy - or, he will expand his horizon beyond the immediate confines and accept the challenge to explore activities that offer some promise of greater gain but grant no immediate return.

This administrative problem arises in all of its manifest distortion when periodic crime rates and reports are issued that call public attention to police statistics. With the passage of time the advocates of the "home rule" philosophy find their hallowed position less tenable as they are more or less compelled to conform certain statistical standards. The trend is toward regional or statewide statistical procedures. The arguments in support of this trend become increasingly formitable as we consider each of the potential advantages of such a program. Among the more obvious benefits would be the following:

A. Larger sample areas for the collection of source data.

B. The availability of more statistical units.
C. Less distortion of data due to unusual incidents.
D. Better classification of data.
E. Greater stability and reliability of the statistical end-product.
F. A more rational basis for comparative statistics.
G. More accurate identification of the nature and magnitude of unique local influcense.
H. True identification of the crime problem as being a product of the local society and not the end result of the efficiency (or lack of efficiency) of any particular law enforcement agency.

— There are many others —

3. Dr. Griffin also commented upon the important fact that before nay major advancements can be made in police statistics - "the attitude of police commanders must be changed to the point that they recognize and accept statistical methods as a useful administrative tool." He further stated that in general police administrators are impatient and view records and statistical processes as a chore.

There is considerable substance to this observation, however, I would urge that the tone of these statements be modified so that fair credit may be given to law enforcement for the significant shift in attitude that has taken place within the last decade. This change in attitude has been notable along the Pacific Coast area and can be amply demonstrated here in California. The manhours of personnel time devoted to records keeping and the status of the statistical duties have advanced at a faster rate that activities in any other phase of police operations

Police "planning and research" units were the subject of academic debate prior to 1950. At this date the existence of such an administrative unit in a mojor police agency is either a fact or one is in the process of organization.

Here I would like to deviate for a moment to express the prayerful hope that the police administrator that has taken the forward step to initiate a Planning and Research Unit within his department will also follow through by securing the employment of statistically qualified personnel. I believe Dr. Griffin will echo the same sentiment.

The movement toward a more extensive application of statistics within police agencies has occurred, in part, due to the need to justify the budgetary request. The results obtained where a proper use has been made of the statistical tool has been encourageing and has paved the way to consider statistical analysis in seeking the solution to other police problems.

4. Dr. Griffin stated that the police service is of the conviction that the "truth cannot be reported".

I am somewhat concerned that this statement might be readily accepted by those who are not acquainted with current police trends and who are equally not aware of the many limitations placed upon law enforcement. I admit that in raising this protest I am speaking from a personal conviction based upon my own limited experience. However, if twenty years of continuous service with a major metropolitan police department is any qualification, and if my working acquaintance with many representatives of law enforcement agencies in the Western states grants any weight to my opinion - then I must suggest that this statement of Dr. Griffin be modified.

There is an old adage that has been bantered about to the effect that " a community gets the amount and quality of police service that it demands". Another statement that is frequently annunciated points out that " the police agency can only apply that degree of enforcement pressure that the public will accept". Both of these statements contain a high measure of truth. They should be kept in mind when any evaluation of police attitudes or administrative policy are under consideration. In the modern era the truth regarding crime and police services cannot be suppressed by the police alone. No police agency can hide the fact in the face of a "free press", an "enlightened public", or where interagency cooperation is the necessary ingredient to daily operation.

5. I am in complete accord with the suggestion made by Dr. Griffin that law enforcement should explore fully four areas of statistical application namely effective presentation, small area sampling, sample surveys, and the use of electronic data processes. Some progress has already been made in certain of these areas. Much remains to be done.

6. Dr. Griffin clearly reveals a principal sore spot in the present use of police statistics. This sensitive area relates to the tendency to use police statistics developed by independent agencies to make comparative analyses.

The very fact that crime statistics are prepared and published invites the reader to use the data for comparative purposes. The use may not be proper and the conclusions may be without substance nevertheless these comparisons go on. This problem became so critical that it was one of the major factors that led to recent changes in the form and content of the Uniform Crime Reports prepared by the Federal Bureau of Investigation from reports received from local law enforcement agencies. The comparisons continue to be made.

The major problem lies not so much in the

fact that comparisons are made as in the invalid and unsupported conclusions or evaluations that are attempted by persons who know little of the true significance of the data but are impressed by the array of numbers. Comparison is inevitable in our dynamic society therefore our every effort should be directed toward the development and acceptance of a properly conceived system of statistical procedure that will serve the operating needs of law enforcement and at the same time allow certain reasonable avenues for the measurement of service efficiency,etc. The present reluctance of law enforcement to use existing data for comparisons flows from the awareness that all police agencies are not exercising the same degree of discretion in the scoring and classification of police incidents.

At this point I wish to conclude my remarks with respect to the paper presented by Dr.Griffin and touch upon a few additional points that I believe should be subject to your consideration at this time. As you are well aware social statistics impose certain limitations upon the analytical processes due to the nature of the basic data under study. This same character of limitation is to be found in police statistics. The problem grows out of the difficulty of establishing valid definitions, identifying characteristics, determining status in a subject undergoing change,etc. May I make specific reference to a few examples.

1. UNCERTAINTY OF BASIC POLICE DATA.

When we attempt to gather and classify police statistics we are identifying and counting units that possess many unknown characteristics. When statistical processes are applied in business and industry the analyst has the advantage of working with unit that identified with a high level of consistency. In addition, the units under study are subject to a measure of restriction and control that is not commonly found in the realm of social statistics. The units are highly similar with respect to the more important basic characteristics. Each of these attributes tend to promote the use of statistical processes. In law enforcement we are frequently required to depend upon social data that is incomplete, inconsistent, and what is more challenging it is highly unpredictible. Our samples are seldom subject to pretest control.

2. LACK OF ENVIRONMENTAL CONTROL.

Our opportunity to control the environment wherein our sample subjects operate is severely limited. The police services deal with human conduct which is as varied as the individuals that are encountered. Our heritage of a "free society" does not encourage police controls even in those instances where the object of the control is known to be dangerous and prone to engage in antisocial acts.

3. THE MEANING OF NUMBERS.

The police administrator must be taught to understand the true meaning of a number and its relation to a statistical report. Too many persons attach an unrealistic significance to a number. It should always be borne in mind that a number is a conventional symbol and has no meaning in and of itself. It has meaning only in relation to that which it describes. It is a semantic medium whereby persons communicate quantitative information from one to another. The number must be distinguished from the incident or class of event that it counts. A number can be in error just as any other descriptive determination. The fact that a number has been assigned to an event or incident does not change the event nor does it prove its existence.

4. A TREND IS INTIMATELY RELATED TO THE BASE.

While it seems that this statement should not have to be made experience has proven that quite frequently statistical studies are attempted and reports prepared only to find that no proper base or plane of reference was established upon which the collected data could be properly arrayed for analysis. This statement also has another important connotation which underlies all statistical processes. That statement is " the end product in a statistical analysis has no more validity that the initial premise upon which the study is based". No amount of mathematical computation will correct for inadequate sampling or the use of an erroneous base.

5. THE PRESENT UNIFORM CRIME REPORT SYSTEM MUST STILL BE IMPROVED AND EXTENDED.

The present system of Uniform Crime Report classification leaves much to be desired. The broad crime groupings are used to score major offenses which differ according to some element in their definition. This is a mechanical separation that does not distinguish between crime that fall into a general definition but vary greatly in some other aspects which are significant to the police. Take for example the following situations:

A. Two juveniles while on school vacation break into a neighbors barn, steal some nails, a hammar, and a saw in order to build a tree house nearby.

B. A prowler sneaks into an unoccupied bedroom and rummages for money or jewelry while the occupants are having dinner downstairs.

C. A team of thugs, armed and with heavy burglar tools force entrance into a office and attack a safe.

Each of the above cases would be scored as one "burglary" according to current police statistical

practices. From the standpoint of practical law enforcement operation each case offers a different chalange both from the standpoint of police protective services and as far as threat of danger to life or property. While these three cases have significan differences in fact they would be indistinguishable in the formal statistical tabulation. Other examples could be given for each of the other major crime classifications - rape, robbery, auto theft, aggravated assault, etc.

6. THE EVALUATION OF PROPERTY.

Another problem that is demanding renewed attention relates to the present practices employed by police agencies in the evaluation of property that has been reported as lost or stolen. Most departments use an arbitrary system that is skewed toward the conservative scale. As long as the evaluation formula is used consistently within an agency no great harm will result. If the formula is subject to frequent or periodic change then the true meaning of property values is lost or subject to serious question. Property values as reported by various police agencies cannot be subjected to comparative analysis under the present procedures. However, the Federal Bureau of Investigation asks for and receives reports of property value from individual departments each month and publishes some of this data. This will continue to be an area of irritation and some complaint until some general formula for the evaluation of property can be developed and adopted by all agencies.

7. PREMATURE RELEASE OF STATISTICAL REPORTS.

One failing found among police officials is the tendency to rush the release of statistical data particularly when that data reflects favorably as to the accomplishments of the organization. This is a practice that can have serious consequences if not held in check. I favor the early employment of statistical determinations as they will tend to be more in harmony with the current events that are demanding control. However no administrator should pass statistical analysis reports relating to his department until he has given the report proper attention. Despite the excellence of the statistical product, its form or findings it will often require some degree of interpretation. This interpretation should be developed by the administrator or under his immediate direction.

Statistical determinations are viewed with indifference by too many administrators. They exhibit interest only when the reports reveal that police incidents are on the decline which they find favorable. When the tabulations show that police incidents are on the rise they become unhappy with the statistical process and tend to discount its significance. The process should be based upon sound statistical practices which will produce valid measures of the occurrence of events and services performed that can be accepted by the administrator whether the trend be up or down.

With these comments I will conclude my part in this program as a discussant for the paper presented by Dr. Griffin concerning "police statistics". I am in general agreement with the greater portion of Dr. Griffin's paper and have tried to touch briefly upon areas where I felt some additional discussion was warranted.

I wish to thank the Chairman for the opportunity to speak before you and wish to thank you for the courtesy extended to me this day.

DISCUSSION

Marie Vida Ryan, California Department of Corrections

Jim McCafferty's paper has been very informative. My background is prison statistics, and, therefore, my comments should be evaluated with respect to my particular interest. We are immediately assured that those collecting and evaluating prisoner statistics are not likely to lack in subject matter, at least through 1980. Mr. McCafferty's prediction of almost 319,000 sentenced adult offenders in state and federal institutions reflects a 50 percent increase in the next 20 years and, therefore, the problems and suggestions he presents should be seriously considered.

Although in the past ten years there has been a noticeable development in the states toward better statistical systems, additional progress is manditory if administrators are to meet the problems and needs of the confined prisoners.

It is extremely pleasing that the National Prisoner Statistics reports on court commitments with trends data since 1942 and on release data are to be published soon. As to the two new proposals for the "NPS" program, I would like to especially comment on the one for the collection of minimal data regarding prior commitment history beginning with the calendar year 1961. The proposed question was: "Has this court commitment ever served a sentence in an adult correctional institution?" - to be answered "yes" or "no". Although a statistical series on the recidivistic patterns of committed persons is needed badly, I have some concern as to the significance of a "yes" answer.

I. Consideration

A. The served sentence could be one day, or 20 years, or more.

1. A "yes" answer would include those who served sentences of one or 10 days for such offenses as drunkenness, disturbing the peace, vagrancy, and would also include those committed to a more severe sentence as 5-life for robbery, assault, or such.

   It seems to me that the answer of "yes" is too inclusive for meaningful evaluation.

2. The words "adult correctional institutions" were in the question. What is an adult correctional institution? We think of a jail as an adult institution, but some juveniles or persons in the younger age groups are sentenced to them. In California we have a Youth Authority. Boys sentenced to the Youth Authority may be transferred to our prisons. Although they were sentenced to a Youth Authority commitment they are being held and trained for an extended period of time in what is normally thought of as an adult correctional institution.

3. If this question is considered immediately at time of admission it is possible that the answer is not available except from the inmate, for there is a time lapse necessary to get prior criminal records, known as rap sheets, from state or federal identification bureaus. Perhaps the answer to this proposed recidivistic history question should be made at time of release from prison or at least some months after admission.

   Those of us in California who use the "NPS" series are glad to see them embarking on new programs.

B. I would like to re-emphasize a point that Mr. McCafferty mentioned which is easily overlooked when using criminal statistics; that is, the non-comparability of prison statistics from one state to another. A person must know the differences among states' laws, practices, and policies, if other than only naive interpretations are to be made. Adult institutions care for juvenile wards. Burglary in one state is called larceny in another. Differences in criminal procedures are such that what is a felony in one state is equivalent to a misdemeanor in another. A sentence for more than one year may mean confinement in a prison in one state but jail in another.

   For some branches of government like the Social Security Act, where the individual states have their own laws -- such as, amount of unemployment insurance to be paid each week and the number of weeks to be covered -- briefs of the state laws have been done in a uniform outline style for each state. At a glance, say vertically, one can see the differences in amounts of payment each week in each state, or reading horizontally one has a brief of each state's law. Some such document in outline form would help us in the comparison of prison statistics from state to state.

   I agree with Mr. McCafferty that "prisoner statistics are a useful tool if used with caution."

II. Consideration - Criminal Career Statistics Program

California has established a small criminal career recording system of men who have been in prison. But the system is extremely limited and is only applicable to parolees or dischargees who remain in California, as our record followup system is based principally upon the California Bureau of Criminal Identification and Investigation reports. At the present time we have no way of knowing what happens to the felons that do not appear on a California Bureau of Criminal Identification and Investigation report. The felon may not be involved in criminal activity or he may be involved out of California.

III. Consideration - Comprehensive Inventory of Law Enforcement-Judicial-Correctional Processes

The Attorney General's Survey of Released Procedures of 1939 and the Wickersham Reports were mentioned. These were two outstanding reports at that time. If similar studies are to be done, where should the leadership and planning be placed? In the federal government? In the Council of State governments? There seems to be a wealth of foundation funds or grant monies available today. Could we get the necessary studies this way? On whom does the leadership for inaugurating such studies rest?

IV. Consideration - Handbook on Criminal and Delinquency Statistics

If such a handbook establishing standards and collection patterns is made in this field, it seems that the task of preparing such a book rests upon the person or persons who are most qualified in leadership, planning, and experienced in the field today. From Mr. George Davis' paper on Court And Probation Statistics, which was just read, there is proof that an actual operation is being carried on and the results are evident.

V. Consideration - Human and Material Resources

Under this consideration it was stated that criminal statistics agencies must turn to colleges for personnel requirements. This is a goal we should work toward, but at the present time, actual experience is the reverse, as some colleges or universities are turning to the agencies for training. In some universities and colleges in California, the professors who have graduate students preparing papers on correctional administration which may require statistics, send the graduate students to us with such remarks as: "Model your report after the style used in Departments of the Youth Authority, Justice, or Corrections, and ask their staffs how and why they used that particular series and arrived at their findings."

We encourage this work with the students. Perhaps persons in other agencies and in other states have had similar experiences with universities and colleges.

More and more administrators in the field of criminal justice are recognizing the need of statistics and research tools. Some states who started years ago in collection of prisoners' statistical series have now started to branch into experimental research in the treatment of sentenced prisoners, while operational research is still growing by leaps and bounds. When operational research is applied to the majority of functions in the correctional field it will necessitate the use of electronic processing machines.

The Departments of the Youth Authority and of Corrections in California have asked the Department of Finance to make a feasibility study with respect to the use of electronic data processing equipment in these departments in the fields of:

(1) research and statistics.

(2) inmate records, reports and controls.

(3) business management including accounting, personnel, and property records.

(4) Correctional Industries' cost accounting, and production and sales order planning and control.

Although most of us are doing some applied research now, we realize that there is a wealth of information yet to be obtained from the data now available, and all of us are watching the progress in each locale, state, and federal government for ideas in the development in our own systems.

As expressed by Mr. McCafferty, I, too, believe that criminal statistics is becoming more and more scientific and not based upon hunches or subjective evaluations.

Thank you

III

STATISTICS OF CRIME AND CORRECTION – II

Chairman, Ronald H. Beattie, California Bureau of Criminal Statistics

---

Recent Developments in Federal Criminal Statistics – Dana M. Barbour, Office of Statistical Standards, Bureau of the Budget

Measurement of Juvenile Delinquency – Peter P. Lejins, University of Maryland

Summary Statement – Ronald H. Beattie, California Bureau of Criminal Statistics

RECENT DEVELOPMENTS IN FEDERAL CRIMINAL STATISTICS

Dana M. Barbour, Office of Statistical Standards, U. S. Bureau of the Budget

Several improvements have been made during the last few years in the criminal statistics--police, judicial and correctional--issued by the Federal government. In this paper I propose to describe briefly the existing Federal series in the field, both those based on data collected from State or local authorities and those relating to Federal offenses, paying particular attention to recent developments and noting some plans for further improvement. I want to point out certain limitations which are not always understood by some of the users of these data.

POLICE STATISTICS

In the area of police statistics we have the Uniform Crime Reports, which attract the most attention and are probably the most widely used of any of the Federal series on crime and correction. While the reports are published by the Federal Bureau of Investigation, the program is a cooperative one with the International Association of Chiefs of Police. Begun 30 years ago the number of reporters has increased over the years until about 7,000 State and local police departments, sheriffs' offices and other police agencies are now sending monthly and annual reports to the F.B.I.

Recommendations of Consultant Committee on Uniform Crime Reporting

Late in 1957 the F.B.I. appointed a Consultant Committee on Uniform Crime Reporting to make an independent analysis of the program. The committee was headed by Dr. Peter Lejins of the University of Maryland. The other members were Dr. Charlton F. Chute, Director of the Institute of Public Administration, New York City, and Col. Stanley R. Shrotel, Chief of Police of Cincinnati. Their report was published in 1958 as a special issue of the Uniform Crime Reports. It contained 22 recommendations, all of which have been accepted at least as ultimate goals by the I.A.C.P. and the F.B.I. Only the major ones can be considered here.

One of the most important recommendations related to expansion of coverage of the statistics on offenses known through arrests, which did not include rural areas at all, and which had represented only about 40 percent of the urban population. Since the 1958 report was issued, Mr. Jerry Daunt, the new chief of the Uniform Crime Reporting Section of the F.B.I., reports some progress, though he thinks the data for rural areas are not yet representative enough to justify publication in the 1959 annual report. The Consultant Committee considered sampling as an alternative to attempting to collect information on crime for the entire country. While not ruling out sampling as a possibility if the F.B.I. is not able to build up its reporting, the committee was inclined to favor complete coverage if that could be attained. The reasons were the interest of police and the public in crime figures for individual communities and States and the almost complete coverage already achieved for a part of the Uniform Crime Reports--the offenses known to police.

Another important change resulting from the work of the Consultant Committee is the use of annual estimates of population instead of the last decennial census figures in computing crime rates. Because of the differential population growth in the United States, the continued use of the 1950 figures had made the crime rates in States like California and Arizona look worse than they were, while having the opposite effect in States or cities with stable or declining populations. A somewhat similar change was the replacement of the old urban-rural classification used in presenting statistics with a three-way classification of standard metropolitan areas (now standard metropolitan statistical areas), other cities, and rural areas.

An important change in publication policy was the shift from semiannual bulletins to an annual bulletin. This recommendation was made to permit more time for verification of data and statistical analysis, and to eliminate a certain amount of overlap between the semiannual reports. The first annual bulletin under the new plan appeared in September 1959 for the calendar year 1958. These annual bulletins are being supplemented by a four-page quarterly release containing percentage changes in the city crime index and preliminary figures by offense for cities over 100,000.

Finally the so-called Class I offenses used in measuring crime trends have been reexamined and some changes made as the result of recommendations of the Consultant Committee. The offenses now included are murder and nonnegligent manslaughter, forcible rape, robbery, aggravated assault, burglary, larceny--$50 and over in value, and auto theft. Negligent manslaughter was dropped partly on the ground that since it was made up almost entirely of culpable but nevertheless accidental traffic deaths it does not seem a proper offense to be used in an index of crime, and, partly, because the police are frequently hesitant in classifying an accidental death as manslaughter and tend to wait for action by the grand jury. Larceny under $50 was also dropped for a number of reasons, the most important being that it is often a minor offense which is not reported to the police. Beginning with 1958 the F.B.I. on its own initiative excluded statutory rape from its Class I offenses on the ground that this was an offense which was seldom reported to the police. The seven offenses remaining are included because they meet several criteria reasonably well. They are serious offenses; they are usually reported to the police (though even among these seven offenses the completeness of reporting varies); their definition is relatively uniform throughout the country; and they are sufficiently numerous to be

statistically significant. They include offenses against both person and property, and they also are of importance to the cooperating police departments.

Uniform Crime Reports as a Measurement of Crime

We may ask how well this revised list of seven offenses measures trends in crime. More specifically, is the F.B.I. justified in stating, as it did in its last annual report, that "crime in the United States showed an over-all increase in 1958 of 9.3 percent over the 1957 level," or even, as the most recent quarterly release does, that preliminary figures for the city crime index in the first half of 1960 show an increase of 9 percent over the corresponding half of 1959? My own opinion is that while the trends in the United States for each of the seven offenses shown are fairly reliable, we are still far short of a valid index of crime as a whole in the United States. Perhaps such an index can never be compiled.

Implicit in the F.B.I. index is the assumption that the seven offenses measure, not merely all types of offenses known to the police, but all crime--reported or unreported--in the United States. While the seven index offenses are certainly important ones, such serious types of offenses as forgery and counterfeiting, embezzlement, drug peddling and drunken driving are not included.

No direct numerical comparison between the total number of offenses known to the police and the total shown for the seven index offenses is possible because (with minor exceptions) the F.B.I. does not collect data on offenses for other than the seven. If, however, we use the figures on arrests by offense (which are less satisfactory because they are less comprehensive), we find that of the total of 2,340,000 arrests in 1958 only 256,000 were for the seven offenses shown in the index of crime. If we subtract from this 2,340,000 total arrests on suspicion, for vagrancy, for disorderly conduct, and for drunkenness (on the ground that these are either not crimes at all or are not very serious ones), we still have almost three times as many arrests as the total shown for the index offenses. Moreover, several of the offenses not included in the index are covert rather than overt offenses, and as such are less likely to become known to the police. This is particularly true of embezzlement, fraud and sex offenses.

We must conclude that in attempting to measure crime we are dealing with an unknown universe, of which the offenses known to police represent only a part. I have not seen any evidence to demonstrate what, if any, correlation of movement exists between these seven offenses and total crime in the United States. As Professor Griffin put it four years ago, "The relationship between known crime and total crime is unknown, and is, in part at least, a function of the intensity of police activity."1/

One more point may be noted before leaving this matter of a crime index. This is the wide range of criminal behavior that is covered by the seven offenses included in the F.B.I. index, and the lack of any weighting by seriousness of the offense. For index purposes a $50 larceny is equated with a premeditated murder.

Use of the Uniform Crime Reports for Interarea Comparisons

Wide use is also made of the Uniform Crime Reports for interstate or intercity comparisons. For the seven index offenses this is facilitated by the publication in the annual bulletin of the rate per 100,000 inhabitants for each offense and all seven combined by States and by standard metropolitan statistical areas.

The dangers of interarea comparisons are recognized by the F.B.I. itself. In the last annual report this statement is underlined: "Caution should be used in comparing crime data for individual cities because the differences in the figures may be due to a number of factors. Such comparisons are not necessarily significant even though the figures for individual communities are converted into terms of offenses per 100,000 inhabitants."2/ The report goes on to say that it is more important to determine whether the figures for a given community show increases or decreases than to ascertain whether they exceed or fall short of those for some other individual community.

The Chairman of this session, Ronald Beattie, Chief of the California Bureau of Criminal Statistics, has devoted considerable attention to use of the Uniform Crime Reports for interarea comparisons in an article in a recent issue of the Journal of Criminal Law, Criminology and Police Science.3/ He notes that the 1958 report shows California has the highest crime rate of any State in the country and Los Angeles the highest of any metropolitan area, and suggests two possible explanations: one is California's broad definition of burglary, which includes acts that many other States would classify as larceny under $50; the other is the high standard of police efficiency in the State, the Los Angeles Police Department in particular being one of the best in the country. These high standards include good recording and reporting systems. Ronald Beattie's own organization, the California Bureau of Criminal Statistics, does a good deal to promote high standards

---

1/ John I. Griffin, "New Perspectives in Police Statistics," Journal of Criminal Law, Criminology and Police Science, Vol. 46, March-April 1956, p. 879.

2/ "Uniform Crime Reports," 1958, p. 79.

3/ Ronald H. Beattie, "Criminal Statistics in the United States--1960," Journal of Criminal Law, Criminology and Police Science, Vol. 51, May-June 1960, p. 49.

of recording and reporting among the California police departments. Of course, the highly urban character of California's population is also a factor.

In his article Beattie points out figures and ratios for other States than California and cities other than Los Angeles that appear suspect. To cite an extreme example there were only five cases of aggravated assault in Vermont in 1958 for a rate of 1.3 per hundred thousand inhabitants. There were more than twice as many murders and nonnegligent manslaughter cases, the rate being 3.2. The national rates, on the other hand, are 65.5 for aggravated assaults and 4.7 for murder and nonnegligent manslaughter. This suggests that police authorities in Vermont are just not classifying assaults properly in their reports.

In fairness it should be said that the F.B.I. has devoted a great deal of effort to obtaining reports from police departments according to its uniform classification of offenses. A "Uniform Crime Reporting Handbook" is given to participating police agencies, and special instruction bulletins on particular subjects are issued from time to time. F.B.I. agents are available to visit local police departments where needed to help in the preparation of reports. There have been instances where the F.B.I. has omitted data for a city with the footnote, "The crime reporting for the city indicated does not meet acceptable standards established by the Committee on Uniform Crime Records of the International Association of Chiefs of Police." (Incidentally, it may be noted that such a footnote has proven a powerful stimulus in improving the reports of the city in question.) Despite these efforts to maintain high standards, it is true that, as the F.B.I. itself puts it, "The basic responsibility for the accuracy of the reports rests with the local law enforcement executives from whom and for whom the data are collected."4/

Particular caution must be used in interpreting the data on juvenile arrests in the tables on arrests by age and offense. The practices of jurisdictions differ widely in their treatment of arrests of juveniles. Some will not report detention of a juvenile as an arrest at all. This makes impossible interarea comparisons. As regards trends, the increasing emphasis on juvenile delinquency and the creation of special units in police departments to deal with juveniles may account in part for the increase in the number of arrests.

Statistics on Federal Crime

We should not leave the area of police statistics without discussing briefly statistics on Federal crimes. These may be divided into two classes: ordinary offenses which become Federal crimes because they are committed on Federal property or involve interstate commerce, and specialized offenses which are peculiarly Federal because they involve violations of Federal laws such as the immigration, internal revenue, postal, customs, and pure food and drug laws. Probably most of the ordinary offenses are included in the F.B.I. Uniform Crime Reports, without being identified as Federal offenses. The theft of an automobile will be reported as such by the local police department without waiting to determine whether a Federal crime may also have been committed because the automobile had been driven across the State line. The F.B.I. also makes an effort to include offenses committed on Federal reservations. However, specialized Federal offenses are not included in the Uniform Crime Reports. Usually such offenses are not reported anywhere until an arrest has been made, which may be done by officers of the Federal agency having the responsibility for enforcing the particular law. While some of these agencies, such as the Immigration and Naturalization Service, do publish data on arrests, ordinarily a peculiarly Federal offense does not become a "statistic" until it is reported by the U. S. Attorney's office.

JUDICIAL STATISTICS

Coming now to the field of judicial or court statistics, we find a large gap. There are at present no national data on the operations of State and municipal courts in processing criminal cases involving adults.

Census Series on State Judicial Criminal Statistics

From 1932 to 1946 there was a series on State judicial criminal statistics compiled by the Bureau of the Census. The coverage of this series was never truly national. In its last year only 25 States were cooperating, and not all counties were included even in those States. In addition to incomplete coverage, there was lack of comparability in the data because of variations among States with respect to the types of courts handling felony cases and because of differences in the offense classifications used. This last was true despite the efforts of the Census Bureau to obtain uniform reporting according to a standard classification. Some of the data were of questionable reliability, largely because most clerks of court, who prepared the reports initially, insisted on using a tally-sheet method of reporting cases. There were long delays in publishing some of the reports. Finally, the Census Bureau appears not to have attached much importance to judicial criminal statistics and to have devoted very little staff or attention to them. In 1946 the series was discontinued by the Census Bureau, after consultation with a committee of experts convened by the Office of Statistical Standards. The experience does illustrate the great difficulties faced by any Federal agency which tries to compile data based on reports from thousands of jurisdictions with varying types of organizations, laws and recordkeeping practices.

Juvenile Court Statistics

National juvenile court statistics have been collected and published by the Children's Bureau for more than 30 years. It was not until 1956, however, that it could be said with confidence

4/ "Uniform Crime Reports," 1958, p. 21.

that the number of cases reported was representative of the country as a whole. Some regions of the country had been underrepresented, others overrepresented, because the statistics depended on the cooperation of the juvenile courts.

The first report based on a national sample of courts was for the year 1956. This sample utilized the Census selection of 230 primary sampling units, each consisting of a county or several contiguous counties, for its own Current Population Survey. The selection was made on a random basis, after stratifying the primary sampling units on the basis of geographical location, population density, rate of growth, racial composition, economic characteristics, etc. It was found that there were 502 courts in these 230 primary sampling units. Estimates of the number and type of cases for the United States are made, based on the reports from the sample courts. The rate of delinquency cases per 1,000 children over age 10 is also computed by type of court (urban, semiurban and rural). These data are supplemented by reports from a number of courts not included in the sample.

The cooperation of the courts included in the national sample has been good, but occasionally there are delays in submitting the reports. The policy of the Children's Bureau has been to work through the State agencies (State Welfare Departments, Youth Authorities or Attorneys-General). While this avoids possible duplicate reporting and strengthens the State agency, it has sometimes been a cause of delays in publication.

Unfortunately, the data obtained from the national sample of courts are meager, being limited to the number of cases disposed of and the method of handling the cases--official or unofficial. Traffic cases have been reported separately since 1957. No information is obtained on the reason for referral or the method of disposition of the cases, nor is there anything on the characteristics of the juveniles brought before the courts. When the national sample of courts was established it was deemed advisable to keep reports to a minimum in order to secure cooperation. The Children's Bureau hopes later to expand the information called for, but has no definite plans at present. It should be noted that last year the Children's Bureau and the National Institute of Mental Health of the Public Health Service made a special study of juvenile delinquency at the request of a subcommittee of the Committee on Appropriations of the House of Representatives. This study involved a number of one-time statistical surveys calling for much more detail than is available in the regular series. The results of these special surveys have been summarized in the published hearings of the House Committee 5/ and are also appearing as a series of pamphlets on juvenile delinquency issued by the Children's Bureau.6/

5/ "Report on Juvenile Delinquency," Hearings before the Subcommittee of the Committee on Appropriations, House of Representatives, 86th Cong., 2nd Sess. (1960), pp. 22-57.

Federal Court Statistics

When it comes to judicial statistics for Federal courts we are in a much better position. The reports of the Department of Justice go back to 1872. Shortly after its creation in 1939, the Administrative Office of the U. S. Courts took over judicial statistics. These statistics are based on a separate card for each case prepared by the clerks of all Federal district and circuit courts. The information is put on punch cards, and quarterly and annual reports are prepared. The annual report of the Director of the Administrative Office of the U. S. Courts shows by circuit and district the number of criminal cases commenced, terminated and pending at the close of the year, and the number of defendants involved; information on cases and defendants by nature of offense; and criminal defendants disposed of, by nature of offense and disposition, including length of sentence for those sent to prison. Separate figures are presented for juveniles. In addition there are a number of tables on probationers under the supervision of United States probation officers, showing the number received for supervision and removed from supervision during the year by districts; the caseload of probation officers; the age, race, sex and offense of probationers; and the number of parolees as well as probationers who were reported as violators during the year.

The Annual Report of the Attorney General also includes some judicial statistics, as well as detailed tables showing by judicial district the number of criminal and civil cases handled by United States Attorneys and their status at the end of the year. These statistics are based on reports from the United States Attorneys.

CORRECTIONAL STATISTICS

The third and last field is that of correctional statistics.

National Prisoner Statistics

The National Prisoner Statistics program has been discussed by Mr. McCafferty of the Bureau of Prisons,7/ so I shall not go into it here. I would like to emphasize two encouraging developments--the progress being made in reducing the backlog and putting the detailed statistics on court commitments and prisoners released from State and Federal institutions on a more current basis, and the use by the Bureau of Prisons of

6/ "Juvenile Delinquency: Facts, Facets," No.___, (subtitle--subject of report), U. S. Department of Health, Education, and Welfare, Social Security Administration--Children's Bureau.

7/ "Prisoner Statistics--National and State," Paper read at the Statistics of Crime and Correction session, Annual Meeting, American Statistical Association, August 23, 1960.

punch cards which are being prepared by an increasing number of the statistical units of State central correctional agencies.

Prisoners in Jails and Local Institutions

The national prisoner Statistics program does not include prisoners in jails and other local correctional institutions. The only coverage in this area is that of the Census Bureau, which enumerates prisoners along with the rest of the population in its decennial census, and publishes statistics on them in its special report on institutional population. The report for the 1950 census was published in 1953,8/ and presented considerable data on the age, sex, color, etc. of persons in local jails and workhouses by State. No information was obtained on offenses or whether inmates were waiting trial or serving sentences. The Census Bureau plans to present about the same type of information for the 1960 census. However, the report on institutional population will probably not be published until late 1962 or early 1963.

Juvenile Delinquents

Mention should be made of the statistics which the Children's Bureau compiles on public training schools for juvenile delinquents.9/ These statistics were put on a regular annual basis in 1956 and include data on movement into and out of the institution, the status of the institutional population at the end of the period, and some information on staff and expenditures of these institutions. Certain supplementary data on training schools have also been obtained from time to time.

Federal Prisoners

Statistics on Federal prisoners are included in the National Prisoner Statistics program. In addition the Bureau of Prisons publishes considerable information about them in its annual report, "Federal Prisons." This includes data on characteristics of the prisoners, length of sentence and average time served by those released by type of offense. One important type of information not now obtained for prisoners in State institutions is gotten for Federal prisoners. This is information on recidivism by type of offense for prisoners committed to Federal institutions during the year.

Statistics on parole of Federal prisoners are included both in the "Federal Prisons" report and the report of the chairman of the Board of Parole contained in the Annual Report of the Attorney General. These show parole decisions and parole grants by offense, average time served prior to parole, and violations by type of offense.

PLANS FOR FURTHER IMPROVEMENT

Thus far in this paper we have discussed what Federal criminal statistics are presently available, noting particularly recent developments. Brief mention might be made of a plan for better-integrated statistics on Federal crime developed by Ronald Beattie when he was in Washington last winter. The system would begin with complaints filed with the U. S. attorneys, and involves getting some additional information about complaints not prosecuted and more detail about defendants. It also involves agreement between the Department of Justice and the Administrative Office of the U. S. Courts on uniform classification of data and uniform presentation of such data. Beattie points out that in almost every Federal case the offender is fingerprinted and given an F.B.I. number. If this number could be put on all documents relating to the offender throughout the criminal process, the problem of identification would be solved and it would be possible to show for X number of alleged Federal offenders, the proportion not prosecuted, the proportion convicted by the trial court, the proportion receiving fines or placed on probation, the proportion sent to prison, etc. It should be emphasized, however, that this plan is still in the proposal stage.

My own office, the Office of Statistical Standards in the Bureau of the Budget, is planning to take the lead in creating an interagency committee of Federal statisticians in the criminal statistics field. We have noted the need for closer coordination in dealing with a number of problems. One of the problems such a committee might consider is the development of a single standard classification of Federal offenses.

CONCLUSIONS

In conclusion, I have tried to indicate that, despite their limitations, there have been a number of encouraging developments during the last few years in the criminal statistics compiled and published by the Federal government, and that there are plans for further improvement. It may well be, however, that filling the major existing gaps in this field must await further progress by the States. As Thorsten Sellin put it ten years ago: ". . . It is more than likely that we can never hope for a further fundamental improvement in the structure of a system of national statistics based on voluntary cooperation until we have strengthened the foundation. That foundation must be laid in the individual states."10/

8/ "Institutional Population," United States Census of Population: 1950, Special Report P-E No. 2C.

9/ Published under the title, "Statistics on Public Institutions for Delinquent Children, (year)," Children's Bureau Statistical Series, No.___.

10/ Thorsten Sellin, "The Uniform Criminal Statistics Act," Journal of Criminal Law and Criminology, Vol. 40, March-April 1950, p. 683.

MEASUREMENT OF JUVENILE DELINQUENCY

Peter P. Lejins, University of Maryland

The relatively short time assigned to this broad topic with its many ramifications suggested the general structure of a catalog of a number of the more important selected issues on which a stand must be taken and many of which should be further explored.

That criminal statistics are "poor" in general, but particularly in this country, especially owing to the fact that law enforcement here is a function of local government, is readily and frequently acknowledged by American criminologists. That juvenile delinquency statistics are even poorer than criminal statistics is also quite obvious, an additional reason in this case being the fact that the special institutional systems for juvenile delinquency control are of very recent origin. Consequently, no uniform standards of procedure have as yet jelled, and no personnel with a uniform background has so far been developed. The result is that data from different communities --and even from one and the same community over a period of time-- are not comparable.

In spite of the fact that the title of this paper would probably convey a rather definite idea to most people in the field, the statement "measurement of juvenile delinquency" can really denote various things. E.g., it could be interpreted as referring to the measuring of the delinquency of a single juvenile in the sense of the intensity of his antisocial attitudes or the persistance of the misbehavior. For the purposes of this presentation, however, the term is assumed to be used in its currently most frequent meaning; namely, that of measuring the amounts of juvenile delinquent behavior in communities for the purposes of comparing these communities in this respect with one another and also of establishing variations over a period of time within one and the same community. In other words, we are dealing here with the question of juvenile delinquency statistics.

First of all, there is the problem of what should be counted or measured, or just what is understood by juvenile delinquency. There appear to be two tendencies in this respect. One is, to consider as juvenile delinquents those juveniles who have been officially acted upon by such agencies as the police and the courts --in other words, those who have a record. This writer suggests referring to this concept as the formal definition or formal concept of juvenile delinquency. The statistics of juvenile delinquency available in this country are, of course, based on this kind of a definition. The Juvenile Court Statistics published by the Children's Bureau are of that nature, as are the arrest figures published by the Federal Bureau of Investigation in the Uniform Crime Reports for the age groups usually considered juvenile. But researchers on juvenile delinquency also usually accept this approach. Thus, to cite but one example, in Unraveling Juvenile Delinquency the Gluecks included only delinquents from training schools into their delinquent group.

On the other hand, there are and always have been serious students of juvenile delinquency who believe that the fact that one juvenile has been officially acted upon by the police or the juvenile court and another one has not been so acted upon, does not make one of them a delinquent and the other not, as long as both of them "were doing one and the same thing". Some authors interpret the presence of an official record of delinquency as a matter of accident and resolutely condemn the use of only "official delinquency" in research. They subscribe to what one might term a "descriptive" or a "content" definition of delinquency. The advocates of this type of definition apparently visualize that the instances of delinquent behavior can be described in the same manner as adult crimes are described in the "statements of fact" in the sections of a criminal code.

As an example of this type of approach might be cited Sophia M. Robison's Can Delinquency Be Measured? (1936), or, to give a more recent example, the use of the concept of "unrecorded juvenile delinquency" by James F. Short and F. Ivan Nye in a study reported in an article which contains this term in its title (1958).

This writer accepts the formal definition of juvenile delinquency as against the content or descriptive definition. First of all, at least at present and in the immediate future, this formal definition seems to be the only kind of concept of delinquency that can be used for the purpose of sizing up the scope of the delinquency problem in larger communities and the country as a whole. Secondly, this writer believes that this formal definition is the proper definition of juvenile delinquency, not only because of its statistical usefulness, but also essentially so. The essence of juvenile deviations which are delinquencies consists in the fact that society finds it necessary to step into the process of socialization of the young as ordinarily carried on by the usual social institutions --the family, the church, the neighborhood, etc.-- and to place the juvenile in the hands of its special agencies --the police, the courts, etc. The very entry into the picture of such a special agency as the police or the court is a social fact which differentiates the deviations of juveniles into delinquent and non-delinquent behavior.

Thus it is suggested that the measuring of juvenile delinquency means statistics of juveniles acted upon by the police and the courts qua juvenile delinquents.

The acceptance of this formal definition of juvenile delinquency makes recognition of the following fact quite imparative: the registry of a case of official juvenile delinquency is a function of at least two factors, namely, the behavior of the juvenile and the policies of the law enforcement agencies. Any variation in juvenile delinquency statistics may be due to the variations in the behavior of the juveniles or in the policies of the special agencies. The fact is that with a change in the person of the juvenile court judge, a tougher policy on the part of the police department, or perhaps an increase in such personnel as probation officers, the statistics may fluctuate drastically.

Thus in the collection of juvenile delinquency

statistics, special care must be taken to provide for a clear distinction between changes in the behavior of the juveniles and changes in the policies of the enforcement agencies. This must be done when setting up the collection of data and in the interpretation of the collected data.

With regard to the policies of the special (enforcement) agencies affecting statistics of juvenile delinquency, one should differentiate between policies in the real sense of the word and what one might refer to as procedural techniques. Policies in the true sense of the word mean determination of the point in the developing problem at which the special agencies decide to step in. Such a policy on the part of the law enforcement agency determines the height of the "nuisance threshold," or expresses the society's conception of what constitutes such a degree of danger to the child that public remedial action is necessary.

The procedural techniques which can and do affect juvenile delinquency statistics mean differences in the practices of record keeping. While one court may keep no record of informal cases, another may include these in the statistics. A hardly noticeable change in the recording technique may mean a difference in figures that is much greater than the supposed trends in juvenile behavior.

The next issue to be considered here is the question which of the various juvenile delinquency statistics should preferably be used for the purpose of measuring juvenile delinquency. This brings to mind the well-known and so frequently quoted maxim of Professor Sellin with regard to adult criminal statistics, viz.: "The value of a crime rate for index purposes decreases as the distance from the crime itself in terms of procedure increases." Does this formula apply also to juvenile delinquency statistics?

It is obvious that many more factors must be considered in the case of delinquency statistics than in the adult field in this respect, because the concept of juvenile delinquency, as our modern societies deal with it, is more complicated than the concept of adult crime. Again, in view of the limited scope of this paper only two of these factors will be briefly pointed out.

1. Communities which have introduced juvenile courts, have on the whole--at least in this country--introduced two elements which are different as compared to the control of adult crime and which are of interest to us here:

   a. the juvenile is being adjudged delinquent <u>in general</u>, rather than found guilty of a <u>specific</u> criminal offense. Because of this, relatively little attention is given to the identification of the offense committed by the juvenile. The focal point of the juvenile court proceedings is, rather, the establishment of the child's need for special treatment which the court has at its disposal and may prescribe--such as probation, an out-patient psychiatric clinic, foster placement, incarceration, etc.
   b. the basis for adjudging a juvenile as a delinquent may be not only criminal code offenses, but also other kinds of behavior which are not considered crimes if engaged in by adults. Here belong, e.g., being ungovernable, association with persons who are known to be immoral or criminal, running away from home, truancy, etc.

Thus statistics of juvenile delinquency, by their very nature, are not the same as statistics of criminal code offenses committed by juveniles. They are more than that, but at the same time they may by-pass some of the criminal offenses committed by juveniles. They are most certainly not concerned with the identification of specific offenses committed by juveniles and therefore cannot be relied upon for precision in this respect. Many a juvenile court decision makes no reference to any specific single criminal act of a juvenile.

All this indicates that in the field of juvenile delinquency it is not the "crimes known to the police", but rather the "juvenile court statistics" that should be resorted to for the purposes of an index of juvenile delinquency, unless, of course, the police assumes and performs the functions of the juvenile court. Thus, Sellin's formula does not apply to the juvenile field.

2. With regard to statistics of criminal code offenses committed by juveniles, the following observations seem to be in order. Although some of these offenses are obviously juvenile offenses from the very moment the law enforcement process goes into action, many are first reported as criminal offenses, and only when the offender has been found does it become known that he is a juvenile. Because of this, statistics of criminal code offenses committed by juveniles are usually considered to be arrest statistics.

In the light of the prevailing conceptions of juvenile delinquency and its control in this country, the policies suggested --and in many cases actually adhered to-- for arresting juveniles differ considerably from those applicable in the case of adults. Thus, if "criminal code offenses committed by juveniles" are to be measured by "arrests," care must be taken to correctly assay the differences which this injects into juvenile data as compared with adult arrests. There are possibilities of both more restricted and more liberal arrest policies with regard to the juveniles. The more restricted juvenile arrest policy is usually due to a set of attitudes prevailing in this country to the effect that juveniles must be shielded from the harshness of the conventional law enforcement process, that criminal action should be avoided whenever possible, that "giving the kid another chance" is desirable, etc. The more liberal use of "arrests" with regard to the juveniles, on the other hand, has its roots in the assignment of preventive and welfare functions to the police in the handling of juvenile problems.

The police contacts with juveniles can be differentiated into four categories:

(1) Contacts based on the juveniles' involvement in criminal code offenses.
(2) Contacts based on the type of behavior which by law (usually special statutes establishing the juvenile courts) may be the basis for adjudication as a juvenile delinquent, but which behavior does not constitute a criminal code offense--i.e., a crime if committed by an adult.
(3) Contacts based on the function of preventing juvenile delinquency which is often assigned to the police by law, by administrative orders, or by the informal but

very definite expectation of the community. This type of contact results from the police anticipating the acts of the first two categories before such acts have actually been committed.

(4) Welfare contacts, made by the police in terms of the concept of the child's welfare, without any implication of any delinquency on the part of the child.

In view of these many-fold activities of the police with reference to "children in trouble," it becomes apparent that if statistics of juveniles arrested for criminal code offenses are desired, a careful differentiation of police contacts with juveniles needs to be made. The current formula might read approximately as follows: if the offense and the circumstances are such that, if the offender were an adult, an arrest would be tallied, an arrest of a juvenile should be reported.

Granted the theoretical soundness of this procedure, one is of course still faced with the fact that the individual policeman's interpretation of his own action becomes the basis for the juvenile arrest statistics. It is of course true that to a certain extent also the statistics of crimes known to the police and the arrest statistics contain this element of danger, but in the case of juveniles the situation is presumably much more complicated in view of the need to differentiate between law enforcement motivation, the preventive motivation, and the welfare motivation in a loosely defined and permissive setting. The problem can of course only be met by:

a. carefully structured uniform instructions to the police to help proper classification of contacts with juveniles;
b. education to insure the policeman's understanding of his complicated role in the juvenile area.

It can be argued that both "statistics of juvenile delinquency" and "statistics of criminal code offenses committed by juveniles" have their justification and should be gathered. The use of both of these analytical tools may presently be justified by:

1. their basic usefulness for a more complete and more detailed analytical picture of the problem of juvenile delinquency.
2. the fact that in this country the issue of the interpretation of the problem of juvenile delinquency is by far not settled yet with two concepts, --the concept of criminal offenses committed by juveniles and the much broader specific concept of juvenile delinquency-- both currently in use.
3. the fact that any comparisons on an international scale require the above differentiation in view of the differing conceptions and policies adhered to by different countries.

Three more issues will be mentioned here very briefly and very generally.

The concept of criminal career records as distinct from agency criminal statistics --as I like to formulate this distinction-- which has been recently brought up frequently with regard to criminal statistics in general, should be pointed up. The need for including the juvenile record into the total criminal career record is obvious. Yet also in this respect the juvenile field offers some additional difficulties.

1. The general aversion to the finger-printing of juveniles makes the identification of the individual more difficult. Hence, assembling of the career record around the fingerprints and the number of the offender is more difficult with regard to a juvenile.
2. The assembly of the more serious elements of a criminal career is more feasible. Juvenile delinquencies, as mostly minor offenses, are not as a rule handled as carefully. Usually it is the record of felonies or indictable offenses that is kept.
3. The policy of protecting the juvenile from a public record offers additional difficulties, especially in this country.

The difficulties in obtaining reliable and valid data on juvenile delinquency in its totality for the United States has recently lead to the exploration of sampling procedures as a substitute. As a matter of fact, the major current undertaking in the area of delinquency statistics, the Juvenile Court Statistics by the Children's Bureau, changed its earlier method to that of "estimates from a national sample", beginning with its 1956 report. It is obvious that not only the feasibility of obtaining the statistical data and the scope of the necessary effort, but also the purpose of the collection should be considered in determining the method. Estimates from samples provide only a general picture for the universe as a whole from which the samples are taken. The statistical picture of the delinquency problem for individual communities, to be used by these communities in the planning and execution of their preventive and control efforts and in making evaluations by comparison with other communities can be provided on a national scale only by appropriate statistics of the universe.

Finally a few words might be said about the currently rather frequently raised issue of national versus state and local programs for the collection of delinquency statistics and, in general, crime statistics. The major point of the current discussion is really the comparative effectiveness of the two procedures for the development of good statistics. Those who advocate concentrating the effort on state and local statistics emphasize as their advantages the better understanding of the true meaning of the data within the given legal, administrative and cultural conditions of the particular state, as well as the greater authority available for obtaining information from the local agencies. They visualize the national statistics, thereafter, as a simple summation of the data collected by the individual states. Those, on the other hand, who emphasize the initiative and activities of the national organizations and federal agencies, claim that the assumption of the development of statistics by the states begs the question, that the national-scale prompting is exactly what is needed to stimulate the states to develop the state statistical systems, and that only the national organization of a statistical system can assure the uniformity of statistical categories.

SUMMARY STATEMENT

Ronald H. Beattie, Chief, California Bureau of Criminal Statistics

This session in the field of criminal statistics, which is the first scheduled by the American Statistical Association for some years, is timely and has given all of us an opportunity to take stock and to consider just where we stand and what must be done from now on to make real progress towards measuring and accounting for crime and criminal offenders in this country.

After listening to this series of excellent papers which have touched upon most of the major problems relating to statistics of crime and correction, it would seem somewhat presumptuous on my part to attempt to summarize or even review the many significant points which have been called to our attention. It may serve a more useful purpose if I attempt to underline briefly some of the continuing major problems in this field of statistics from the point of view of our California experience.

As far as I know, the California Bureau of Criminal Statistics is the only state bureau in the country that has been given a maximum responsibility for the collection of all available information relating to crime and delinquency within the state served. Over the past fourteen years we have had some measure of success in developing state crime statistics, due primarily to the genuine interest and concern of both state and local agencies engaged in criminal law enforcement and correction who have supported, encouraged, and fully cooperated with us in all the efforts that have been undertaken to collect statewide information in their fields. Despite this favorable situation, we candidly admit that we are far from the point of having developed adequate, acceptable, and useful criminal statistics.

I would like to suggest that the sources of data for all phases of criminal statistics are to be found in the agency records of three operational areas: law enforcement, administrative process of criminal justice, and corrections.

Police Statistics

In my opinion, our weakest statistics are in the police or law enforcement field. This is most unfortunate because it is in the primary law enforcement agency to which offenses are first reported that the crime problem first comes to public attention. These agencies take the first steps towards investigations of offenses and the apprehension of persons who are criminal offenders. It is here that the most complete picture (in terms of what is known about crime) is possible. Yet our information of crime at the police level has been, for the most part, put together in gross general terms, obscuring a real understanding of either the volume of specific kinds of crime that are occurring, or the kinds of persons who are first brought into the criminal process as alleged offenders.

The organization of our efforts to control criminal justice in this country has a great deal to do with these limitations. Law enforcement is primarily a local responsibility. There are literally hundreds of policing agencies in most states, and probably over 10,000 in the nation, but it is the records and information in the hands of these agencies that must supply the primary information concerning criminal offenses and criminal offenders. To obtain uniformly defined and classified data from these innumerable sources is a difficult task. The efforts that have been made in this direction on a national basis are well known as the Uniform Crime Reports series. Since 1930 the Federal Bureau of Investigation has sought to collect information on a certain series of crimes reported, from many thousand local police agencies, on a monthly basis in summary form. In addition, other information on arrests is collected, usually on an annual basis. This collection has severe limitations simply because, despite uniform directions, it is impossible to collect information from so many sources with the uniformity desired. The differences that occur in definition and classification in all of the various states only add to the problem of obtaining information in common terms. Yet, this is the only method that has been used to date. This method is used also by the California Bureau of Criminal Statistics to obtain information on crimes reported and felony arrests from the 400 law enforcement agencies that exist in the State of California.

Not only is summary information, under these circumstances, limited and restricted as to its comparability, but it does not permit description and analysis of the data in detailed terms. Individual offense or arrest reports are always superior as a method of obtaining information. It would seem that if we are ever to know in realistic and accurate terms how much crime we have, what its real nature is, and what kinds of persons are involved in these offenses, the next step is to devise methods of obtaining this information on an individual reporting basis. There should be another another thorough-going study of police records today, such as was made in the late Twenties preceding the development of Uniform Crime Reports. It is our hope that in California during the next decade we can carry on experimentation of this nature and take definite steps in the direction of individual crime classifications.

Process of Criminal Justice Statistics

In the area of the process of criminal justice, we, today, in 1960, are not as advanced

as was the case during the decade of the Twenties when some of the great major crime surveys demonstrated how to account for criminal offenders from the point of apprehension through the various steps of the selective process that we call the administration of criminal justice. Mortality tables were developed that showed how many and what kinds of offenders were released by law enforcement agencies and each of the successive steps of prosecution. It will not be possible to develop this kind of information unless there is an individual accounting of each offender at the law enforcement level and a procedure established whereby he can be traced through the process of hearing, trial, and conviction. We do have, today, an accurate accounting of individuals who were prosecuted in the general trial courts of California--the superior courts--but we cannot identify these same people at the point of arrest or in preliminary stages of prosecution.

A project is now being carried on in the State of California which will serve as an experiment in this direction. Because the narcotic problem is one of extreme concern, and facts relating to narcotic offenders are meager, the Bureau of Criminal Statistics was authorized a year ago to set up a reporting system which would account for all persons arrested by law enforcement agencies on any narcotic charge. The Bureau obtains from law enforcement agencies an individual copy of the original arrest report and any supplemental reports pertaining to each offense and offender. Also, received is a copy of the criminal record that is developed by the State Bureau of Criminal Identification and Investigation upon receipt of the arrestee's fingerprints. This is a project, therefore, that begins with the individual arrested. The attempt is made to obtain all further information as to how this person was handled and what was the final outcome or disposition of the charges. A major difficulty that is encountered here, and will be encountered in any attempt to obtain information on the outcome of prosecutions is the lack of any systematic reporting by courts handling misdemeanor offenders. This is a problem that is common, I know, to all states and areas. It is much easier to build a reporting system at the felony level in a state than it is to cover dispositions from the numerous courts that dispose of misdemeanants. Such information, though, is most essential, not only to account for the final disposition of a person arrested, but also to be transcribed to the identification records so that when a person's record is issued by the FBI, state identification bureau, or local police agency, such entry indicating an arrest will carry also information concerning the disposition made of the arrest. Far too often this information is never reported and the criminal record will contain pages of arrests with little indication whatever of what happened to the individual in many of the arrests reported.

Because of the decentralized nature of the administration of criminal justice--law enforcement agencies first investigating and arresting offenders, prosecuting agencies being responsible for the formal accusation of offenders, courts being responsible for the trial of offenders and the sentencing of those convicted, and many correctional agencies, both state and local, for handling offenders after conviction--it is not easy to identify and follow an individual offender in this process from one stage to another. Mere name and charge are insufficient as both may be changed at different levels of the action. In order to insure a thread of positive identification that will make it possible to follow a person from arrest through final disposition and discharge, some additional record item is necessary. It is suggested that the identification number that is first issued by a central identification bureau, the FBI, a state bureau, or a local law enforcement agency, be the basis for such a record, and that this number be used on all official reports and papers relating to an offender throughout the total process. As most offenders are fingerprinted and identified at the point of arrest, it would not seem to be too difficult to record and carry some identification number through each successive stage of activity with respect to the offender.

Correctional Statistics

In the field of institutional and correctional treatment more information regarding the offender has been available than in any other phase of criminal process. Particularly, this is true of those who are sentenced to prison on whom case records which are fairly complete in terms of their individual case history, offense, and prior criminal record are developed. Also, the accounting of persons sentenced to long-term imprisonment is carefully carried out with exact data available as to the length and term of imprisonment and parole. Thus, prison statistics have been, on the whole, more complete and accurate than any others in the field of criminal justice. Because of the accessibility of information on prisoners and also the possibility of making special studies of persons under long-term incarceration, there has been a great deal of research and statistical analysis carried out, particularly in the area of parole. Such research has established rather definitely that persons who are committed to prison and thereafter placed on parole or released directly from prison, show varying patterns of subsequent criminal behavior. These patterns indicate that the type of offense, and the prior criminal conduct, are very strong predictive factors as related to involvement in future criminal behavior.

However, despite many attempts to study the effects of different programs, different kinds of treatment, different lengths of imprisonment, or different parole situations, there has been little evidence developed, so far, to suggest that these differences affect or change very much the general probabilities of criminal behavior already established with respect to offense and recidivism. It would seem that one of the reasons for not making further progress along

this line arises from the almost complete lack of knowledge of the kinds of offenders (aside from offense and recidivism) who are being studied. It is difficult to classify persons in terms of personality, early development, and measure the impact of past social experience, but it would seem that more must be done before there can be a further refinement in terms of predicting and understanding the subsequent probable behavior of persons who have been convicted of crime.

Again, it should be pointed out that a real weakness arises from our lack of knowledge of the offenders who are first originally arrested and charged with crime. If only one out of ten or fifteen is committed to long-term imprisonment, then to study only this group does not give an adequate picture of what kinds of persons and what kinds of dispositions were made of the other nine to fourteen who were charged with criminal offenses. The building of more adequate statistical data on persons coming into the crime picture through arrest and original charge is essential to the further development and understanding of the selective processes of criminal justice and of the evaluations that can be made of different kinds and methods of correctional treatment.

Juvenile Delinquency Statistics

And now a word regarding the juvenile delinquency statistical problem, which like the adult problem, involves all three source areas. There is no question that we are generally confused as to definitions of juvenile delinquency and as to our aims and objectives in handling persons alleged to be delinquent. From the standpoint of criminal statistics, it would seem necessary to separately account for those juveniles who have engaged in specific criminal conduct for which they would have been criminally charged if they had been adults. Further, because of the tremendous variations in attitudes towards handling juveniles, it would seem that if we are ever to have an index of juvenile crime or delinquency that we will have to establish a uniform reporting system relating to offenses above a certain level of seriousness. These would include offenses that for adults are considered felonies plus some of the misdemeanors, such as theft and assault, against which society demands protection. It would seem utterly impossible to obtain uniform information on all children contacted by law enforcement or even arrested by law enforcement for a host of activities which are generally described as waywardness, incorrigibility, truancy, lack of parental control, etc. Very little has been accomplished to date in narrowing the reporting to specific offenses, but it would appear that this is the direction that must be taken if we are ever to have statistics on juvenile delinquency that have real comparability.

Conclusion

Those of us who are working in the field of criminal statistics wish to express our appreciation to the Social Statistics Program Chairman of the American Statistical Association for the opportunity this year of having these valuable section meetings in our area of effort. The papers and discussion of these meetings certainly point up the need for a great deal more effort to be directed toward the building of definitions, more descriptive classifications of offenses and of kinds of offenders, and the accounting for individuals who are arrested and prosecuted under our criminal justice system, if we are to ever develop adequate measures of crime and delinquency.

IV

VALIDITY AND INTERVIEWER VARIANCE

Chairman, Joseph F. Daly, Bureau of the Census

A Study of Validity in Reporting Medical Care in Michigan — Robin Barlow, James Morgan, Grover Wirick, University of Michigan

Two Studies of Interviewer Variance of Socio-Psychological Variables — Leslie Kish and Carol W. Slater, Survey Research Center, University of Michigan

Discussion -- Eli S. Marks, National Analysts, Inc.

A STUDY OF VALIDITY IN REPORTING MEDICAL CARE IN MICHIGAN

Robin Barlow, Survey Research Center, The University of Michigan
James Morgan, Survey Research Center, The University of Michigan
Grover Wirick, Study of Hospital and Medical Economics, The University of Michigan

Sampling error, bias in estimating aggregates, interviewer errors, coding errors, and tabulation errors are all under the potential control of the surveyor through sample design and training and supervision of personnel. In most cases, however, response error must remain not only beyond his control, but largely unknown as well. In some cases it is even difficult to estimate the order of magnitude of this type of error with any confidence. It is a rare occasion when independent measures are available against which survey results can be checked, even by aggregates, and the opportunity to check each observation against the actual measure being investigated is even rarer.

This paper reports four types of evidence on the reliability of information obtained from respondents in a recent survey of medical experience and health insurance coverage of the people of Michigan. This Survey was part of the larger University of Michigan Study of Hospital and Medical Economics which in turn stemmed from the Governor's Study Commission on Pre-Paid Hospital and Medical Care Plans. The Survey was conducted jointly by that Study and the Survey Research Center. The entire Study is financed by the W. K. Kellogg Foundation.

To concentrate the available time on the discussion of the validity checks, no attempt is made to summarize previous work in this area nor to discuss the objectives, survey design, interview schedule, and conclusions of the Survey, except where they are relevant to the reliability of reporting.

The Survey was carried out on a multi-stage probability sample of about 2,000 dwelling units, chosen so that every non-institutionalized Michigan family had a known probability of selection. Interest was centered on two relatively small sub-groups of the population: the aged, and families and individuals with large medical expenses. To increase the precision of estimates for these two groups without greatly enlarging the sample, a randomly selected three-quarters of the original sample was screened by "doorstep" questions, and interviews were taken only if the dwelling unit contained someone 65 years of age or older, or someone who had been hospitalized during the previous year. The remaining one-fourth of the original sample was interviewed completely as a basic probability sample. This resulted in a sample for the special sub-groups that would have required interviewing 2000 units instead of 1000 if screening had not been used to eliminate some of the young and healthy.

RELIABILITY OF DOORSTEP SCREENING

The first evidence of reliability of reporting to be discussed was a check on the effectiveness of the doorstep screening process used to reduce the proportion of young healthy families interviewed. The test involved comparing the estimates of the proportion of families of each type with presumably more accurate estimates obtained from interviewing a simultaneous, but separate, probability sample.

The results are shown in Table 1, which compares the proportions of eligible families--someone 65 or over and/or someone hospitalized within the previous year--for those families in the small sample, which were fully interviewed, with the proportion in the larger sample where determination of the two criteria was by doorstep screening questions. The question on age yielded almost identical estimates of the proportion of families with someone 65 or older in the two samples, 16.5 per cent and 16.4 per cent.[1/] The samples gave disparate estimates of the proportion of families with someone hospitalized during the year; 39.3 per cent when determined by an extensive interview, and 27.2 per cent when determined by the screening question. Even allowing for sampling variation it appeared evident that the screening question on hospitalization experience had not been entirely successful. There was apparently substantial underreporting of hospitalization experience in response to an initial brief screening question. This does not preclude the possibility that even with the full interview there was additional underreporting: a question that will be turned to later.

Working on a hypothesis that the difference in hospitalization proportions was due to respondents forgetting to report stays with certain characteristics, the two samples were compared as to length of hospital stay, number of stays during the year, and time of year stays were experienced. Only dwellings where no one was 65 or over were compared since the others were completely interviewed in both samples and there was no chance for the particular bias under discussion to arise. When the two samples are compared, as shown in Table 2, we find the discrepancies concentrated in families with a single short hospitalization during the year, but it apparently did not matter whether the stay occurred early or late in the year.

The importance of having a complete probability sample where the full interview was given to all units is that it not only permitted the screening biases of the other sample to be estimated, but also largely to be eliminated by proper weighting as shown in the last column of Table 1.

[1/] These percentages are the sum of the percentages for the 65 or over hospitalized and the 65 or over not hospitalized; similarly for the following percentages on hospitalization.

The other evidence of reliability of survey information is contained in the three remaining studies of validity which consisted of attempts to verify, from other sources, individual reports on health insurance coverage, on hospital stays and expenses, and on the type of doctor or other medical practitioner used. We take the last of these first.

RELIABILITY OF REPORT ON TYPE OF MEDICAL PRACTITIONER USED

It must be pointed out that the type of practitioner was not directly asked of the respondent, although it was noted when volunteered, and the name and address were asked in all cases. For some of the practitioner-patient contacts, the name was not clearly enough reported to enable it to be found in any medical directory. Hence the evidence of error was clean only in those cases where the name was given and could be identified properly and where information on the type of practitioner was volunteered.

Table 3 shows the results of this check for all except dental practitioners who were not verified. For about 45 per cent of practitioner-patient contacts the respondents specified the type of practitioner. Except for M.D. specialists who were frequently referred to as general practitioners, and a slight tendency to confuse M.D.'s with clinics (which may have resulted from real ambiguity of definition[2]) respondents were generally correct in the specification of practitioner where positive verification could be made. A practitioner was considered positively identified only if he could be found in some directory which specified his type and specialty.

Two other facts emerge from this table. The 55 per cent of contacts referred to only as "doctor" were positively verified in 85 per cent of the cases, and 82 per cent of all practitioner mentions were positively identified.

HEALTH INSURANCE COVERAGE

Another validation study concerned reported health insurance coverage. The respondent was asked not only the name of the insurance carrier, but policy and contract numbers and other identifying information. These were given in most cases. Over 200 insurance carrier designations were named by respondents, although about one-third of these were obviously names of groups through which the insurance was obtained (usually employers). An attempt was made to contact the insurance carrier and to determine whether the policy was effective during the survey year, and the details of the coverage it provided. This report is concerned only with the fact of coverage and its verification, not the degree.[3] Underreporting was clearly not checked in this process since it was impossible to ask all companies whether they had issued policies covering anyone in the sample. The estimates of overreporting may be exaggerated, too: A person may have had coverage, but verification might fail because of incorrect identification of carrier or failure of the carrier to identify the particular individual.

As indicated in Table 4, a little less than half of the 1507 reported policies were Blue Cross/Blue Shield, of which nearly all were verified. About one-fourth of the private policies were something other than health insurance, and no response was received from the company in respect to one-eighth. Of the balance, about three-fifths were verified leaving about 18 per cent of the total which couldn't be verified, and seven per cent where identification of the company was inadequate.

The design of the interview schedule leaned in the direction of getting too much, rather than too little, information on insurance. It is possible that some people failed to report coverage, and others named the wrong company, but it appears likely that there may be an overall net overreporting of health insurance by people who confused it with disability and other types of insurance. Higher income people were found to report other types of insurance as though they were health insurance more frequently than did low income persons. This might be because they reported, and had, more insurance of all kinds.

HOSPITALIZATION

The final, most interesting, validity check is that of hospitalizations. For all respondents reporting Blue Cross coverage, whether or not any hospital stays were reported, Blue Cross records were examined for hospital stays, so that both over- and underreporting could be found. Where Blue Cross coverage was not reported, records of the hospitals in which stays were reported were checked. In the latter case, underreporting could not be found if the person had had additional stays in other hospitals not reported by him. On the other hand, hospital stays of persons who reported Blue Cross coverage could not be found if they were not covered at the time of the stay, so the Blue Cross check may have exaggerated the estimates of overreporting. Hence, cases checked against hospital records provide a reasonably clean measure of overreporting, while those checked against Blue Cross provide a good measure of underreporting. For verified stays, however, both sources provide good checks on length of stay, size of total bill, and service received.

2/ It is difficult to determine, in some cases, whether a group of physicians practicing in one office should or should not be considered a clinic.

3/ Degree of coverage refers to the type and amount of protection provided by a policy. A discussion of the attempt made to scale the degree of coverage will be found in the report of the Michigan Study of Hospital and Medical Economics, to be published soon.

Reports of hospital stays are multidimensional, involving not only whether or not a stay was experienced, but also the time and length of stay, what services were received, the amount of the bill, and the sources of payments. These factors are subject to various degrees of reporting error due to memory loss or misunderstandings, or unwillingness to report. Hypotheses were formulated that reporting error would be related to such variables as whether the respondent was reporting for himself or someone else, age, education, and whether or not covered by insurance, and they were separately tested for hospital stays checked through Blue Cross and hospital records.

Table 5 indicates quite small frequencies of under- and overreporting (about 4 per cent each), particularly if one assumes that reported hospital stays not found in Blue Cross records might be genuine (person not covered by Blue Cross at the time), and that the hospital records checked are most unlikely to reveal all underreporting. Table 5 also indicates that both under- and overreporting are more frequent when the hospitalized person was not the respondent. Table 6 presents discrepancies between reports on kinds of services received, and they seem to be more frequent for older people. There is also an interesting tendency for older respondents to overreport checkup and diagnoses, and to underreport operations, surgery and fractures somewhat more than other age groups.

Table 7 is restricted to the confirmed stays checked against Blue Cross or hospital records, and indicates that the respondents reported the correct number of days a majority of the time, and on the average tended to report stays slightly longer than they actually were. Most of the average discrepancy, however, was the result of one extreme case. Neither education of the head of the family nor whether the stay was reported by the hospitalized person or someone else made any appreciable difference in the accuracy of the report of length of stay.

Table 8 shows for verified cases a very small average overestimate of the total hospital bill (about 5 per cent of the average bill of $250 for hospitalized persons), and some tendency for bills for females to be overestimated more than those for males. The mean discrepancies, however, suggest little pattern in respect to either age or sex. It is possible that the strikingly small average overestimate of the hospital bill is due to respondents considering as hospital expenses other legitimate charges which were treated in some other way in the hospital or Blue Cross records. Examples would be blood transfusions and charges for special patient services such as television and commissary purchases. Table 9 reveals only minor differences between insured and uninsured persons, and between hospitalizations reported by the person involved or by someone else.

A companion study, conducted by the Survey Research Center in conjunction with the National Health Survey, designed to provide a measure of underreporting of hospitalizations and its causes, will be reported at some future meeting. It might be added that the interview schedule used in this study was simpler than those used by the National Health Survey and the N. O. R. C. studies of health care, expenses, and debt.

SUMMARY

In summary, this report portrays a series of validity studies of varying degrees of precision. They do not indicate substantial difficulties in accepting survey data, except in the use of doorstep screening questions to determine eligibility of respondents. Even in this case, the criterion for eligibility seems to influence the effectiveness of the process. (Similar checks in a Minnesota study of disability, using mail and telephone screening questions, found similar problems.)[4/] By the very nature of things, the more complex the data, the more difficult it is to provide a symmetrical and complete validity check. Although one must always reserve judgment while accumulating as much evidence as possible from a variety of tests, the results reported here indicate a rather close correspondence between survey results and source data, especially for simple measures which are not graduated too finely.

[4/] See Minnesota Studies in Vocational Rehabilitation. Part V. Methodological Problems in Rehabilitation Research (Bulletin 25; Minneapolis: Industrial Relations Center University of Minnesota, December 1958).

Table 1: Age and Hospitalization Experience in Occupied Dwelling Units within Sampling Procedure

Age and Hospitalization Experience in Dwelling Unit	Unweighted Data: Completely Interviewed Sample	Unweighted Data: Doorstep Screened Sample	Unweighted Data: Total Sample	Weighted Data[1/]: Total Sample
Someone in Dwelling Unit hospitalized; someone 65 or over	6.4	4.2	4.7	4.8
Someone hospitalized; no one 65 or over	32.9	23.0	25.2	32.1
One[2/] hospitalization of 3 days or less	8.5	3.7	4.8	8.4
One[2/] hospitalization of 4 days or more	16.2	11.9	12.9	16.0
Two or more hospitalizations	8.2	7.4	7.5	7.7
No one hospitalized; someone 65 or over	10.1	12.2	11.7	11.9
No one hospitalized; no one 65 or over	50.6	60.6[3/]	58.4	51.2
Total	100.0	100.0	100.0	100.0
Number of cases	425	1468	1893	1003[3/]

1/ Weighting corrects for four separate sources of potential bias:

(a) the completely interviewed sample and the doorstep screened sample have different overall response rates
(b) response rates differ according to size of place
(c) use of screening questions in the doorstep screened sample apparently led to an under-representation in this part of the sample of dwelling units where there had been only one hospitalization during the year; seriously in cases where the hospital stay was short, less seriously in cases where the stay was long
(d) dwelling units in the doorstep screened sample where no one was hospitalized and no one was aged 65 or over were not interviewed.

2/ Maternity cases counted as one hospitalization, not two.

3/ 890 dwelling units in the doorstep screened sample where it was reported that no one was hospitalized and no one was 65 or over were not interviewed, and are not included in the total weighted sample.

670:28:3:76; Weighting Sheet

Table 2: Distribution of Occupied Dwelling Units Where Someone Was Hospitalized but No One Was Aged 65 or over by Hospital Experience within Sampling Procedure

Hospital Experience in Dwelling Unit[1/]	Completely Interviewed Sample	Doorstep Screened Sample
One short early stay	12.7	8.3
One short late stay	13.4	8.0
One long early stay	23.9	24.7
Two or more short early stays	0.7	2.1
One long late stay	26.1	27.4
Two or more short late stays	1.5	3.6
Two or more long early stays	6.0	5.4
Two or more long late stays	15.7	20.5
Total	100.0	100.0
Number of cases	134	336

1/ "Short" means that the longest hospital stay in the dwelling unit was of three days or less; "early" means that the latest hospital stay was in March 1958 or earlier.

Note: This table is meant to show the reasons for the relative underreporting of hospitalization at dwelling units in the doorstep screened sample. The categories of "hospital experience" are so arranged that the types of hospitalization most likely (by hypothesis) to be underreported in the doorstep screened sample are shown first.

670:27:3:76; and hand tabulation from verified hospital records

Table 3: Verification of Medical Practitioners: Distribution of Positively Verified Types of Practitioners Seen within Type of Practitioner Reported, Excluding Dentists[1/]

Positively Verified Type of Practitioner	Reported Type of Practitioner: Total	Clinic	M.D. General Practitioner	M.D. Specialist	Osteopath	Chiropodist	Chiropractor	Doctor (not otherwise specified)
Clinic	8.8	57		1	-	-	-	-
M.D. General Practitioner	29.2	9	44	1	-	-	1	37
M.D. Specialist	31.0	6	37	75	-	-	-	35
Osteopath	10.5	1	5	1	86	-	-	11
Chiropodist	0.5	-	-	-	1	40	-	1
Chiropractor	2.3	-	-	-	-	-	82	1
Sub-total	82.3	73	86	78	87	40	83	85
Not verified	17.7	27	14	22	13	60	17	15
Total	100.0	100	100	100	100	100	100	100
Weighted per cent of sample	100.0[2/]	15.0	15.1	6.6	3.7	0.4	2.1	55.3

1/ Dentists and other dental practitioners were excluded from this table. Time limitations did not permit verification of 1,165 dental contacts, and it was felt that the error in designating type of practitioner would be small in this case, inasmuch as respondents were asked specifically about dentists. The table includes a total of 3,771 patient-practitioner contacts other than dental practitioners.

2/ Includes other practitioners reported, 0.2 per cent, and type of practitioner not ascertained, 1.6 per cent.

Table 4: Distribution of Reported Health Insurance Coverage by Result of Verification

Results of Verification:	Number of Policies			Unweighted Percentage Distributions		
	Total	Blue Cross	Private	Total	Blue Cross	Private
RECORDS CHECKED						
Policy verified as health insurance: policy is in force or lapsed during survey year	927	615	312	61.5%	94.5%	36.4%
Policy verified as health insurance: policy lapsed before survey year	14	2	12	0.9	0.3	1.4
Policy verified as not health insurance: policy is loss-of-time, accident and health or life insurance	178	-	178	11.8	-	20.8
Policy verified as not health insurance: type of policy other than those described above	8	-	8	0.5	-	0.9
Policy verified as not health insurance: type of policy not ascertained	4	-	4	0.3	-	0.5
Policy not verified: insurance company unable to identify policy (names or numbers)	183	32	151	12.1	4.9	17.6
RECORDS NOT CHECKED						
Policy judged from respondent's description to be not health insurance	26	-	26	1.7	-	3.0
Policy inadequately described by respondent: no insurance company or insurance company unidentifiable	58	*	58	3.9	*	6.8
Policy adequately identified by respondent, but no reply received from company	109	1	108	7.2	0.2	12.6
Total	1507	650	857	99.9	99.9	100.0

*Some of the policies where the companies were unidentified may be Blue Cross/Blue Shield policies, of course. In 28 cases no company was named at all, and in the remaining cases 22 different names of companies or associations were given, but the actual insurance carrier could not be positively identified.

Table 5: Distribution of Hospital Stays[1] by How Stay Reported within Type of Records Checked, Whether Hospitalized Person Was Respondent, and Length of Stay

How Stay Reported	Blue Cross Records Checked: Hospitalized person was respondent: All cases[2]	Length of stay: 1-2 days	3-9 days	10 or more days	Hospitalized person was not respondent: All cases[2]	Length of stay: 1-2 days	3-9 days	10 or more days
Stay reported by respondent and confirmed by records	85	81	86	83	81	74	86	74
Stay reported by respondent, records checked but stay not confirmed	13	12	12	15	12	20	9	13
Stay reported by respondent, no records checked	*	*	*	*	*	*	*	*
Stay not reported by respondent, but found from records	2	7	2	2	7	6	5	13
Total	100	100	100	100	100	100	100	100
Number of cases	261	24	169	62	204	54	86	44
Weighted per cent of hospital stays	29.9	2.8	19.5	6.9	23.0	7.2	9.3	4.4
Hospital Records Checked								
Stay reported by respondent and confirmed by records	94	94	96	87	86	89	91	84
Stay reported by respondent, records checked but stay not confirmed	3	2	2	4	6	2	3	11
Stay reported by respondent, no records checked	2	2	1	9	7	9	5	5
Stay not reported by respondent, but found from records	1	2	1	*	1	*	1	*
Total	100	100	100	100	100	100	100	100
Number of cases	253	45	159	46	158	43	72	38
Weighted per cent of hospital stays	28.8	5.0	18.6	4.9	18.3	6.0	7.7	4.1

*Less than 0.05 per cent

1/ All hospital stays excluding (a) stays of newborns who left the hospital before or at the time the mother was discharged; (b) stays which were found to be more than a year prior to date of interview.

2/ Including cases where length of stay was not ascertained.

S670:61:6:74x61x59,MH-3

Table 6: Distribution of Hospital Stays[1] by Discrepancies between Reports of Services Received within Type of Records Checked, Whether Hospitalized Person Was Respondent, and Age of Respondent

Discrepancies between Reports of Services Received During Hospital Stay[2]	Blue Cross Records Checked						Hospital Records Checked					
	Hospitalized person was respondent			Hospitalized person was not respondent			Hospitalized person was respondent			Hospitalized person was not respondent		
	Age of respondent			Age of respondent			Age of respondent			Age of respondent		
	Under 35	35-54	55 & over	Under 35	35-54	55 & over	Under 35	35-54	55 & over	Under 35	35-54	55 & over
Respondent reports services not mentioned by records:												
Checkup, diagnosis	6.2	7.5	13.2	5.5	6.8	12.9	2.9	11.9	13.1	3.9	16.7	13.4
Pregnancy, delivery	*	*	*	*	*	*	1.0	*	*	*	*	*
Operation, surgery, fracture	3.6	5.4	1.7	*	1.1	*	1.0	1.5	2.7	1.3	*	4.5
Treatment, other services	7.4	11.4	11.9	4.4	2.9	4.6	2.2	2.7	1.3	*	*	18.3
Records report services not mentioned by respondent:												
Checkup, diagnosis	1.0	*	*	*	*	*	1.0	7.0	1.3	*	*	*
Pregnancy, delivery	2.5	2.8	*	*	2.0	*	0.7	*	*	*	*	*
Operation, surgery, fracture	2.6	6.2	10.4	3.1	4.9	8.7	2.4	4.4	4.0	1.3	5.4	18.3
Treatment, other services	2.6	2.9	6.6	8.9	14.2	4.6	0.7	11.2	9.2	2.6	11.5	13.4
No discrepancy between reports	74.5	51.6	48.9	62.2	38.8	50.6	86.4	62.9	68.6	80.2	66.1	54.9
No services reported either by respondent or by records[3]	10.7	25.7	19.2	18.0	39.2	23.2	5.5	8.3	10.4	15.9	17.2	8.9
Total	111.1	113.5	111.9	102.1	109.9	104.6	103.8	109.9	110.6	105.2	116.9	131.7
Number of cases	110	94	57	81	101	22	114	65	74	69	68	21
Weighted per cent of hospital stays	12.4	11.0	6.1	9.2	11.6	2.2	13.6	7.6	7.6	7.9	8.2	2.3

*Less than 0.05 per cent

1/ All hospital stays excluding (a) stays of newborns who left the hospital before or at the time the mother was discharged; (b) stays which were found to be more than a year prior to date of interview.

2/ Columns add to more than 100 per cent because some stays involved more than one type of discrepancy.

3/ Including cases where (a) reports about hospital stay from both respondent and records existed, but one or the other was N.A. as to services received; (b) hospital stay was reported by respondent but not by records, and vice-versa.

S670:63:6:74x75x44,MH-4

Table 7: Discrepancy between Reports of Length of Hospital Stay within Whether Hospitalized Person Was Respondent and Education of Head

Discrepancy between Reports of Length of Stay	All Confirmed Stays [1]	Hospitalized Person Was Respondent — Education of Head: Grammar school or less	High School [2]	College [3]	Hospitalized Person Was Not Respondent — Education of Head: Grammar school or less	High school	College
Mean discrepancy in number of days between respondent's and record's report of length of stay [4]	0.6 (0.3)	-	-	2 (0)	1	-	-
Per cent of stays for which respondent overestimated length of stay	24.1	25	29	19	19	22	24
Per cent of stays for which respondent correctly estimated length of stay	52.6	53	53	55	42	52	63
Per cent of stays for which respondent underestimated length of stay	14.3	15	15	15	9	16	8
Per cent of stays for which either respondent or records did not give length of stay	9.0	7	3	11	29	10	5
Total	100.0	100	100	100	100	100	100
Number of stays	734	131	221	86	78	171	47
Weighted per cent of stays [5]	84.0	14.3	26.2	9.9	8.6	19.2	5.7

1/ This table shows only those stays which were reported by the respondent and confirmed by hospital or Blue Cross records. Of the confirmed stays it does not show (a) stays of newborns who left the hospital before or at the time the mother was discharged (b) stays which were found to be more than a year prior to date of interview, (c) stays for which education of head was not ascertained.

2/ Includes cases where high school was not completed, also high school plus non-college, grammar school plus non-college.

3/ Includes cases where college was not completed.

4/ Positive means indicate respondents overestimated length of stay. Means in brackets excludes one case with a discrepancy of 151 days.

5/ As a percentage of all stays except those of newborns and those which took place more than a year prior to date of interview.

Table 8: Discrepancy between Reports of Total Hospital Bill within Age and Sex of Hospitalized Individual

Discrepancy between Reports of Total Bill	All Confirmed Stays 1/	All Males' Confirmed Stays	All Females' Confirmed Stays	Age and Sex of Hospitalized Individual: Males Under 5	Males 5-14	Males 15-44	Males 45-64	Males 65-69	Males 70 & over	Females Under 5	Females 5-14	Females 15-44	Females 45-64	Females 65-69	Females 70 & over
Mean discrepancy between respondents' and record's reports of total bill 2/	$14 ($12)	$7	$17 ($15)	$11	$-6	$26	$-29	$10	$33	$22	$-1	$15 ($11)	$28	$65	$4
Per cent of stays for which respondent overestimated total bill	28.2%	25.8	29.5	21	15	33	22	35	29	17	29	33	23	31	32
Per cent of stays for which respondent correctly estimated total bill	15.0%	12.4	16.3	6	14	18	12	8	4	23	13	15	18	19	19
Per cent of stay for which respondent underestimated total bill	19.7%	18.8	20.1	25	18	12	19	31	21	27	10	22	14	25	24
Per cent of stays for which either respondent or records did not give total bill	37.1	43.1	34.1	48	53	37	47	26	46	33	48	30	45	25	25
Total	100.0	100.0	100.0	100	100	100	100	100	100	100	100	100	100	100	100
Number of stays	762	261	501	35	38	79	59	26	24	30	50	287	93	16	25
Weighted per cent of stays 3/	86.9	29.4	57.4	3.9	4.8	8.9	6.8	2.6	2.4	3.1	5.9	34.1	10.3	1.6	2.5

1/ This table shows only those stays which were reported by the respondent and confirmed by hospital or Blue Cross records. Of the confirmed stays, it does not show (a) stays of newborns who left the hospital before or at the time the mother was discharged, (b) stays which were found to be more than a year prior to date of interview.

2/ Positive means indicate respondent overestimated total bill. Negative means indicate respondent underestimated.

3/ As a percentage of all stays except those of newborns and those which took place more than a year prior to date of interview.

( ) Means shown in bracket are figured without one case with a discrepancy of $1136.

S670:62:6:77x78x50,51-53,MH-20

Table 9: Discrepancy between Reports of Total Hospital Bill within Whether Hospitalized Person Was Respondent and Whether Has Verified Insurance

Discrepancy between Reports of Total Bill	All Confirmed Stays 1/	Hospitalized Person Was Respondent			Hospitalized Person Was Not Respondent		
		Has no insurance 2/	Has verified insurance	NA whether has health insurance 3/	Has no insurance 2/	Has verified insurance	NA whether has health insurance 3/
Per cent of stays for which respondent overestimated total bill	28.2	25	30	49	21	23	33
Per cent of stays for which respondent correctly estimated bill	14.9	9	16	13	19	15	17
Per cent of stays for which respondent underestimated total bill	19.7	26	21	18	22	15	15
Per cent of stays for which either respondent or records did not give total bill	37.2	40	33	20	38	47	35
Total	100.0	100	100	100	100	100	100
Number of stays	761	100	306	51	50	227	27
Weighted per cent of stays 4/	86.9	11.1	35.7	5.7	5.6	25.5	3.3

1/ This table shows only those stays which were reported by the respondent and confirmed by hospital or Blue Cross records. Of the confirmed stays it does not show (a) stays of newborns who left the hospital before or at the time the mother was discharged, (b) stays which were found to be more than a year prior to date of interview.

2/ Reported no insurance, or all reported insurance verified as not health insurance or lapsed.

3/ Respondent reported insurance, no verification possible; also per cent of coverage N.A.

4/ As a percentage of all stays except those of newborns and those which took place more than a year prior to date of interview.

S670:82:6:73x69x50,MH-24

TWO STUDIES OF INTERVIEWER VARIANCE OF SOCIO-PSYCHOLOGICAL VARIABLES

By: Leslie Kish and Carol W. Slater
Survey Research Center, University of Michigan

Introduction

We report results obtained in two surveys in which respondents were randomized among interviewers to permit the valid estimation of the interviewer variance as a component in survey errors. In each study, done by the Survey Research Center of the University of Michigan, the blue-collar workers of a plant were asked many socio-psychological questions about their jobs, and company and such.

We suspect that few people worry at all about the interviewer variance. They, however, are apt to fear that for "vague" psychological and attitudinal questions the effects must indeed be large. Our results may hold at least one surprise for everybody.

On the one hand, the interviewer effects are not very great: they compare well with effects on "factual" items, and, because of this we were unable to separate different classes of items - the "soft" psychological items from the "hard" factual items. On the other hand, even these small or moderate effects on individual interviews can have important effects on the sample means. As a final dramatic effect, a happy ending: even great effects on the means of the entire sample are reduced for subclasses and the effects usually seem to disappear completely from the comparisons of subclasses.

Here we present a summary of our findings; details and references to related literature will appear in an article already submitted for publication.

In the First Study in 1948, at a large unionized auto plant in the Midwest, we selected with equal probability a stratified random sample of individual employees, of whom 462 gave interviews. The names and addresses of the selected employees were typed on cards which were then shuffled and assigned randomly to interviewers at the beginning of each day. The interviews were taken in the respondents' homes and lasted an average of an hour and a half. Open-ended questions were used to gather information about attitudes towards foremen, stewards, the union, higher management and various aspects of their jobs. The 20 interviewers were selected, screened and hired specifically for this study. All had had some previous experience in interviewing, not necessarily in survey work. A week of training was carried out before the study began and was augmented, as needed, by individual supervision and group sessions.

For the Second Study, we selected in 1958 with equal probability a stratified random sample of individual employees, with a final n = 489. The interviews were conducted, in 1958, in offices provided by the company, and lasted an hour on the average. After the interview, each respondent was also asked to fill out a paper-and-pencil questionnaire in the presence of the interviewer, which took roughly an additional three-quarters of an hour. From the list of respondents available during each week, random assignments were allocated to the interviewers working during that period. Completely open-ended items were few: almost all questions included in the written questionnaire and many of those used in the interview involved asking the respondent to choose from a prepared and pretested list the alternative coming closest to his own viewpoint. The nine interviewers were members of the Center's field staff with several years of interviewing experience.

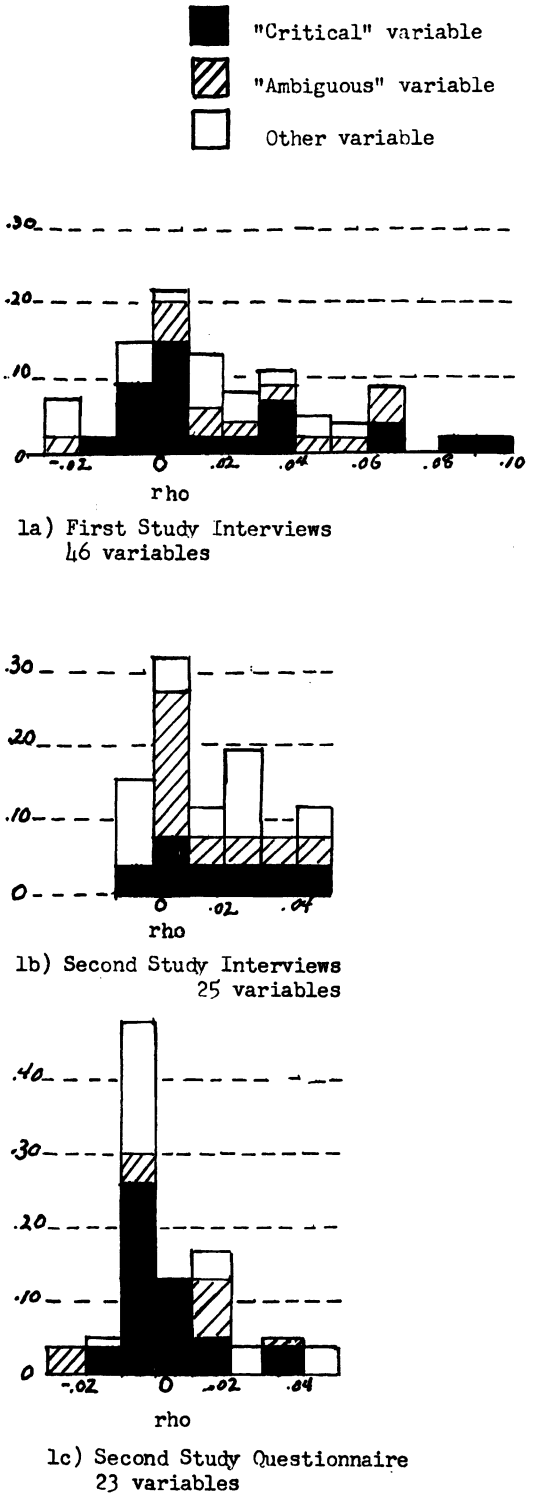
The Measurement of Interviewer Variance

Besides sampling errors proper--those arising in selection or estimation procedures--survey results are also affected by errors which occur in the course of the observation (measurement), recording and processing of the data. These errors fall into two broad types, having very different effects on the summary results (such as means or totals) of a survey. The first includes the "biases" or "systematic errors" imposed by the "essential survey conditions": the average or "expected" deviations of sample estimators from their estimands, the population values. These, although important, were not the subject of our research.

The second type consists of variable errors: those not fixed by the "essential survey conditions." Some variable errors are uncorrelated among the elements, and, unless replicate measurements are taken on the respondents, these cannot be distinguished from sampling error among respondents. We are not here concerned with them and generally they can be regarded as random errors which increase the variance of estimators with contributions which enter automatically into the estimate of the variance. Some other variable errors, however, involve the correlated effect that each interviewer's bias can impose upon the respondents (the elements) making up his workload. Insofar as the individual interviewers have different average effects on their workloads, this "interviewer variance" contributes to the variance of the sample mean. This contribution of the interviewer variance to the sampling variance is our present concern. The contribution, as we shall see, can be large and its neglect can lead to serious underestimation of the total survey variation.

Our model assumes the random selection of a sample of a interviewers from a large pool of potential interviewers, that pool defined by the "essential survey conditions." Each interviewer has an individual average "interviewer bias" on the responses in his workload; we estimate the effect of a "random sample" of these biases on the variance of the sample mean. This effect is expressed as an "interviewer variance" which decreases in proportion to the number (a) of inter-

CHART 1 - Three Distributions of Relative Frequencies of Rho's for Different Variables



1a) First Study Interviews
46 variables

1b) Second Study Interviews
25 variables

1c) Second Study Questionnaire
23 variables

viewers. Its contribution to the variance of sample means (s_a^2/a) resembles other variance terms, being directly proportional to the variance per interviewer and inversely to the number of interviewers. This increase in the variance may be substantial; failing to take it into account (as when estimating the variance simply by s^2/n) results in neglecting a potentially important source of variation actually present in the design, introduced by the sampling of interviewer's biases.

The interviewer variance s_a^2 should be viewed as a component of the total variance, denoted as

$$s^2 = s_b^2 + s_a^2$$

where s_b^2 is the variance without any interviewer effect, and all three terms are measured per element. It is convenient to take the interviewer component s_a^2 relative to the total variance, and to denote this ratio by the ratio of homogeneity, often called the coefficient of intraclass correlation:

$$roh = s_a^2/s^2 = \frac{s_a^2}{s_a^2 + s_b^2}$$

The individual roh's are subject to very great variabilities; the values of s_a^2 are computed with 9 degrees of freedom in the First Study and 19 in the Second Study. As a rough guide we consider the values of the First Study subject to coefficients of variation of 0.5 and those of the Second Study to about 0.3. Nevertheless, the results are useful, particularly when considered in the aggregate over many items.

Primary Results and Implications

How are these values useful in planning surveys? First, they show that it is feasible to obtain responses with rather low interviewer effects on what appear to be ambiguous and emotionally loaded attitudinal items, if the interviewers are carefully selected and well-trained. The low values of roh on these items speak well for the prospects of obtaining attitudinal, socio-psychological data of this kind with reasonable reliability. The variability for these attitudinal interview items appear to be generally not much, if any, higher than responses to "factual" items obtained in a good Census--expect probably for the simplest items like age and sex. They compare favorably with some other results relating to "factual" items.

The primary results appear on Chart 1; the First Study in 1a, the interviews and questionnaires of the Second Study in 1b and 1c, respectively. Each of these presents a distribution of the relative frequencies (percentages) of occurence of rohs in size classes of .01. (The total height of each class is divided into three to separate "critical", "ambiguous" and other items.)

Second, this kind of analysis can distinguish

items for which the interviewer variances appear unexpectedly high, and by so doing, lead to corrective actions either through better training or by changing the survey operations. Extension of this kind of analysis may also be used to separate interviewers who make undue contributions to the variances.

Third, we can distinguish in the three tables concomitants of different interviewing situations. The results of the First Study (1a) came from newly hired and trained interviewers taking open-ended interviews; the roh's range, in the main, from zero to .07, with an average of .02 or .03. In the Second Study we see expert interviewers taking a more structured interview (1b); the roh's vary mostly from zero to .04, with an average of .01 to .02. For written questionnaires we find (1c) that the a priori hypothesis of zero effect is generally acceptable (with the exception of three puzzling items).

Fourth, our results indicate the difficulties involved in making judgments beforehand about the degree of interviewer variance associated with what may seem a priori to be different kinds of items. In each of the three parts of Chart 1 the areas corresponding to "critical", to "ambiguous" and to "other" items do not appear to have very distinct distributions. Even informed intuition, it seems, needs considerably more conceptual and empirical tools than are now available to evaluate the relative susceptability of survey items to interviewer bias.

Fifth, we find that interviewer variance, although it appears small, definitely exists. Furthermore, it can exert important influence on the total variability of survey results, since even a small roh, when multiplied by moderate or large interviewer workloads, can have large effects. This effect on the variance is about $[1 + \text{roh}(\frac{n}{a}-1)]$. Let us consider an increase in the variance by a factor of 1.5 as "serious" and by 2 as "critical"; these correspond to increases in the standard errors of $\sqrt{1.5} = 1.22$ and $\sqrt{2} = 1.41$. With $n/a = 22$ in the First Study, roh becomes serious at .025 and critical at .045 categories which include 16 and 8 items respectively. In the Second Study, with $n/a = 52$, roh = .01 is serious and roh = .02 is critical, thus including 13 and 10 items respectively. In the case of element sampling; these effects can and should be included in the variance by computing the interviewer's load as if it were a "cluster." (In the case of actually clustered samples, where the interviewer is confined to a single primary selection, such as a county in a national sample, the usual computation automatically includes this effect.)

Sixth, analysis of this type makes it possible to include interviewer effects in considering the economic aspects of survey designs. If the ratio of the cost of hiring and training an interviewer to the cost of a single interview is $\frac{C_a}{C_b}$, then the most economical plan - least total variance ($s_a^2/a + s_b^2/n$) - results from the optimum workload size of $\frac{n}{a} = \sqrt{\frac{C_a}{C_b}\frac{s_b^2}{s_a^2}} = \sqrt{\frac{C_a}{C_b}\frac{1-\text{roh}}{\text{roh}}}$.

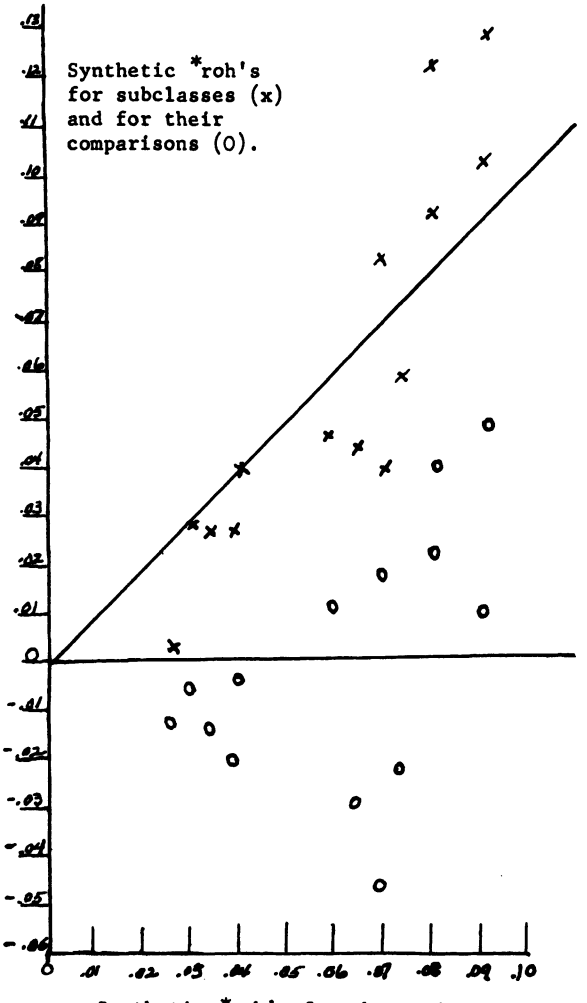
For example, if it costs ₹180 to train an interviewer and ₹10 to take an interview, then $\frac{C_a}{C_b} = 18$. For roh's of .02 this gives an optimum workload per interviewer of $n/a = 30$. The actual workloads in our two studies were in this neighborhood.

Effects on Subclasses and Their Comparisons

Current models of response errors deal mostly with the effects on the mean for the entire sample, but applying the model and methods to the means of subclasses is straightforward.

The data support our hypothesis that the effects of interviewer variance on the variances

Chart 2 - The Effects of Interviewer Variability on Subclasses (x) and on Their Comparisons (0) Plotted Against the Effects on the Entire Sample. (The effects are measured as ratios to the total variance per interview - as synthetic equivalents of roh's.)



of subclass means tend to decrease in the same proportion as the average workloads of the subclasses per interviewer decrease. The effect on the variance is approximately $[1 + roh(n^*/a-1)]$, where n^* is the sample size of the subclass. This effect decreases if roh remains constant, where roh expresses the interviewer contribution per element. That roh remains fairly constant for the subclasses is evidenced by the proximity of the values marked by x to the 45 degree line on Chart 2; this line denotes equality for the roh's of the subclasses and of the entire sample. The x points mark the values of roh in subclasses against the roh for the entire sample for the same variable. Actually the ordinates denote the average of the roh's for two subclasses into which the entire sample was divided.

We also investigated the effects of interviewer variance on the comparisons of pairs of subclass means. These are even more important, research workers often say, than the estimation of individual means. Because of the considerable effort required we had to limit the extent of this investigation. For strategic reasons we chose for investigating both the subclasses and their comparisons seven of the most critical variables from the two studies: those for which the effect on the means were greatest. For each variable we used two different ways of forming subclasses and this gives rise to the fourteen comparisons marked O on Chart 2 (as well as the fourteen subclass averages marked x).

The results show that the effect of interviewer variance on comparisons between subclass means is reduced drastically to the neighborhood of zero. This important result seems to hold roughly and on the average in our investigation. As evidence, note on Chart 2 that the O marks for interviewer effects on comparisons fluctuate around the horizontal line denoting zero effect. These come from plotting the effects per element of the comparisons against those of the entire sample. These data show a great deal of fluctuation, the causes of which should be sought in later investigations; nevertheless, the tentative working hypothesis of zero average effects appears to be a good working hypothesis on the average, and better than any alternative we could form. This should apply also to comparisons of any two (or more) samples which have been randomized over the same set of interviewers. Important examples arise from the comparisons of periodic samples assigned to the same set of interviewers; such comparisons of periodic surveys should tend also to be free of the effects of the interviewer variance that affects a single sample. (Similar results were obtained on the very different data, with very high initial roh's.)

This result gains added significance in combination with the likelihood that the systematic biases of comparisons are often also less than the biases of the individual means. In other words, if the interviewers' biases affect the subclasses equally (corresponding to lack of "interaction" between interview bias and subclass) then both the systematic bias and the interviewer variance tend to disappear from the comparisons of subclasses.

Some Remarks on Research Strategy

Research on interviewer variability may be designed to different degrees of symmetry and completeness. A very complete design might call for simple random selection of equal workloads; the effects of other sources of errors, especially coding error, would be included in a symmetrical, clean (orthogonal) design; the questions could be chosen to test various hypotheses about them. Our studies lack these virtues.

Our randomization procedures were designed to minimize costs and interference with field operations. We sacrificed chiefly: (a) equal size workloads which would have resulted in somewhat simpler computations and more efficient estimates; (b) eliminating the complications arising because the randomized set (the workload for a day in the First Study and for a week in the Second Study) is difficult to treat exactly in the analysis; (c) and the possibility of separating the components of the variance due to coding variability by randomizing coders in a neat design. Perhaps we most regret lacking the means and persuasiveness to achieve the last of these three improvements--which a modest disposal of means could have brought. In defense we plead that the choice was between little or nothing--as it often is. The procedure for assigning interviewers to coders does not depart enough from random to interfere seriously with our analysis: the distribution of coders against interviewers was checked and found about as even as a random assignment would have made it.

With relatively modest extra means it is possible to get a little closer than we did to a more symmetrical and complete design. Nevertheless, we are convinced of the desirability and economy of allocating near the lower end of the scale the limited resources available for research in interviewer variability. This is not merely post hoc justification for our research, but a belief based on the expectation that a few--because expensive--"crucial experiments" will not yield definitive evidence about a small set of "basic parameters"-- because that small set does not exist. It is more likely that interviewer errors differ greatly for various characteristics, populations, designs and resources--this last including questionnaires, nature and training of interviewers, etc. Therefore, knowledge about this source of variation, as with sampling variability, can be accumulated only from a great deal of empirical work spread over the length and breadth of survey work. This implies, together with the necessarily limited total means for this kind of research, that most research in this area must be done at marginal cost, as appendages to the main aims and designs of surveys.

Therefore, general strategy should call for many investigations of modest scope and that these be widely communicated.

APPENDIX

The response from the j-th individual to the i-th interviewer is expressed as $y_{ij} = y'_{ij} + A_i$, where A_i is the average "effect" of the i-th interviewer. Any constant (or "systematic") biases of the interviewers are not distinguished and we assume that the sum of the interviewer effects is zero for the population of A interviewers, from which the actual a interviewers are a random sample.

Each response is viewed as composed of two components, the sampling variance of the individual response and the component due to the variable interviewer bias, or interviewer variance:

$$s^2 = s_b^2 + s_a^2 \quad .$$

The "ratio of homogeneity" is

$$\text{roh} = s_a^2 / s^2 = s_a^2 / (s_a^2 + s_b^2). \qquad (1)$$

Assuming that of the n respondents n_i were assigned with simple random sampling to the i-th interviewer, we have in terms of the usual "anova" table for computations:

Source of Variation	Degrees of Freedom	Sum of Squares (SS)	Mean Square	Components of the Mean Squares
Among interviewers	a-1	$\sum^a y_i^2/n_i - y^2/n$	$V_a = \frac{SS(a)}{a-1}$	$s_b^2 + ks_a^2$
Within interviewers	n-a	$\sum^a \sum^{n_i} y_{ij}^2 - \sum^a y_i^2/n_i$	$V_b = \frac{SS(b)}{n-a}$	s_b^2

Here $n = \sum^a n_i$, $y = \sum_i^a y_i$ and $y_i = \sum_i^{n_i} y_{ij}$ and $s_a^2 = (V_a - V_b)/k$ where

$$k = \sum^a n_i^2 \frac{1/n_i - 1/n}{a-1} = \frac{n}{a-1} - \frac{\sum^a n_i^2}{n(a-1)} = \frac{n}{a} - \frac{1}{n}\left[\frac{1}{a-1}\sum^a (n_i - \frac{n}{a})^2\right] .$$

To measure the effects on the differences between the means of two subclasses we had to improvise approximate methods. To compare with the preceding we began by computing the "effect" of interviewer variance as the ratio of the actual variance to the simple random variance for the entire sample:

$$e(\bar{y}) = \frac{\text{var}(\bar{y})}{s^2/n} \quad \text{where} \quad s^2 = \frac{1}{n-1}\sum_i^a \sum_j^n (y_{ij} - \bar{y})^2$$

and $$\text{var}(\bar{y}) = \frac{1}{n^2}\frac{a}{a-1}\left[\sum^a y_i^2 + \bar{y}^2 \sum^a n_i^2 - 2\bar{y}\sum^a y_i n_i\right].$$

This last is the variance of the ratio estimator $\bar{y} = y/n$ of a randomly selected clusters. The computed effect on the variance is then equated with $[1 + {}^*\text{roh}(n/a - 1)]$ and this yields the synthetic ${}^*\text{roh} = [e(\bar{y}) - 1]/(n/a - 1)$. (2)

Column 4.

We expected and found that formulas (1) and (2) gave very similar results. We then computed using (2) the values of $e(\bar{y}_1)$ and $e(\bar{y}_2)$, the effects on the variances of the means of two subclasses. We computed actually $e(\bar{y}_1) = \text{var}(\bar{y}_1)/(s^2/n_1)$, and similarly for the other subclass; that is, we did not bother to compute separately s_1^2 and s_2^2 because they would not have differed enough from s^2 to make that extra labor worthwhile.

The synthetic *roh's for the two subclasses were averaged and these were plotted as x in Chart 2 against the roh's for the entire sample.

The variance of the difference $(\bar{y}_1 - \bar{y}_2)$ was computed, taking into account the correlations within the workloads of the a interviewers, as:

$$\text{var}(\bar{y}_1 - \bar{y}) = \text{var}(\bar{y}_1) + \text{var}(\bar{y}_2) - 2\,\text{cov}(\bar{y}_1, \bar{y}_2).$$

As with the variances, the covariance is computed for the ratio estimator of a randomly selected clusters:

$$\text{cov}(\bar{y}_1, \bar{y}_2) = \frac{1}{n_1 n_2}\frac{a}{a-1}\left[\sum^a y_{1i}y_{2i} + \bar{y}_1\bar{y}_2\sum^a n_{1i}n_{2i} - \bar{y}_1\sum^a y_{2i}n_{1i} - \bar{y}_2\sum^a y_{1i}n_{2i}\right]$$

From these we computed the effects on the difference

$$e(\bar{y}_1 - \bar{y}_2) = \frac{\text{var}(y_1 - y_2)}{s^2(1/n_1 + 1/n_2)}$$

Finally, we computed the "synthetic *roh" as

$${}^*\text{roh} = \left[\frac{\text{var}(\bar{y}_1 - \bar{y}_2)}{s^2} - \frac{1}{n_1} - \frac{1}{n_2}\right] \div \left[\frac{2}{a} - \frac{1}{n_1} - \frac{1}{n_2}\right].$$

These values appear as the 0 points on Chart 2 plotted against the values of roh for the entire sample.

DISCUSSION

Eli S. Marks, National Analysts, Inc.

All three of the papers presented in this session exemplify the growing methodological sophistication in the areas of social surveys and interviewing. We are moving away from rule of thumb design of interviews and away from the naive assumption that an answer entered in a designated spot on a printed (or multilithed or mimeographed) schedule represents the literal and exact "truth".

The paper by Kish and Slater differs in approach from those of Lansing and of Barlow, Wirick and Morgan. The latter 2 papers are attempts to isolate factors contributing to respondent and interviewer variance while the paper by Kish and Slater concentrates on measuring the overall effect of interviewer variance. The measurement of overall effect of response error is essential to any decision regarding the amount of research effort it is wise to put into the investigation and control of response error. If an expenditure of $10 per interview will yield results subject to negligible response error, it would be unwise to spend an additional $10 per interview to measure the response error or to reduce it still further.

Both Kish and Slater and Lansing have discussed the question of the direction to be taken by further research in the field of response error. On this point, I agree strongly with Dr. Lansing's emphasis on the importance of basing research on some specific theory of response error. Dr. Lansing cites in support of this view the great variety of experimental techniques which can be considered and the fact that, in the absence of a response error theory, we must resort to hunches and "common sense" to guide our selection among the numerous possibilities. Dr. Lansing's position is even more cogent in view of the difficulties and cost of work in this field.

On the side of "difficulties" we have an excellent illustration in the paper by Barlow, Wirick and Morgan. The technique they use is that of "record check". I had considerable acquaintance with record checks in connection with the evaluation of the quality of the U. S. 1950 Census of Population. It is a very intriguing technique and one which the Census Bureau is using extensively in connection with its 1960 evaluation program. Yet record checks frequently involve the expenditure of much time and effort merely to end up with the conclusion that we can find no record corresponding to the datum we are trying to check. In some few cases absence of a verifying record is (at least presumptive) evidence of an error in our data. More often, absence of a verifying record proves nothing about the truth or falsity of the datum we are trying to check. Failure to find my birth certificate in the file for the place and year in which I claim to have been born certainly doesn't prove that I wasn't born and isn't even very good evidence of an error in the reported date or place of birth. Many of us have, I am sure, had ample acquaintance with errors and misfilings of birth certificates and other records. In the study of Barlow, Wirick and Morgan, for example, the "unverified" (i.e., unlocated) cases are, in most instances, more numerous than those cases where a record was located which indicated a possible error in the respondent's report.

On the side of expense of response error investigation, it must be noted that adequate measurement of interviewer error may require a rather large number of interviewers. This is very strikingly illustrated by the data in the paper by Kish and Slater. While the studies they report involved a substantial number of respondents (462 and 489), they involve relatively few interviewers (20 and 9). One of the points emphasized by the authors is the "happy ending" that the effects of interviewer variance are less for means of subclasses than for the entire sample and these effects seem to disappear completely from the comparisons of subclasses. In fact, examination of Kish and Slater's values of roh disclose 3 negative rohs reported for the subclass means (one of these is - .025) and eight negative rohs for the comparisons of subclass means. While it is possible for the population values to be negative, it is more likely that the estimates are subject to large mean square errors.

In any event it should be noted that the "happy ending" of Kish and Slater is subject to certain reservations. Of the 8 negative rohs for differences between subclass means, 6 correspond to situations where neglecting the interviewer effect in estimating the variance of the difference will result in overestimating by 30% or more. The effect on research conclusions of 6 overestimates of variances may be as serious as the effect of 6 underestimates.

VI

EDUCATION

Chairman, Fred S. McDonald, Stanford University

The United States Office of Education Statistical Program — Virgil R. Walker, Office of Education

Project Talent-Progress Report — John T. Dailey, University of Pittsburgh

Projecting Enrolment at the University of Toronto — Gilbert de B. Robinson, University of Toronto

SOCIAL STATISTICS: PRESENT CONDITIONS, FUTURE NEEDS AND PROSPECTS

R. T. Bowman, Alexander Gall, Israel Rubin, Bureau of the Budget

In his preface to the American edition of Herbert Spencer's The Study of Sociology [1] Professor Youmans notes that Spencer, prior to starting work on his three volume Principles of Sociology, foresaw a difficulty that would arise in working out these principles, namely, "a lack of the data or facts necsssary as a basis of reasoning upon the subject." Professor Youmans then calls attention to the clarity with which Spencer saw the need for facts by quoting from an article written by Spencer in 1858 for the Westminster Review, entitled "What Knowledge is of Most Worth." We commend to your attention the full text of the article cited by Professor Youmans, and give here only a brief excerpt to illustrate the kinds of data which Spencer felt were needed:

> "We want all the facts which help us to understand how a nation has grown and organized itself . . . not only the nature and actions of the central government, but also those of local governments, down to their minutest ramifications . . . Let us be informed of the control exercised by class over class . . . customs which regulated the popular life out-of-doors and in-doors including those concerning the relations of the sexes and the relations of parents to children . . . what was the connection between employers and employed . . . the intellectual condition of the nation in its various grades . . . the kind and amount of education . . . the progress made in science . . . the degree of aesthetic culture . . . the daily lives of the people - their food, their homes, and their amusements . . . These facts, given with as much brevity as consists with clearness and accuracy, should be so grouped and arranged that they may be comprehended in their ensemble, and contemplated as mutually-dependent parts of one great whole. The aim should be so to present them that men may readily trace the consensus subsisting among them, with the view of learning what social phenomena coexist with what others."

Today, more than a century later, Spencer's catalog of information needs still has much that is relevant to the problems with which this paper is concerned. There would be fewer gaps in our social statistics if, as Spencer thought possible, the facts could be gathered and organized according to his prescription - "so grouped and arranged that they may be comprehended in their entirety." But such "descriptive sociology" has proved difficult to achieve and in part must be recognized as an illusion. Yet, quite properly sociologists today are as keenly aware as Spencer of unfilled needs for data. It remains true that "facts" about society are stubborn, hard to capture, and even harder to utilize in ways that are relevant to and adequate for the issues of the times.

This paper was undertaken for one major purpose, to provide a setting for discussion concerning social statistics. It is our hope that this discussion will focus on what social statistics are, their adequacy or inadequacy today, the need for improvement and extension, the possible guides to establishing priorities to meet generally recognized needs, the special problems involved in developing organizations, methods, and procedures for the collection of social statistics of high quality, and, within the climate of interest and need now prevailing and anticipated, the prospects for better social statistics.

Social Statistics Distinguished

A precise definition of social statistics will not be attempted here. Instead we shall try to distinguish the area of social statistics empirically, rather than in accordance with a formula or preconception. It seems to us that even a strict definition of social statistics would probably have to provide criteria that would permit the same body of statistical information to be included as within or outside social statistics depending upon the special uses of the data. For example, to be unemployed is a condition affecting the individual, his family group and perhaps social actions in his community. In this sense data on unemployment, when related to the individual or family, are certainly social statistics. Unemployment, particularly when viewed as an aggregate, is also an indicator of economic activity, and as such is clearly part of the body of economic statistics.

In general, statistics which provide distinguishing characteristics of individuals, families, households, young people, older people, wealthy people, etc., are being used to delineate social (and economic) conditions, and in this sense are social statistics. There are also certain characteristics of individuals and social groups that are clearly not economic even though often influenced by economic or political factors. These aspects such as health conditions, education level and educational aptitudes, marital status, fertility, migration rates, dependency, delinquency and crime, recreation habits, social status, race, citizenship, nationality, are clearly social characteristics on which statistical information is desired. With respect to many of these as well as other significant areas of social statistics, we agree that available data are often inadequate or fail to meet appropriate statistical standards.

We do not attempt in this paper to chart a detailed program to rectify existing deficiencies or serve anticipated needs, but seek rather to suggest factors to be taken into account in the formulation of such a program and in efforts to put it into operation. Our discussion of these

factors is organized under three major headings, as follows:

1. Special aspects of the collection, dissemination and analysis of social statistics

2. Present Federal social statistics evaluated

3. Prospects for the future

Special Aspects of the Collection, Dissemination and Analysis of Social Statistics

One of the first factors to take into account when considering the status of Federal social statistics programs is that most of the operating responsibilities in this field are those of State and local governments. This presents a problem of securing data at the Federal level which meet desired standards without interfering unduly in the exercise of the responsibilities of other governmental units or of private organizations. In this category are to be found for example, education statistics, marriage and divorce statistics, birth and death statistics, and criminal and judicial statistics (except with respect to violations of Federal laws). We do not suggest that a satisfactory set of national data has not in many cases been developed under such circumstances but merely that the difficulties are intensified. These difficulties are further increased if the data reported to the States or local governments involve legal prescriptions which cannot easily be made uniform. In some instances even the resistance to change from inertia alone is a significant factor, particularly with respect to data about people.

A second factor that must be borne in mind about social statistics is that they generally concern information about individuals or family groups ordinarily not recorded in books of accounts or other records. The information, at least in the first instance, has to be obtained from individuals rather than business firms. Non-statistical agencies and organizations, such as schools, hospitals, prisons, courts, welfare organizations, employment offices, ordinarily assemble selected information in the course of their operations, and may be able to provide it for statistical purposes at the Federal level; or information may be obtained directly by enumeration of samples of individuals or households. In either case techniques for gathering such information are an important factor. Our use and knowledge of the techniques required and the institutional arrangements necessary are developing. It is still true, however, that many problems remain concerning appropriate Federal, state and local inter-relationships, the validity of questions asked and the recorded responses, the memory span and recall for information where records are not readily available, and the biases that may be introduced by either or both the respondent and the interviewer.

A third set of factors which are extremely important in evaluating a program of social statistics arises from the uses to which such statistics are put, the local character of problems, the myriad special demands which general purpose data cannot meet adequately, and the lack of unifying principles to guide the design of an overall body of social data. In the economic area, national accounting has come to be a guide for better integrating economic facts, highlighting important gaps, and providing a frame of reference or model to assist in analysis. Demography has provided some of this for the better understanding of population growth, but for the body of social statistics as a whole there is at this time no guiding design comparable to that provided by the national accounts for economic data. Considerable interest has been shown in attempts to develop overall indexes of "levels of living" or separate indicators of different aspects of levels of living but little agreement has been reached concerning the validity or usefulness of such approaches or their value as guides for determining priorities for the collection of social statistics. Not unique with social statistics but also highly important is the fact that much social data are accumulated for administrative purposes and are not designed for analytical uses or broad policy guidance. Since administrative responsibility for many social programs, for example education, resides in the States and local governments, a choice of how to obtain data has to be made. If the responsible operating units are to be the principal sources, then generalized data requirements must be integrated with local needs. If direct inquiry of individuals or institutions is used, a significant problem often arises; it would be generally desirable to collect the data on a sample basis to yield national totals and thus valid data would not ordinarily be provided at the local levels for local policy guidance.

Present Federal Social Statistics Evaluated

It is our intention here to touch briefly on the principal areas of social statistics, indicating major strengths and weaknesses. We shall not attempt to describe the numerous programs contributing to our fund of information concerning social phenomena or problems, but rather to select from them areas which seem to us most illustrative of factors on which attention should be focused. For a brief but nevertheless fairly comprehensive description of our principal social statistics programs, we commend to your attention Part II (pp. 27-77) of the booklet on Statistical Services of the United States Government (revised edition, 1959), prepared by the Office of Statistical Standards of the Bureau of the Budget. We also call your attention to two new publications by the Department of Health, Education, and Welfare. The first is titled Health, Education, and Welfare Indicators, published monthly, and the second Health, Education, and Welfare Trends, published annually. They are in response to the growing interest in social statistics. It is our hope that their use will be instrumental in generating the types of criticism that are effective

for real improvement in statistics. The world-wide interest in social data and the response to that interest are indicated by the effort now being made by the United Nations to develop a Compendium of Social Statistics as recommended by the Statistical Commission at its last session.

Population Statistics

A host of social data are provided in the population and housing censuses. This body of information has grown and improved over the decades and it may well be regarded with pride. In methods, in the items covered, and in the care taken in publication and in making special tabulations available on a cost basis to interested localities and research groups, we are on the way to quality statistics on the characteristics of our country's people. This is not to say that much cannot still be done.

One of the values of the decennial censuses is that they provide comparable data at the local as well as the national level. At the present time some groups are advocating that we should have a "head count" census in years ending with 5 as well as a complete census in years opening each decade. This has merit largely because of the population information it would provide at the local levels. The problem facing us is to assess the value of such interim population data to the States and localities, to estimate proper allocation of governmental financing, and to consider the value of other statistical needs that could be met by an expenditure of about $50 million or, over the years of the decade, $5 million per year.

Census-type data are also collected currently for the nation as a whole on various social and economic characteristics of the population through a monthly sample of 35,000 households. While the Current Population Survey is best known for the information it produces on labor force, employment and unemployment, it is also the basis for information on selected demographic characteristics--such as the population living on farms, migration, fertility--and for social characteristics such as income and school enrollment. Most important, by combining economic, demographic, and other social statistics inquiries in one survey and by analyzing the data for both individuals and for family groups, this Survey provides a wealth of information for the social scientists. From time to time supplementary questions are added to illuminate a particular research or program problem. Here we have an illustration of a design for social statistics along modern lines. It avoids the difficulties of numerous jurisdictions. It does not, however, produce data for small geographic areas nor tie in with the administrative data which arise from social program administration at either the national or local levels. Nor can it continue to meet the increasing demands for supplementary questions on both social and economic characteristics. Serious consideration is now being given to the possibility of establishing a second panel of similar size to carry some of the burden of additional inquiries.

Few analysts seem to be aware of the wealth of CPS data already provided. We cannot forebear mentioning that one of the most frustrating occupational hazards of the Government statistician is that of being publicly chided for the lack of statistics which he has lovingly designed, collected, tabulated, published, and, he thought, publicized. We agree that the fault is largely ours. Obviously, the last step has not been sufficiently well done. If the discussion here can suggest better ways to bring the statistician's product before the appropriate users, we should be grateful.

We readily admit that the Current Population Survey does not measure all the characteristics of the population which the social scientist would find significant. In some cases, the affected population groups are too small for reliable measurement (migratory agricultural laborers, for example) or too difficult to identify in a household interview (who will admit to having a juvenile delinquent in the family?). In other cases, concepts and measurement techniques are not yet sufficiently well-defined to permit general agreement on survey design, for example in the measurement of social status and social mobility. We have yet to obtain agreement in the U. S. on an acceptable scale for assigning "socio-economic status", an endeavor which has absorbed considerable time at various international statistical conferences during the past decade.

In spite of the general state of excellence of our population statistics, there are a number of attributes of the Nation's inhabitants which are not measured at all or measured with indifferent success. In each such case the reasons are not far to seek, but perhaps further efforts would be repaid in more reliable statistics to serve the interests of sociologists. Among these areas, we suggest the following as of considerable importance:

Religion - Information on religious preference was collected on an experimental basis in the Current Population Survey in 1957 but was not added to the 1960 Census because of public disapproval of the inquiry.

Ethnic stock - Although sociologists are almost uniformly agreed on the importance of information on ethnic stock, and medical scientists are developing an interest in the data in studying the incidence of certain diseases, methods of measurement developed to date for use in the decennial censuses are generally inadequate. Neither "country of birth of parents"

nor "mother tongue" seem happy solutions.

Handicapped persons - Information on the numbers, location, and the personal, economic, and family characteristics of handicapped persons, especially the blind, are in great demand. Such information might appear to be suitable for census of population inquiries except for the apparently insurmountable technical difficulties of adequately identifying these people under "census conditions". At present, we "make do" with information on persons served by various governmental and private programs without knowing whether or not they are representative of the unknown totals of the afflicted groups. Some national information may become available from the National Health Survey in later years when the accumulation of data will permit analysis for such relatively small groups.

Employment and Unemployment

Although aggregative measures of employment and unemployment are generally considered economic indicators rather than social statistics, labor force status is certainly an important social characteristic. It seems unnecessary to enumerate here the variety of social characteristics of individuals and families which are cross-classified in the Current Population Survey with employment status. The variety, frequency and reliability of these data are unique in the annals of statistics. In addition to the fact of employment, or of seeking work, the information simultaneously obtained on occupation, class of worker (employee, self-employed, or unpaid family workers) and hours of work are important social attributes.

Nevertheless, important questions remain unanswered which seem suitable for statistical investigation. The measurement of underemployment, for example, in terms of training or capability; the mobility of the unemployed; changes in labor force participation rates and their relation to various social conditions as well as economic events. From the point of view of possible governmental programs, these are questions which are particularly relevant to "depressed areas" and persistent unemployment of individuals. Some exploratory work has been done on some of them in occasional studies of particular areas. For example, the Department of Agriculture has cooperated with a number of local groups in studies of unemployment and underemployment among farmers and their families; the Bureau of Labor Statistics has engaged in studies of the job-seeking experiences of laid-off employees and of youths leaving school before graduation; the Bureau of Employment Security has sponsored a number of inquiries by State employment security agencies into the characteristics of the unemployed and the occupational skills of the labor force. At the moment, these disparate surveys do not seem to add up to a "program". Social scientists might well join with labor force analysts in suggesting desirable lines of investigation and testing research methods.

Very different in nature, and with implications for both economic and social analysis, are the statistics collected on working conditions - accidents, work stoppages, collective bargaining provisions, and labor standards. Some of these have long been a part of our statistical repertoire. In general, they suffer from limited coverage, or insufficient financial support, since they appear central neither to economic or social analysis nor (except for labor standards) to Federal "action" programs.

Social Insurance and Social Welfare

In the fields of governmentally-operated insurance, pension and assistance programs, very considerable progress has been made, especially in the development of "operating" or workload statistics. A look back at the situation up to and largely through the period of the thirties provides major contrasts and indicates significant progress. We now have considerable information about our pension and dependent population segments. Recently, realization has been growing that such statistics are insufficient or of the wrong kind, that related information is needed on the affected individuals and their families. Furthermore, many families may be simultaneously affected by more than one program. Sporadic attempts have been made to fill these needs, for example, by adding occasional supplementary questions about veterans to the Current Population Survey; by taking surveys "in depth" on the resources, living arrangements and scale of living of old-age beneficiaries; by collating a sample of assistance recipients against old-age and survivors' records to determine the number of families receiving income under both programs; by instituting special surveys of the expenditures of unemployment insurance beneficiaries before and after loss of employment, etc.

It is probably true that these types of information have not as yet been properly integrated with other bodies of information about persons and families and one of our major efforts in the future will be to pull the various aspects of social statistics into a more coherent body of information. For this task guidance is necessary from social analysts generally since no clear guide lines exist at present. At the same time it seems evident that for many "program-planning" purposes, information is required in far more variety and depth than can be supplied on a repetitive mass statistical basis. There will always be room for the small-scale intensive exploration of emerging problems, using statistical procedures which approach the case study method. More atten-

tion should be devoted to questions of when the use of such methods is appropriate and how they can be made to yield results which will permit reliable analysis and generalization.

Family Living

Historically studies of family living have provided a whole range of social and economic data, including information about the cultural and social characteristics of the groups studied. It is true, however, that statistical studies of family living have focused most attention on those characteristics which are closely related to economic condition. Thus much emphasis has been devoted to statistics of family incomes and expenditures.

Family-living studies can be approached in two ways - indirectly, from sources exterior to the family and by direct inquiries involving members of the family.

In the first category are most of the comprehensive data relating to consumption and saving which appear in the national accounts. These are compiled largely from business and Government administrative sources.

Family-living statistics which are directly collected from individuals or families include first, the largest of all regular "surveys" in the country, the Federal income tax return, data from which are utilized for a wide variety of purposes both economic and social. They also include the sample surveys of consumer expenditures made by the Bureau of Labor Statistics and the farm family living surveys of the Department of Agriculture.

The BLS Survey of Consumer Expenditures last made in 1950 will again be carried out in 1961 and 1962. This survey of 11,500 families is designed both to provide weights for the Consumer Price Index and to provide data on family income, savings and expenditures in as much detail as feasible. A proposed innovation in this survey is its extension to farm and rural nonfarm families in cooperation with the Department of Agriculture.

Statistics on the expenditures of farm families are available from a number of studies of the Department of Agriculture. In 1955 the Department carried out a nation-wide survey of living expenses and farm operating expenses of farm families. Although a basic purpose of this study was to furnish weights for the Index of Prices Paid by Farmers it provided a wide variety of additional information on family levels of living.

Continuing data on farm-families (and sometimes rural non-farm families) have been made available by recent studies of the Department of Agriculture's Agricultural Research Service. These studies are made on a "spot" basis, usually covering a few counties in a State and are designed to give not only expenditure data, but an analysis of some special topic, e.g. the job-related expenditures of working wives.

Mention should also be made of the nation-wide food consumption studies of the Department of Agriculture, last made in 1955, which provided estimates of food consumption for all population groups in the country, urban, farm, and rural non-farm.

With many workmanlike surveys furnishing statistics of family living it may seem paradoxical to maintain that in terms of reliable totals for large aggregations of families they add little to the data obtained from indirect sources. For aggregate economic data, as Broida points out, "consumers are a high-cost source of information of a quality that is often poor," [2] and this for several reasons.

The skewness of the distribution of consumers according to many variables (e.g. income, corporate stockholdings, etc.) introduces a difficult sampling problem and the difficulty may be compounded if the response rate or the accuracy of replies varies with the variable. It has been the consistent experience of technicians in the field that the refusal rates in such surveys tend to be higher among high-income families. Under-reporting of income is also common and there is evidence that understatement of income from interest, dividends, and similar sources is greater than for other types and therefore is more serious for higher income families.

But most important as a single source of error is the fact that families just do not keep organized records of at least a large part of their expenditures, so that errors of recall, both as to items purchased and time of purchase become quite important.

We cannot undertake here a summary of the errors which affect, in different proportions, data compiled from indirect sources, particularly businesses and those compiled from the families themselves. Research in both fields is needed.[3] Briefly we may point out certain broad considerations.

(1) Data from indirect sources furnish statistics concerning consumers in broad aggregations which are for many purposes superior to corresponding data which are furnished by the existing direct surveys.

(2) Research is needed which is directed toward integrating information from both types of source. At the present time studies producing detailed family expenditures also use appropriately broad aggregated totals corresponding to the detailed data. Further research may well make it possible to utilize indirect sources for broad categories and the more detailed direct surveys for relative dis-

tributions within those categories. Such a development would have an integrating effect upon all statistics of family living, however derived.

Education, Science and Culture

Statistics about the educational process and its achievements are today one of the areas of social statistics being subjected to major adverse criticism. The expressions of dissatisfaction have justification but they should be tempered with a realization that the newly awakened interest must also be constructive. Suddenly, we as a people are demanding that achievements in science and technology be accelerated. We demand such success from the educational process, and we want more education and better education immediately. We want to know more about education to these ends.

More should be known about education and its products and this knowledge should have been insisted upon much sooner. But statistical information about education must not be thought of as an end in itself. It must be designed to provide the guide lines to educational policy that will promote science, develop the arts and produce the educated man. These results do not come quickly or with great certainty.

Our educational system is not an easy one for which to produce both the general purpose type of data and the more searching information which throws light on the educational process and the potentials of our people. Since it does not now exist, a program must be designed to these ends. Efforts to achieve such a program must involve more than carping criticism, they must be unrelenting for the achievement of constructive advances in information about education, the educational process and the success of this process.

By a program for education statistics we mean to include (1) recurring general purpose statistics, (2) one-time statistical surveys, (3) statistical estimates and projections, (4) statistics to plan, administer and evaluate operating programs, and (5) statistics arranged for under contracts and grants for studies in depth. Until a recognizable, understandable, and defensible integration of these various aspects of educational statistics can be formulated into an action program, efforts will remain diffused and subject to unanswerable criticism. Two elements appear essential if progress is to be made. First, the Office of Education should organize so as to give high priority to the function of the analysis and planning of its statistical activities. Second, it should strengthen its staff of statisticians so as to improve the application of statistical standards to obtain quality statistics.

As we see it there is no magic formula (not even more money) for the development of educational statistics to meet major uses. The demands of users if unanalyzed and unrelated to statistical feasibility and best resource utilization provide no proper guide of what should be done. The instituting of more direct national surveys to secure repetitive type data, will not meet all of the needs either. One important element of the problem requires the establishment at State and local levels of a core of information available for both local and national purposes and consistent over the nation as a whole. There are subject areas, however, in which national data needs can be met by national surveys utilizing sample sizes which are easily manageable.

As we have noted, what Spencer called "the intellectual condition of the nation" embraces not only "the kind and amount of education," but also "the progress made in science" and "the degree of aesthetic culture." The latter areas have received relatively little emphasis in efforts in this country to measure social phenomena and behavior.

Health and Vital Statistics

Before adoption of the National Health Survey Act in 1956, national statistics of health and medical care were fragmentary, consisting mainly of data on morbidity covering the most common communicable diseases, compiled by the National Office of Vital Statistics, results of special studies and surveys conducted by the Public Health Service, and data arising from operating programs of the Public Health Service and such agencies as the Veterans Administration, Railroad Retirement Board, and departments of the Army, Navy and Air Force. Now, however, new and valuable national data on a variety of health subjects are provided by the National Health Survey, which serves as an outstanding demonstration of the advantages of the direct survey method as contrasted with indirect methods depending on the cooperation of autonomous units of State and local governments. The scope and value of the data obtained under this program have only begun to be realized, and their further development may be anticipated.

As the term "vital" statistics suggests, data on occurrence of such major events in the life of the individual as birth, death, marriage and divorce are recognized as basic elements of social statistics. In this area, at least with respect to compilation of national figures on births and deaths, a considerable degree of success has crowned years of patient effort to surmount difficulties inherent in the diffusion of responsibilities among State and local agencies for registration of vital events and compilation of data concerning them. We may anticipate a parallel development of marriage and divorce statistics in the future, with the gradual extension of the scope of registration areas and the introduction this year of a new program for direct processing and tabulation by the National Office of Vital Statistics of a national sample of the report forms (which for 1960 will be related to decennial population census data).

Perhaps the foremost unmet need in our vital statistics programs is for greater depth of analysis of available data. To some extent the feasibility of types of analysis that would be valuable for many social purposes is restricted, at least on a national basis, by the difficulty of integrating discrete elements of data from widely scattered sources. That such difficulties are not insurmountable, however, is illustrated by a paper being presented at another session of these meetings by Philip Hauser and Evelyn M. Kitagawa, on a study of social and economic differentials in mortality in the United States, which is to be conducted by the University of Chicago with the aid of a research grant from National Institutes of Health (also a unit of the U. S. Public Health Service). Another type of need will be served when NOVS completes work now planned on preparation for the 1960 census period of unabridged U. S. life tables and related actuarial tables, as well as sub-national life tables for geographic regions and States.

Future improvements in the area of health and vital statistics, particularly in analytical work, may be expected from the amalgamation of the National Office of Vital Statistics and the National Health Survey in a single unit of the Public Health Service, the newly created National Center for Health Statistics. The concept underlying this change is "to bring together those activities in which national vital and health statistics are collected and analyzed, as self-contained programs of national and international significance. Additional objectives are to increase the scope and effectiveness of such programs; /and/ to capture and disseminate more widely the significance of their data."

Other Subjects

In this rather impressionistic sweep over fields of social statistics we have omitted many subjects that could well be explored in depth. With respect to the important field of criminal and judicial statistics the omission is explained by the fact that a paper on this subject is being presented at another session of these meetings by Dana M. Barbour, also a member of the staff of the Office of Statistical Standards. As Mr. Barbour notes in his paper, there are major gaps in our data on crime and criminals, although significant improvements have been made and some further progress may be anticipated. The difficulties here are very great and much greater efforts on the part of specialists in criminology will be needed.

We have mentioned housing only incidentally, as one of the subjects on which a substantial amount of information is available from our decennial census program. Much more could be done in this field, however, by way of relating housing data to significant socio-economic characteristics of the population.

Although we have mentioned some of the important individual components of generally accepted measures of standards and levels of living, we have not undertaken an integrated approach to measures in this field parallel to that followed in recent years in efforts to develop international agreement on concepts and content of components and indicators of well being. We are not certain of the usefulness of such an approach for the United States.

Prospects for the Future

There is an old saying that "If wishes were horses, beggars might ride." Certainly if wishes were prospects the future of social statistics would be bright. It can be said with confidence that there is a rising tide of interest and that more attention will be devoted to this area of statistics. But at least three things must be achieved if there is to be a major development of social statistics generally.

First, better guides must be developed for establishing priorities within the various branches of social statistics and for social statistics generally. The analytical uses which establish these priorities must also be more clearly indicated. To this end Government efforts must be better guided by social scientists providing constructive and specific suggestions broader than the special research interests of individuals.

Second, methods must be developed to bring about consistent and meaningful information at the level of local program direction which is useful for national purposes as well as for local policy guidance.

Third, ways must be developed to integrate the various bodies of information from samples of households and individuals with data from records of operating agencies so that the unique advantages of each type of statistics may serve to reinforce and enrich the other. Operating records can be made to provide reliability and detail at relatively low cost, with individual and household sample surveys furnishing a variety of social characteristics essential for analysis.

Officially we have responsibilities for encouraging and supporting actions to achieve appropriate solutions to the problems here formulated. Some may even presume that the answers should all be known to us. In some instances we do think we know a few of the answers. It is not in a sense of modesty that we admit less than perfect knowledge of solutions but purely as a matter of necessity. It is our hope that from the discussion we stimulate here and from more intensive discussions which we plan to arrange for subsequently, necessary improvements in the whole field of social statistics can be accomplished.

1/ Herbert Spencer, The Study of Sociology, D. Appleton and Co., New York and London, 1921.

2/ Broida, A. L., Consumer Surveys as a source of information for social accounting -- the problems; paper delivered at Conference on Research in Income and Wealth, November 1959.

3/ A competent summary of many of the factors affecting both types of statistics is given by Broida, op. cit. Part II, Sources of Error in Consumer Surveys.

DISCUSSION

Harry Sharp, The University of Michigan

Dr. Bowman and his associates have given us a highly competent review of the state and prospects of social statistics in the United States. Their comment about the presence of a wealth of useful statistics, particularly in the Current Population Surveys, which is not used because the researcher does not know of its existence, is quite true. I would argue with the authors' contention, however, that the fault lies primarily with government researchers. Generally, I feel that the users of government statistics do not employ nearly enough effort to seek out the data which are available to them.

My main feeling about the inadequacy of CPS data was touched on in the paper. Many of us would like to see the government provide such information for relatively small areas. In my own case, CPS data on the ten or twelve largest communities in the United States would be accepted most gratefully.

Perhaps purposefully, Bowman, Gall and Rubin appear to me to be overly critical of the scope of the social statistics which are presently provided by the federal government. For example, they cite the need for data bearing on such important sociological concepts as social status, social mobility, ethnic status, and migration history. It seems a bit unfair to hold the government even partly responsible for definitions of these terms when social scientists, who use such concepts with complete impunity, cannot agree on generally acceptable definitions.

The authors list several areas wherein a greater depth of data is needed, and imply that the government may be criticized for not supplying this information. However, I do not feel we can blame the Bureau of the Census, for example, if recently it has become more cautious in the selection of new areas around which to collect data.

The census people worked industriously to include a question on religious preference in the 1960 decennial census. As the authors indicate, the "public disapproval" of the inquiry, although expressed by only a few well-intentioned groups, was sufficiently organized so that this important question was not asked in 1960. Moreover, publication of CPS data which did relate to this question was severely restricted, through no fault of the Bureau of the Census.

I agree with the authors that a large body of important data remains to be collected, but I doubt that, at least in the near future, we can look to the government to supply a major share of such information. Nongovernmentally financed research probably will continue for many years to give us social statistics that necessarily will be unobtainable through government agencies. The question of costs is an important one, of course. Without a tremendous increase in the allocations for data collections, the only way that more data of the type described in this paper can be made available is to cut-back on the data now being supplied. It would be difficult indeed to obtain agreement on just what forms of presently available data should be dropped, so that different forms can be added. In addition, the very fact that a governmental agency wishes to collect many badly needed types of data will make such action politically, if not economically, unfeasible.

Finally, I sympathize with the authors for not attempting to define "social statistics." But I am not completely clear as to what is or is not social statistics. I thought at first that they were distinguishing between "economic" and "social" statistics. Later I was not so sure. Perhaps they would agree with this definition paraphrased slightly from their paper: social statistics are any statistics which are "used to delineate social...conditions."

SUMMARY OF DISCUSSION

Maurice I. Gershenson, California Department of Industrial Relations

The paper by Bowman and his colleagues brought before those of us who have been long-time users and producers of social statistics, a panorama of the great progress that has been made in our own working life time.

It's easy to take for granted the many excellent statistical series now available and to forget the days, not so long ago, when social scientists struggled with a pitifully small amount of usable data.

We must not become complacent. We must be aware of the many gaps that still exist. I want to emphasize one particular area where much remains to be done.

These words of Spencer quoted by Youmans sound the central theme of what I would like to stress, "...not only the nature and actions of the central government, but also those of local governments, down to their minutest ramifications..."

The authors have touched upon the need for more local area data and on the difficulties of obtaining such data.

Their primary concern, however, has been with national statistics. The growing wealth of national data points up more sharply the gaps in regional information. For a proper understanding of what is going on in the nation, it is necessary to know what is happening to its parts.

We have gone a long way in many states in producing a varied array of social and economic statistics for local areas. I am very proud of the accomplishments of my own State, California.

The quantity and quality of state and local area statistics varies considerably throughout the nation. But even in those states that have made the greatest progress -- where the quantity and quality is high -- most of the series are those that come out of a repetitive mass statistical operation and are in terms of aggregates.

What we need, and need badly, are more statistics on characteristics of individuals. We need for state and local areas more of the kind of information of the type produced by the Current Population Survey.

I am aware that the decennial census gives us a great deal of data on the characteristics of individuals and families for state and local areas. But in a state that is growing as fast as California, much of the data becomes obsolete even before publication.

I would like to see an expansion of the size of the Current Population Survey sample so that data can be developed for states and local areas.

Who should finance this is a question that immediately arises. Perhaps some arrangement for joint financing -- Federal and State -- can be made.

As social scientists we must press hard to fill this important gap. We must press to make some of our wishes come true so that some beggars may ride.

DISCUSSION

Ronald H. Beattie, Chief, California Bureau of Criminal Statistics

The general situation in criminal statistics is such that there are many areas to which a great deal of attention and work must be devoted to improve our present knowledge. Briefly, I would like to mention just five of these areas at this time.

First, a great deal more effort must be given to the uniform definitions, classifications, and descriptions of criminal offenses. At the present time, most of our criminal statistics are gathered in terms of broad offense groupings. These do not distinguish some of the significant differences that must be observed if we are to evaluate our criminal problem. For example, the term burglary covers all kinds of offenses that include breaking and entering, and in some instances merely entering an enclosed area with intent to steal or commit some other offense. This means that in the burglary classification there are literally thousands of cases that are relatively insignificant violations of the law together with very serious cases involving safe cracking and breaking and entering with large-scale theft of property. Until we can subclassify and account for the more serious offenses in these general classifications as distinguished from the rather minor and less serious offenses, we will not be in a position to observe the changes and trends in serious crime with which we are chiefly concerned. The need for these subclassifications involves all offenses, not just burglary. In all of them the problem of what is serious, what is run-of-the-mill, and what is minor is not now well distinguished in our collections of criminal statistics.

A second area that requires a great deal more study and exploration, and which is related to more exact definitions of crime, is that of accounting for the activity in our criminal law control at the police level. Law enforcement is primarily a local affair. This means that in California, for instance, we have over 400 independent police departments and sheriffs' offices. To obtain uniform and complete coverage of the data on offenses and offenders from so many agencies poses a real problem. Today we have summary reporting, but such a method does not permit any assurance that the material reported from all of the sources is in exactly the same terms despite uniform instructions. Much more accurate information will be obtained when it is possible to gather at one central location, individual reports on offenses and offenders so that a consistent uniform classification can be given to all the information, and the many details relating to the offense and the offender can be classified, summarized, and analyzed.

A third area where a great deal of progress can be made, particularly if the first two discussed are sufficiently developed, is the accounting for each person charged with a crime individually throughout the process of criminal justice until the final disposition. Then, it will be possible to analyze and determine what kinds of persons involved in what kinds of offenses, are released at the police level, or are eliminated from the process during the period of prosecution, or are given different types of criminal treatment. Furthermore, such a development would make possible the use of mortality tables in terms of what happens to persons arrested and charged with crimes throughout the total process.

A fourth point that will be just briefly mentioned relates to examining more carefully the social and enviornmental factors surrounding criminal behavior. In recent years a great deal of emphasis has been given to psychological and personality factors which seem to be inherent in persons who commit crimes. A great deal more attention should be given, however, to the enviornmental conditions and other factors outside the person that may relate to criminal conduct. We are all aware that carelessness in leaving automobile keys in cars will contribute to the greater possibility of auto theft. What of the many other situations in the criminal field that have some effect on the volume and kind of offenses that may be committed? Particularly do we recognize how loose conditions contribute to the spread of worthless checks, and there are undoubtedly conditions of this type that contribute to the great amount of theft and burglary that exists.

My final point is to give emphasis to the matter of developing state responsibility in the collection of criminal statistics. Each of the fifty states of this country is sovereign in its control of crime. Each has its own set of criminal codes and its own established criminal procedures. It is almost impossible to conceive the collecting of national statistics from these fifty states, the District of Columbia, and the special Federal jurisdiction by one agency unless each of these sovereign divisions takes the full responsibility for developing complete and comprehensive reporting within each state or jurisdictional border. Our failure to obtain satisfactory national statistics of crime to the present time basically is because the states have not, to any large extent, assumed the full responsibility of developing the reporting, summarization, and analysis of the data on crime and delinquency within their own boundaries.

V

REVIEW OF SOCIAL STATISTICS

Chairman, Frederick F. Stephan, Princeton University

Social Statistics: Present Conditions, Future Needs and Prospects — R. T. Bowman, Alexander Gall, Israel Rubin, Office of Statistical Standards, Bureau of the Budget

Discussion — Harry Sharp, University of Michigan

Discussion — Maurice I. Gershenson, California Department of Industrial Relations

THE UNITED STATES OFFICE OF EDUCATION STATISTICAL PROGRAM

Virgil R. Walker, U. S. Office of Education
U. S. Department of Health, Education, and Welfare

From the standpoint of sheer numbers alone, the magnitude of the Nation's educational program is overwhelming. In the Nation more than 45 million people, or more than one person in every four is either attending some type of school or is engaged in supporting activities. Education is a multibillion dollar business with a huge investment in staff, plant and equipment. Because of the importance of education, the volume of quantitative data that should be collected, analyzed, and interpreted in order to understand and improve the system is enormous. When the need for qualitative data is superimposed upon the need for quantitative data, the total volume of required information begins to defy comprehension.

This brief introduction to a potential statistical program in education indicates that the subject of this paper and its implications for education should be of vital interest to all of us. The current status of our educational system, its problems already defined or requiring definition, the research leading to a better understanding of these problems, and the recent developments and trends in education are all items of far-reaching implications for our future. They are also items provoking statistical inquiry at all levels of responsibility--Federal, State, and local.

The United States Office of Education and in particular the Educational Statistics Branch, with which I have recently become associated, has primary responsibility for the Federal statistical program in the field of education. This is a responsibility beset with many challenges for its proper fulfillment.

During this critical period when our Office must assume an ever-increasing leadership role in education, we face many problems in our statistical program. How well has this program been designed to meet nationwide needs? Does it accord with our 1867 mandate ... "to collect statistics and facts showing the condition and progress of education ..."? Does it meet administrative operating program needs of the Office itself as well as those of other Federal agencies and of private and public institutions? What is the best way to ascertain the magnitude and nature of these needs? How can we determine the priorities of these needs? How can we best unify and coordinate the program?

What personnel and machine resources do we need? How should we organize these resources? Should we develop a functional or a subject-matter approach in the assignment of responsibilities? Centralized or a decentralized approach?

From whom and from what records shall data be obtained? By what means? What can be done to develop and foster the acceptance of uniform definitions? What is the best way to encourage the establishment of uniform and efficient record systems among suppliers of data? What is the optimum utilization of sampling techniques?

How can data best be verified and processed? What about the non-response problem, especially since most of our surveys need to depend on mail questionnaires? What level of analysis should be applied to the data before publication? What preliminary, tentative, or partial findings should be released?

How can desires and needs of users of data be ascertained with respect to content and manner of presentation? How can the format of our presentations be improved? How can we assure ourselves of rapid and appropriate distribution to those who need the data most?

If we do not satisfactorily answer these and many similar questions in order to become fully aware of the status of education and the magnitude and nature of emerging problems, a national crisis in education may come upon us for which we will be unprepared.

Before proceeding to describe briefly what we have done to answer many of the above questions and what we plan to do in the future, it would appear desirable to make several observations. First, we must recognize the many statistical attainments of the Office of Education. It has served the Nation well for many years. It has developed several series of publications which furnish multipurpose statistics in some of the most important fields of education. For many years the best known of these was the Biennial Survey of Education. More recently, other publications such as those concerning degrees and opening fall enrollment, have also become widely accepted and utilized. These series, combined with several special purpose studies, have provided basic statistics concerning education from the kindergarten through the graduate level.

Second, it is necessary to observe that the U. S. Office of Education is an agency with diverse and broad functions involving many ongoing statistical and operating programs. These functions must be continued uninterruptedly while administrative and technical improvements are made in the statistical and research programs. Other responsibilities involve the administration of grant programs and the performance of vital advisory, legislative, and policy-making functions.

The Office of Education is proud of the fact that it has developed a very high level of rapport with respondents to its questionnaires and has achieved some excellent response rates in a purely voluntary reporting structure.

Without the utmost cooperation of local and State school officials and of responsible personnel in institutions of higher education, we could not have been successful in obtaining valid and reliable data in many sectors of education. Through personal visitation of staff members, this level of rapport is constantly being safeguarded and enhanced.

In verifying and processing returns, personnel in the Office of Education have exercised exceptional care. Careful checks are made for internal consistency. Comparisons are made with previous returns and with a wide variety of available documents. Significant discrepancies are checked with respondents by means of letters, telegrams, telephone conversations, and, wherever feasible, personal visitations. A very well organized follow-up system has also been developed.

We are constantly mindful of the advantages of sampling and are expanding the utilization of samples for many studies. Just recently, our sole mathematical statistician reported that he had worked on 10 different projects involving sampling theory and techniques during the past month. We have also developed or obtained several directories which serve as universes for sampling. Typical of these are the universes of approximately 40,000 school districts and 25,000 public secondary day schools. We are planning to improve and extend our sampling techniques by utilizing the findings from completed studies and from "Project Talent" which is being carried on by the American Institute for Research and partially supported through the Cooperative Research Branch of the Office of Education. We are also planning to develop universe distributions and parameters to aid in our future sample designs.

As many of you are aware, the Educational Statistics Branch conducts many recurring general-purpose statistical surveys of a cross-sectional nature. Recently it has also completed two depth studies concerning beginning teachers and teacher turnover which we hope may become longitudinal. In addition to these functions, the Branch provides statistical and research advisory services and editing and processing operating services for educational specialists in other divisions of the Office who conduct studies. Recent acquisitions to our staff and some reassignment of functions have enabled us to strengthen both of these services. By developing better understandings and closer relationships between these specialists and the statistical personnel in our Branch, we are becoming more rigorous in the application of statistical standards and in extending the principles of sound research design throughout the Office. Much remains to be done, since our research needs have expanded rapidly and statistical services need to be strengthened similarly.

Fortunately we have also recently strengthened the quality of our staff by employing statisticians well versed in the field of sociology and economics. These individuals are already improving the quality of our products in the Reference, Estimates, and Projections Section. The most recent addition to the staff of the Educational Statistics Branch is our assistant director who is here today. He is a most competent analytical, survey, and mathematical statistician who has had many years of sample survey experience as Chief of the Statistics Branch for the Bureau of Old-Age and Survivors Insurance, Social Security Administration.

Many of the questions posed earlier in this paper remain unanswered. Some have been only partially answered. We must continually strive for a more unified program which eliminates overlaps, reduces the volume of questionnaires, and fills the major gaps. We must develop greater awareness for uniformly high standards of quality by means of inservice training and the establishment of standard operating procedures. Our data, in addition to being analyzed and interpreted in greater depth, must be made available in a form most appropriate to users, with less delay than at present, and with a fuller explanation of their nature and qualifications.

At this point some of you will perhaps be interested in more definite facts about the size of the Office's current statistical program. In the fiscal year ending last June about 50 different statistical surveys, both of a recurring and special type, were processed. This number represented an increase of about one-fourth in our survey volume in the last two years. Owing to the expansion of the Office's specialist staff as a result of new programs begun under the National Defense Education Act of 1958 and other recent legislation, the number of statistical surveys is rapidly increasing. More than 80 survey instruments are expected to be approved during the present fiscal year. Proposed surveys for fiscal year 1962 indicate that the 1961 load will continue and that the average size of the individual survey will expand.

The problem being created by the magnitude of this workload becomes very significant when one considers the small professional staff of the Educational Statistics Branch. Presently the staff consists of about 30 professional members including administrative personnel, only a few of whom are professional statisticians. Materialization of our plans to strengthen the mathematical statistical staff, we hope, will alleviate our problem of keeping pace with the increased demand for educational statistics.

The net result of the above expansion will inevitably be to require a new level of statistical operation not contemplated in the Office a few years ago. Consequently, the Office is presently studying the feasibility of making a radical change in the approach to compiling educational statistics. This plan envisages the collection initially of "basic" items concerning education through an organized system

of data flow. This will include data on students and pupils, faculty members, school and school districts, and institutions of higher education. We hope that these items will create in the U. S. Office of Education a permanent "Bank of Information" that will eventually contain a large number of essential facts about education. By utilizing modern computer capabilities and performing "automatic data processing" we will be able to produce statistical interrelationships never before possible under our present system of questionnaire studies which normally ask for derived data. Time permits but a very limited discussion of the potentialities of this system. As statisticians you can readily visualize the kind of data that could be produced regarding staffing the institutions of higher education in our Nation if you had such basic items as sex, marital status, rank, birthdate, title, highest degree held, salary, major and minor field of preparation, workload, subjects taught, and length of service for every faculty member or even for a sample. Then too, consider the rosters and mailing lists that could be produced rapidly through ADP processing. This system, we hope also will provide almost unlimited possibilities for the storage and retrieval of such types of educational information as the titles, authors, and abstracts of research in education. There is also much to be done in the field of investigation, for at present it is only on the drafting board.

We are not certain that the educational profession appreciates the need for statistics that can be generalized to the Nation as a whole, since many of our past data collections have been on specific localized problems. Most data that are gathered by associations and institutions tend to be highly oriented toward specific needs. It is our hope that the "basic items" approach will supply data which will make many generalizations possible.

The U. S. Office of Education has set in motion a strong effort to improve its statistical program, in fuller use of sampling, via ADP processing and in modern statistical standards. We will welcome your suggestions during the discussion period of this meeting or at any other time. We hope that many of you will be able to visit our office in Washington. We would like the opportunity of having each of you personally see our operation and meet our staff.

PROJECT TALENT - PROGRESS REPORT

John T. Dailey
University of Pittsburgh
American Institute for Research

In Project TALENT, 450,000 students were tested in 1357 schools. Each student was tested for two days and each school filled out approximately 50 pages of questionnaire material regarding its characteristics, activities, guidance program, and counselors. Close to one hundred individual test scores are available for each student. The students also filled out an extensive questionnaire of several hundred items of information about background, aims, experience, aspirations, etc. Several hundred items of information are available about each school. All of these test and school measures will be related eventually to a number of target or goal criterion variables which will be measures of important educational outcomes and life outcomes. A basic goal of the analyses for Project TALENT is to study the interrelationships of these data in order to determine the residual or unique relationships of each test or school variable and each goal criterion variable, with all other test and school variables held constant. While this may be done eventually through curve fitting and the like, the initial approach will be through the basic procedure of multiple-regression analysis. Here, for large numbers of the independent variables (test and school measures) versus a single dependent variable (goal criterion variable) at a time, partial r's and multiple R's together with multiple-regression equations will be computed. Other types of covariance or residual variance analyses will also be done.

Simple zero-order correlations between an independent variable and a dependent goal criterion variable are usually highly deceptive and do not really represent the true unique relationship between a pair of variables. A striking example of this occurred in a study of Naval recruiting. It was found, for the various recruiting areas of the country, that there was a high negative correlation between the proportion of families with TV sets and the recruits obtained from the area per thousand population. However, a multiple-regression analysis showed that there was actually a zero partial correlation between TV sets and recruit productivity. This statistical artifact was a result of TV sets (at that time - 1950) existing mainly in the industrialized areas of the country. There was a true unique relationship between degree of industrialization and recruit productivity. This remained as a strong partial correlation when about 40 variables were analyzed simultaneously, whereas the TV correlation vanished.

It is planned that such multiple-regression analyses will be carried out for groups of 100 to 200 independent variables at a time against a specific goal criterion variable. Such criteria will include graduating from high school, going to college, grades made in high school, grades made in college, the occupation entered, the salary received after N years, whether the student has migrated, how well he likes his occupational choice, the types of courses he takes in college, whether he becomes a scientist, or leader, or delinquent, or something else.

It will be possible to include many questionnaire items as variables in the multiple-regression equations by dichotomizing the questionnaire items when necessary.

Group Comparison Studies

As a first approach in exploring the domain, item analyses will be carried out where means and standard deviations on all variables will be computed for large numbers of special groups such as:

a. Occupational groups
b. Parental occupational groups
c. Regional groups
d. Unusually successful groups (creative)
e. Unusually unsuccessful groups (delinquents, psychiatric cases, etc.)
f. Migrants
g. Types of school
h. Groups with specific attitudes towards business, saving, spending
i. Groups planning military careers

After the first round of item analyses has indicated potential clusterings of the variables' interrelationships, more sophisticated matched group comparisons will be made. Here, means and standard deviations for all variables will be computed for matched groups (groups that are alike in N dimensions but are different in dimension N + 1). Examples:

a. Students alike in sex, grade, region, socio-economic status, high school grades, and total Information Test score but different in planning to attend college (or skipping a grade, failing a grade, wanting to become a scientist, ability to write a paragraph, self-concept, basic values, etc.)

b. Schools alike in type, size, and parental occupation mix, but different in student-teacher ratio (or per pupil expenditure, average teacher salary, number of books in library, whether in a specific experimental curriculum group, age of principal, whether multiple or single track, per cent of teachers who are men, courses offered, etc.)

After the completion of the multi-variate analyses it should be possible to apply more

sophisticated mathematical procedures of the operations research type to the data to maximize and minimize certain functions which will be represented by key goal criterion variables. It might be desired to maximize the average reading score of the graduates of schools. Once the unique relationships between the independent variables and the criterion have been established, as well as the interrelationships of the variables, then it should be possible to determine an optimum mix of school characteristics which would result in maximizing the criterion function. Many of the basic concepts of systems analysis and linear programming might be applicable here once the basic unique interrelationships become known. This could help lead to possible scientific decision making in regard to school problems as suggested by Joseph A. Kershaw and Roland N. McKean, Economics Division, The Rand Corporation, Santa Monica, California, in the May 1960 issue of School Management.

Because of the large size of the sample, the data in Project TALENT collected in the spring of 1960 cannot be fully scored, punched, edited and processed into a master tape before the spring of 1961. However, a random sample of four per cent of all students drawn from all schools was drawn and preliminary analyses were made of it by the 650 computer at the University of Pittsburgh. Because of time limitations the scores were put on tape with only a minimum amount of editing. The primary purpose of the partial analysis of the four per cent sample was to study the problems that will exist in editing fully the data as they are made into a master tape later on. Nevertheless, the results did seem consistent and meaningful, and certain aspects of them may be reportable.

The data were analyzed separately for ninth grade boys and girls and twelfth grade boys and girls. The means and standard deviations were satisfactory from the point of view of psychometric characteristics of the tests. The general level of intercorrelations was satisfyingly high where it was expected to be high and satisfyingly low where it was expected to be low. The pattern of intercorrelations, means, and standard deviations seems quite consistent with the results obtained on the experimental forms of the tests earlier administered to a sample of about 6,000 high school students in the spring of 1959.

Most tests showed considerable difference by grade, as was to be expected. The differences between boys and girls occurred in both directions and formed a meaningful pattern. The boys tended to be noticeably superior in total information and information in the areas of Social Studies, Mathematics, Physical Science, Biological Science, Aeronautics and Space, Electricity and Electronics, Mechanics, Motors, and Sports, while the girls were markedly higher in information about Home Economics and Cooking. The boys also tended to score more highly in the tests of Mechanical Reasoning, Visualization, Creativity, Abstract Reasoning, and Mathematics. On the other hand, the girls scored more highly in the various Memory tests, the Foreign Language Aptitude test, and the English test. The girls were very markedly higher in highly speeded tests, such as Arithmetic Computation, Table Reading, Clerical Checking, and Object Inspection. There was no difference between sexes in Reading Comprehension. On the Student Activities Inventory, the girls tended to score higher in Sociability, Social Sensitivity, Tidiness, and Culture.

For one sample, data are also available showing the distribution of total scores on the Information Test distributed by 36 categories of the occupation of the father. The average Information Test score varied quite widely from category to category of father's occupation indicating a very substantial correlation between father's occupation and over-all Information Test score. The children of such categories as clerical worker, unskilled laborer, farm worker and farmer tended to score on the low side; whereas children of officials, owners of businesses, and professional people tended to score quite high. The children of such categories as skilled worker, salesman, manager, and technician scored in between. An important finding here was that the children of the "don't know" category tended to score quite low along with laborers, farm worker, clerical worker, etc.

These data indicate that the father's occupational category may be very useful as a control variable in certain types of studies where the correlation between father's occupation and cognitive-type test performance needs to be partialed out or held constant. Since the correlation between father's occupation and cognitive test performance is quite high, this would mean that schools whose students are the children of such categories as officials, owners, and professionals will tend to score very highly on measures of aptitude and achievement even though they may vary greatly in their treatment of the students. On the other hand, if the schools have students who are predominantly the children of farm workers, unskilled laborers, etc., the aptitude and achievement performance of them is going to be very low even though the schools vary greatly in their treatment of the students.

By assigning weights to each father's occupation category, it will be possible to metricize this variable and use it as a continuous control variable. This may possibly be superior to using the conventional social-economic scale for similar purposes.

Each student in the study was required to write two essays and was given five minutes for the writing of each. The topics for the essays were: 1) "My Ideas about an Ideal Occupation" (describe what you would most like to do with your life), and 2) "What High School Means to Me." While these essay questions cannot be scored, at least immediately, for all of the students in the study, they will be held and scored later for samples of students of particular interest, such

as those who may become statesmen, scientists, or who were failures in life in spite of having many talents.

A pilot study is now under way to develop a scoring scheme for the scoring of the essays. A random sample of one hundred themes was drawn from the four per cent sample. These were classified and scored by two raters on a nine point scale. Inter-rater correlations were computed for subgroups and these range from .62 to .70. The Spearman-Brown formula would estimate a poteptial reliability between .85 and .90 for an average score from two themes each graded by two raters and averaged over-all. The themes were also scored for the misspelling that occurred. Here the girls were noticeably better spellers than the boys and were noticeably better in over-all quality of writing. This is very consistent with the results obtained on the Spelling and English Tests.

One interesting sidelight is that a considerable number of the girls wanted to be secretaries and a large proportion of them misspelled secretary. One of the most important objectives of the themes, however, is that of obtaining measures of the basic value systems of the students. A thematic scoring procedure has been developed for this in which each theme is rated on a go--no-go basis on a number of value factors. This approach shows promise of yielding a measure of such factors as altruism, scholarliness, hedonism, money orientation, religiosity, affiliation, security, achievement drive, and the like. The following examples of these themes show graphically the emergence of important values.

"Ideal Occupation" Themes

a. Boy aged 19. "The ideal peiece of work I would enjoy having would be something in line of social working among deliquents in New York or maybe a job as an artist. These jobs are both creative jobs and I think more suited for this type of work, for it gives me a chanch to express my self and maybe helping a job or two like myself."

b. Boy aged 17. "a millionire"

c. Boy aged 16. "I would like to go and work for my father and help him. But I would also like to be a Artist so I have not made up my mind as yet. And ideal occupation is one that can make a nuf money to pay the employer and to pay the owner for running it so he don't go bank rupt."

d. Boy, age unspecified. "My ideal of an ideal occupation is to do something important such as a research chemist. Also I would like to be a man of the world."

e. Boy aged 19. "I would like to go to college, play football, get married, have a family and a good job as a coach."

f. Boy aged 17. "I believe that the ideal occupation is a lawyer. A lawyer doesn't have to get paid by the week or month, but by the case. Some cases yield him more than the average persons total wages for a year. But of Course there are all different types of lawyers, such as criminal, Business--"

g. Boy aged 15. "I would like to go in the Air Force and let uncle Sam take care of me. I like the air force because it is one of the most active and talked about sirveses. I would like to be a bombidier if I am qualified. This job pays well."

h. Boy aged 15. "I'd like to go to college and be a dentist or a proffesional baseball"

i. Girl aged 17. "I would like to attend a beauty school and learn all the trades. Then I wouldlike to go out on the stripe or to Hollywood to work. Then I would like to get married."

"What High School Means to Me" Themes

a. Boy aged 19. "If I were to be dishonest, I would say I obtained much from high school; but alas, to my dismay, my honesty provides that I should tell the truth. It meant a good time!"

b. Boy aged 17. "High School means a great deal to me but the under-paid teachers a grouchy sometimes but it turns out in the long run & I like it preatty well but I will be glad when I can graduate and leave this community and everything in it. I learned a lot in school, more than I would have at home."

c. Boy aged 17. "High school is a means to prepare my self for college not only scolarly but socially. although I expect high school to prepare me, as much as possible, for the rigors of college classes, I also expect it to prepare me socially. A fine mine is very good, but even geniusus may become a nervous wreck."

d. Boy aged 17. "If I get hight school edcation it will make it easy to get a job if you have your hight deplomant to show that you went through school and to have a edcation you can get a good job in a plant being a boss or somthing."

PROJECTING ENROLMENT AT THE UNIVERSITY OF TORONTO

Gilbert de B. Robinson - University of Toronto

1.1 The report on expansion problems in the University of Toronto, June 5, 1956, was based on a projection of University enrolment made by Prof. B.A. Griffith. It is the purpose of the present Study to bring these statistics up to date and to refine them so far as the Faculty of Arts is concerned. The writer would express his thanks to Prof. J.H. Chung for his able assistance, and to Professor N. Keyfitz and Professor D.B. DeLury for their advice in deciding various matters of procedure. Thanks are also due to Prof. R.W.B. Jackson of the Ontario College of Education for providing statistics of school populations, and to the Registrar's Office for statistics relative to the University.

1.2 The Griffith Projections. In Professor Griffith's study five distinct projections were made each based on different assumptions. It is interesting to note how closely subsequent enrolment figures for the University of Toronto have followed the projection E(T) based on the total population of Ontario, assuming a linearly increasing percentage of the age group 18-21 attending university. The projection E'(T) based on a fixed percentage was unrealistic while E''(T) based on the York County population was calculated under the disadvantage that 1956 was a census year and the author could not predict the large local increase in population then taking place. If the proper statistics had been available, E''(T) would have been almost identical with E(T) in the period 1956-1960.

Prof. Chung has brought up to date the projection E(T), using the same method of calculation, and his results are as follows:

Projected Enrolment E(T) Based on Linearly Increasing Percentages
1960 - 1970

Academic Year	E(T)
1960-61	16,500
1961-62	17,800
1962-63	19,100
1963-64	20,500
1964-65	22,000
1965-66	23,500
1966-67	25,000
1967-68	26,700
1968-69	28,300
1969-70	30,100

1.3 A new projection for the Faculty of Arts. In projecting the enrolment of the University of Toronto on the basis of the population of Ontario in the age group 18-21, we are dealing with a large number of people who would not be eligible to darken its doors. If we restrict our attention to those who would be eligible, i.e. to students in Grade 13, our estimates will be dependent on educational policy and practice in an explicit manner and changes in these will have predictable effects. In order to arrive at such figures the total enrolment in each grade of all the schools in Ontario in 1958 was taken from the Report of the Minister of Education for that year. By comparison with the corresponding figures of the preceding years, 'passing' percentages were computed between grades. Assuming these percentages to remain constant, Prof. Jackson projected the school populations in Grade 13 up to 1965. Of course such an assumption implies that immigration into the area will also remain constant and that there will be no change in educational policy or practice. We have extended his projection to 1970, basing the results on the known school population for 1958.

If we insert the school populations in the various grades from the Minister's Report for 1959 we note some decreases which lead to a slight lowering of the 'passing' percentage. Such lowerings reflect the known decrease of immigration into the Province as reported in D.B.S. interim statistics. Though insignificant when applied to one year only, such decreases become important when compounded by projection to subsequent years. In view of this sensitivity of the method it seems worth considering past enrolments in Grades 8-13 as far back as 1947. During this whole post-war period we can see the gradual rise of these percentages (e.g. for Grade 8-9 from 78 percent in 1948 to 94 percent in 1958) and, assuming a uniform educational policy and practice, we may attribute this to the high birth rate in the post-war years and to immigration. In view of the noted drop in 1959, one might ask whether a prediction for the future should not return to the percentages of the past as immigration diminishes. On the other hand, according to the Gordon Commission Report a permanently higher level of immigration is to be expected and indeed encouraged. Also, the presence of greatly enlarged classes

at a lower grade will tend to prevent a decrease in the passing percentages. Balancing these considerations it seems worth while to introduce a slight decrease of at most 2 percent in the 1959 rate in any one grade over the years in question. These two figures might be taken as maximum and minimum figures for a grade in any one year.

In order to extend these calculations to the University of Toronto, the numbers of full-time Ontario students in each of the four years of the Faculty of Arts were obtained for 1957, 1958 and 1959 and the passing percentages calculated. It is of interest to note that the passing percentages from Grade 13 into the First Year are almost constant, certainly not increasing. The anomaly is easily resolved by comparing the numbers of students in Grade 13 with totals for the age groups in Ontario in two successive census years.

	Total Ontario population aged 18-21	Grade 13 population	Percent
1951	271,500	8,827	3.3
1956	288,143	11,487	4.1

Thus it appears that the increase in percentage of the age group 18-21 attending university arises in Grade 13, and the implication may be drawn that this is partly due to the demand of employers for this badge of attainment. If employers could be persuaded to accept a Grade 12 examination in lieu of Grade 13, the latter might return to its status of 20 years ago as primarily a preparation for the university.

Since it was impossible to obtain data for years prior to 1957 without undue labour, and on the assumption that University policy and procedure is less subject to pressures of immigration than that in the schools, the passing percentages have been taken as constant. In fact, the assumed percentages coincide for the second and third years with the overall percentages in the Faculty of Arts in 1959. On this basis, two figures have been obtained for the total population of the Faculty from Ontario based on those of Grade 13. An average of these two is given in column (1) of Table I.

If we denote an entry in column (1) by x, we may subdivide the total enrolment t in the Faculty as follows:

x = no. of full-time students from Ontario
y = no. of full-time students from elsewhere along with special students not in the General Course (Extension)
z = no. of part-time students in the General Course (Extension)

so that

$$x + y + z = t.$$

In column (3) of Table I we enter the known value of t for the years 1957, 1958, 1959 and calculate the percentage t is of x in column (2). From the regularity of increase of these percentages we infer their future behaviour and obtain a prediction for t from that for x.

It could be argued that the three years 1957-1959 provides a slim basis for this projection, but the Registrar was able to supply the information contained in Table II, which relates z to t in the period 1950-1959. Since we know that y is small and not likely to increase very rapidly, we may compare the conclusions of Tables I and II by plotting the percentages in column (5) of Table I along with a linear projection of those in column (3) of Table II. That the line from Table I is almost parallel to that from Table II, the difference in position and slope being due to y, provides a check on the validity of the projection in each case.

1.4 Comparison of projections. Comparing the results obtained in §1.3 with the projection E(T) of §1.2 we have:

	Faculty of Arts				Univ. of Toronto
Year	x	y	z	t	E(T)
1957	3589	475	1732	5,796	13,574
1958	3913	531	2153	6,597	14,402
1959	4128	662	2460	7,250	15,400
1960	4680	780	3140	8,600	16,500
1961	5320	840	3940	10,100	17,800
1962	6040	860	5000	11,900	19,100
1963	6600	900	5900	13,400	20,500
1964	7010	960	6830	14,800	22,000
1965	7630	1010	7960	16,600	23,500
1966	8360	1040	9400	18,800	25,000
1967	8970	1030	10800	20,800	26,700
1968	9360	1040	12000	22,400	28,300
1969	9550	950	13000	23,500	30,100

It will be observed that the difference E(T) - t remains approximately

constant. The reason for this is that the rapid growth in the Metropolitan Area is influencing t more than E(T), which is based on the province as a whole. Thus the projection for the Faculty of Arts may be high if this growth is not maintained or that for the University as a whole may be low if the Metropolitan Area continues to expand as at present. The other Faculties in the University, however, will likely continue to select and restrict much as they do now, so no great deviation from these predictions is to be expected.

1.5 Expansion at Ontario Universities

In this final section I record observations made on visits to the major Ontario universities and the implications of their expansion plans for us in Toronto. The chief question to which I sought an answer from each institution was: to what extent did they expect to expand, when and in what manner.

In Table III these predictions are used to estimate the number of full-time students which will have to be accommodated by Toronto, York or other new foundations. Columns (2) and (3) of Table III are taken from DBS statistics. Note that in enumerating full-time students at Ontario universities several small institutions are included which raise the total for 1959 from 26,448 (major universities + Toronto) to 29,200. The basis of calculating columns (5) and (8) is similar to that used in §1.2, but we concentrate attention here on full-time students.

Since $E(T) - z$ of §1.4 is approximately equal to $E_F(T)$ and $E'_F(T)$ up to 1965, it follows that the expansion planned at other Ontario universities will accommodate their share of the general increase up till then, but thereafter this will no longer be the case and Toronto, York or other new foundations will have to shoulder the burden of at least 10,000 full-time students over and above the number contemplated in §1.3 and §1.4. ($E'_F(T) - (E(T)-z) = 11,070$ in 1970).

If our facilities for higher education in Ontario are to provide for these full-time students, and also for part-time students, the number of whom is rapidly increasing, heroic measures will have to be taken.

TABLE I

Year	(1) Ont. enrolment in Fac. of Arts x	(2) 100t/x	(3) Total enrolment in Fac. of Arts t	(4) t-x = y+z	(5) 100(y+z)/t %
1957	3589	162	5796	2207	38
1958	3913	169	6597	2684	41
1959	4128	176	7250	3122	43
1960	4680	183	8600	3920	46
1961	5320	190	10100	4780	47
1962	6040	197	11900	5860	49
1963	6600	204	13400	6800	51
1964	7010	211	14800	7790	53
1965	7630	218	16600	8970	54
1966	8360	225	18800	10440	56
1967	8970	232	20800	11830	57
1968	9360	239	22400	13040	58
1969	9550	246	23500	13950	59

TABLE II

Year	(1) Total enrolment in Faculty of Arts t	(2) General Course (Extension) z	(3) 100z/t %
1950	5043	771	15
1951	5044	897	18
1952	4763	987	21
1953	3987	836	21
1954	4384	1060	24
1955	4819	1352	28
1956	5226	1523	29
1957	5796	1732	30
1958	6597	2153	33
1959	7250	2460	34

TABLE III

Full-Time Enrolment at University of Toronto in Relationship to other Ontario Universities

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Academic Year	P_o	$E(U_o,T)$	$\frac{E(U_o,T)}{P_o}$	$E_F(T)$	$\frac{E_F(T)}{P_o}$	$E(U_o)$	$E'_F(T)$
1953-54	273,600	19,563	.0715	10,240	.0374		
1954-55	274,000	20,470	.0747	10,660	.0389		
1955-56	273,100	21,489	.0787	10,840	.0397		
1956-57	288,143	22,869	.0794	11,350	.0394		
1957-58	296,800	25,000	.0842	11,840	.0399		
1958-59	306,700	26,964	.0879	12,250	.0399		
1959-60	322,700	29,200	.0905	12,940	.0401		
1960-61	338,000	31,670	.0937	13,520	.0400	18,560	13,110
1961-62	353,600	34,260	.0969	14,140	.0400	20,690	13,570
1962-63	369,400	36,940	.1000	14,780	.0400	22,640	14,300
1963-64	385,400	39,770	.1032	15,420	.0400	24,410	15,360
1964-65	401,600	42,730	.1064	16,060	.0400	26,000	16,730
1965-66	418,100	45,820	.1096	16,720	.0400	27,420	18,400
1966-67	434,800	49,050	.1128	17,390	.0400	28,650	20,400
1967-68	451,700	52,350	.1159	18,070	.0400	29,700	22,650
1968-69	468,800	55,830	.1191	18,750	.0400	30,580	25,250
1969-70	486,200	59,460	.1223	19,450	.0400	31,290	28,170

Note: The dotted line separates the experienced data from 1953-60 from the projected figures for 1960-70.

P_o = the population of Ontario in the age group 18 to 21 years.

$E(U_o,T)$ = the full-time enrolment at all Ontario universities (including Toronto). (Basis for projection from 1960-70 : Linearly increasing proportions in Col. 4).

$E_F(T)$ = the full-time enrolment at Toronto. (Basis for projection from 1960-70 : Approximately constant proportions in Col. 6).

$E(U_o)$ = the projected full-time enrolment at all Ontario universities excluding Toronto, obtained on the basis of estimates provided by seven of the largest universities.

$E'_F(T)$ = $E(U_o,T) - E(U_o)$.

VII

THE 1960 CENSUS OF POPULATION AND HOUSING

Chairman, Conrad Taeuber, Bureau of the Census

The Concept of "Usual Residence" in the Census of Population — Henry S. Shryock, Jr., Bureau of the Census

Discussion — Carl M. Frisen, California Department of Finance

Concepts, Coverage and Trends in the 1960 Housing Inventory — Wayne F. Daugherty, Bureau of the Census

Discussion — Leo Grebler, University of California, Los Angeles

Some Preliminary Results of the 1960 Census of Population -- Howard G. Brunsman and Charles P. Brinkman, Bureau of the Census

THE CONCEPT OF "USUAL RESIDENCE" IN THE CENSUS OF POPULATION

By: Henry S. Shryock, Jr., Bureau of the Census

Where people are counted in the United States Census of Population, that is, the area in whose statistics they are included, is important from many standpoints. Of greatest consequence is the effect on the apportionment of representatives to the Congress and to State legislatures. There are also other legal uses of population counts that are affected by the way in which persons enumerated are assigned to a geographic area. These uses include, for example, the distribution of certain State funds to municipalities on the basis of population and the requirements in State laws that a municipality attain a specified population before it is granted certain powers. Moreover, the description of an area's population and the levels of per caput rates that are computed for it will be influenced by the enumeration procedure followed; and these statistics, in turn, affect the use of the data in administration and in scientific research. As this audience probably well knows, the size of some types of areas determines whether certain kinds of statistics will be published, not only in Census reports but also in other statistical reports.

The choice of population concepts and of specific procedures and rules obviously affects the count for States, cities, and other geographic subdivisions. Even the national total is affected, although the effect is relatively slight in most circumstances. At the peak of activity in World War II, however a de facto count would have yielded about 9 million fewer persons than a de jure count.

The Population Census of the United States uses the concept of "usual residence," and there may be some question as to whether that is strictly equivalent to the de jure concept. The Act of 1790 providing for the first decennial census states "...Be it enacted, That every person whose usual place of abode shall be in any family on the aforesaid first Monday in August next, shall be returned as of such family; and the name of every person, who shall be an inhabitant of any district, but without a settled place of residence, shall be inserted in the column of the aforesaid schedule, which is allotted for the heads of families, in that division where he or she shall be on the first Monday in August next, and every person occasionally absent at the time of the enumeration, as belonging to that place in which he usually resides in the United States."1/ Although the rather archaic term "usual place of abode" still remains in the law, the synonym "usual place of residence" was introduced for greater clarity in the 1940 instructions to enumerators.

The Constitution does not mention the type of census to be taken, and Congress has left the definition of "usual place of abode" to those responsible for taking the census; nowadays, the Secretary of Commerce has the authority for making such decisions but customarily delegates this authority to the Director of the Bureau of the Census. This arrangement implicitly recognizes the need for technical competence and experience in this matter, as well as in such matters as the particular population questions to be asked and the definition of the terms used in these questions. Statistical surveys were in a primitive state at the beginning of the nineteenth century, and no written instructions were published for the census until 1820 and no printed schedules until 1830. As far as the term "usual place of abode" was concerned, however, it is likely that few questions of interpretation would arise in an essentially agricultural society with relatively little physical mobility except for the migration of entire families. The discussion of "usual place of abode" did not cover any points that were not already spelled out in the Act of 1790 until 1850, when the instructions to the marshals specified that,

> "By place of abode is meant the house or usual lodging place of a person. Anyone who is temporarily absent on a journey, or for other purposes, without taking up his place of residence elsewhere, and with the intention of returning again, is to be considered a member of the family which the assistant marshal is enumerating."
>
> "Students in colleges, academies, or schools, when absent from the families to which they belong, are to be enumerated only as members of the family in which they usually boarded and lodged on the 4th day of June."

The instructions to enumerators in the Census of 1880 contain the following: "The census law furnishes no definition of the phrase, 'usual place of abode'; and it is difficult, under the American system of a protracted enumeration, to afford administrative directions which will wholly obviate the danger that some persons will be reported in two places and others not reported at all. Much must be left to the judgment of the enumerator, who can, if he will take the pains, in the great majority of instances satisfy himself as to the propriety of including or not including doubtful cases in his enumeration of any given family. In the cases of boarders at hotels or students at schools or colleges, the enumerator can, by one or two well-directed inquiries, ascertain whether the person concerning whom the question may arise has, at the time, any other place of abode within another district at which he is likely to be reported. Seafaring men are to be reported at their land homes, no matter how long they may have been absent, if they are supposed to be still alive. Hence, sailors temporarily at a sailors' boarding or lodging house, if they acknowledge any other home within the United States, are not to be included in the family of the lodging or boarding house. Persons engaged in internal transportation, canal men, express

men, railroad men, etc., if they habitually return to their homes in the intervals of their occupations, will be reported as of their families, and not where they may be temporarily staying on the 1st of June, 1880."

The 1960 Enumerator's Reference Manual has twenty numbered sections on "How to Determine 'Usual residence'." It is remarkable that the three basic principles laid down in the Act of 1790 are still maintained, but the types of problems covered and their solution show some variations in the instructions for successive censuses. With a very few exceptions, the residence rules used in the numerous special censuses taken at the request and expense of local governments, as well as the rules used in sample surveys such as the Current Population Survey, are the same as those used in contemporary decennial censuses.

In modern times, a small proportion of people are listed in one area but are counted in the statistics for another area. This allocation is accomplished by a transfer of forms for persons or families and a check to see whether or not they had also been listed at the claimed place of usual residence. Such a procedure is possible only in a census that covers the entire Nation.

Although a number of statistically advanced countries have always taken their census on the de facto basis, i.e., counting persons where they slept on a specified "census night," there does not seem to have been any serious proposal to adopt this type of census for the United States.2/ Not only would this type of population be contrary to the Census law but the very notion of a de facto count appears to be contrary to American ways of thinking. The Bureau of the Census in order to secure supplemental information and to check on the completeness of coverage has frequently had occasion to ask for a listing of persons who were present at an address on a given day. No matter how this question is worded, it continues to be mystifying, laughable, or offensive to a surprisingly large part of the general public.

There is a more widespread demand for a closer approximation to a count of the de jure population, or for defining "usual residence" with a connotation of greater permanence. The Statistical Office of the United Nations has stated, however, that an allocation of persons to geographic units by their usual place of residence within the country does yield a de jure distribution of the population.3/ Neither "usual residence" nor "de jure" is otherwise defined. In fact, these concepts are more difficult to define than de facto or "population present"; and, to the extent that national census reports discuss the terms, they characteristically do so by giving a list of rules. In its 1960 Enumerator's Reference Manual, the U.S. Bureau of the Census introduces the subject with the following paragraph:

> "Usual place of residence is, ordinarily the place a person regards as his home. As a rule, it will be the place where he usually sleeps."

It may be seen that this general explanation does not by itself constitute an operational definition.

Possibly from a literal translation of de jure, or from other considerations, it is sometimes argued that people should be counted at their legal residence. One difficulty with this approach is that a person has no unique legal residence in this country. He may have certain rights or duties (voting, public assistance, admission to a public institution, jury duty, certain taxes, etc.) in one State or community and other rights or duties in another State or community. A citizen who has recently moved may not have some of these rights in any State.

The type of legal residence that is probably most commonly cited or implied is voting residence because it seems peculiarly appropriate to an enumeration that will be used to determine the number of representatives to be elected to legislative bodies. Although no accurate estimate exists, it can safely be said that millions of adults are not registered anywhere or otherwise entitled to vote.4/ It would be necessary to elicit information on voting residence from the respondent not only in terms of a State but also in terms of a specific address so that the person could be counted in a specific enumeration district and political area. Certainly, at least one extra question would have to be added to the schedule. To the large number of children and unregistered adults, we should have to add the number of adults for whom an adequate voting address could not be obtained. In what area would these be counted? The chief objection to using legal residence remains, however, that it is not equivalent to the basis prescribed in the law, namely, "usual place of abode."

Another alternative would be to let respondents determine their usual place of residence themselves without any guidance from the Bureau of the Census or its agents. In all likelihood such a procedure of handling a rather complex concept would not yield very uniform results.5/ Furthermore, many persons will report an area where they have not lived continuously for a long time but which represents a more or less permanent base of operations to which they can return occasionally. For the unmarried, this place is frequently the parental home; for the married man long absent from his wife and children, it may be the place where these are living. In some cases, it is where the person votes or owns property or even where he was reared. Thus, in many cases, there will be a tendency to report an area to which there is an emotional tie but in which the person does not now live or to which he is unlikely to return except for brief visits.

Some critics contend that the daytime population of an area would provide a more realistic and useful body of statistics than the nighttime population. Although statistics on this basis would undoubtedly be a useful supplement to the conventional ones, it is highly unlikely that they will replace them.

Before I describe how the decennial census actually handles some of the principal types of problem cases, we may find it useful to consider some of the criteria that should be used in making the "usual residence rules." I propose the following:

(1) the rules should be understandable by enumerators and respondents
(2) they should provide for a complete and unduplicated count
(3) they should produce useful statistics

With regard to the last point, the statistics should, for example, give a description of the population of an area that will meet the greatest number of needs in administration, business, and scientific research. They should articulate with statistics from other sources. It is desirable that the method of counting should keep together as a unit the people who actually live together in households, families, and similar groups. In some cases, of course, a rule made to meet one of these ideals will be antithetical to another ideal.

One class of the population for whom the proper basis of enumeration is often a matter of dispute is members of the armed forces. Although "seafaring men" were enumerated at their homes ashore according to some of the earliest available instructions, the first explicit mention of members of the armed forces (in the 1880 instructions) provides that, "All soldiers of the United States army, and civilian employees, and other residents at posts or on military reservations will be enumerated in the district in which they reside, equally with other elements of the population." In the Census of 1890, military reservations, navy yards, and vessels were covered by special resident enumerators engaged from among the staff; and members of the armed forces have continued to be counted where they are stationed. There is no record that they were ever allocated back to a place of preservice origin.

Although the quotation earlier from instructions for the Census of 1850 shows that college students were at one time counted where they lived while attending college, they were in many intervening censuses counted at their parental homes instead. The earlier method was restored in 1950.6/ The present method is deemed to be more consistent with the usual residence principle. Students who go away to college are away from home for a longer period, on the average, than military draftees of similar age; about one-fifth of them are wholly self-supporting and another two-fifths partially self-supporting 7/; and one-quarter of them are married. Moreover, most of those who go away to college will never again live regularly with their parents. Leaving the parental home is a gradual process for many young people, and going away to college is often the most significant event in this process. According to the Current Population Survey of October 1958, the address at which college students would be enumerated would be unchanged for at least 55 percent of them if there were a reversion to the 1940 rule. The rule chosen has an important effect upon the population size of college towns but relatively little effect upon other areas.

Where persons in institutions, hospitals, sanitaria, and similar group quarters are counted depends basically upon whether the length of stay is characteristically long or short.8/ This seems to be a procedure of long-standing. Unfortunately, for those who are interested in current population estimates, death rates, etc., the customary procedure in the field of vital statistics is different. Since 1949, the practice has been to allocate the deaths of most institutional inmates back to the area of residence at time of commitment, regardless of whether this is where the inmate would go upon his release.9/ It must be admitted, on the other hand, that an area with a large proportion of its resident population living in an institution will present a most peculiar statistical picture, not only as to the health of its inhabitants but also as to their demographic characteristics, educational attainment, income, and so on. In the 1960 Census, a novel effort has been made to give a better picture of the character of the ordinary population. This takes the form not of allocating the persons in group quarters back to their areas of previous residence but of publishing a table for the areas most affected on some of the characteristics of persons living in households, excluding persons in any type of group quarters.10/

Perhaps the most difficult decisions to be made concern persons with more than one residence. These include those who change their homes with the seasons. We think first of the idle rich who move from Newport to Palm Beach and of the hotel employees and other service workers who move because of them. But, in our increasingly affluent and mobile society with its long vacations and early retirement, the occupancy of more than one home during the year is increasingly a mass phenomenon. Of course, there have also long been classes of workers who changed their residences seasonally with their jobs—lumbermen, fishermen, agricultural laborers, canners, etc. In the cases that involve an annual cycle, the ordinary rule is to choose the residence where the person lives the greatest part of the year—although there is always the temptation to play safe and count him where you find him. Another class with dual residence consists of persons who work and live away from their homes and families, perhaps returning on weekends. In their case the need for meaningful family statistics clashes with the need to include persons in the area

where they are living most of the time. Obviously, we must take into account the length of stay, not only past but prospective. In the end, as mentioned in the instructions of 1880, there are some complex cases that must be left to the enumerator's judgment.

That persons with no fixed place of residence are to be counted where they are found by the enumerator is an instruction that has continued since the Act of 1790 itself. One of the reasons for a "T-night" and an "M-night" on specified dates is to facilitate a complete and unduplicated count of such transients.11/ It has been suggested that such persons be counted as a class apart in the United States population and that they should not be allocated to geographic areas. By law, however, they must be included in the apportionment population; and the tabulation of other population statistics on a different basis is probably a luxury that we cannot afford and hardly require.

A few remarks may be made concerning the fairly large number of Americans who have their usual residence--by Census definition--abroad and hence are not included in the count for the United States. To round out the description of our national population, data on basic demographic characteristics were first obtained in 1950 for the following classes living abroad: (1) members of the armed forces and their families; (2) civilian federal employees and their families; (3) crews of vessels in foreign ports or on the high seas.12/ In the 1960 Census, we are attempting to extend both the coverage and the number of characteristics. An inventory of our manpower resources is incomplete that leaves out the fairly numerous technicians on long-time foreign assignments. This information on Americans abroad is being obtained through the Department of Defense, the State Department, etc. Whether we have secured a sufficiently complete count of such groups as American employees of international agencies, employees of private firms, students at foreign universities, expatriates, etc., and their dependents remains to be seen. Statistics for those groups that are adequately covered will be published in a special report.

In conclusion, it seems likely that important statistical needs that are not now satisfied will be met not by changing the "usual residence rules" but rather by providing supplementary information in the tabulations. Some innovations along this line have been made in the 1960 Census. Additional data could be made available by special tabulations, or by introducing new questions in future surveys or censuses.

Footnotes

1/ Most of the historical information in this paper is taken from: Carroll D. Wright and William C. Hunt. The History and Growth of the United States Census, Washington, Government Printing Office, 1900.

2/ The United Nations has recommended that, for the country as a whole, the de facto population be obtained except that military and diplomatic personnel should be counted in the country to which they belong. This population is called the "international conventional total." See: United Nations. Handbook of Population Census Methods, Vol. III, "Demographic and Social Characteristics of the Population." (Studies in Methods, Series F, No. 5, Rev. 1), New York, 1959:57-58.

For a further comparison of American practices and international standards, see: Henry S. Shryock, Jr. "The concept of de facto and de jure population: The experience in censuses of the United States" in Proceedings of the World Population Conference, 1954, United Nations, New York, 1955, Vol. IV, pp. 877-888.

3/ Op. cit.

4/ According to an article in Washington Evening Star of 30 July 1960, Frank Thompson, Jr., chairman of a newly formed Registration of Voters Committee, estimated that from 30 to 40 million persons were otherwise eligible to vote but had not registered.

5/ Survey statisticians have learned that respondents' replies to such seemingly straightforward questions as "Are you employed?", and "Do you live on a farm?" do not yield satisfactory statistics.

6/ College students have continued to be counted at their parental homes in the Current Population Survey, largely because of the fact that a household remains in the panel for four consecutive months, some of which may fall in the academic vacation period. The feasibility of changing to the decennial census procedure is now under review.

7/ From unpublished data collected in the Current Population Survey of October 1959.

8/ An exception to this principle is that prisoners in jails are counted there despite the usual short sentences. There has been some doubt that their relatives would be inclined to report them.

9/ National Office of Vital Statistics. Vital Statistics of the United States: 1950, Volume I, Washington, Government Printing Office, 1954:36.

10/ This will be Table 31 in Series PC(1)B.

11/ T-night covers large transient hotels and motels whereas M-night is for missions, flophouses, jails, hobo jungles, and all-night movies. Possibly another special date for migratory agricultural workers would prove helpful.

12/ A wartime innovation in current population estimates that has persisted is a national series including the armed forces abroad--but not any other Americans abroad. That much-misunderstood toy, the "Census Clock", shows this population rather than the population resident in the United Census, which is the subject of the regular decennial statistics.

DISCUSSION

Carl M. Frisén, California Department of Finance

In his paper, Dr. Shryock has presented the historical pattern of definitions of residence as they have evolved in the decennial censuses, he has then discussed some alternative definitions, set down criteria for rules used in applying official definitions, and finally has commented on major problem cases experienced by the Census Bureau in applying current definitions.

At the start, two general comments are in order. First, the paper demonstrates both the strong feeling of continuity and comparability and also the nature of changes which have developed from using the same basic definition through eighteen decennial censuses. Second, any discussion of the use of a de facto definition would amount to an academic exercise, since there exists no basis for expecting a change insofar as the official definition is concerned. This does not preclude the possibility of its use as an auxiliary definition.

Our major interest then may be directed toward the problem of considering the present definition. Dr. Shryock has said that "there is a more widespread demand for a closer approximation to a count of the de jure population, or for defining "usual residence" with a connotation of greater permanence." I would raise the question as to whether the latter is possible when a population is becoming increasingly mobile.

The problem cases which have been discussed are perhaps relatively new only in the sense that they have now grown large and troublesome enough to raise questions as to the adequacy of census practices of classification. Military population is a case in point, for in California the military personnel stationed in the State numbered less than 50,000 in 1940, were more than 170,000 in 1950 and may be expected to exceed 300,000 in the 1960 Census count. Similarly, persons with more than one residence, in the categories mentioned by Shryock, have probably increased markedly in the past two decades.

Looking ahead, these problem cases may be expected to become larger and still more troublesome in the future. Our role in world affairs, for example, should lead to increasing numbers of Americans abroad, both in civilian and military capacities. Dr. Shryock's comment on the number of technical assistants and consultants abroad points up a relatively new area which challenges our population recording system. In quite a different field, the increasing number of house trailers is an indication of existing or potential mobility in sizeable segments of our population, at least in states which offer retired persons an opportunity to follow the weather, or call for large numbers of construction workers who follow the jobs.

We might ask: are these and other problem groups large enough to have an unfavorable effect on the quality of census data? We don't really know, because at present we have no measures of their aggregate size.

We assume that we can safely say that most United States residents on April 1, 1960 had a single, clearly recognized place of residence and were enumerated at that place. To determine how many posed complex problems of classification and definition by the enumerators would carry us into that area of which Shryock says: "the tabulation of other population statistics on a different basis is probably a luxury that we cannot afford and hardly require."

I would like to suggest, nevertheless, that consideration be given to an approach that may prove useful both in assessing the data we now receive and also in indicating the degree of mobility of the population. It involves additional coverage and also additional tabulation, hence it poses financial and other problems. Nevertheless, it may have some utility.

First, we need a measure of the effect of international movement. We now have monthly estimates of the United States military personnel abroad and, as Shryock has pointed out, the 1950 Census included military personnel and some civilian categories overseas. But civilian coverage was and is incomplete and it appears that a considerable undercount may occur. Granting the many uses to which census data are put, it would appear that those persons temporarily abroad with usual residence in the United States should be allocated to places of residence for the original purpose of legislative apportionment. Those Americans overseas for extended stays, whose residence may be defined as abroad by the Census Bureau, should be reflected in the national total if not in a residence allocation within the national boundaries.

Similarly, we need more precise measures of those persons temporarily in the United States whose usual residence is abroad. In California we know that there are many Mexican farm laborers in labor camps within the State whose usual residence is in Mexico. Large metropolitan centers inevitably have many visitors to our shores who are counted where found since they have no residence within the United States.

A combination of these data, dealing with Americans abroad either on short-term or long-term stays and with persons temporarily in this country but with usual residence abroad, would help in defining the extent of movement across our national boundaries.

Second, how many people reported in the Census are counted where found because they have no usual place of residence?

Third, how many persons were transients on Census day: not enumerated at their place of usual residence?

I am suggesting here for exploratory purposes a tabulation that would show for an area on the enumeration date:

1. The number of residents present at their

place of residence and enumerated at that place.

2. Residents absent, i.e., temporarily elsewhere in the United States or abroad.

3. Military personnel present. (The problem of assigning military personnel to place of pre-service residence is not considered in this particular analysis).

4. "Migrants"--persons present but with no usual place of residence.

5. "Transients"--persons present with a usual place of residence elsewhere.

In addition there is needed a tabulation of all Americans residing abroad, military and civilian.

Further refinements may be called for, especially in terms of populations residing in group quarters. But this approach would be a start toward analyzing problem cases and at the same time would provide a better perspective both of mobility patterns, and perhaps of the evolving meaning of the term "residence".

As Dr. Shryock has mentioned, one criterion for definition is whether it produces useful statistics. Every producer or supplier of population statistics knows the variety of needs at least insofar as they are defined by the requests received. We feel that as yet we have not been able to answer adequately a request for California population estimates "by sex, male, female, etc."

Many requests involve legitimate needs and pose questions as to alternative definitions. The "daytime" population was pointed out as one such, and others may be noted. In many states there are problems related to "peak" populations in beach and mountain recreational areas. Similarly, seasonal agricultural activities lead to considerable variations in population numbers.

But, while means are needed to obtain these types of data, no one of them offers a basis for improving on "usual place of residence" as the basic definition in the decennial census. Most of them imply a measure of change over time, usually of cyclical nature, rather than fitting the concept of the census as a cross-sectional measure.

The task ahead is to continue to function within the established framework, devoting particular attention to the questions and problems which we can see--and most of them appear related to our increasing mobility. They foreshadow further difficulties in improving upon and clarifying the definition of "usual place of residence".

CONCEPTS, COVERAGE, AND TRENDS IN THE 1960 HOUSING INVENTORY

Wayne F. Daugherty, Census Bureau

Two years ago at the American Statistical Association annual meeting I discussed some of the problems associated with developing basic concepts for the 1960 Housing Census and efforts to improve coverage of all private living accommodations. At that time we had not decided on the best way to improve our concepts, nor had we agreed upon procedures necessary to secure better coverage.

Since then we have completed both our plans and the Census. Now, while waiting for the results, we are speculating about the results of our plans. Recently we learned that the preliminary count for 1960 is 58.4 million housing units. This is an increase of more than 12 million units above the count enumerated in 1950—46,137,000 units for what is now 50 States. Nevertheless, the 1960 preliminary count is somewhat below the anticipated figure. We had expected to count about 60 million units based upon information from pretests and other census programs—the CPS household estimates and the 1956 National Housing Inventory. This expected level did include an increase attributable to improvement in our concept of the unit of enumeration which was designed to expand and improve coverage of all private living quarters. We had, at various times, estimated that such improvements might add 1/2 to 1-1/2 million housing units. Further, we anticipated that the general procedure of a two-stage census in 1960 would improve coverage over that obtained in 1950.

In a paper entitled "Living Quarters and Household Concepts in the 1960 Census" given at last year's ASA meeting, Mr. Hugh Rose described at length our attempts to improve the unit of enumeration. The concept was simplified and the criteria for identifying private living quarters was made more comprehensive.

In brief, we hoped to eliminate the complicated applications of a basically simple concept by removing most of the exceptions to the rule. We retained the fundamental concept and even the basic wording of the old dwelling unit. We went one step further, however, to make more explicit our definition of living quarters. We said "a housing unit is separate when occupants do not live and eat with any other household and when there is either (a) direct access through the outside or through a common hall, or (b) kitchen or cooking equipment for the exclusive use of the occupants." In essence, almost everything that was explicit or implicit in the 1950 dwelling unit definition has been retained. But our procedure was strengthened by the either-or aspect of the two fundamental criteria for determining a housing unit.

By far, I think, our greatest reform came in the application of these criteria. We conclusively demonstrated to ourselves, by the research conducted in Washington and pretests in other cities, that enumerators who had to record the presence of access and kitchen by a mark on the census schedules did a more consistent and complete job of enumeration than enumerators who applied the memorized definition of the housing unit without having to record the two criteria on the schedules.

Housing technicians will be keenly aware that this is the first housing census for which we have a record of access and kitchen in connection with the defined housing unit. Users of our data made us painfully aware that in 1950, when we asked the enumerator to apply the definition of the dwelling unit, we had absolutely no inkling of how effectively he applied these concepts, what was the access-kitchen status of dwelling units that were enumerated, nor did we know how many units were in structures enumerated as non-dwelling-unit quarters and thus completely omitted from the housing census. In succeeding years we did learn, both through the results of the post-enumeration survey of the 1950 Census and from housing technicians studying local conditions, that the application of the dwelling unit definition in the 1950 Census left much to be desired. Housing users should not have this complaint about the 1960 Census. Since we have eliminated most exceptions to the rules and strengthened the enumeration procedures, any lack of coverage of housing units will be due to other facets of census operations.

Some minor conceptual problems still remain. For example, we instructed the enumerator not to record vacant trailers as housing units. Future developments in the use of trailers may compel us to modify or abandon this restriction. Someday it may become necessary to include all such structures if evidence develops that vacant trailers or mobile homes are predominantly for occupancy as a home for either an individual or a family.

Another point which creates some concern, particularly in small areas such as census tracts and blocks, is our exclusion from the housing inventory of structures occupied by five or more lodgers whose quarters cannot be defined as individual housing units. The number five was chosen rather arbitrarily. Of more importance is the question whether these structures belong in the housing inventory of whether such use of space should remain in the same category as dormitories, nurses' homes, and barracks.

Aside from these problems, I believe that from both the conceptual and procedural standpoints, we have come a long way toward complete coverage of private living accommodations and their characteristics. Beyond this point we are at the mercy of the coverage that our general procedures succeed in obtaining during a census.

It is in conduct of the census itself that we must look for the reason that the 1960

Census housing unit count has fallen below our expectations. On the other hand, it is possible that we over-anticipated the impact of the change in definition on the number of housing units. On the surface, the lower count seems to negate our efforts to clarify concepts and improve techniques used by the enumerator. During the census, however, we did not find evidence of expected improvements in enumeration that we had anticipated. The problems of coverage in large cities have not been solved. Further, we are aware that, in the processing of census schedules, there are instances where the enumerators' work is so incomplete that our tabulating equipment cannot recognize the entries made for the purpose of including such cases in the housing unit count. This type of loss occurs particularly in connection with vacant houses. Evaluation studies now in progress should permit us to estimate the extent of these problems.

Despite the lower than anticipated preliminary count, there is clear evidence that significant changes in the Nation's housing supply have occurred. The most startling fact has already been widely disseminated through our earliest figures, i.e., a majority (15) of the 25 largest cities in the United States have declined in population while the housing supply of nearly all of these cities has increased. Although not necessarily implying an improvement in the housing situation in every case, a strong argument can be made that the housing situation at least has eased if not improved, for the great majority of these cities.

At the county level, the changes also form an interesting pattern, although at the State level these patterns are lost to some extent. Starting with the pattern of change at the State level you will note that each of the 50 States had some increase in the housing supply.

The average increase in the housing supply for all 50 States in the Nation is in excess of 26 percent. This net increase of more than 12 million units does not fully account for the activity within the housing supply. New construction of houses and apartements, conversion of single family homes, and shifts from nonresidential to residential use added considerably more than 12 million units. Some of these were offset by losses—demolition due to governmental programs as well as private action; abandonment of homes, particularly in rural areas; mergers of two family houses into single family houses; and shifts from residential to nonresidential uses.

The increase in the housing supply from State to State ranges from as little as 2 percent to as much as 99 percent. Significant losses in rural areas and small towns have been overcome in every State by growth in and around the large urban centers. As a result, the State pattern of growth blurs the trend of changes occurring in rural and isolated places, the extent of our suburban growth, and even the changes occurring in the large cities. But even among States the impact of housing changes varies greatly. Five States of the Nation, one of which is Alaska, have increased their housing supply by more than 50 percent. The housing inventory of eleven additional States increased more rapidly than the national average. The increases in these 16 States, which account for only a third of our housing supply, were sufficiently large to balance the slower growth of the remaining 34 States in the Nation. The large increases occurred in the seven Southwestern States, and in Connecticut, Delaware, Florida, Maryland, Michigan, New Jersey, Virginia, and two new States, Alaska and Hawaii. At the other extreme, the minimum net growth in our housing supply occurred in Arkansas and Mississippi.

Turning to changes in the housing supply at the county level, about one-fourth of the counties in our Nation experienced a net loss of housing during the decade. Downward changes in the net supply of housing in an area mean that housing is lost through demolition, abandonment, and merger to a greater extent than produced by new construction and conversion of existing structures. However, there are only a few counties in which the loss exceeds 10 percent for the decade. In contrast, we find that some counties more than doubled their housing supply.

For purposes of discussion, I have grouped the 3,000 counties in the Nation into three strata. First, are counties with a significant decrease in their housing supply; second, are those which remained relatively static (with less than 2.5 percent increase or decrease), and third, those which have had a significant increase in supply. On this basis, about 20 percent of all counties lost housing; 10 percent remained relatively static, and 70 percent of the counties gained. Of the 2,000-odd counties having an increase, the rate of growth exceeded the national average in only 600 counties.

A majority of the counties that have lost housing are located in about 20 States, and are rural in character. On the other end of the scale, those counties showing growth in excess of the national average are scattered throughout the Nation but closely associated with regions of high urban concentration. As you may note from glancing at the map, practically all of the counties on the seaboards have had tremendous growth regardless of whether located on the Atlantic, the Gulf, or the Pacific Coasts. Counties on the Northwestern coastline of Oregon and Washington tend to be an exception.

It is particularly interesting to note that this pattern of change for the decade 1950 to 1960 is not new. During the decade 1940-1950 many counties sustained similar losses in their housing supply while others remained relatively static. Counties which lost or remained static differed little from the pattern shown for the earlier decade. Examining the county-by-county pattern for the two decades, over three-fourths of the counties showed the same pattern of gain or loss, although the degree of loss or gain frequently differed. Nearly half of our counties had the same relative proportionate amount of gain or loss in both decades. Examining this pattern geographically, we find that the 900 counties which sustained a loss of housing in either or

both decades are concentrated in the agricultural plain States and in the Southern cotton belt States. Losses also are evident in counties in the Appalachian Mountains and in sparsely populated counties throughout the Western States. Although these trends are based on preliminary counts, the final reports are likely to show the trends I have outlined today.

With respect to the characteristics of the inventory, our statistical knowledge of 1960 Census results is very sketchy. In checking our processing operations, we have had occasion to examine a few tabulations for reasonableness, consistency, and completeness. Based on fragmentary results in 14 States, I thought you might enjoy predictions of what we think is happening to the character of our housing supply.

An examination of the few housing items which are repeated from census to census, and a look at some of the new items, support the impression that the character of our housing supply is anything but static. The evidence of change from 1950 to 1960 is striking and widespread. For example:

About one-fourth of our housing inventory will be reported as built during the 1950's.

The average size of households continues to decrease in almost all States.

Home ownership has reached its highest level.

The size of the housing unit, in terms of rooms, is larger except for the New England States; this reverses the decline in the 1940's.

Our supply of vacant housing has doubled in 10 years.

The decline in farm housing is so great that it is becoming unimportant statistically. This is a result of the far-reaching changes occurring in agriculture as well as change in the definition used to measure farm housing. (All vacant rural housing is classified as nonfarm housing.)

The number of dilapidated housing units is somewhat less than the 4.3 million we had in 1950 while the proportional decline is sharper due to the larger supply of housing.

There are other changes which imply higher standards of living as well as greater flexibility in the use of our supply.

Housing on the market for rent or sale is of much better quality than in 1950 in terms of availability of plumbing facilities.

Hot and cold running water is available to many more homes than in 1950.

Likewise, the existence of toilet and bathing facilities for exclusive use of the household has increased markedly.

The number of homes without piped water supply, toilet or bathing facilities is about half the 1950 figures.

The number of homes with more than one person per room has declined substantially.

PERCENTAGE CHANGES IN HOUSING SUPPLY OF COUNTIES IN THE UNITED STATES: 1940 TO 1950 AND 1950 TO 1960

Percentage change: 1940 to 1950	Number of counties	Percentage change: 1950 to 1960							
		Decrease			Increase				
		9.5 or more	2.5 to 9.4	0.1 to 2.4	0.1 to 2.4	2.5 to 9.4	9.5 to 29.4	29.5 to 99.4	99.5 or more
Total.....	3,073	210	381	177	173	586	1,014	483	49
Decrease:									
9.5 or more.....	173	56	44	12	10	16	21	13	1
2.5 to 9.4.....	309	69	116	33	24	35	29	1	2
0.1 to 2.4.....	158	15	48	25	23	22	14	11	...
Increase:									
0.1 to 2.4.....	158	13	35	21	22	45	18	4	...
2.5 to 9.4.....	535	32	72	47	46	179	136	22	1
9.5 to 29.4.....	1,192	19	61	33	41	252	636	142	8
29.5 to 99.4.....	496	5	5	4	6	36	150	265	25
99.5 or more.....	52	1	...	2	1	1	10	25	12

NOTE: Based on preliminary counts for 1960.

Among the new items included in the census, there is one considered fundamental to evaluating the quality of our housing stock. Having learned from the 1950 Census that dilapidated housing is a very small proportion of our supply, we have added another gradation to the "Condition" concept so that not dilapidated housing is now reported as either sound or deteriorating. Even with this additional gradation we can expect nearly three-fourths of our inventory to be classified as sound. For those particularly concerned with inadequate housing we are providing another measure—the number of units sharing facilities. At the moment, it seems unlikely that more than five percent of households are sharing bathing and toilet facilities, kitchens or access to unit.

All in all, these last 10 years have been ones in which extensive changes have occurred in the field of housing. Although I have been able only to discuss changes in the most sketchy form, the evidence indicates that the housing used by the American people is better—both in quantity and quality. When the final reports are analyzed, it is hoped that they will aid in identifying weaknesses in our housing supply and in the markets for housing, essential facilities and equipment, thereby furnishing the basis for action for better housing for all.

DISCUSSION

Leo Grebler, University of California, Los Angeles

As usual, we are indebted to Wayne Daugherty for a highly informative and candid paper. All Census users who have had a glimpse of the conceptual and operational problems in the Census of Housing, and of the budgetary and other pressures on the Bureau, will share a sense of gratitude for the high quality of work performed by its professional staff. The introduction of the "components of change" concept in the National Housing Inventory of 1956 and its repeated application in the 1960 Census represent a major innovation. This innovation has already changed markedly our notions of the recent level of new housing construction and the methods of collecting current data on housing starts.

I was keenly interested in Wayne Daugherty's preview of Census results, as I am sure you were. The evidence of great improvements in the quantity and quality of the housing inventory will make it exceedingly difficult to perpetuate the notion that we are still faced with a general housing "crisis." On the other hand, it will become more and more important to identify precisely the population groups and areas with remaining deficiencies and to concentrate on remedial policies for these. The 1960 Census should make a real contribution to such an effort.

It would be premature, however, to comment on our current first glimpse of the Census findings. Instead, I shall address myself to a few conceptual and related points. In doing so, I have an eye on the future rather than the 1960 Census. I am using Wayne's excellent paper as a starting point rather than as a text. And if there are any challenges in what I have to say, they are offered to Census users, the Congress, and perhaps the Bureau of the Budget, as well as to Census officials.

Concerning the definition of a housing unit, it is all to the good that we now have a count of units that eliminates or reduces subjective judgments by enumerators. I find it personally refreshing to see the Census emphasize the criterion of "separatedness" in the definition of housing units when there is so much fashionable talk of "togetherness." It is all to the good that Census technicians are worrying about such matters as the inclusion of vacant trailers or of structures occupied by five or more lodgers when their quarters cannot be defined as individual housing units. But there are still more important though not wholly new questions with which we must wrestle.

For example, is the housing unit, however defined, a satisfactory unit of measurement? To paraphrase Gertrude Stein, is a housing unit a housing unit a housing unit? To be sure, the Census furnishes certain quality characteristics including the number of rooms per dwelling unit which, we just learned, increased in the past decade. But is this adequate? In some foreign countries, census reports include information on square footage. Here, analysts of the 1950 Census have had considerable trouble in arriving even at the total number of rooms in the housing inventory, segregated for renters, owners, and the various income, rent, and value groups. To my knowledge, the 1960 Census does little to relieve them of these pains.

Another problem is a firmer determination of the number and characteristics of seasonal units owned by households in addition to the year-round unit they occupy. Currently, the Census gives us only a general and somewhat blurred outline of this phenomenon. The apparent growth of week-end and vacation cottages is a startling development with many marketing, planning, and social implications. These are difficult to assess without better data on the units, their utilization, the income and social characteristics of the people who own them, and the characteristics of their year-round housing.

These comments suggest a more general observation. Are we too hidebound by the concepts and questions developed in past housing censuses, and are we too easily satisfied with relatively small increments to knowledge and minor improvements? Would a bolder approach be more fruitful even at the price of reduced comparability? To borrow a phrase from one of the present contenders for the Presidency, are there "new frontiers"?

I have no grand design to offer, but let me sketch some of the directions a new approach may take. One would be a decennial census of much more limited scope, supplemented by more frequent intercensal surveys with a minimum of standard items and additional varying questions. Another would be experimentation in selected localities with items that will advance knowledge of land use and community structure. Much of the steam behind the Census of Housing has come from the Federal housing programs. These have more and more been augmented by Federal aids for community development. This trend, and the severity of problems posed by future urban growth, would seem to justify some reallocation of resources in favor of data needed for metropolitan planning purposes, at least on a selective and experimental basis. An example that comes to mind is the journey to work. The 1960 Census of Population will for the first time furnish data on this item, but the areas of place of work are defined so broadly that the information will have very limited usefulness. One wonders whether more intensive Census surveys in a few areas would be more profitable in the long run. The 1960 solution may be a case of compromise that pleases hardly anyone.

A second set of comments refers to the tabulation and publication program. The block statistics in the form of published tabulations are undoubtedly useful, but one wonders whether it would not be sufficient to make them available

upon request and at cost, and use the savings for other things. By introducing the price mechanism, the Census would also be in a better position to assess the strength of demand for block statistics.

As I said earlier, the Census results indicate that we should increasingly concentrate on specific policies to remove remaining housing deficiencies. In this connection, one hopes that the tabulation process in 1960 will provide more adequate information on two important aspects of housing policies and programs: housing for low-income families and for the elderly -- partially overlapping groups, of course. Each of these groups, however defined, is composed of widely heterogeneous elements, and there is great question whether one type of housing program could or should do anything for all of them. But it was difficult from available tabulations in 1950 to match housing conditions with the income and social characteristics of the various subgroups. In the case of the elderly, analysts were also frustrated by the limitation of many Census data to the households headed by an elderly person. It would obviously be revealing to have more complete data on the households that include an elderly person, whether he heads it or not.

Finally, there is reason for concern over Wayne Daugherty's reference to the quality of enumeration. We have been spending a great deal of money on better and faster machines but it seems we have done less to improve or even maintain the quality of enumerators. This is a good reminder of the fact that the machines are still largely at the mercy of what is being put into them. I do not know of any solution to this problem except more money, better training, greater selectivity, and perhaps a shift of the Census period to a time when our large student population is more readily available for temporary work. We would not wish to recommend that the monetary and fiscal authorities engineer a major slump at the time of each decennial census so that the Bureau may pick up more qualified personnel, and some of us would even question their ability to administer the shock at the proper time and with the proper doses.

SOME PRELIMINARY RESULTS OF THE 1960 CENSUS OF POPULATION*

Howard G. Brunsman
and
Charles P. Brinkman
Bureau of the Census

The 18th Decennial Census of the United States revealed a total of 179 million persons in the 50 States and the District of Columbia on April 1, 1960. This preliminary figure represents an increase of 27.7 million over the count of 151.3 on April 1, 1950. Numerically, this is the largest increase ever reported in the United States. This increase of more than 27 million exceeds the total population of the United States at any census prior to that of 1860. The relative increase of 18 percent between 1950 and 1960 exceeds the increase in all of the decades since 1910 but is less than the relative increase in any decade before 1910. It is substantially less than the decade increase of more than 30 percent shown in all periods before 1860.

The 1960 preliminary figure of 179 million is based on field counts of the population, and includes in addition an estimated 1.1 million persons who were enumerated by various special procedures. The subsequent discussions of changes in the internal distribution are based on the field count figure (177,874,042). It is unlikely that the inclusion of these additional groups will alter any of the conclusions drawn in this discussion.

The increase in population between 1950 and 1960 represents an increase through immigration of nearly 3 million, and a natural increase of approximately 25 million--the difference between about 41 million births and 16 million deaths. Although tabulations of the characteristics of the population have not yet been made, it seems probable they will show that our population has grown older--and younger. The size of the group entering the 65 and over age group during this decade is considerably larger than the corresponding groups for previous decades. The continued decline in mortality also contributes to the increase in the upper end of the age distribution. At the same time, the unrelenting fertility of the 50's has served to produce a population under 10 years of record size. Thus, it is expected that both the group 65 and over and those under 10 will show more rapid increase than the population as a whole. On the other hand, a less rapid increase--or even a decline--is to be expected in the age group 20 to 34 as a reflection of the relatively small number of births in the period 1925 to 1940.

Within the country the rates of most rapid increase tended to be concentrated in the Southwest. Thus, California, Nevada, Utah, Arizona, Colorado, New Mexico, Texas, and Louisiana all had rates of increase in excess of 20 percent. Florida, with a 76 percent increase, however, outdistanced California with a 47 percent increase. There were also rates of increase in excess of 20 percent in Michigan and Ohio, and in Delaware and the States which might be described as suburban to either Washington or New York City--Maryland, New Jersey, and Connecticut. The 2 new States--Alaska and Hawaii--showed substantial rates of increase. In all, there were 17 States with rates of increase of 20 percent or more.

In terms of absolute amount of increase however, the roster of States was somewhat different. The 17 States that ranked highest in absolute increase, to be sure, included--California, Connecticut, Florida, Ohio, Michigan, Texas, New Jersey, Maryland, Louisiana, and Arizona, in which the rate of increase was high, but the highest ranking States also included large States such as New York, Illinois, Pennsylvania, Indiana, North Carolina, Virginia, and Wisconsin, in which the rate of increase was somewhat lower. In absolute increase, for example, New York State, with an increase of 1.8 million, was outranked only by California and Florida. Taken together, the 17 States with the highest ranking absolute increases accounted for more than 80 percent of the total increase during the decade. The shifting pattern of growth among the States was not sufficiently great to change the ranking by population of the larger States. Thus, the 9 highest ranking States in 1960--New York, California, Pennsylvania, Illinois, Ohio, Texas, Michigan, New Jersey, and Massachusetts, were also highest ranking in the same order in 1950. Florida however jumped from the 20th State in order of size in 1950 to the 10th State in order of size in 1960, and there were some additional minor changes in ranking among other States.

There were 3 States--Mississippi, Arkansas, and West Virginia--that lost population during the decade. Two of these States--Arkansas and Mississippi--also lost population during the 1940-1950 decade. West Virginia, on the other hand, showed a slight increase between 1940 and 1950. The District of Columbia also lost population during the 1950-1960 decade, but in this context it is more meaningful to consider the District as a city rather than a State.

Just as the population growth in the decade 1950-1960 was concentrated in certain specific parts of the country, it was also concentrated in certain types of areas. More than 80 percent of the 26.5 million increase between 1950 and 1960 occurred in standard metropolitan statistical areas; that is, in cities of 50,000 or more and the suburban areas surrounding them. This concentration of population increase in metropolitan areas was simply a continuation of the pattern observed in the decade 1940-1950 when nearly 80 percent of the increase occurred in the same type of areas.

* The authors are grateful to Henry D. Sheldon for assistance in the preparation of this paper.

In terms of growth rates, there was again a similarity. Between 1950 and 1960 the population of standard metropolitan statistical areas increased by about 25 percent and between 1940 and 1950, about 22 percent. The percentage increase for the remainder of the country outside metropolitan areas was 8 percent for the decade 1950-1960; it was 6 percent for the decade 1940-1950. The difference in these levels of increase reflect the difference in the percentage gains in the country as a whole--18 percent between 1950 and 1960, and 14.5 percent between 1940 and 1950.

Within metropolitan areas, however, there were appreciable differences. In the decade 1950-1960, the population of central cities increased by only 8 percent, whereas in the previous decade the corresponding increase was nearly 14 percent. In the suburban ring, however, the 1950-1960 increase was nearly 50 percent as compared with 35 percent in the preceding decade. Considered as a proportion of the countrywide increase, the suburban ring accounted for about two-thirds of the total 1950-1960 increase, but slightly less than one-half of the 1940-1950 increase.

The metropolitan-nonmetropolitan pattern of increase varied considerably among the various regions. The population in metropolitan and nonmetropolitan areas of the Northeast increased at about the same rate (12 and 13.3 respectively), central cities decreased by about 4 percent, and the suburban ring increased by slightly more than one-third. In the North Central States, the rate of increase in metropolitan areas was about 4 times that of the nonmetropolitan areas (23 vs. 6 percent), central cities showed a modest increase of about 3 percent, and the suburban ring increased by something more than one-half.

In the South, the population of standard metropolitan statistical areas increased at a rate 8 times greater than the population living outside such areas (33 vs. 4 percent), that of central cities increased by nearly one quarter, and that of the suburban ring by about one-half. In the West, the population of metropolitan areas increased at more than twice the rate of the population of nonmetropolitan areas (46 vs. 21 percent). The rate of increase for central cities was about 28 percent and the ring about 65 percent.

These findings suggest a sort of evolutionary hypothesis with respect to the pattern of population growth in this country. In the mature Northeast, central cities have achieved a high density, are generally speaking unable to extend their limits, and thus tend to lose population, and the exchange of population between metropolitan and nonmetropolitan areas has reached a state of dynamic equilibrium. In the North Central Region, the decline in the population of central cities is foreshadowed by the low rate of increase, but there is still evidence of the movement of population from nonmetropolitan areas to metropolitan areas, the net gain appearing in the ring. In the South, there is the same net movement into metropolitan areas, but this net increase appears to be more evenly distributed between central city and ring--evidently the limits of central cities were less rigidly fixed or initially the incorporated area of central cities was relatively large in relation to population.

The West is unique in having an overall rate of increase more than twice as great as that of any other region. Like the North Central Region and the South, but to a lesser degree, there is evidence of movement from nonmetropolitan to metropolitan areas. The growth rate of the ring is somewhat greater in relation to the growth rate of central cities than in the South, but less than that in the North Central Region.

The results of the 1960 Census which have attracted the most attention, not all of it favorable, have been the losses of population in a considerable number of our large cities. Of the 225 central cities of standard metropolitan statistical areas, 72 lost population between 1950 and 1960. Of the 5 cities of a million or more--New York, Chicago, Los Angeles, Philadelphia, and Detroit--only one, Los Angeles, showed a gain; the others had appreciable losses.

In one sense, the alarm and dismay with which the losses have been viewed is not justified. To be sure, New York City had about a 3-percent decline in population, but the New York Standard Metropolitan Statistical Area increased by about 10 percent. The corresponding figures for Chicago were -3 percent and 19 percent, respectively; for Philadelphia, -5 percent and 17 percent; and for Detroit, -10 percent and 25 percent. In short, although the legally defined cities declined in population, the complete areas increased in population.

Thus, if we consider the natural concentration of population independent of artificial and arbitrary boundaries, then these concentrations have grown. It is clear that an urban aggregate which is growing, grows more rapidly at its periphery than in its central parts, and as it spreads outward it sooner or later overflows arbitrary boundaries from the past. At the same time, the growth of the total area increases the need for nonresidential use of land in the center of the city to accommodate increased business activity, increased traffic, and the like. Thus, if city limits are fixed, there comes a time when a decline in the population of the central city may be taken as evidence of the economic well being of the whole area.

Between 1950 and 1960, the central cities which lost population were in the main mature cities, large cities, and cities which in 1950 had a high density. Generally speaking, they were located in the Northeast, although there were notable exceptions--Baltimore and Washington in the South and San Francisco in the West. The one characteristic which they had in common was a political and legal situation which made it nearly impossible to expand their boundaries.

The importance of annexation in determining whether or not a central city gained or lost population is demonstrated by the fact that the central cities of SMSA's in which there had been any annexation to central cities during the decade increased by 19 percent, whereas in those areas in which there had been no annexations decreased by 4 percent. This classification, based on the sheer presence or absence of annexations, is relatively crude since it does not take into consideration the amount of territory and population annexed. Thus, in the Northeast, SMSA's with annexations showed a greater decline than those without annexations. In the San Francisco-Oakland SMSA there was a small annexation to Oakland which put the SMSA's in the "with annexation" column even though both cities are essentially hemmed in and lost population.

The effects of annexation are perhaps most clearly demonstrated in the case of those SMSA's which showed modest to sharp increases for the whole area, but losses in the suburban ring. Thus, for example, Evansville with only a gain of 3 percent for the whole area was able to produce a gain of about 9 percent in the central city and Tucson, Arizona, by a series of annexations during the decade, a gain of 360 percent in the central city, and a loss of 46 percent in the ring.

It is clear then that for a given rate of growth in the entire SMSA, the rate of growth of the central city will depend in large part on whether or not the municipal corporations are able to recapture the "flight to the suburbs" by annexation or on whether or not by "big thinking" at some earlier period they extended the city limits well beyond the area of heavy settlement. In this situation, the cities of the older and more settled parts of the country are at a disadvantage in that they are frequently ringed around by places incorporated in their own right, further expansion would involve annexations into adjoining counties, and generally speaking annexation is more difficult. In Ohio, for example, an annexation can be made only on the petition of the inhabitants of the area to be annexed, whereas in Texas a municipality can annex adjoining unincorporated territory practically at will. Thus, in terms of sheer population growth, the losses experienced by many American cities in the decade 1950-1960 are illusory, but in terms of the problems of municipal administration they are, in many instances, very real.

Although the population growth during the past decade was highly concentrated in metropolitan areas, the population outside such areas increased by about 4.9 million, or 8 percent. This increase was by no means evenly distributed. For the country as a whole, counties that in 1950 contained urban population increased by 10 percent. On the other hand, those with no urban population in 1950 suffered a population loss of about 4 percent.

This same relationship was shown for the North Central States and the South, with appreciable gains occurring in the counties with urban population and losses occurring in counties that were entirely rural in 1950. In the Northeast and in the West, both groups of counties had increases, but the increases for the counties with urban population were greater.

Throughout the country, for counties with urban population in 1950, the percentage gain increased with size of county from counties of 10,000 to 20,000 to counties of 100,000 or more. For the class of counties containing urban population but with less than 10,000 inhabitants in 1950, the rate was higher than that of the class containing 10,000 to 20,000 inhabitants. The same pattern was observed in each of the regions except the West where there was a consistent pattern of increasing percentages from the smallest to the largest classes.

The heavy population losses were concentrated in the counties that were completely rural in 1950, primarily in those of less than 20,000. The very small number of larger rural counties showed an increase at the national level, but a somewhat erratic pattern of change among the regions.

In summary, the preliminary results of the 1960 Census showed a continuation of the concentration of population growth in metropolitan areas. They also indicate that outside metropolitan areas there has been a continuation of the relatively modest rate of growth characteristic of the decade 1940-1950 and the concentration of this growth in the larger counties. The most notable, but not unexpected, result of the 1960 Census is the decrease in the population of a considerable number of our larger cities. About one-third of all cities of 100,000 or more had decreases during the decade and of the 10 largest cities of 1950, 9 lost population.

VIII

POPULATION RESEARCH

Chairman, Dudley Kirk, The Population Council, Inc.

Social and Economic Mortality Differentials in the United States, 1960: Outline of a Research Project — Philip M. Hauser and Evelyn M. Kitagawa, University of Chicago

SOCIAL AND ECONOMIC MORTALITY DIFFERENTIALS IN THE UNITED STATES, 1960: OUTLINE OF A RESEARCH PROJECT

Philip M. Hauser and Evelyn M. Kitagawa
University of Chicago

The declines in death rates and increases in longevity have, without question, been among the most significant and dramatic consequences of man's activity as a culture-building animal. A high death rate is in the contemporary world one of the best indicators of a low level of economic development. In the economically advanced areas, death rates have been pared by approximately three-fourths in the modern era; and despite their older populations, economically advanced nations tend to have general death rates approximately one-fourth the level of death rates in the economically underdeveloped areas of the world.

The great declines in mortality have been the product of a combination of forces of increased productivity, political order, environmental sanitation, personal hygiene and modern medicine. Yet, despite the great advances that have been achieved, large mortality differentials are still to be found even in the most advanced nations. To obtain a comprehensive picture of these differences among the various areas and subgroupings of the population in the United States in relation to personal, social, economic, and housing characteristics is the objective of the research now under way which is described in this paper.

The findings of this study, in revealing the populations within the United States which have achieved the lowest and highest death rates, respectively, will point to the areas and groups in which substantial reductions in mortality can yet be achieved and, also, will indicate the levels to which death rates can be further reduced. The findings of this study thus can serve as guides to public and private health programs in continuing efforts to increase longevity.

Many studies have been conducted on the death rates of the population of the United States and of other nations.[1/] But nationwide information on mortality differentials by social, economic and housing characteristics has been restricted by reason of the limited information provided by the official record of mortality. That is, the death certificate as the official local legal record of death, a copy of which is made available to the National Office of Vital Statistics for statistical purposes, is limited to a few descriptive items about the person and the cause of death. The only item on the death certificate which may be considered a measure of the social-economic status of the deceased person is occupation. But it has not been possible to do much in the analysis of mortality by occupation because of the great discrepancy between occupational return on the death certificate, which provides information for the numerator of the death rate, and that reported in the population census from which the denominator of the death rate is drawn. In fact, even in respect to the other items on the death certificate such as marital status and color or race, the death rate is subject to error because of the lack of correspondence in the information reported on the death certificate and census schedule, respectively. Because of these discrepancies in the reporting of occupation on death and census records, the preliminary findings of a 1950 study of mortality of white males 20-64 years of age are restricted to only five occupational levels. Even with this limitation, however, there is an inverse relation between occupational level and mortality incidence, the death rate of laborers being twice as high as that of professional workers.[2/] A good measure of the error in 1950 occupational death rates calculated by this method will soon be available.[3/]

Important information has been obtained on mortality differentials by social-economic status through an indirect form of analysis possible in cities and metropolitan areas where census tract data are available. In census tracted areas, it is possible to measure differentials in mortality, utilizing the social and economic characteristics of the census tract in which the decedents resided. A study now being completed in Chicago, for example, estimates 1950 mortality rates by social-economic status for the City of Chicago by allocating 1950 deaths to the 935 census tracts into which the city is divided, and using the regular census tabulations of population by census tracts as the base.[4/] In this study, age-adjusted death rates are calculated for several "income levels," and also for several "quality of housing" levels by assigning death and population data for each tract as a unit to one of the levels on the basis of the median family income, or the proportion of substandard dwellings, in the tract. This method although feasible for census tracted areas, cannot be used to analyze differential mortality for non-tracted areas and, therefore, the entire country. Moreover, the method is subject to the error of ecological

---

[1/] Dublin, L.I., A.J. Lotka and M. Spiegelman, Length of Life (New York: Ronald Press, 1949). Spiegelman, Mortimer, Significant Mortality and Morbidity Trends in the United States Since 1900 (Revised) (Philadelphia: American College of Life Underwriters, 1960).

[2/] Moriyama, I.M. and L. Guralnick, "Occupational and Social Class Differences in Mortality," in Milbank Memorial Fund, Trends and Differentials in Mortality (New York, 1956).

[3/] The Scripps Foundation for Research in Population Problems has undertaken a study of the correspondence of reporting of occupation on death certificates and census schedules for a sample of persons who died in the four-month period following the 1950 census.

[4/] Hauser, P.M. and E.M. Kitagawa, "Differential Mortality in Relation to Quality of Housing" (Paper prepared for Western Branch, American Public Health Association Program, June 2-5, 1959, San Francisco).

correlation, in necessarily involving the assumption that a summary index of social-economic status for an entire census tract applies to individual residents of the tract.

The study which is described below is designed to provide information on differential mortality despite the difficulties in doing so arising from the way in which official death data and population census materials are made available.

The Design of the Study

The data for the study will be information on death certificates for approximately 400,000 deaths that occurred from May through August, 1960; and the returns for these decedents and the entire population as obtained in the 18th Decennial Census of the United States as of April 1, 1960, in the Censuses of Population and Housing. The total number of deaths which occurred in the nation during the four-month period was about 500,000. All deaths of white persons under 65 years of age, and deaths of all nonwhite persons will be included in the study. A random selection of half the deaths for white persons 65 years of age and older will be excluded from the study, since such a large proportion of all decedents are 65 or older. This will provide, in all, a universe of approximately 400,000 deaths, for each of which there will be available a copy of the death certificate.

The serious problem posed by the limitation of personal, social and economic information about the decedent on the death certificate, and the lack of correspondence between the reporting of particular characteristics on the death certificate and on the population schedule for the same person, will be resolved by matching each of the 400,000 death certificates to the 1960 census schedules. Such a matching of the death certificates to the census returns, moreover, will provide not only all of the census information for the decedent but, also, for the decedent's family, household and dwelling unit. The matching procedure assures, automatically, comparable information for the numerator and the denominator of the death rate.

The matching operation will involve the following steps:

Step 1. Xerox prints of the 400,000 deaths included in the study will be made from the microfilm copies of death certificates forwarded by the states to the National Office of Vital Statistics. These Xerox prints will be sent to the Census Bureau. A duplicate set of the National Office of Vital Statistics punched mortality cards for these deaths will also be forwarded to the Census Bureau.

Step 2. The Census Bureau will determine the 1960 E.D. (enumeration district) number of the place of residence of each deceased person, using the address shown in the Xerox print of his death certificate.

Step 3. The Census Bureau will search the 1960 census schedules for the schedule containing the name of each deceased person. The E.D. number assigned in Step 2 will direct the searcher to a file containing all of the census schedules for that E.D.

Step 4. The Census Bureau will collate the information obtained from the death certificates and the census schedules, and will summarize this information on computer tape, which will be used to compile tabulations of matched deaths by characteristics of the deceased persons as reported on their April 1960 census schedules.

The major limitation in this procedure arises from the incompleteness of the matching which can be achieved. Past experience indicates that about 20 percent of the deaths will not be matched by reason of incomplete addresses on the death certificate, the mobility of the population resulting in the changing of address between the census date and date of death, incompleteness of census enumeration, illegibility of returns on death certificates and census schedules, etc. A matching of only 80 percent of the death certificates to the census schedules would undoubtedly produce biased results in understating death rates and in making questionable the representativeness of the decedents for whom the census information is successfully obtained. Fortunately, it has been possible to devise procedures which will eliminate or bring into reasonable control this possible source of bias.

To control this source of bias a special sample survey is being conducted in advance of the large-scale study in order to provide a basis for estimating the characteristics of the decedents for whom no census match will be possible. A questionnaire is mailed to the informant who provided the information on the death certificate for a sample of approximately 10,000 deaths for the same four-month period. The questionnaire calls for most of the information which was obtained in the 1960 census. To assure as high a response rate as possible provision was made to mail the questionnaires as soon as possible after the death records are received in the National Office of Vital Statistics. Provision was also made for follow-up requests by certified mail to persons not replying to the first mailing. Previous experience of the National Office of Vital Statistics indicates that information can be collected for about 85 percent of the deaths in such a mail survey. To control the possible bias of non-response at this point, personal interviews will be conducted with all, or an adequate sample of, non-respondents to the mail survey. It is anticipated that a response rate of about 90 to 95 percent will be achieved.

It may be gathered that approximately 20 percent (2,000) of the questionnaires collected in the sample survey will provide information about deceased persons not matched to the census schedules. Which fifth of the 10,000 decedents in the sample survey will represent unmatched cases will, of course, not be known until after the large-scale matching operation is completed. But the availability of the sample survey results will make possible, eventually, the estimation of the characteristics of the decedents for whom death certificates could not be matched to census schedules. Moreover, the sample survey results will include about 8,000 cases for whom the matching will have been completed. This information can be used for quality check purposes in analyzing the correspondence of information obtained on the death certificate, the census schedule and the survey questionnaire.

Needless to say, the handling of the individual death certificates and the census records in the matching operation will be done entirely by the staffs of the National Office of Vital Statistics and the Bureau of the Census. The confidentiality of the returns, both on the death certificates and the census records, will be scrupulously maintained. The participating governmental agencies after performing these processing operations will make available to the study directors, for analysis, the results of tabulations of magnetic tape containing the collated data.5/ Moreover, the base population statistics needed to calculate death rates will be tabulated by the Bureau of the Census in regular and special tabulations.

The Specific Objectives of the Study

The analytical objectives of the study are limited by the sampling design introduced into the 1960 Censuses of Population and Housing. For the characteristics which were collected in the Censuses on a 100-percent basis, mortality differentials can be calculated utilizing the full universe of deaths. The census characteristics collected on a 100-percent basis were, however, restricted to age, sex, color or race, marital status, relation to head of household (for family and household composition), tenure, occupancy status, condition of structure, number of rooms, type of housing unit, water supply, toilet facilities, bathing facilities, kitchen cooking equipment. All other items in the Population Census were collected on a 25-percent sample basis, and items in the Housing Census on a 20-percent or 5-percent basis. Thus, most of the population and housing information will be available for 25 percent of the decedents included in the study. As a result, there will be a sample of approximately 100,000 deaths, including about 80,000 matched deaths for whom the detailed census information will be available. Similar information will be available--from the Special Sample Survey questionnaires--for about 2,000 of the 20,000 deaths not matched with census schedules but presumably included in the 25-percent census sample. Detailed cross tabulations of mortality by social and economic characteristics of the decedent will, therefore, be affected by sampling variance, and especially will this be true for cause of death tabulations. Despite this limitation, however, it is clear that for the more important specific causes much valuable information can be obtained.

With these restrictions in mind the anticipated specific objectives of the study can be outlined as follows:

1. Geographic differences in mortality. Mortality differentials will be obtained for meaningful geographical areas for which analysis is not now possible by reason of the difficulties of achieving comparability in the classification of geographic areas from residence as reported on the death certificate and census schedule, respectively. The use of the geographic classification coded in the 1960 Census records, will permit the calculation of death rates for urban and rural areas as defined in 1960, for Standard Metropolitan Areas subdivided into central cities and rings, for urbanized areas, for groups of cities classified by size of population, and for farm areas. Thus, mortality analysis will be available for the new and significant areal categories for which rich census materials are available.

2. Nativity and ethnic differentials in mortality. The place of birth of the person and of his parents will be used to calculate death rates for first and second generation groups, by ethnicity. It will, therefore, be possible to analyze in a comprehensive manner for the nation as a whole not only the extent to which ethnic groups differ in mortality but, also, in at least broad categories ethnic mortality by cause of death.

3. Income differentials in mortality. Differential death rates by income will be possible on the basis of the 1959 income of the individual decedent, and the family income of the family of which he was a member. For example, mortality rates by family income can be computed for all those living in family groups (92 percent of the population in 1950); mortality rates by individual income can be computed for unrelated individuals (those not living in family groups); mortality rates by individual income can be computed for all men and women with an income; mortality rates for wives (and children) can be computed by income of husband (father).

4. Occupational differentials in mortality. The occupation and "class of worker" for each person at work during the week preceding the 1960 Census will be available and, also, the last occupation held by persons not at work the week before the census but who have worked some time during the past ten years. This information can be used to obtain mortality rates for both men and women with work experience since 1950; and mortality rates for women (and children) by occupation of husband (father).

Differential death rates by occupation will, in the main, reflect social and economic status differentials of which the occupation is an index rather than "occupation mortality risk." Although it may not be feasible, given the number of deaths in the study, the possibility of obtaining some measure of mortality risks for certain kinds of occupations will be investigated. The list of such occupations would, of course, be confined to those in which long apprenticeships, restrictions for admission, and long tenure permit the conclusion that it is one in which the person has spent all, or an appreciably large part, of his working career.

5. Education differentials in mortality. For persons 25 years of age and older, the census information on years of school completed can be used to calculate differential death rates by educational level.

5/ The study is being conducted through the Population Research and Training Center at the University of Chicago. Principal investigators are Philip M. Hauser and Evelyn M. Kitagawa. The study design is the product of consultation with a number of people, including Iwao Moriyama, Lillian Guralnick and Monroe Sirken of the National Office of Vital Statistics, Charles Nam of the Census Bureau, and Harold Dorn and Mortimer Spiegelman.

6. Mortality differentials by family status. Census information will permit the calculation of death rates not only by marital status of the deceased but, also, by family status of the deceased, including presence and age of spouse, and number and age of surviving children; by parity (number of children ever born) for all women who have ever been married.

7. Mortality differentials by housing and environmental factors. Death rates can be computed by various characteristics of the dwelling unit of the decedent, including rent, value of dwelling unit and quality of housing. To some extent, thus, environmental factors can be related not only to mortality in general but, also, to cause of death.

8. Social and economic differentials in mortality for selected causes of death. To the extent that the size of the sample permits, deaths by each of the social and economic variables described above, will be studied separately for selected causes of death. This will include, for example, income differentials in cancer mortality, occupational differentials for deaths from diseases of the heart, etc. It is anticipated that some social-economic differentials in the causes of mortality can be obtained for at least the 10 to 15 leading causes of death.

9. Interrelationships among the social and economic factors. The interrelationships among the various social and economic factors will be explored by analyzing mortality differentials with respect to each factor holding constant one or more of the other factors. For example, occupational differentials in mortality will be analyzed holding income constant, and ethnic differentials in mortality will be analyzed holding income constant. Similarly, the effect of income will be analyzed while holding occupation and education constant, etc.

10. Methodological studies. The investigation will permit a number of methodological studies. One to which reference has already been made relates to variability in response, possible by comparing the returns for the same person on the census schedules, the death certificates, and the special sample survey. It will also be possible to study the ecological correlation problem mentioned earlier, by comparing social and economic differentials in mortality based on direct cross tabulations of a sample of decedents, with results obtained from the same sample using the social and economic characteristics of the census tracts in which they resided. At a number of points operational studies can be made on the linkage of records from independent sources, which can throw light on the feasibility of matching studies of various types of documents.

Aspects of the Analysis

In addition to comparisons of general mortality rates for the diverse geographical areas and population groupings possible, various types of refined mortality measures will be employed. For one thing, mortality analysis will be made, as far as possible, for significant age groupings, namely, for decedents 25 to 44 years, 45 to 64 years, and 65 years of age and older, respectively. Analysis will also be made of mortality at younger ages but, because of problems affecting the denominator for infants, there is not much that can be done in this study in the way of infant mortality analysis.

In addition to analysis for the significant age groupings, age-adjusted mortality measures will be used throughout the study. Furthermore, largely for purposes of graphic and convenient summarization, abridged life tables will be calculated by a number of characteristics not hitherto possible, for example, by education, income, ethnicity, marital status and the like.

To highlight differential death rates, mortality ratios will be calculated relating death rates of subgroupings to the general or lowest death rates used as a base.

Two types of adjustment will affect the results of the entire investigation. One will involve the adjustment of the data for the unmatched deaths. It will be possible to estimate the characteristics of the decedents for whom certificates have not been matched to the census schedules, using data from the Special Sample Survey for a sample of these decedents. With such an adjustment, mortality rates can be calculated so as to make the results representative of the entire United States.

Finally, since the deaths are for four months of the year, the problem of differential mortality by season must also be taken into consideration. For the universe as a whole, seasonal adjustment can readily be made by relating mortality for the four-month period of observation to the national mortality data available for the year as a whole. Seasonal adjustments by characteristics, however, will, of course, be restricted to the items that are available on the death certificate itself, including age, sex, color or race, marital status, cause of death, place of residence. Problems of seasonality may conceivably affect differential mortality by various other social and economic factors. Preliminary investigation, however, indicates that serious distortions as a result of this problem are unlikely.

Phasing and Cooperation

This study of differential mortality is necessarily a cooperative venture. It would not be possible without the unstinted cooperation and interest of the National Office of Vital Statistics, the state and local Vital Statistics Offices and the Bureau of the Census. These agencies have indicated their interest in participating in the investigation and have already, in the essential preliminary work, enthusiastically provided the necessary assistance.

The entire study was made possible by a grant from the U.S. Public Health Service and the National Institutes of Health.[6/] The grant is for a five-year period which began in May, 1960. It is expected that the coding, matching and collation of the data will be completed by the end of 1962, and tabulations should be available by the end of 1963. This leaves the

---

6/ The total amount of the grant is $1,017,000, about 90 percent of which will be subcontracted to the National Office of Vital Statistics and the Bureau of the Census to collect the basic information.

principal investigators approximately two years for the analysis and writing of the reports in which the results will be published.

Needless to say, in a comprehensive and large scale study of this type, a number of supplemental and special investigations will be possible. Various agencies and individuals have already indicated an interest in special tabulations and analytical possibilities. The investigators themselves have in mind a number of extensions of analysis of the materials which may become desirable as the study reaches its later stages.

The investigators regard the data which will be collected as being largely in the public domain and are prepared to cooperate as best they can in helping to make the data available for special studies, within the limits of available resources and sharing of costs. Further reports on the status of this study will be made from time to time.

IX

UNEMPLOYMENT AND THE LABOR FORCE

Chairman, Margaret S. Gordon, University of California, Berkeley

The Labor Force and Unemployment in Mid-1960 — Robert J. Myers, U. S. Department of Labor

Characteristics of the Insured Unemployed: A Tool for Economic Analysis — Louis Levine, U. S. Department of Labor

Labor Force Projections for California, 1960-1975 — Maurice I. Gershenson, California Department of Industrial Relations

Discussion — Joseph W. Garbarino, University of California, Berkeley

THE LABOR FORCE AND UNEMPLOYMENT IN MID-1960

Robert J. Myers, Bureau of Labor Statistics
U.S. Department of Labor

It was about 20 years ago that we began to think of the labor force as a dynamic segment of the population and that we first developed a satisfactory method of measuring the labor force and enumerating the unemployed. It is difficult to believe that we could have been without such statistics during so great a part of our national life. Over a period of two decades, however, we have grown accustomed to counting the unemployed monthly, ascertaining all manner of intimate information about them, and announcing our findings in all their gruesome detail to the entire world.

We were not the first country to obtain and report dependable information on the level of joblessness, and it is gratifying to note that many countries now follow this policy. We have found it to be sufficiently satisfying that we can recommend it to those remaining countries which still attempt to hide their unemployment behind denials of the obvious.

I. Recent Trends

Nineteen-sixty has been a rather puzzling year for analysts of the labor force. As it began, we wondered whether it would see a continuation of the 1959 recovery in employment which was interrupted by the steel strike. More than half-way through, we're still not entirely sure. Industrial production and the GNP have reached new high levels. But productivity customarily leaps up in recovery periods, so that high-level production is not sure to be translated into high-level employment. And if employment continues to rise, will it rise enough to absorb the expected additions to the labor force?

In recent years the labor force has been growing at a relatively slow rate, adding only about 500,000 workers per year in contrast to an expected annual increase of almost double that number. In the first quarter of 1960, the labor force was only 300,000 higher than it was a year earlier (after allowing for the inclusion of Alaska and Hawaii). Economic activity in the first quarter had been slowed by unusually bad weather, widespread illness, and a late Easter. In the second quarter, however, the labor force rebounded sharply to show an average increase of 1 million over the second quarter of 1959. Present indications are that the total labor force will average about 73 million for the year 1960 as a whole, or some 800,000 more than in 1959.

A striking feature of the evolution of the labor force has been the increase in the proportion of women. In the coming decade, for the first time in our peacetime history, women will average over one-third of the total labor force.

Prior to this year, middle-aged women dominated the expansion in the female labor force. An important factor in this expansion was the growing tendency for married women to return to work after their children had reached school age. The labor force participation rates of women between the ages of 45 and 65 rose dramatically in the 1950's. Thus far in 1960, however, there has been no further rise in the participation rates for these women. In contrast to other recent years, most of those added to the labor force in 1960 were young men and women under 25 years of age. In part, this was attributable to increased numbers of young people reaching working age and completing or leaving school, but it also reflected a rise in the participation rates for young women over the year. We are not yet sure whether new trends in labor force participation among women are emerging or whether the developments in the first half of 1960 may prove to be merely transitory.

Total employment has been at record levels in most months of 1960, thus far averaging about 1 million more than in 1959. In most peacetime years, an increase of 1 million in employment would be more than enough to absorb the increase in the labor force. But this has been an unusual year. In the second quarter the growth in employment did not fully keep pace with the expansion in the Nation's labor force, so that in June unemployment was higher than in the corresponding month of 1959.

All of the gains in employment have been in nonfarm industries; the long-term decline in agriculture has been continuing in 1960. The number employed in agriculture this year will account for only about 8 percent of the total employed; 5 years ago they accounted for 11 percent, 20 years ago for about 20 percent.

Employment in trade, finance, services, and State and local government continued to expand in 1960. Jobs in most of these sectors were increasing steadily throughout the 1950's, even during periods of recession for the economy as a whole. Employment in service-producing industries surpassed employment in goods-producing industries in 1949 for the first time, and since that time has continued to expand much more rapidly. In fact, manufacturing employment in 1959 was still 1 million lower than in the peak year of 1953.

Unemployment declined encouragingly in the early months of 1959, reflecting our emergence from the recession and the build-up of inventories in preparation for the steel strike. The seasonally adjusted rate of unemployment was 6 percent as the year 1959 opened, declined to 5 percent by April, and then remained at about that level until July. The last half of that year, however, was rather disappointing. Secondary layoffs resulting from the steel strike helped push the unemployment rate to 6 percent in late fall before the strike was enjoined in early December. In the first half of 1960, the unemployment rate has again come down to average about 5 percent. It has reached 5.5 percent on two occasions due to temporary

situations (bad weather in March and an unusually large influx of teenage jobseekers in June).

So far in the present recovery period, unemployment rates have failed to return to the 3 percent we attained after the 1949 recession, or even to the 4 percent we reached after the 1954 recession. The former rate, however, does not afford a reasonable comparison. It was attained during the Korean War period and greatly influenced by the mobilization of manpower and other resources for military purposes. The years 1955-57, when the rate was a little above 4 percent, provide a more realistic benchmark against which to evaluate the present rate.

II. Unemployment--Comparisons With Other Industrial Countries

One cause of discomfort over the level of our unemployment is the suggestion made from time to time that other countries are doing better. This cannot mean simply that other countries have fewer unemployed; if the comparison is to be meaningful, it must take the size of the labor force into account. Only the rate of unemployment provides a suitable basis for such comparisons.

Before proceeding to an examination of unemployment rates in the various countries, however, it is necessary to take a look at the definitions used in determining who is unemployed.

There is an international standard definition of "the unemployed," which was approved in 1954 by the Eighth International Conference of Labor Statisticians. It is a good definition, and I am glad to note that the definition observed by the United States in the Monthly Report on the Labor Force conforms rather closely to this standard. So do the definitions used by Canada, Japan and Puerto Rico, which regularly produce unemployment statistics similar to our own, and those of several other countries with systems of unemployment statistics that differ from ours.

Many countries, however, do not have access to a precise and specialized means of measurement, but depend instead on administrative statistics compiled primarily for other purposes; e.g., the number registered at employment exchanges, the number receiving benefits under an unemployment insurance system or the number of unemployed reported by trade unions.

Such statistics are sometimes incomplete. They may exclude important groups who are not covered by the unemployment insurance system, others who have exhausted their benefits, jobless workers who neglect to register at the exchange, new workers entering the labor force, or former workers re-entering it. On the other hand, administrative statistics may enumerate certain workers who under the United States definition would not be counted as unemployed; for example, persons registered at the employment exchange while still employed, persons receiving unemployment insurance benefits in a week in which they also do some work, and persons who work part of the survey week but are not employed on the single day on which the count of the unemployed is based in some countries.

It is commonly believed that the unemployment statistics of foreign countries are less comprehensive than those of the United States and that in any dependable comparison our statistics would have to be scaled down considerably or the figures of the other countries adjusted upward. So far as the statistics of the underdeveloped countries are concerned, this is probably true. But as applied to industrial countries of the free world, it is not a safe generalization.

In recent years a number of the industrial nations have made occasional or regular labor force sample surveys very similar to our own MRLF. These have afforded comparisons with the unemployment statistics obtained from administrative operations. France and Sweden have made labor force sample surveys and have come up with unemployment figures appreciably greater than those resulting from their regular administrative systems. In Germany and Italy, however, labor force sample surveys have yielded unemployment figures lower than the number based on the regular system. Canada has two well established series of unemployment statistics, 1/ one based on a sample survey very similar to our own, the other based on statistics from the compulsory unemployment insurance system. In 1959 the former produced an unemployment rate of 6.0 percent and the latter a rate of 10.9 percent.

The United Kingdom depends upon a system very different from ours; it is based on registrations with the Employment Exchanges and the Youth Employment Offices. A preliminary analysis of the differences between the two systems suggests that the U. S. statistics are more comprehensive in some respects while the United Kingdom statistics are more comprehensive in others. There is evidence that the respective "errors" may about offset each other. 2/

1/ As is noted below, Canada does not issue official statistics on the "unemployed." The figures quoted are published by the I.L.O., based on appropriate official series.

2/ One recent analysis of United Kingdom unemployment statistics in terms of American standards identified some 80,000 persons in various categories regarded as unemployed according to the United States definition, but not so classified under the United Kingdom system. At the same time, the U. K. classified as unemployed an estimated 70,000 who were out of work on the single day of the U. K. count but who would have been counted as employed under United States practice because they had had some work during the survey week. Further research regarding the comparative systems of unemployment statistics in the United Kingdom and the United States is to be undertaken during the coming year.

It would be misleading to suggest that any studies made to date will permit a precise comparison between the unemployment rates of the United States and other industrial countries, but the general position of the United States is abundantly clear from a review of the published data in the light of what is known about the various statistical systems. In 1959, when the unemployment rate for the United States was 5.5 percent, the comparative levels of unemployment 3/ were unquestionably lower in the Federal Republic of Germany (2.6), France (1.3), Sweden (2.0), Switzerland and the United Kingdom (2.2). But the rate of unemployment that year was somewhat higher in Belgium (5.9), Canada (6.0), and Italy (8.7); also, among nonindustrial countries, in Ireland (8.1).

We can well devote further study to differences in scope which obstruct comparisons with other countries, and cooperate with the International Labor Office in urging adherence by all countries to the approved international definition of unemployment. But at the same time it is worthwhile to consider social and economic factors which help explain why unemployment is relatively more common in this country than in some other countries. Without pretending to cover this subject fully, it may be suggested that the following characteristics of our economy and of the American people tend to result in higher unemployment rates than are found in European countries. To begin with, there are several characteristics which are not evil in themselves but may be very beneficial, but which tend to increase unemployment rates.

1. The greater mobility of American labor, which results in increased loss of time between jobs.

2. The relatively small proportion of our labor force in agriculture, where unemployment rates tend to be low.

3. Our higher incomes, which sometimes enable the unemployed to hold out longer while searching for the most favorable job available.

Another important factor, for which few will say a friendly word is:

4. The greater susceptibility of our economy to cyclical fluctuations.

Still another characteristic, which has both favorable and unfavorable aspects, is:

5. The greater freedom of American employers to determine for themselves how many workers they need in periods of expansion, and to reduce employment in times of slack business--a freedom which in European countries is severely limited by law or convention.

3/ Figures in parentheses give unemployment rates, where available, as published in the "Statistical Supplement" to the International Labour Review. In most cases these rates are not comparable with the United States rate without adjustment.

The above comments relate to differences between the United States and other democratic countries. Under circumstances in which the State decides what is to be produced, where the employer is told whom he shall employ and the worker where he shall work--that is, under an effective dictatorship--unemployment can indeed be reduced to a minimum. Even under such circumstances, however, some frictional unemployment is inevitable as workers are shifted from one job to another or from one occupation to another. Surplus workers tend to accumulate in declining industries and localities. At best this unemployment can only be disguised. It is thus that some of the totalitarian states, while denying the existence of unemployment, have found it necessary to introduce unemployment benefits. Yugoslavia, a country which is more candid in such matters than most other countries with a high degree of state control, has regularly reported unemployment, which in 1959 attained a rate in excess of 6 percent.

III. Types of Unemployment

But now let us come back to our own unemployed here in the United States. In the interest of brevity, I shall refrain from discussing the characteristics of the unemployed, 4/ but I should like to comment on the types of unemployment.

It is customary to refer to the various types of unemployment--i.e., cyclical, seasonal, between-jobs, etc.--but we are only beginning to assess the relative importance of the different types, and even our present crude assessment does not cover all of them. Efforts to determine the magnitude of unemployment in its various forms are worthwhile because of the guidance we gain for planning action programs and the suggestions we may receive as to future trends in unemployment.

The most striking fluctuations in the level of unemployment, of course, are due to cyclical changes in the production of goods and services. A good deal of the unemployment we endure must be blamed on the business cycle, and we can only assume that this will continue to be the case for some time in the future. Each recession, however, has a character of its own, and I am sure you do not expect me to say anything very useful about the volume of cyclical unemployment. It is perhaps sufficient to note that the rate of unemployment averaged 25 percent in 1933, 5.5 percent in 1949, and 6.8 percent in 1958 as compared with rates of 4 percent or less in a number of prosperous peacetime years. It is reassuring to reflect that most economists agree that no future depression is likely to reach the depths of the depression of the thirties.

Unemployment in all forms other than cyclical is sometimes referred to as "frictional unemployment." In a study 5/ completed by the Bureau of

4/ A brief discussion of this topic which was included in the original paper has been omitted from the present condensed version.

5/ U. S. Bureau of Labor Statistics, The Extent and Nature of Frictional Unemployment, Study Paper No. 6, Joint Economic Committee, Congress of the United States, Government Printing Office, 1959.

Labor Statistics some months ago for the Joint Economic Committee of Congress frictional unemployment was defined as "that level of joblessness that could not be reduced significantly in the short run by increased aggregate spending." This study, based on the period 1955-57, when cyclical unemployment was at a minimum and total unemployment averaged slightly below 3 million, throws further light on some of the forms of frictional unemployment and gives a general idea of their magnitude.

Voluntary job mobility, i.e., changing jobs in order to improve employment status in some way, was found to account for about 10 percent of all unemployment in 1955. Improved communications are likely to encourage such job changes in the future, but the growing emphasis on seniority in job assignments and job security, the development of equities in pension funds, etc., will tend to discourage them. On the whole, it is difficult to foresee any clear indication of change in this area.

Roughly one-fifth of the unemployed in 1955-57--an average of about 550,000--were entrants or re-entrants into the labor force who became unemployed before finding jobs or withdrawing. Present estimates are that new entrants into the labor force will increase by about one-half in the decade of the '60's, thereby providing some upward pressure on the number of unemployed.

A minimum of one-fourth of the unemployment existing in 1957 was attributed to seasonal factors. Seasonal unemployment is particularly important in construction, agriculture, and certain manufacturing industries such as lumber and food processing. Perhaps it is reasonable to expect some future reduction in seasonal unemployment. Some of the industries in which such unemployment is severe have been declining in relative importance. Seasonal fluctuations are mild in many service and trade industries, which have been growing.

Another factor of potential significance is the changing sex and age structure of the labor force. An increase in the proportion of women workers, other things being equal, would tend to increase the level of unemployment, since the unemployment rate for women is typically higher than that for men. An increase in the proportion of young workers would have a similar effect. The Bureau of Labor Statistics study found that from 1948 to 1956 the net effect of the sex-age changes in the labor force tended toward a slightly lower over-all level of unemployment because the influence of the increasing proportion of women was more than offset by that of a declining number of young workers. Looking to the future, however, and assuming that recent unemployment rates will persist in each age-sex group as the distribution of the population changes, it has been estimated that this factor alone would increase the number of unemployed by roughly 200,000 persons during the next 15 years.

Structural changes in the relative importance of industries can also affect the unemployment rate. The shift out of agriculture into the other sectors tended to increase unemployment slightly from 1948 to 1956. This movement is expected to continue to exert upward pressure on unemployment in the future, though with only mild effect since the number of workers remaining in agriculture is already so small. The increasing proportion of workers entering the service industries will exercise an offsetting influence and may, indeed, be more important.

No attempt has been made to measure the effects of automation or other technological change on unemployment. This is, in fact, a causal factor of another category whose effects cannot be clearly distinguished. Increasing productivity resulting from changes in technology is both a cause of unemployment and a stimulant to increasing employment, and therefore it does not seem possible to assess the net effect of improved technology (as reflected in productivity change) on unemployment.

Reviewing the history of economic cycles in the United States since the days of the first peacetime crisis in 1819, one competent analyst has concluded that "the record to date suggests no tendency to an increase in the unemployment rate." 6/ In the face of this conclusion it may be hazardous to suggest that the increase which has been avoided so long is at last about to occur. The foregoing observations, however, have certain implications for the future which, however murky the crystal ball, appear to call for a conclusion.

Two influential forces will make for a higher rate of unemployment in the years to come: (1) the prospective substantial increase in the number of entrants into the labor force and (2) the changing age-sex distribution of the labor force. These factors may be offset in part by a decline in the influence of seasonality, but the effect of this factor is speculative. Certain other developments are expected to have divergent influences, the net effect of which is uncertain. The implication is that unemployment will tend to increase and that even "containment" of our old enemy "unemployment" will call for a determined national effort.

IV. Some Problems of Measurement

A number of problems remain to be solved before we can have full confidence in our measurement of the labor force, employment and unemployment. The final report of the Senate's Special Committee on Unemployment Problems, moreover, referred specifically to "additional data and measurements necessary to guide the Congress in the development of programs." Before concluding my remarks, I should like to mention a number of problems which the BLS has recently faced in the field of manpower statistics and certain new statistics which may be introduced in the future.

The basic concepts and definitions in this field are always under review. As recently as

6/ Stanley Lebergott, "Economic Crises in the United States"; document submitted to the Joint Economic Committee of the U. S. Congress, Hearings on Employment, Growth and Price Levels, Part III, Government Printing Office, 1959.

1957 an appreciable change was made in the definition of the unemployed, and it would not be difficult to identify several features of our present definitions which are not entirely satisfactory.

One change which has been suggested for consideration 7/ would reduce the count of unemployed slightly by omitting jobless persons who are not looking for work but "who would be looking for work except for the belief that no work is available." On the other hand, the AFL-CIO Research Department has recommended consideration of several changes which would somewhat increase the count of the unemployed. In addition, the Research Department has recommended monthly publication of estimates, distinct from the unemployment count, of the "full-time equivalent of involuntary part-time unemployment." Several other groups have also expressed interest in such estimates.

The Canadians have steadfastly refrained from issuing figures identified as measuring the "unemployed," although they regularly publish the number "without jobs and seeking work" and collect separate information on additional groups covered by the United States definition. In recent months there has been great pressure from the Parliament and from other quarters to develop and publish an official series on the unemployed, and an interdepartmental committee which has been studying the question is soon to issue its report. It will be of great interest to know whether the Canadians actually adopt a definition of the unemployed and, if so, to what extent it parallels our own.

Although the various suggestions for change will continue to receive study, it seems unlikely that the major concepts and definitions in use in the United States will be appreciably changed at an early date. There has, however, been growing interest in the definition and measurement of underemployment. Our chief current information on this problem consists of statistics on part-time work.

With respect to our present organization for measuring the labor force, an increase in the sample for the MRLF has long been under consideration. The present sample includes about 35,000 households, or about the same number covered by the Canadian labor force sample. While the probable error of our published results is generally quite low, some of the detailed results obtained cannot be published because the sample is too thin. For the same reason we are barred from undertaking special inquiries which would have considerable significance.

Current collection and tabulating costs are in excess of $1 million yearly, however, and it is unlikely that sufficient funds for any sizable expansion of the monthly survey will soon be forthcoming. BLS officials are more optimistic that it may become possible to double the sample occasionally, say, one month in the year, thus yielding considerably greater detail for that period. Such a proposal has been included as one of our objectives for an early year.

In the more immediate future we hope to obtain data supplementing presently available information on characteristics of the unemployed. In one approach we may be able to bolster up the present sample of the unemployed by accumulating over a period of three or four months all of the unemployed identified in successive MRLF inquiries. Thought is also being given to studying the health problems of the unemployed, perhaps through some further identification of the unemployed turned up in the National Health Survey. Another possibility is to follow up on the experience of a small sample of persons whose unemployment is presumed to stem from automation.

The next few years may witness the appearance of new employment series, perhaps overlapping some of the present series, but produced to meet special needs. One series in which we are especially interested is a series on the employment of construction workers. The present series is limited to employment in "contract construction," yet it is often used to represent all construction employment, or erroneously compared with the volume of construction put in place. Substantial numbers engaged in force account construction, working as self-employed, etc., are not included, though these workers are reported elsewhere, in other industries. Thus, while our current series for those employed on contract construction averages 2.5 million to 3 million per year, our estimates suggest that the number engaged in all construction is in excess of 5 million. We hope to issue a first rough report on this in the very near future and to prepare periodic estimates hereafter, at least once a year.

Additional attention has been given to the question of seasonality in our various series of the labor force, employment and unemployment; new adjustment factors were developed at the beginning of this year and are now in use. But we are already considering a different, and we believe better, approach to the question of seasonality, in which different seasonality factors will be worked out separately by sex and age groups.

Finally, I should mention important work under way on the series of employment statistics based on establishment reports. These statistics provide the only comprehensive information on employment, hours and earnings by industry, and are vital economic indicators. Conversion of the industry data to the new Standard Industrial Classification codes is now far advanced, and employment data with the new codes will be available early in the next calendar year. New benchmark adjustments, to improved benchmarks, will be introduced at the same time to make correction for systematic error which may have affected the series since the time of the last benchmark adjustment, 1957. Important improvement in the quality of the series is expected to result from stratification of the reporting establishments by size. Continuing research

7/ Albert Rees, "The Measurement of Unemployment," in Studies in Unemployment, U. S. Senate, Special Committee on Unemployment Problems, Government Printing Office, 1960.

on the volume of employment in certain statistical problem areas, such as employment by churches and clubs, is suggesting other revisions which should improve the accuracy of this useful series.

In conclusion, the prospects are for more accurate, detailed and sensitive information on the labor force, employment and unemployment than we have had in the past.

CHARACTERISTICS OF THE INSURED UNEMPLOYED: A TOOL FOR ECONOMIC ANALYSIS

Louis Levine, Bureau of Employment Security, U. S. Department of Labor

Problems of unemployment and its measurement will attract increasing attention in the decade of the sixties. The experience of the 1930's destroyed, for all practical purposes, the complacent belief that automatic market adjustments were adequate for maintaining economic equilibrium. The necessity for assuring a satisfactory volume of job opportunities for all who can and want to work has become a major tenet of public policy. This fundamental change in economic and political thinking set in motion a train of events that affect, in some degree, the lives and the livelihood of everyone in this and succeeding generations.

Most economists are agreed that the economy is not likely to experience again the massive unemployment of the Great Depression. Avoidance of depressions and unemployment of disaster proportions is now generally accepted as an obvious and minimum objective of economic policy. We have gone well beyond that minimum in the Employment Act of 1946. Our national policy provides that government shall promote and stimulate maximum employment. 1/

It may be useful for a better understanding of the characteristics of unemployment to analyze briefly some of the factors underlying the general development of our economy and to appraise realistically their impact on the labor force. Recent optimistic forecasts are based essentially on two assumptions. These are, first, that the unprecedented increase in population since the end of World War II is resulting in vastly larger needs for everything that goes into making up the American standard of living; and, second, that the new products of our rapidly advancing technology will stimulate widespread desire accompanied by effective consumer demand for their purchase. This combination, it is held, will produce such a large and increasing volume of demand that sustained levels of peak prosperity are reasonably assured for the indefinite future. It does not necessarily follow, however, that technological advances will be quickly reflected in a roundabout increase in job opportunities. On the contrary, such advances reduce direct labor requirements in the industries experiencing such advances, and some period of time is likely to elapse before the productivity gains are reflected in greater employment opportunities. In any event, the new jobs may well be in occupations for which the displaced workers are not qualified or in areas far removed from the localities where they reside.

We have just completed a decade in which both factors, sharply rising population and rapid technological advance, were operating. In addition, economic stimulation was provided by the large reserve of purchasing power and the need to make up the depression- and war-induced shortages of housing, industrial capacity, and commercial and public facilities of every kind. Job opportunities reached record levels. Nevertheless, unemployment problems were still in evidence at the end of the decade.

While periods of business recession since the end of World War II have been relatively short, a gradual trend toward rising unemployment seems to be shown by the record of the last few cycles. In addition, unemployment rates among new entrants in the work force, displaced older workers, the unskilled, and nonwhite workers have been persistently and markedly higher than for the labor force as a whole. Certain industries and areas have also developed apparently chronic problems of unemployment. 2/

The nature of the problem has changed over the years, but unemployment as a serious cause for concern is still with us. If "Mass unemployment of the thirties largely gave way to class unemployment of the fifties," 3/ in Prof. Dunlop's colorful phrase, what may we reasonably anticipate for the sixties? The occupational hazards that lie in wait for unwary forecasters are well known. Nevertheless, we may review the general outline of the picture of coming events with a minimum of risk. Labor force projections for the decade ahead can be made with reasonable accuracy. Every member of the labor force of the 1960's, actual or potential, has already been born. Possible changes in marriage and birth rates may influence participation rates for certain groups in the population without significantly affecting the validity of the projections.

Some dimensions of the future labor force problems and the objectives of private and public policies and programs were set forth in the U. S. Labor Department's publication the Manpower Challenge of the 1960's. The projected increase in the labor force during the 1960's, about 13.5 million, will be by far the largest for any 10-year period in our history and 50 percent greater than the expansion during the fifties. Along with this huge increase will come major shifts in the composition of the work force. Young workers under 25 years of age will increase by about 46 percent to a total of over 20 million, while the number of older workers--those 45 years and over--will rise by 20 percent to reach a level of about 33 million. These figures do not reveal the total magnitude of the impending shifts in the labor

1/ For a discussion of the role of government in promoting employment opportunities, see Louis Levine, "Problems in Labor Market Organization and Administration," in Manpower in the United States: Problems and Policies, William Haber, et al., editors, New York, Harper, 1954.

2/ Cf. United States Congress, Joint Economic Committee, 1959 Joint Economic Report, Washington, Government Printing Office, 1959, Chapter 6, pp. 161 ff.

3/ John T. Dunlop, "Public Policy and Unemployment" in Studies in Unemployment, Special Committee on Unemployment Problems, United States Senate, Washington, Government Printing Office, 1960, p. 1.

force, however. The 1970 labor force will reflect the entrance of 26 million young people over the decade.

Moreover, these enormous changes will be accompanied by an accelerated rate of technological innovation, changes in consumer preferences, and geographical shifts in industrial location, whose magnitude and impact no one can now accurately foresee.

We immediately recognize that the major increases in the labor force of the 60's are concentrated in two groups--new entrants and older workers--who are already disadvantaged in the job market. Our experience of the 50's is not especially reassuring as we face the future. Over the decade just past, the net number of young workers added to the labor force was a mere 700,000, reflecting the low birth rates of the depression years, and the postwar trend toward earlier marriage and childbearing which reduced the labor force participation of young women. With economic activity at generally record levels, and with relatively few young people coming into the job market, employers turned increasingly toward the older groups, and especially to the large numbers of women in the middle and older age groups who were quite willing to exchange household tasks for paid employment. Yet despite this generally favorable situation, the special employment problems of both younger and older workers were clearly revealed in the statistics of unemployment as the decade came to a close.

What are the implications of the manpower challenge of the sixties for the future of unemployment statistics? It may be appropriate to recognize that along with the record level of wealth achieved in this country, we have also succeeded in producing a wealth of statistics. We have become a very statistically oriented people with a high degree of economic literacy. It was not always so. Most of the development of current economic statistics and analysis has come within the recent past. Thirty years ago the few available economic statistics were generally relegated to the technical journals and the financial pages. Economic literacy was limited to a few professional economists and experts in the business world. Today our broad range of economic indicators, from national income to housing starts, are regularly featured in the newspapers, not infrequently with front page billing. This is indicative of the widespread public interest and use of economic data.

Of all the available economic indicators, however, unemployment continues to be the most widely recognized and most generally understood. It is the figure that enables us to evaluate all the others, in terms of direct human impact and human values. Imprtant as it is to know that personal income and the index of industrial production are reaching into new high ground, we will not be satisfied that they are increasing fast enough if at the same time unemployment is too high or is actually increasing.

It is sometimes difficult to realize that the economic collapse after 1929 found us with no measure of unemployment and without even a generally acceptable definition of the term. Nearly ten years passed before the concept of the labor force was evolved and reasonably fast and accurate methods of measuring it were developed. We have come a long way since then in both the collection and analysis of unemployment data.

As the first measures grew out of a condition of mass unemployment, primary emphasis was on the total number of unemployed. With increased knowledge of the behavior of the labor force and unemployment there has been a growing awareness of the differential economic impact of changes in the economy. Overall unemployment figures have become less meaningful, and the need for data on the specific nature of the unemployment problem is expanding. Professor Dunlop stated the situation well when he wrote:

"The leveling influence of a single count neglects very significant differences among individuals and classes of workers. Thus the same head count of..... unemployed may constitute quite different human needs and represent widely different impacts on the economy depending on the class composition of unemployment. The appropriate public policies are likely to be significantly different for various compositions of unemployment with the' same total head count.

"The conclusion is that the structure of unemployment now deserves as much--if not more--attention than the level of unemployment. Public policies on unemployment always need to be conceived in terms of twin related policies: those designed to influence the level and others directed toward the structure of unemployment." 4/

This shift in emphasis has important implications for the measurement of unemployment. There are now two major sources of unemployment data, the Monthly Report on the Labor Force (MRLF) and the statistics of insured unemployment. Each of these is related to somewhat different concepts. Unemployment has different meanings to different expert individuals. The distinctions between employment and unemployment, and between unemployment and non-labor force, cannot be so sharply drawn that they are beyond dispute. Differing concepts may have equal validity, depending upon the source of the data and the uses for which they are intended. 5/ The two types of unemployment data current-

---

4/ Dunlop, Op. cit., p. 2.

5/ "The term 'unemployed' encompasses a variety of meanings. It may describe a condition--that of being not at work; an 'activity'--that of seeking work;an 'attitude'--that of desiring a job under certain conditions; and a 'need'--that of needing a job. The term also has other connotations and various shadings and combinations. For example, should a definition of unemployed include individuals who do not have jobs and who are not looking for work but who would accept jobs under certain conditions? Is a person unemployed who is in need of a job but because of home responsibilities is unable to look for or to accept a job? Should only those persons be counted as unemployed who are without jobs but who are breadwinners of their families? Obviously, the definition used determines the resultant count." Levine, Op. cit., p. 325.

ly available are therefore properly regarded as complementary.

The monthly survey of the labor force is a sample survey of households. It covers the entire population and utilizes the most inclusive definition of unemployment. 6/ It therefore provides current estimates of total unemployment and, as interest and emphasis have shifted to the analysis of unemployment in depth, increasing amounts of data on the personal and economic characteristics of the unemployed.

Probably the best known measure of economic developments that is derived from employment security operations is the count of insured unemployed workers. In fact, this figure has been included in the monthly publication, Economic Indicators, issued by the Joint Committee on the Economic Report, since 1954. Insured unemployment represents the number of benefit claimants who have completed a week of total or partial unemployment; and the count is a regular administrative by-product of the employment security system. The statistics measure the number of workers covered by the State unemployment insurance programs, unemployment compensation for Federal employees, and the programs for ex-servicemen who are certifying to partial or total unemployment. Also included in the overall count are those workers covered by the Railroad Retirement Act.

Insured workers account for over four-fifths of nonfarm wage and salary workers and, of course, for the bulk of the unemployed. Insured jobless workers are those who have become unemployed following separation from a job, and are sometimes designated as the "disemployed unemployed." It is this group that is generally considered to have the greatest direct impact on the economy through the loss of purchasing power.

Insured unemployment figures as an economic indicator have certain distinct values. First, there is their timeliness--the figures are reported on a weekly and monthly basis. Second, they are the only current sources of unemployment information by geographic location. Third, the data are based on a face-to-face meeting of the unemployed with the employment security interviewer. Fourth, the figures represent actual counts of unemployment insurance claims taken by public employment offices. Fifth, the statistics are obtained as a by-product of operations of the unemployment insurance program, making the collection process one of the least costly ways of obtaining such information. Thus the claimant information not only comes from the person best able to provide it but also the categorization of that information has the added flavor of professional interpretation.

To understand the character of the data it is necessary to describe briefly the unemployment insurance processes that yield these data. An individual desiring to receive benefit payments must file an initial claim certifying that he has become unemployed, and in subsequent weeks file continued claims certifying to the completion of one or more weeks of unemployment. Initial claims measure new or emerging unemployment, while continued claims which are the basis of insured unemployment measure continuing joblessness. 7/

The number of workers in covered employment in an average month has increased from 21 million in 1938 to 43.1 million in 1959. Of this total, railroad workers numbered about 1.1 million in 1938 and 900,000 in 1959. State coverage which averaged 19.9 million in 1938 had increased to 35.4 million by 1954, and has since risen to 39.8 million in 1959. The remaining 2.4 million workers in 1959 were civilian employees of the Federal Government. In addition, members of the Armed Forces were covered, averaging 2.6 million in 1959.

The main groups not covered by the unemployment insurance program are agricultural labor, domestic workers in private homes, employees of most State and local governments, employees of nonprofit organizations, proprietors and self-employed. In addition to the general groups excluded from coverage, workers in small firms are not covered in many States. Twenty States, however, have extended coverage to include workers in smaller firms.

Aside from these broad limitations stemming from the coverage provisions of State laws, some groups of covered workers who are unemployed may not be eligible for benefits and therefore not included in the insured unemployed figures. These consist of unemployed covered workers who are not available for work because of temporary illness; those unemployed persons whose previous jobs were in covered industries but who did not earn sufficient wage credits or were not employed the required length of time to be eligible for benefits; persons disqualified for non-monetary causes, such as voluntary separation from work without good cause, discharge for misconduct, unavailability for work, refusal of suitable work, and direct involvement in a labor dispute.

6/ The sample survey of households, collected and tabulated by the U. S. Bureau of the Census for the Bureau of Labor Statistics, provides a comprehensive measure of the labor force, i.e., the total number of persons 14 years of age and over and who are employed or unemployed. The information is obtained from a scientifically selected sample of about 35,000 interviewed households in 333 areas throughout the country and is based on the activity or status reported by surveyed persons for the calendar week ending nearest the 15th day of the month. The unemployed total includes all jobless persons who were looking for work, regardless of whether they were eligible for unemployment insurance. Also counted as unemployed are persons waiting to be called back to jobs from which they had been laid off; those scheduled to start on new jobs within 30 days (except students); and those who would have been looking for work except that they were temporarily ill or believed that no work was available in their line of work or in the community.

7/ For a full description see "Insured Unemployment, Employment, and Wage Statistics" in the Labor Market and Employment Security, U. S. Department of Labor, Bureau of Employment Security, March 1960.

Finally there is also a group of workers who are excluded from the insured unemployment figures because they have exhausted their benefit rights. In a period when unemployment is substantial and of long duration the volume of exhaustions may have an important bearing on the magnitude of the insured unemployment levels.

Despite these limitations, data on the insured unemployed are of major economic significance since they cover virtually all of the most sensitive and volatile sectors of the economy. The intrinsic value of the insured unemployment figures is enhanced by the inclusion of information not only on a national but also on a State and local level.

National figures on unemployment, though important in providing a general indication of the status of the economy, conceal wide variations in unemployment conditions in different areas of the country. Though the availability of data on the insured unemployed plays an important part in meeting the needs of information on a geographic basis, their expansion to include economic and social charactersitics of the unemployed was necessary to the development of programs for a meaningful attack on the problems. For example, because there are wide variations in the industrial composition of individual areas, even a minor change in the economic position of an industry nationally may have a significant impact on the economy of one or more areas where the industry is important locally. 8/

Congressional concern for the unemployed (most recently, the creation of the Senate Committee on Unemployment Problems) indicated a continuing and current need for information about who are the unemployed--their age, sex, occupational skills, industry attachment, and length of unemployment. Similar interest is evident in many State governments. Questions are being raised with increasing frequency on what kinds of workers are exhausting their benefit rights, who are the hard-core unemployed, and what are the more severely affected groups.

In addition, widespread demands for such information on the characteristics of the unemployed, nationally, and on a State and area basis, came from a variety of sources, including the Council of Economic Advisers, the Advisory Board for Economic Growth and Stability, and other Federal agencies, government officials in States and communities, employer associations, labor organizations, employment development committees, and other public groups.

The Bureau of Employment Security and its affiliated State agencies have undertaken to provide for the collection of these data. Into the 1,800 full-time local public employment offices come the jobless workers who, when certifying to their availability for work, make known their personal and economic characteristics. As a result, without incurring the costs of a household survey, it is possible to supplement the data on the unemployed with information on their industry, occupation and other economic and personal characteristics. The fact that the data are essentially a by-product of operations, which can be had for only the cost of processing, was not overlooked by cost-conscious administrators.

The next logical development was to fashion a reporting system to extract this wealth of information. At the outset, a basic decision had to be made between adoption of a case record, or continuous measurement approach, in which the unemployed worker would be followed through his spell of unemployment, or a "snapshot" approach in which information on the unemployed was obtained for a point in time on a regular basis. The "snapshot" approach was selected for the current characteristics program because it provided administrative simplicity, and low cost and speed in obtaining and processing the data so that they may be released on a current basis. There is also due recognition, however, of the need to exploit the kind of information that can be developed only from a case history approach. Some thinking on this score is highlighted in a subsequent portion of this paper.

Initiated in January 1960, the current survey is based on a sample of the insured unemployed claiming benefits or waiting-period credit during a 1-week period each month. The sample represents the insured unemployed for the same week as that covered by the Monthly Report on the Labor Force and, therefore, is comparable in time reference with the national estimates of total unemployment. The sample is confined to continued claimants under the State UI programs and covers age, sex, occupation, industry, and duration of current unemployment. In addition, the same characteristics are reported separately for claimants exhausting their benefit rights; at present the latter data are valid only for developing national totals.

The survey sample was designed to permit easy comparison of data among the States. Sample sizes range from 1 to 50 percent, varying in accordance with the volume of State claims.

The data which flow from this program are to be found in "The Insured Unemployed," a monthly publication of the U. S. Department of Labor's Bureau of Employment Security. Contents of this publication include on a State basis 1) Age and sex of the insured unemployed; 2) distribution of the insured unemployed by industry division; 3) distribution of national totals only, by two-digit manufacturing industry; 4) distribution of the insured unemployed by major occupational group (quarterly, national totals by selected three-digit occupations); 5) distribution of the insured unemployed by length of current unemployment; 6) age, sex, and length of current unemployment of claimants exhausting benefits (national totals); and 7) industry division and major occupational group of claimants exhausting benefits (national totals).

Insured unemployment statistics differ in several respects from those for total unemployment

---

8/ Vladimir Chavrid and Gabriel Cherin, "Use of Standard Metropolitan Statistical Areas Concept in Government," in American Statistical Association, Business and Economic Statistics Section, Proceedings, Washington, American Statistical Association, 1959, pp. 300 ff.

prepared by the Bureau of Labor Statistics in the Monthly Report on the Labor Force. The MRLF count of total unemployment includes all jobless persons who did not work at all during the survey week and were looking for work, regardless of their eligibility for unemployment insurance. As mentioned earlier, several groups of unemployed workers are excluded from the insured unemployed under State programs. On the other hand, the count of insured unemployed includes some persons who worked during the week but whose earnings fell below a specified amount (usually the weekly benefit amount).

In addition there are differences in the occupational and industrial classification. The occupational classifications used in the data on insured unemployment are assigned by trained employment service interviewers and are based on the total work experience and qualifications of job applicants. These classifications are defined in the 1949 edition of the Dictionary of Occupational Titles (DOT) and its supplement published by the Bureau of Employment Security. The MRLF occupational information is derived from household interviews and is arranged according to the classification groupings used by the Census of Population.

Definitions of some of the occupational groups in the two series are fairly comparable, e.g., professional and managerial, clerical and sales, and service. Craftsmen are roughly comparable to workers in the skilled category in the DOT, operatives to semiskilled, and nonfarm laborers to unskilled. However, a significant number of jobs included in the craftsmen category are classified in the semiskilled group in the DOT system; certain jobs in the operative category will be found in the skilled or unskilled group, while some types of nonfarm laborers are identified among the semiskilled.

The industrial information used by the State agencies for insured unemployment data differs in source from that used for total unemployment in the MRLF. While both series use the Standard Industrial Classification (Bureau of the Budget, 1957), the classification of establishments covered by the unemployment insurance system is based on detailed nature-of-business information obtained from employers. In contrast, the industry attachment information reported in the MRLF is that supplied by the respondent in a household interview.

Data on insured unemployment from the sample surveys, together with data on covered employment, are used to obtain unemployment rates by industry. The rates represent insured unemployment as a percent of the average covered employment during a 12-month period. This concept is also different from that used for the total unemployment rates in the MRLF. MRLF rates by industry represent the number of unemployed whose last full-time job was in the specified industry as a percentage of the total currently employed in that industry plus the number of unemployed from the industry.

Experience with the collection and analysis of the data on the characteristics of the insured unemployed, using the current BES approach, is little more than a half year old. Consequently, it has not yet been possible to build up a time series for comparative analysis. Already, however, the data have produced a wealth of information which has gone a long way toward illuminating the pattern of unemployment in this country. In April 1960, for example, we knew not only that six industries accounted for nearly three-fifths of all manufacturing unemployment, but also that most of this unemployment was concentrated in a very few States.

In addition to the expanded statistical data that will be available for analysis, the system allows for increasing the sample to compile valid data on a local area basis. Local information on the characteristics of the insured unemployed is of invaluable assistance to communities and industry in promoting and providing jobs to utilize fully available manpower and skills, and in developing suitable training programs for jobless workers. The prospect of obtaining such information for labor surplus areas is already attracting the attention of students of labor market problems.

Knowledgeable persons in the field will recognize that there are many gaps in our understanding of labor market behavior, the answers to which are needed for policy formulation, program evaluation, and guidance of the Congress and the State legislature in developing unemployment and other social legislation. Interested groups such as labor unions, employer associations, and universities also have a vital stake in obtaining more elaboration of the characteristics of the work force.

One of the major areas for development is the gathering of information, State by State, on the characteristics of all covered workers. Here the planning is directed to a selection of a sample of workers covered by the unemployment insurance program, which will provide not only information on the personal and economic characteristics of these workers, but also will relate the information to their employment and unemployment history--a reservoir of data which will provide a complete picture on the characteristics of the most significant component of the labor force. This aspect of research contemplates the use of the case method or continuous sampling approach.

Most of the necessary information is now available in the records of the State employment security agencies. When properly organized and analyzed they will provide information to answer such questions as: What are the variations in unemployment among employees of small firms as compared with those in larger establishments? What is the extent and variation of workers' attachment to the labor force? What are the variations in the employment pattern of workers in different industries? What are the industry, occupation, age, sex, and other relevant characteristics of workers with limited or irregular employment as compared to those with substantial employment? What are the characteristics of workers with some wage credits but with insufficient earnings to qualify for benefits as compared with those who do qualify? What are the differences, if any, in the characteristics of unemployed workers in depressed areas as compared with those in other areas? What are the differences, if any, between the employment and wage

patterns of claimants as compared with all covered workers?

If information were made available from employer records as to weeks of employment, through a special survey of employers, it would be possible to answer such additional questions as: What is the relationship between a worker's weekly benefit amount and his average weekly earnings? What would be the effect of using given numbers of weeks of work as measures of labor force attachment?

Even with the availability of information on covered workers and subject employers discussed in this paper, we are touching only the periphery of what we need to know about the insured unemployed. There are a host of other questions to which we need answers. For example: family wage earner status (sole, primary, secondary); distribution of claimants by number of dependents; adequacy of the weekly benefit amount; and although a good deal of work has already been done, continuing research is needed on the labor market experience of claimants after the exhaustion of benefits. Periodic or one-time studies would be needed in order to obtain the necessary information.

The economic analyst, like the military commander in time of war, is likely to feel that he never has quite all the material he needs to achieve the desired objective. And so it is with those of us who are engaged not only in the study of labor market developments, but also in the formulation of programs and policies designed to deal with the impact of these problems on the individual workers.

Nevertheless, those who are familiar with the insured unemployment statistics program, including information now produced and data potentially available, cannot help being impressed and encouraged by this remarkably facile tool. Having successfully over the past 25 years devised a measure of the unemployment volume and its national characteristics, we are now on the threshold of gaining fuller insight into where the unemployed are, and who they are within those narrower geographical confines. We have, in essence, added a rifle to our shotgun arsenal. While national policy will always be basic to a solution of the unemployment problem, our ability to pinpoint and define the problem locally will enable us to marshal our resources at the grass roots level where efforts frequently can be most effective.

LABOR FORCE PROJECTIONS FOR CALIFORNIA, 1960-1975

Maurice I. Gershenson, California Department of Industrial Relations

In June 1959, Governor Brown asked the members of the State Interdepartmental Research Coordinating Committee[1] to prepare projections of socio-economic trends in California through the year 1975 to be used in long-range planning for the needs of the State.

Under the plan developed by the Committee, each member was assigned a field covered by, or related to, his department. Among the fields assigned for the 1960-1975 projections were the following:

Population, by age and sex
Water requirements -- urban and agricultural
Agricultural production by commodity groups
Cash farm income by commodity groups
Births, deaths, marriages, and divorces
Labor force, by age and sex
Employment, by industry
Social welfare recipients, by type of aid
Patients in state-operated and state-subsidized mental hygiene programs
Prison and parole population
School and college enrollment
Personal income
Appropriations for construction of flood control and reclamation projects
Capital expenditures for water developments
State revenues
Highway expenditures
Motor vehicle registrations
Electric power load

The assumptions underlying all of the California projections, although independently arrived at, are very similar to those adopted by the Bureau of Labor Statistics for the national labor force projections.

Among the California assumptions were the following:

1. Present world tension will continue. There will be no major war or other cataclysmic event.

2. The total volume of defense expenditures will not change significantly, and the shift from missiles to aircraft will not be to the detriment of the California economy.

3. There will be no appreciable increase in governmental controls. Personal and business taxes will remain at present levels in relation to incomes.

4. Economic activity will continue at a relatively high level.

Joint responsibility for the labor force and employment projections was assumed by the State Departments of Industrial Relations and Employment.

Population

California's future labor force will be shaped by changes in the population. We can turn this statement around and say that the future population will reflect changes in the labor force and economic developments. We chose to break into this circle by accepting as a starting point the population projections for California prepared by the population research unit in the State Department of Finance under the direction of Carl Frisen. These are based upon estimated natural increase plus projections of future immigration.

California's total population is expected to increase from 15,830,000 in mid-1960 to 22,090,000 in 1970 and 25,600,000 in 1975 -- a 15-year gain of almost 10 million. This represents a rate of increase about 2½ times that for the United States as a whole.

In the United States' projections, net immigration is considered a minor factor. Not so in California. We expect more than half our population growth in the next 15 years to come from migration into the State.

Labor Force

The California labor force projections assume that the labor force participation rates by age and sex will tend to be the same as for the United States, except for the younger age groups where the California rates have always been significantly below those for the nation.

Accordingly, the labor force participation rates used by the United States Bureau of Labor Statistics for the U. S. projections[2] were applied to the California population projections by age and sex with the following modifications based on the 1950 relationship between labor force participation rates in California and in the United States: For males 14 to 17 years, the United States rates were cut approximately 12 percent. For females 14 to 17 years, the United States rates were decreased by 18 percent; and for the 18 to 19 year group, the rates were dropped 11 percent.

The national labor force projections and participation rates are in terms of total labor force, which includes the armed forces. In making the California projections, the national participation rates, with the modifications indicated above, were applied to California population projections that included the armed forces.

This yielded projections of the total labor force for California comparable with the labor force projections for the nation. The armed services component was then subtracted to arrive at civilian labor force projections for California. Thus we have two sets of labor force projections for our State.

California - United States Comparisons

I would like to turn from this brief discussion of methodology to a review of some interesting comparisons between our projected trends for California and those for the United States prepared by the Bureau of Labor Statistics.

The total labor force in California is expected to increase from an average of 6,330,000 in 1960 to 9,002,000 in 1970 and to 10,609,000 in 1975. This means a rise of 68 percent in the next 15 years which contrasts with an expected increase of 29 percent in the nation. At present, California accounts for about 8.6 percent of the country's labor force. By 1975 we expect this ratio to increase to 11.2 percent.

Age

In terms of age, the largest increases in the California labor force will be among those under 25 years of age. This parallels the national trend. A marked difference, however, is anticipated for the age group 35-44 years. The California projections indicate an increase of 24 percent in this age group, which is in sharp contrast to the projected decrease in the nation.

The net result of the changes in the respective age distributions of the labor force in the next 15 years will be that California will have a younger work force than the United States as a whole. California will have a larger proportion of its total labor force in the ages 20 to 44 years and a smaller proportion in all other age groups.

It is interesting to note that although we projected the same or lower labor force participation rates for the individual age-sex groups than for the United States, the overall rates come out higher, in general, in California than in the nation, as follows:

TOTAL LABOR FORCE PARTICIPATION RATES
(Percent)

	Total		Male		Female	
	U.S.	Calif.	U.S.	Calif.	U.S.	Calif.
1960	58.0	58.6	80.9	82.0	36.4	36.2
1965	58.1	58.7	79.1	80.8	37.6	37.6
1970	57.8	59.0	78.4	80.4	38.3	38.5
1975	57.9	59.3	78.4	80.5	38.5	39.0

This is so because of differences in the relative weights of the various age groups in the United States and California projected populations. California is expected to have a larger proportion of its population in the age groups with the highest labor force participation rates.

Women workers

As in the nation, the number of women workers will increase faster than men in the next 15 years, but the differential growth will not be as great in California as in the United States. Between 1960 and 1975, the female labor force in California will increase 75 percent and the male labor force 64 percent. The corresponding increases for the United States will be 38 and 25 percent respectively.

Employment by industry

Projections of employment by industry are extremely difficult to make under any circumstances. But in a fast-growing state like California during a period of rapid technological changes, such a venture borders on crystal ball-gazing.

We have attempted some industry employment projections which, in part, represent what we believe we must achieve if the total labor force grows as we have projected it, and if unemployment is to be kept low. Basically, California will need to provide more than a quarter of a million new jobs a year for the next 15 years.

We assumed an unemployment rate of 4 percent of the labor force and distributed the other 96 percent as employment among the various industry divisions on the basis of past trends, relationship of the industry divisions to one another and to the total, relationship to population, and anticipated future trends. There is an extremely large measure of subjectivity in these estimates which we regard as a first approximation.

We foresee a different pattern of industry growth for California than for the United States as a whole.

As in the nation, the largest relative growth is anticipated in the finance, insurance, and real estate division.

One of the major differences between the California and United States projected trends concerns manufacturing. In California, we anticipate that employment in this industry will have the second largest relative increase of any industry division, well above the average rate for all industries combined. The national projections indicate a less-than-average rate of increase for manufacturing. Despite the rapid rise of manufacturing employment in California in the past 15 years, this industry division still accounts for a smaller proportion of total employment than in the nation as a whole and than in a great many other states. As the population of the West continues to increase relative to the rest of the country, we expect an acceleration of the shift of manufacturing operations to California to serve the expanded population.

We see a continued sharp increase in government employment to provide services for California's growing population. In terms of relative gain, this division ranks third in California and fifth in the United States.

Another interesting difference concerns construction. Employment in this industry presently is at a relatively high level in California. Percentage changes from 1960, therefore, are measured

from a high base. In terms of percent change to 1975, construction ranks fourth among the various industry divisions in California as against second for the nation.

Our employment projections for services may be low. We rank this division fifth. In the national projections, it is third.

Trade is in sixth place in California and fourth in the United States.

The transportation and public utilities and mining divisions are expected to have less-than-average increases in the next 15 years in both California and in the nation.

The percentage increases between 1960 and 1975 in projected total employment in California for each of the industry divisions are as follows:

	Percent
1. Finance, insurance, and real estate	86.5
2. Manufacturing	77.5
3. Government	75.7
4. Construction	75.3
5. Services	73.3
6. Trade	72.7
7. Transportation and utilities	46.8
8. Mineral extraction	14.3
Total[a/]	67.5

a/ Includes agriculture.

It is not possible to make direct comparisons between California and the nation as to percentage change for the individual industry divisions because the United States' projections of employment by industry are in terms of nonagricultural wage and salary workers. The California estimates are in terms of total employment.

Employment by occupation

We do not, as yet, have any projections by occupations. We foresee very much the same trends for California as for the country as a whole -- largest increases in the occupations requiring the most education and training.

Because of the heavy concentration in California of activity in the space-age technologies, we may have an even greater rise relative to the United States in the number employed in professional and technical occupations.

Summary

The projections of the California labor force to 1975 seem very steep; and they are steep. If the projections are realized, it will mean an addition to the California work force of more than 4¼ million -- equivalent to the combined present total population of 8 of our states.

The projected rates of increase for both population and labor force are quite in line with past experience. Since 1860, California's population has been doubling about every 20 years. The projected population growth for 1960-1975 is at a somewhat lower rate than this.

Labor force growth has paralleled population growth. The projected rate of increase in the California work force is also less than the rate of growth of the past.3/ Our labor force projections rest squarely on population projections which were made independently of estimated labor force growth. If the population projections prove wrong, the labor force estimates will be wrong.

The opposite approach, being taken by the Stanford Research Institute, is first to project the probable growth of the labor force by estimating future employment. The Institute is making studies of past trends and prospects for individual industries. On the basis of these industry employment projections, the total labor force will be estimated.

These labor force estimates will then be the basis for projecting the population of California. The results of the Stanford Research Institute study have not yet been released. It will be interesting to compare their projections with ours.

In our approach, we say, "If there will be this much population, there will be this much of a labor force." The Stanford Research Institute method says, "There probably will be this much employment and, consequently, the population will be of this size."

We have taken the much simpler approach. The great difficulty in attempting to project employment opportunities is how to allow for the unknowns -- the new industries which will inevitably spring up in the next 15 years.

I am certain that no matter how wide our projections may be of the mark, there will be startling differences between what will happen in California in the next 15 years and what will happen in the nation as a whole.

FOOTNOTES

1/ The State Interdepartmental Research Coordinating Committee was established in 1945 by an executive order of the Governor. The members, appointed by the Governor, are the heads of the research and statistics agencies of the various departments of State government. For an outline of the committee's functions and organization, see The American Statistician, June 1957, Vol. II, No. 3, p. 4; also California Statistical Abstract, 1958, State Printing Division, Sacramento, 1958, p. 1.

2/ Population and Labor Force Projections for the United States, 1960 to 1975, U. S. Department of Labor, Bureau of Labor Statistics, Bulletin No. 1242.

3/ For an excellent study of population and employment trends, see Employment Expansion and Population Growth, The California Experience, 1900-1950 by Margaret S. Gordon, University of California Press, Berkeley, 1954.

ESTIMATED LABOR FORCE OF CALIFORNIA
1950-65, 1970, 197

(In thousands)

Year (annual average)	Total	Civilian
Estimated:		
1950	4,687	4,514
1951	5,099	4,722
1952	5,415	4,971
1953	5,561	5,141
1954	5,543	5,202
1955	5,738	5,402
1956	5,998	5,664
1957	6,172	5,851
1958	6,293	5,973
1959	6,420	6,100
Projected:		
1960	6,631	6,331
1961	6,860	6,560
1962	7,096	6,796
1963	7,341	7,041
1964	7,596	7,296
1965	7,859	7,559
1970	9,302	9,002
1975	10,909	10,609

Source: Division of Labor Statistics and Research, Department of Industrial Relations and Research and Statistics Section, Department of Employment.

ESTIMATED CIVILIAN LABOR FORCE OF CALIFORNIA,
BY SEX, 1950-65, 1970, 1975

(In thousands)

Year (annual average)	Total	Male	Female
Estimated:			
1950	4,514	3,169	1,345
1951	4,722	3,287	1,435
1952	4,971	3,430	1,541
1953	5,141	3,552	1,589
1954	5,202	3,584	1,618
1955	5,402	3,706	1,696
1956	5,664	3,869	1,795
1957	5,851	3,973	1,878
1958	5,973	4,038	1,935
1959	6,100	4,105	1,995
Projected:			
1960	6,331	4,242	2,089
1961	6,560	4,382	2,178
1962	6,796	4,526	2,270
1963	7,041	4,675	2,366
1964	7,296	4,830	2,466
1965	7,559	4,989	2,570
1970	9,002	5,904	3,098
1975	10,609	6,953	3,656

Source: Division of Labor Statistics and Research, Department of Industrial Relations and Research and Statistics Section, Department of Employment.

Discussion

Joseph W. Garbarino, University of California, Berkeley

As the sole discussant of the three papers presented on this panel, I find myself confronting a formidable task, and one that is made more difficult by the nature of the presentations. Not only are they reports of studies and analyses that have been carefully formulated and critically analyzed by experienced professionals before being undertaken, but the reports themselves are tight presentations that leave few loopholes open to criticism. As a result, my plan of action will be to address myself not only to specific points in their papers but to related questions that are suggested by them.

First, with regard to Mr. Myers' paper -- I was interested to see him point up the "puzzling" character of the labor force developments of 1960 and the immediately preceding years. These seem to center around the behavior of the labor force participation rates of various age-sex groups in the population and to suggest that the validity of extrapolations of trends in participation rates may be in question. In the past, peacetime participation rates have seemed to be changing slowly and consistently but it is always possible that this pattern may alter in a relatively short period. In view of the substantial changes in the occupational and industrial composition of the demand side of the labor market, it may be that a modification of past trends is in prospect. Changes in the type of work to be done, the locations of this work and the conditions under which it is performed also may induce a different pattern of labor force behavior in the future. Labor force analysis tends to treat participation rates as independent of the age-sex distribution but this may not be true in the long run.

A comment on Mr. Myers' section on international comparisons of unemployment rates seems appropriate. We are all aware of the problems of differences in unemployment estimates that have arisen in the past in the United States. I am dubious of using the Monthly Report of the Labor Force techniques to check the comparability of the unemployment statistics of foreign countries with our own. Particularly at low levels of unemployment, a substantial part of the total is fairly elusive. As a concept, the more inclusive definitions of unemployment make it almost a cultural phenomena, and international comparisons based on these definitions seem to me to be suspect. Let me remind you that the MRLF count of the unemployed in the Census month in 1950 produced a substantially higher unemployment rate than did the Census enumeration in the same month. While the sources of the difference are not particularly mysterious, it serves as an illustration of the problems involved.

Turning to Mr. Levine's paper -- I agree with him that the unemployment record of the 1950's has some omnious implications for the 1960's, and I agree with his forthright prediction that the problem of unemployment will attract increasing attention in the coming decade. This might well be true in any event but the recent and the impending changes in the size and composition of the labor force make it even more likely.

Even if the character and the dimensions of the problem were to remain unchanged from the 1950's, the program of data collection on the characteristics of the insured unemployed that he describes would be of major importance. If a reference to one of my more neglected works is permissible,[1] in 1954 I worked up a paper analyzing the industrial and geographical distribution of unemployment revealed by the 1950 census data on the assumption that the nature of the unemployment problem in the future would be one of special areas and industries. Of necessity, the study relied on data collected in one specific month some years earlier and could only illustrate the problem in a very special context. The BES data now being collected represent a major advance and will be of great significance in years to come.

One comment might be made. Mr. Levine speaks of using the data in various ways, most of which center around the development of programs to raise the level of employment. Another aspect might be stressed. Our unemployment compensation and other security programs are being liberalized in a variety of aspects. As the benefits of the programs rise above the relatively low levels of a few years ago, it becomes increasingly important to tailor these programs to meet specific needs of specific groups of the unemployed in order to economize on resources, to control abuse and to realize the purposes of the programs. This will call for more information of the work history type as well as what Mr. Levine calls the "snap shot" approach. Past earnings records and the degree of attachment to the labor force are two obvious types of information that will be needed.

Mr. Gershenson's paper on the California labor force is of particular interest to many of the audience. Here I would like to repeat a point made in commenting on Mr. Myers' paper. Such a projection requires the use of labor force participation rates whose trends may not be as clear cut as we have tended to assume. As an example, rates are almost certainly going to reflect any change in the work week that may occur and it is difficult to believe that a major change will not occur before 1975 if the basic assumption of high employment is realized.

This point aside, forecasting for California is even more precarious an occupation than forecasting or "projecting" in general. In the case of California, this type of analysis depends on the projections of inmigration that are adopted. Mr. Gershenson has adroitly avoided this problem by using the population forecasts of another state

agency. This probably is the crucial question in the procedure. To what extent is immigration dependent on economic growth rather than vice versa? Some persons, including our chairman, Margaret Gordon, would argue that immigration is dependent on economic expansion to an important extent. If Mr. Levine's forebodings about the buoyancy of the labor market turn out to be correct, the volume of immigration might well reflect this fact.

Once Mr. Gershenson adopted the population estimate, calculated the age-sex distribution and applied the participation rates, he was left with a monumental labor force to allocate by industry. Part of the solution was to project an expansion of manufacturing employment that was large both absolutely and relative to the trend for the United States as a whole. Prediction of a major expansion in California manufacturing has always seemed on shaky ground. In the past, California has lived on income generated by federal expenditures, extractive industries such as oil, lumber and agriculture, and tourism. Except for the non-market demands for aircraft, missiles and perhaps ships, California has always found it very difficult to compete in national markets for manufactured products. Major changes in production and transportation conditions are likely to be required before this situation is modified.

Once again, let me say that these papers were stimulating and useful and, as has been demonstrated, difficult to criticize on their own premises.

1. J. W. Garbarino, "Some Implications of Regional and Industrial Differences in Unemployment," Reprint No. 75, Institute of Industrial Relations, University of California, Berkeley, 1955.

X

METROPOLITAN AREAS AND URBAN GROWTH - I

Chairman, Harry Sharp, University of Michigan

Basic Economic Materials for Metropolitan Research — John I. Griffin, The City College of New York

The Metropolitan Area and External Lines of Communication — David L. Glickman, The Port of New York Authority

The Flight to the Suburbs Slackens — A. F. Parrott, Consolidated Edison Co. of New York

BASIC ECONOMIC MATERIALS FOR METROPOLITAN RESEARCH

John I. Griffin, The City College of New York

The shift in the United States from a predominantly rural to an urban people has been comparatively recent. It was not until the census of 1920 that over half of the population was reported as urban. The increasing urbanization in the last forty years has given rise to three great waves of research activity concerned with urban problems and having as its principal focus the economic aspects of urban growth as contrasted with the purely social or political aspects. The first research wave, which followed the dislocations of World War I, was closely identified with the city planning movement. The economic base studies were largely designed to provide a scientific foundation for planning urban growth. The objective was stated in the monumental Regional Survey of New York and Its Environs, published in the late nineteen twenties, as "How can the American City obtain beauty and order without too much constraint of individual liberty and too much interference with rights of persons and property? How can it obtain these things without impairing, but rather while improving, the quality of the home, the spaciousness of the parks and playgrounds, the efficiency of industry, the relief of traffic congestion and the equitable distribution of advantages and opportunities for all its citizens?" The second research wave was a consequence of the Great Depression and the expansion of the role of the Federal Government in the economic life of the nation. The National Resources Committee in its report on Our Cities, published in 1937, found that "the city has become one of the primary problems of the Nation's economy....it is in the Nation's cities that the shadow of economic insecurity is darkest....Subject to continuing unemployment, lacking the rural reserves of shelter and subsistence, the city worker is seriously handicapped in the struggle for existence." The most recent research wave, and the most massive, has been the proliferation of urban research in the recent past, influenced by the acute problems of central city decay rather than growth, suburbanization, transportation inefficiencies and population group shifts. The post World War II phenomenon of abundant research grants has caused an unprecedented amount of support for urban research on all levels. Much of this current research effort is descriptive and the researchers, while better scientists, are not fired by visions of the "city beautiful" or the "common adventure in pioneering on the frontiers of a new social world" as motivated the urban research of the nineteen twenties and the nineteen thirties. In his foreward to the New York Metropolitan Region Study, directed by Raymond Vernon, Edward S. Mason wrote "As a study of these underlying currents, this project is neither a blueprint for action nor an analysis of metropolitan government. It has no recommendations to make about the physical structure of the Region or about the form or activities of the governmental bodies there,,..Its end product is an analysis of the Region's probable development, assuming that the economic and demographic forces in sight follow their indicated course and assuming that the role of government is largely limited to existing policies."

Urban research has perennially been beset by the lack of suitable statistical data and much of the resources of researchers has been absorbed by the necessary but yet preliminary task of obtaining data. The study, published in 1927, by the New York Regional Plan, on Major Economic Factors in Metropolitan Growth and Arrangement, remarked that "The unique character and scope of the survey has involved experimentation in methods of collecting and analyzing facts at all stages." The most important single source of information utilized as the records of the factory inspection departments of the three states, New York, New Jersey, Connecticut. Some nine industries were surveyed in detail. No subsequent research project has made use of the factory inspection records. In the National Resources Committee report it is lamented that "the available information regarding cities grows less and less. There has been a sharp decline in the reporting of urban information since 1890, and in consequence we know comparatively little about cities at a time when the need for accurate, complete, and fully analyzed data is most urgent." One of the recommendations of the National Resources Committee was that "the Central Statistical Board should give special consideration to the inadequacies in the existing urban data and the shortcomings in the methods of collecting them, and the possibilities of a program for collecting such important census data as place of work or place of daily activity as well as place of residence, and such important current information as employment and unemployment." While the reports of the New York Metropolitan Region Study, now being published, reflect the substantial progress in basic urban data, the gaps are still evident and a number of key sets of data are estimates prepared on a "one shot" basis by the staff of the study. The amount of reworking of the existing statistical data without any contribution to the improvement of basic data or the originating of new data, raises a fundamental question that must concern the statistical profession. The very growth of metropolitan and urban research projects which replow the same infertile fields and leave no lasting mark upon the quantity or quality of currently available statistical data, serves to raise the question of the optimum allocation of research resources. If the third wave of urban research should recede in the nineteen sixties without a continuing contribution to urban data, a great opportunity will have been lost.

There is, fortunately, some evidence that statistical data for states, cities and metropolitan regions are being assembled and published on a more systematic basis than ever before. It is particularly important that these developments are taking place under local sponsorship rather than as an outgrowth of Federal Government programs. On this basis local interest is more likely to be sustained. In fact pressures for continued support of local statistics, which are a by-product of national programs, are likely to be politically most effective when expressed by local interests. The vigorously expressed demands for continued publication of certain area data for the New York Metropolitan Region in con-

nection with the 1960 Census produced desired results. The economist of the New York City Department of Commerce and Public Events recently noted critically the fact that the Economic Report of the President, prepared by the Council of Economic Advisors, presented almost no facts on regional and metropolitan centers. Such a policy in respect to metropolitan statistics "complicates the efforts of mayors and governors to enlist public support for a policy of determined self-help where communities are threatened with industrial emigration" said Martin W. Wilmington. Another interesting example of the expression of local support for the publication on a continuing basis of local statistics is the criticism of the policy of the Federal Reserve Bank of New York to emphasize international and national data in its Monthly Review. The success of the Bureau of Labor Statistics - State employment statistics program is no doubt a consequence of the use of cooperating state agencies in the program. This BLS - State program now originates much of the fundamental economic statistics for metropolitan areas and cities. A number of statistical abstracts are being published by state agencies, among which may be mentioned the Pennsylvania Statistical Abstract, which contains 105 tables and 22 charts. Many European cities have long issued statistical yearbooks, which are comparable in the number of tables and in the variety of subject matter to those included in the Statistical Abstract of the United States. The City of Birmingham, England, publishes an Abstract of Statistics which contains 182 tables. Three years ago New York City began the publication of a Statistical Guide which brings together assorted statistical information ranging from the number of specimens in the New York Aquarium to the number of telephone calls made to WEather 6-1212. This publication should develop into a statistical abstract.

The publication of an increasing variety of urban data on a continuing monthly basis has a great impact because, once a series is established and widely disseminated in an area, its continuance is reasonably assured. In terms of impact upon metropolitan research, resources allocated to a program which will develop and publish continuing series, probably have greater value at this stage than any other form of statistical program. Since so much that is important in economic analysis is necessarily possessed of the time dimension, adequate monthly data for cities and metropolitan regions are essential. Of course, the research bureaus of many universities throughout the country as well as the Federal Reserve Banks and commercial banks have long issued valuable regional and local indexes and other statistical series. The Pittsburgh Business Review and the Monthly Summary issued by the Security First National Bank of Los Angeles are excellent examples. Arthur H. Cole in his guide to business indexes lists 150 regional and local measures of change. No doubt, since 1952, more have been published. However, the assembling and systematic publication of monthly statistics for states or cities extends beyond the preparation of a business index and ultimately would result in the publication of a local counterpart to the Survey of Current Business. In fact, the New York State Department of Commerce at Albany has published since 1945 Current Business Statistics. The latest issue of this publication includes: general business indicators, production statistics, domestic trade, foreign trade, employment, hours and earnings, construction and real estate, income, savings and finance, price levels, and miscellaneous measures. This monthly publication contains more than 250 statistical series obtained from 31 different sources, both public and private.

Organized efforts to develop current publication of city and metropolitan area economic data are proceeding in many parts of the United States. The sponsorship of such efforts frequently involves local chapters of the American Statistical Association. In the New York Area Chapter an active Committee on Metropolitan Area Statistics has long been functioning in this field. This committee publishes a monthly table of New York Metropolitan Area Economic Indicators, which shows 18 statistical series. Because of the complexity of the New York Metropolitan Area these series, where available, are shown for New York City, Nassau and Suffolk Counties, Northeast New Jersey, Westchester, Rockland and Fairfield Counties. This monthly table is set up by Waite S. Brush of the Consolidated Edison Company. It contains a number of series which, without the activity of the local committee, would not be available for general publication. Such series include total originating telephone calls and construction data from the F. W. Dodge Corp. This table is received monthly by about 200 organizations interested in metropolitan area statistical data. The committee of the American Statistical Association Chapter has prepared the draft of a factbook for the New York Metropolitan Area and plans for publication are under discussion. As is true in most major cities, the local public utilities are a strong support for the New York publication. Special analyses have been circulated showing possible uses of the electric energy sales figures as a measure of change in industrial and business activity. It would appear that, as the monthly bulletin of the Federal Reserve Bank of Chicago indicated in its March 1960 issue, a series on production worker man-hours and kilowatt hours utilized by manufacturing firms may together dependably measure short-term changes in manufacturing activity in local areas. In addition to the development and distribution of the monthly table of Economic Indicators, the Chapter of the American Statistical Association was instrumental in the creation of a statistical program for the City of New York which, although now inactive in the developmental sense, led to a re-examination of procedures for estimating the current population of the city and the establishment of a new continuing quarterly series of reports based on sales tax data.

The City of New York, through its Department of Commerce and Public Events, began the publication of New York City Monthly Statistics in March 1958. This publication now contains information on employment and earnings, construction and real estate, retail business activity, general business activity, transportation, Port of New York, and miscellaneous series, cost of living, population trends, weather summary, and activity of city government agencies. There are over 200 individual series assembled in this monthly publication. Clearly this effort is a significant contribution of basic economic materials in respect to the New

York City economy. A great proportion of the data usually published by cities and countries consists solely of "workload" statistics, which are not of general interest. Since 1932 the Cleveland Real Property Inventory has conducted an annual census of dwelling units and other buildings by census tracts. Various commercial interests, often in cooperation with local city planning agencies, conduct property inventories but the results are usually not on a continuing basis. In the search for more city and regional economic indicators, some new series have been developed. For example, help-wanted ads in five New York newspapers, is one of the series published in New York City Monthly Statistics. The help-wanted series shows considerable sensitivity to the business cycle, and to labor market conditions. The B. K. Davis & Bro. Advertising Service of Philadelphia publishes a monthly report called the Help-Wanted Trend.

In order to obtain the data necessary for continuing publication of basic economic data for cities and metropolitan areas, changes will be needed in existing statistical procedures and the coordination of much that is now done by distinct public and private agencies. There is little doubt that for most metropolitan areas, given suitable coordination, monthly reports, similar to those which have been discussed, could be assembled and published. In the staff study prepared for the Joint Committee on Washington Metropolitan Problems, Stuart A. Rice and his associates in the consulting firm Survays & Research Corp., proposed the establishment of a statistical agency to have the following functions: (a) responsibility for the continuing development of a statistical program for the area; (b) coordination of relevant area data now collected by Federal, State and local agencies of government and by private firms and organizations; and (c) maintenance of a central repository and index of statistical information relating to the area. Such an agency would no doubt be appropriate in those metropolitan areas which do not have any recognized focal point for such centralized statistical work. The emphasis in the study of a metropolitan statistical program for Washington was, quite properly, upon the need to have a continuing statistical unit, not just another research project. This study identified the particular problems of transportation and highways, water supply and pollution, and economic development. The statistical unit, however, would function as a central data clearing house.

A careful inventory of immediately evident as well as hidden statistical data will frequently reveal that the statistical gaps are not as large as sometimes thought. Since the development of new statistical series to fill gaps is usually a very costly procedure, inquiry should be made to determine the possible existence of data which may serve to answer the research questions. Perhaps the questions which may be answered by new data may be trivial and not worth answering in light of total available data. Ingenuity in solving problems by discovering data and using previously unexploited sources of data is as necessary a qualification for the metropolitan research worker as the knowledge of specific statistical techniques. Unfortunately, statistical detectives are scarcer than statisticians. Hidden data, in certain areas, have been successfully exploited. For example, in the making of population estimates between census dates, much experience has been gained in using such indirect sources of information as school enrollment figures, residential electric meters, rapid transit turnstile counts and the like. Alfred F. Parrott's analysis of the slackening flight to the suburbs is an example of the use of data from a variety of sources to detect a significant new trend. While considerable resources could be expended in the development of direct evidence in respect to his hypothesis, it is probable that the conclusions he reaches would not be substantially altered. Frequently data can be developed in respect to characteristics of an area which for a combination of reasons can not be determined directly, thus a recent study of the estimated Jewish population of the New York area, used the Yom Kippur method for estimating the size of the Jewish population. This involves measuring the drop-off in school attendance between a normal school day and on the Jewish holyday. Other types of ethnic data may be developed from information on burial certificates. In some cities, while crime data are not classified by race, statistics on prisoners in jails will provide this information. In a study of the completion of new residential buildings, where completion dates could not be determined without expensive field studies, it was found that the city department of water supply dept a daily record book showing water connections which were made only when the builder reported completion of the structure. Unless a trained statistician saw these data in the context of the problem, these records would have remained a purely administrative record. The statistician or economist may not possess, as part of his professional equipment, the ability to know where to look, to recognize what he has found and to place the newly discovered data in the proper relationship to other basic materials. If the research staffs working on metropolitan research projects throughout the United States gain this type of knowledge, then it will no longer be possible to say, as did Jervis J. Babb, of the Committee for Economic Development, that "we do not know very much about the economies of our 'little economies'. Facts are inadequate, statistics obscure or incomplete, the framework fuzzy."

References

Cole, Arthur H., Measures of Business Change, Chicago, Irwin, 1952.

Congress of the United States, Joint Committee on Washington Metropolitan Problems, "A Metropolitan Statistical Program for the National Capital Region," Dec. 1958.

Griffin, John I., Industrial Location in the New York Area, New York, (City College Press) Arco Publishing Co., 1956.

Hoover, Edgar M., and Vernon, Raymond, Anatomy of a Metropolis, Cambridge, Harvard, 1959.

Perloff, Harvey S., Regional Studies at U. S. Universities, Washington D. C., Resources for the Future, Inc., 1956.

THE METROPOLITAN AREA AND EXTERNAL LINES OF COMMUNICATION

CONTAINERSHIP OPERATIONS: A PRELIMINARY EVALUATION OF AN EVOLVING FORM OF OCEAN TRANSPORTATION

David L. Glickman, The Port of New York Authority

Few recent developments in the fields of transportation and communication have aroused as much discussion and concern as the trend towards integration of different modes of transportation in the form of piggyback for land movements and containerships - or fishyback - for oceanborne movements. In both instances, the key to the development of these services is the trailer portion of the conventional tractor-trailer truck form of movement, with or without the chassis. In the piggyback operation, truck and rail movements are integrated by separating the trailer from the tractor and then placing the trailer on a railroad car for movement to its destination either directly or by a subsequent tractor hook-up; in the containership operation the container or trailer is similarly removed and placed aboard ship for subsequent movement. In both instances the cargo, once placed in the container at point of shipment or at consolidation station is not rehandled until it reaches point of destination except for less than container or trailer loads which must be segregated prior to ultimate delivery to destination.

These two related developments, or perhaps better said different aspects of the same development, already give promise of major changes in the economics of transportation and in the distribution of cargo volumes among the major ports of the country. This afternoon, I should like to explore with you in some depth one of these forms, namely, the containerization of cargo for oceanborne movements. I should like first to discuss the different forms of containerization of oceanborne cargo and then to attempt a preliminary evaluation of the outlook for this integrated mode of transportation in the domestic coastwise and in foreign trades.

The basic materials for my presentation are derived from an extensive study now underway at The Port of New York Authority. Here it must be noted that although this study was in the first instance undertaken to measure and evaluate the impact of container and containership operations on the Port of New York and on its future requirements for marine terminal facilities, the nature of the problem is such that both the problem and its potential impact are equally applicable to other ports, including those here on the Pacific Coast. In fact, of the ports which until now have been most directly concerned with either actual container or containership operations or in plans for such operations, three are located here on the Pacific Coast, viz., San Francisco, Seattle and Los Angeles. On the Atlantic Coast, the port most directly involved has been and is the Port of New York.

FORMS OF CONTAINERSHIP OPERATIONS

Until now, I have used the terms "containerization" and "containerships" without differentiating between the two. In either case, the cargo to be transported is what is designated in the transportation industry as general cargo or general merchandise, rather than bulk cargo. The former is cargo which is generally packaged, crated, bundled or boxed; the latter is cargo which generally moves in bulk and loose form over specialized facilities.

Not all general cargo, however, is necessarily or feasibly containerizable. While no hard and fast lines can or should be drawn, it may be said that in order for cargo to be physically and economically containerizable, it must generally possess certain characteristics with respect to stowage factors, dimensions, value, susceptibility to damage, breakage and pilferage, size of shipment and materials handling techniques.

There is however, a fundamental distinction that needs to be drawn between the possibility of containerization of cargo and the development of full containership operations. Containerization per se involves the movement of containers of varying sizes as substitutes for another type of packaging. The size variations are influenced primarily by the composition of the cargo and the extent of consolidation and break bulk, and direct point to point movements. As a general proposition these containers do not exceed 8 or 10 feet in overall length and are frequently smaller. Many steamship lines, including Bull, American President, Moore-McCormack, United Fruit, etc., use some containers at the present time. They handle these containers in the same manner as they handle any other type of boxed or crated cargo, over their regular facilities at the ports they serve.

Containerships on the other hand, involve the construction and use of ships which are designed to carry only containers. Generally, these ships are cellular in construction and cannot accommodate any other type of cargo, except for one qualification which will be noted shortly. The containers used in such operations are always of trailer size, designed to fit the chassis of the tractor-trailer, and the cells in the ship are designed specifically to accommodate them. Here too, there are variations in the size of the containers, these variations stemming primarily from the lack of uniformity in trucking standards and also, in one instance that I am familiar with, from an extensive study on the optimum size of containers best suited to meet the requirements of a particular trade

route. I refer here to the study conducted by Matson Navigation Company to determine the optimum container size for its Pacific Coast-Hawaii trade which, after evaluating container lengths ranging from 12 to 40 feet came to the conclusion that a 24 foot container would best meet its requirements. The two most frequently used sizes, however, are 17 and 35 foot containers. In this context, it is pertinent to note that the lack of uniformity of trailer sizes has been a matter of some concern to the trucking industry generally, and that the American Standards Association recently recommended that sizes be standardized in 10, 20 and 40 feet modules.

In between the container and containership there is a transitional form of operation in which the ship is designed to handle both conventionally packaged or crated cargo and containerized cargo. In such instances, the design of the ship may be conditioned by the requirements of the trade route served by the shipping operator, as in the case of the planned Grace Line Service between New York and the West Coast of South America, or may, as in the Matson case, be conditioned by the capital investment in existing ships and the need to phase out the construction of entirely new ships in light of the operator's ship replacement program.

Where an operation begins from "scratch," with no investment in existing fleet, this latter consideration does not apply and the operator can begin with either fully converted ships or entirely new ships. This was the case with two coastal operators out of the Port of New York, Sealand Service Inc. and Erie and St. Lawrence Corporation. Grace Lines, however, on both its North Coast South American run for which two fully converted containerships are already available and for its planned West Coast South American service, for which three containerships are to be constructed, had to phase out these ships in accordance with its overall ship replacement program. Here on the Coast, Matson, which also has substantial capital invested in its existing fleet of ships, began its container service in 1958 with deck loads of containers on six of its standard cargo ships, and then progressively brought in a converted combination container-bulk sugar ship[1/] and an entirely reconverted containership, each of cellular design, with more of each type planned.

The containership itself may be of one or two types: a roll-on/roll-off or a lift-on/lift-off. The first, the roll-on/roll-off type requires that the container be <u>driven</u> into and aboard the ship and placed in position. The container may first be detached from the chassis and carried into the ship by special capacity fork-lift trucks, as in the Erie and St. Lawrence operation, or the container resting on its chassis may be driven on to and positioned by a tractor unit, as in the military application of containership operations. The second, the lift-on/lift-off operation, on the other hand, requires that the container be detached from its chassis, lifted by specially designed gantry cranes and then positioned in a cell aboard the ship. In the first of these types, the holds, the interior cargo areas of the ships, are horizontally open to provide for the vehicular movement; in the second, the holds are fitted with the cells, arranged vertically to provide for tiering of the containers.

One comment is in order here on the military application of the containership and the specific form which it takes. Military logistics planning for both peace and war conditions includes a requirement that about 25 per cent of the total cargo volume be wheeled cargo. This includes tanks, artillery and other field pieces, construction equipment, trucks, staff cars, jeeps, etc. Under conditions of active military conflict roll-on/roll-off operations are deemed essential in order to provide for speed of movement and thus a reduction of time required in port or unloading area. Such equipment can be rolled off the ships in the various theatres of operation by ordinary military personnel, whereas crated or containerized cargo requires longshore operations. Military sources stress that since the roll-on/roll-off type is considered essential for combat operations, this criterion has, therefore, been accepted as the norm for military peace time operations. This judgement is strengthened by the fact that the cargo composition of the military movements even under peace time conditions also includes a 25 per cent ratio of wheeled to total cargo. Stated differently, the military view is that unless these operations can be provided for during peace time they will not be available during war time conditions, hence the clear emphasis on roll-on/roll-off rather than lift-on/lift-off operations.

Turning back to the commercial application of containership operations, it may well be asked what benefits it offers over existing forms of ocean shipping, and by whom these benefits are derived. From the shipping operator's point of view the key factors are (1) that approximately 50 cents of each revenue dollar is expended for cargo handling purposes on dock and aboard ship - this is a rough average for the industry as a whole, and (2) that the time required to load and unload a ship limits the number of round trip voyages - and hence utilization - of a ship or a fleet during any given period of time and thus increases the number of ships required to service a given trade route or complex of routes.

To illustrate these factors, consider the following pattern in existing operations. At the

---

1/ This ship was originally a bulk ore carrier, which was lengthened by 70 feet by insertion of a new mid-section. As is noted subsequently, (page), there is a heavy sugar movement from Hawaii to the Pacific Coast. The ship is therefore designed to serve a double duty purpose: carriage of 205 loaded containers in each direction and approximately 16,500 tons of sugar eastbound. The cargo hold areas utilized for each of these movements are compartmentalized, and are not interchangeable.

present time, the typical shipment is handled and rehandled from 9 to 15 times between shipper's point and consignee's location, with most of these handlings taking place at the steamship operator's terminal facilities and aboard ship. The average rate of loading and unloading a ship is in the 450-500 ton range per day in port. Keep in mind that this is an overall average and that there are of course individual operators whose cargo handling rates are substantially higher. And there are also those whose rates are lower. In any event, it takes approximately 9-10 eight hour working days to load and unload 5000 tons of conventional general cargo at the present time. A ship serving a trade route requiring 20 days round trip sailing time thus requires 40 days for a full round trip, assuming that the cargo handling rate at both ends is similar and that 5000 tons are handled at both ends. On an annual basis, this is equivalent to approximately 9 complete round trips.

The ship operator's efforts to reduce costs therefore are focused on two areas: reduction of cargo handling costs and reduction of ship's stays in port. Both these objectives can, it is believed, be achieved by containership operations. The same 5000 tons of cargo, when containerized, can be loaded and unloaded within two working days, reducing the ship's stay in port and making possible a faster complete round trip turn-around. The same round trip will now take 24 rather than 40 days, thus increasing the number of sailings which a ship may make per year to 15 as compared with 9 for the conventional ship. This in turn, has the effect of producing a corresponding increase in the ship's annual cargo carrying capacity and revenue earning power.[1] On a fleet basis, this example may well result in reducing the number of ships required to serve a given trade route or complex of routes. The particular experiences of different steamship companies will vary because of somewhat different ratios in ship utilization, cargo carrying capacity and earning power depending on a variety of operational characteristics.

As against these potential savings in operating costs, there must be equated the capital costs required to put such an operation into service, including not only those for the ships but also for the containers and ancillary equipment, and the additional maintenance costs involved in such operations. Together, these run to substantially more than the capital and maintenance costs of a conventionally designed cargo ship. This is a simplified - in fact oversimplified - version of the comparative cost structures of the conventional versus the containership type of operation and should obviously be used with caution.

From the point of view of the shippers and consignees of cargo, stress is most frequently placed on the following three potential benefits: 1) delivery in the least possible time, 2) delivery in the best possible condition, 3) delivery in the least expensive manner.

The first is made possible by the fact that total shipment time is considerably reduced. Many of the multiple handlings that are necessary under current conventional methods are eliminated by the simplified handling that is possible with the container service. Faster loading and unloading of cargo reduces total elapsed time. In addition, truck waiting time in making delivery to or pick-up from steamship berth is almost entirely eliminated.

With respect to the second point, namely, arrival of cargo in best possible condition, the elimination of the multiple handlings reduces the possibility of damage or other losses resulting from mishandling, bad stowage, weather, contamination, etc. For refrigerated cargo, additional benefits may be derived from the constant temperature control made possible with sealed containers equipped with refrigeration units.

Reduction in costs come about as a result of less rigid and bulky packaging requirements which, in turn, reduce both overall weight and cubic measurement which are the basis for assessing shipping charges. In addition, elimination of truck waiting time and reduction of losses, results in additional economies. Reduction in losses may in turn, lead to lower insurance premiums. While insurance rates are based upon actual loss ratio experiences, it is believed that some reductions in premium rates may be obtained for door-to-door movements of sealed containers, and, as experience improves, for less than container load movements. Further, there is also the possibility that in foreign trade the containers will be considered as integral elements of the ship and their weight not made subject to customs duties. This is in contrast with current practices under which the entire weight of the shipment, including the packaging, is frequently subject to duties assessed on a weight basis.

PROSPECTS OF CONTAINERSHIP OPERATIONS

What now can be said of the outlook for containership operations in the oceanborne trade of the United States? Here it is necessary to begin with a clear differentiation of the two major segments of what has until now been called "oceanborne trade." These are the domestic coastwise trades - note the plural because there are three rather dis-

---

1/ This is based on the assumption that the containership is designed to carry an equivalent amount of cargo as the conventional ship. A conventional ship converted to carry containers would necessarily have a smaller cargo carrying capacity because of the loss of space required for the cellular construction and the broken stowage factor involved in loading the containers with cargo. Further, it must be noted that shipping charges are assessed on a stowage conversion factor of 40 cubic feet per long ton, with the shipping operator fully entitled to assess charges on either basis.

tinct components - and foreign trade. The three components of the domestic coastwise services are (1) coastal trade, that is trade within a single or adjacent coastal areas; (2) intercoastal trade between the Atlantic and Gulf Coasts on the one hand and the Pacific Coast on the other; and (3) the off-shore trade typified by trade between the Pacific Coast and Hawaii, and between the Atlantic and Gulf Coasts and Puerto Rico. While there are some differences in the factors effecting the development of containership operations in each of these components, there are nonetheless a number of significant considerations affecting all in common, which are radically different from those affecting the development of such operations of foreign trade.

Under existing law the domestic services may be provided only by United States flag carriers whereas in foreign trade shipping services are of course provided by both American and foreign flag carriers, the latter generally operating with lower cost structures. Further, while the domestic flag carriers operating in the domestic services are not eligible for either construction or operating differential subsidies, domestic flag carriers operating in foreign trade are eligible for both types of subsidies in order to enable them to compete with foreign flag operations. Third, whereas there are no customs and regulatory agency controls exercised in the domestic trades other than those which are generally applicable to land forms of transportation, such controls do exist in foreign trade and pose important operating problems. Finally, whereas in foreign trade the containership operation must meet competition only from conventional water carriers, both United States and foreign flag, in the domestic coastal and intercoastal services, the competition is primarily with the two established modes of land transportation, the railroad and the truck. While the rate structures of both ocean and land modes of transportation are subject to regulation, the controls over the domestic services, both water and rail, are subject to far more extensive and absolute controls.

Coastal and Intercoastal Trade

The development of containership services in the United States has until recently been confined almost entirely to the coastal trades, notably between the Port of New York and Florida and Gulf Coast ports and to a more limited extent on the Pacific Coast between Seattle and Alaskan ports. The latter, however, actually possesses more of the characteristics of an off-shore trade since there are no direct rail connections between the West Coast and the 49th State and highway connections are at best limited.

In both instances the institution of containerized services - full ship out of the Port of New York and container out of Seattle - benefited from experiences derived from an earlier and specialized form of containerized operations. This was and still is the transportation of rail cars in oceangoing vessels. In New York, there were two such services, provided by Seatrain and Newtex, the latter no longer in operation, the former converting increasingly to trailer carriage; here on the Coast, Alaska Steamship Company has for some years provided service for the water movement of rail cars but it too is turning increasingly to the movement of truck containers. In both cases the shift has been prompted by three major considerations: the increasing role played by the truck in land haulage of freight particularly for short and medium length hauls; the division of joint rail-water rates; and the increase in the cargo capacity of the ships resulting from the transport of containers as compared with rail cars which must necessarily be transported intact, complete with undercarriage.

Prior to World War II, there were substantial coastal and intercoastal services serving the Atlantic and Pacific Coast ports. At the Port of New York, it has been estimated that during the late 1930's these services accounted for about a third of the total volume of oceanborne general cargo. During the post-war period, however, the coastal services of the conventional general cargo carriers have all but disappeared, while in the intercoastal trade only one carrier, Luckenbach Steamship, currently operates a continuous two directional service. The remaining carriers in the intercoastal trade tend to be specialized carriers for the movement of lumber from west to east and iron and steel mill products in the reverse direction.

The re-establishment and expansion of coastal and intercoastal services, particularly of containership operations, depends on the ability to provide frequent and steady sailing schedules at equitable rates.

Seatrain in New York is already, as has been noted, converting increasingly to trailer-container movement. In addition, in 1956 an entirely new service was established by the McLean interests, now known as SeaLand Services Incorporated, to provide service to Florida and Texas ports. The measure of success of these two operations lies in the fact that there has been a combined increase of almost 700,000 tons in the past three years in the volumes carried by these two operators. This is of actual weight tons and not ton-miles.

There is some evidence to the effect that the utilization of these services has resulted in some shifts in source of supply - consumption market relationships, with the areas being served substituting in some measure for prior linkages with other sources of supply and consumption markets. As a result of these developments, one other carrier, Erie and St. Lawrence Corporation, has already had two entirely new ships constructed and is soon to enter into coastal containership operation between New York and Florida and South Carolina

ports.[1] And the outlook is that additional carriers will, in the not too distant future, also seek to enter into operations between New York and South Atlantic and Gulf Coast ports.

As for the possibility of instituting containership operations in the intercoastal service, it must first be noted that the comparative distances of haulage are roughly 3000 miles by land and 5260 nautical miles by sea to San Francisco through the Panama Canal. Two steamship companies, American-Hawaiian and again Sealand Services, have indicated that they would like to establish new services and have applied to the Federal Maritime Board for mortgage guarantees on funds to be borrowed for construction of their proposed vessels. Whether they actually will be established or whether Luckenbach will consider converting cannot now be answered with any definitiveness. These are possibilities but cannot at the present moment be viewed as probabilities.

Off-Shore Trade

Turning now to the third segment of the coastwise services, the off-shore trade, there seems little doubt that this will within a relative short time be converted largely to containership operations. In New York, such a service is already provided by SeaLand Services with a sailing schedule currently providing for three sailings per week in each direction. In addition, Bull Lines has for some years been carrying containers and is said to be considering a full containership operation.

Here on the Coast, Matson Navigation has already instituted deck load, combination container-bulk sugar and full containership movements between the Pacific Coast and Hawaii, with terminals in Seattle, Los Angeles and the San Francisco Bay area. Both the East Coast and West Coast services, it should be noted, benefit substantially from the fact that the trade with the off-shore areas served has two major components; inbound sugar trade which is not generally containerizable because of the handling problems of the commodity itself and secondly, other types of cargo in and outbound which are in large measures extremely suitable for containerization. It is for these reasons that it is suggested the off-shore general cargo trade, with the exception of sugar, will move almost completely in the direction of containership.

Foreign Trade

The future of containership operations in foreign trade is, as already noted, affected by complex of factors radically different from those in the domestic trade. Three additional factors of major importance must now be added. These are the characteristics of trade route composition, the differences in trucking and highway standards and systems abroad as compared with United States, and the length of ocean haul. It will be helpful to indicate briefly the manner in which these factors affect our problem at hand.

1. There is a tendency in some quarters to minimize the problems involved in customs and other regulatory agency controls as these affect containerized movements. While I personally believe that, in time, these issues will be resolved, it will be necessary to determine where, for example, the examination of containerized cargo for customs and health purposes is to take place, that is, whether at point of unlading or at the point of ultimate delivery, whether the weight of the containers is to be taken into account in assessing customs duties, whether charges are to be assessed on the containers for inland movements away from ports, etc. These issues are equally as pertinent here in the United States as they are abroad. Some of these issues have been resolved in the movement of containers across boundary lines in Western Europe. As I have indicated, I believe that these privileges will eventually be extended to transatlantic traffic, whether originating in this country or abroad.

2. While there are certain types of cargo which are generally considered as containerizable, the volumes of such commodity movements vary considerably from trade route to trade route and from country to country. Further, even when cargo, originating in any particular country abroad, is viewed as containerizable there are serious questions as to the probability of concentrating these movements in a limited number of ports. This is particularly important since it is accepted as a general criterion of effective containership operations that the number of ports served should be limited in order to make possible quick turnaround and hence increased carrying capacity of the ship itself, and to limit dispersion of the containers. As a case in point, cargo originating in Italy may be loaded in Sicily or in Naples, Genoa and at a number of Adriatic ports. Even assuming that there is a sufficient volume in the Italian-U. S. trade which would make possible a containership operation, there is some question as to whether any one of these ports would have sufficient volume to warrant such a service. The same is true for example, in France, with cargo being handled at Mediterranean, Atlantic and Channel ports. These examples could be multiplied many times. Further, since cargo ships generally service trade routes touching upon a number of countries, the cargo capacity of the entire route must be evaluated.

3. Trucking and rail practices and the highway systems in other countries will probably limit - at least initially - the expansion of containership operations in foreign trade. Re those countries which have adopted American trucking practices, e.g. in the West Indies, North Coast of

---

1/ First ship sailed from the Port of New York on August 19, 1960.

South America and the East Coast of South America, the possibility of ultimate adoption of containership services - based solely on this criterion - is more likely with these areas than with Western Europe, where trucking, railroad and highway standards are quite different from those in this country. This is not to be interpreted, however, as suggesting that some modification may not be brought about to make it more possible for the large size American containers to be used in Western Europe. The point is rather that this presents problems which should not be minimized.

4. The longer the ocean voyage the smaller the likelihood of development of full containership operations. There are three basic reasons for this: (a) the pattern of distribution of general cargo with the more distant trade routes is for these routes generally to have smaller volumes, e.g. Australasia, South and West Africa, the Indian Ocean area. In those instances where some of the more distant trade routes do have substantial cargo volumes, there tends to be concentration of these movements in a limited number of commodities. For such routes it is important to maintain flexibility in cargo hold space. An exception to this is the trade with Japan, which is both large in volume and diversified in character. (b) Ships serving the more distant routes generally service a substantial number of ports. To service all of them by containership would require extensive capital investment and tend to disperse the containers over wide areas without the concentration which is deemed necessary for successful operation. (c) The longer the trade route the less the possibility of achieving major savings in turn-around time and hence an increased annual carrying capacity of the ships. As the length of voyages increase the economies attained in reducing stays in port are proportionately reduced. For example, a trade route requiring 40 days round trip sailing time benefits less from reduced port stay than a voyage requiring 20 days sailing time. Where the optimum point is reached with respect to length of voyage cannot be pinpointed with any degree of certainty. Each trade route will have to be evaluated in terms of its own characteristics.

If, however, I were asked to speculate as to the trade areas which are likely to be candidates for the institution of either complete or part containership services, I would suggest that the most likely candidates would be - and here I trust you will understand why I must limit myself to the Atlantic side of the country - (1) the North-Atlantic-Caribbean routes; (2) North Atlantic-North Coast and West Coast South America; and (3) the several North Atlantic and Gulf Coast - European trade routes, including the entire Mediterranean area. Possibly also North Atlantic and Gulf Coasts to East Coast South America. The remaining major trading areas, including all of Africa south of the Mediterranean, Australasia, the Far East and the Indian Ocean area are not regarded as prime possibilities at this point. One word of caution, however. This is not a forecast of the shape of things to come, nor is it meant to exclude other United States coastal areas or foreign trading regions.

One final word on the outlook for containership services in foreign trade. It is quite probable that American flag lines rather than foreign flag lines will be in the forefront in developing containership operations. This outlook is related to the ship replacement programs of the American carriers. In turn, it raises the question as to how much of the total volumes which may be involved will be carried by American flag lines without distinction as to whether these lines will operate the conventional cargo ships or containerships or combination conventional cargo-containerships. At the present time, American flag lines carry less than 40 per cent of total United States general cargo exports and less of total United States general cargo imports. On this score there is also some question as to whether the governments of some of the leading maritime nations with national flag lines will permit - through various devices - any major changes in the distribution of cargo trade with the United States.

I should like to conclude this presentation with a general observation on the implications of containership operations as they affect the distribution of oceanborne volumes of commerce among the ports of the nation. It has already been noted that it is considered a sine qua non of effective containership operations that the number of port calls made by the ship be limited and also that the containers be concentrated in as few ports as possible. It follows from this that a containership operator will tend to use one major port in a range of competitive ports rather than as it is today the common practice to call at each of these ports.

As an example, the five major port areas on the North Atlantic Coast - Boston, New York, Philadelphia, Baltimore and Hampton Roads Area - all compete for oceanborne foreign trade cargo originating or terminating in the interior of the country. It would be uneconomical for a containership operator to call at all of these ports on a single voyage as is the customary practice at the present time. Instead it is much more likely that the operator will concentrate his movements at one of the ports. This will have the effect of improving the competitive position of the port chosen as the base of operations vis-a-vis the other ports in question. What has just been said also applies to the domestic offshore trade.

Much the same impact is likely to be felt among the competing ports on the Gulf Coast and probably to a more limited extent here on the Pacific Coast. My reason for suggesting a possibly more limited impact here on the Coast is that there are three fairly distinct groups of port areas - Columbia River-Puget Sound, Bay Area and Southern California - separated by sufficient distances to make it more likely that such oper-

ations as are instituted will be concentrated within a single port in each of these areas rather than in one port, or port area, for the entire Pacific Coast range of ports. This will not, of course, minimize the competition among the individual ports in each of the port areas; each port will seek to derive the advantages of this evolving forms of integrated transportation and to maximize its benefits therefrom.

In large measure the outlook for both coastal and intercoastal operations will take the same form. There is a qualification, however, which must be noted in this respect stemming from the fact that the areas behind the ports which can be served effectively by containership operators is much more limited than in the case of foreign trade. Nevertheless, even in these instances there is likely to be a fanning out behind the port areas of service to the point of serious encroachment on the immediate hinterland of competing ports.

It is our feeling that the Port of New York will benefit substantially from these anticipated developments. As the largest direct consumption market, both personal and industrial, on the North Atlantic and also as the largest individual area of generation of outbound movements it begins with a major advantage over its competing ports in having substantial cargo volumes immediately available for containership movements. With such a base to begin with, it will, we believe, become the favored base of operation for both domestic and overseas services. The fact is that it is today the only port on the North Atlantic which has containership services and for which additional such services are currently in the planning stage.

As we look forward to the future, therefore, we are optimistic that this new form of integrated transportation will be of substantial benefit to the Port of New York. It will be an important factor in increasing the cargo volumes which we will be called upon to handle and in improving our overall competitive position for oceanborne cargo volumes.

THE FLIGHT TO THE SUBURBS SLACKENS

A. F. Parrott, Consolidated Edison Co. of N.Y.

Since the war, the face of much of the country has been remade. The much publicized flight to the suburbs has affected not only the cities and suburbs, but also the exurbs and even the country beyond. However, the greatest effect has been felt upon the cities and suburbs, and here the most pressing problems have arisen.

This flight to the suburbs is really as old as the country. From early Colonial days, the large eastern cities have grown through a series of outward movements. As the area of original settlement became crowded, the more prosperous would seek more room and more congenial surroundings by moving towards the outskirts. The space they vacated was filled by newcomers, a great many of whom were immigrants. In New York, this outward movement was first northwards in Manhattan. Later it spilled over into the other boroughs and, later still, into the surrounding suburban counties.

Every large city of the country has grown in much the same manner. However, there is one excellent reason for using New York as an example besides the fact that I am more familiar with it than any other. Since it is by far the biggest city, we would expect any changes in this trend to appear first in New York. If the tide has already turned in New York, and I believe it has, it should not be long before it begins to turn in the other large cities.

First let us examine the extent of this movement and the problems it has posed for both cities and suburbs. Before the war many of the people who moved outwards still stayed within the city limits. After the war almost universal ownership of automobiles loosened the ties which had held people close to public transportation lines. Now almost any potato or fruit farm within forty miles of a large city became fair prey for a builder of mass developments. Under the V.A. and F.H.A. programs, huge numbers were able to buy houses with practically nothing down. In addition, the very low interest rates and long repayment terms that prevailed often made it cheaper to buy a house than to rent. It is no wonder that, after the war, this flight to the suburbs became the great mass migration of our times. Its extent may be judged from the fact that such migration figures as we have indicate that in just seven years, from 1950 to 1957, between a million and a million and a quarter people moved out of New York City to the suburbs. This is just one city. As the recent census figures have made painfully clear, the same thing was happening across the country, in small cities as well as large during this past decade. However, it was most marked in the large cities. Of the ten largest cities, the eight east of the Mississippi all lost population. Only Los Angeles and Houston gained. Even in the fast-growing West, San Francisco and Oakland declined.

Yet the central cities were not depopulated. While one-third lost some population, in toto they gained some eight per cent in the past decade. Part of the loss was made up by natural increase but, in most of the large cities of the country, the biggest factor was a large influx of poor negroes from the rural areas of the South.

This, too, was a large migration. Harlem is still the largest and best-known negro community. But two articles* in U.S. News and World Report and one in Architectural Forum# show the striking increase in the negro population of the northern and western cities since 1950. Here are a few examples of their fairly recent estimates of the percentage of negroes in northern cities: Cleveland about 25%; Chicago 19%; Detroit 22%; Philadelphia 27%; Dayton, Ohio, over 19%; Newark, New Jersey almost 35%; Gary, Indiana, 36; New York, 12.6%. While the percentage is lower in the western cities, the negro communities in Los Angeles, San Francisco and Oakland are among the fastest growing in the nation. In fact, the 1960 census may well show that Los Angeles has now a higher percentage of negroes than New York. These examples are sufficient to show that this is a general problem common to most large cities.

In New York, we also had a large influx of Puerto Ricans. According to the last estimate I have seen, there are now about 650,000 in the city. In other eastern cities, the Puerto Rican communities, though much smaller, are increasing rapidly. Here in the West, you have had a considerable influx of Mexicans. Both of these groups speak a foreign tongue and have a different culture.

Here is the crux of the problem. Those who moved out of the cities were, as a rule, members of the fairly prosperous middle class. They were replaced by poor people, unused to urban living, with comparatively little education and few skills, many of whom spoke foreign tongues. Thus the cities' expenses for welfare, relief, police protection, etc., all skyrocketed at a time when a large percentage of its wealthier citizens, who were best able to pay these added costs, were moving out to the suburbs. That this trend may continue unabated has become a nightmare for the leaders of many of our cities.

This mass exodus has also created serious problems in the suburbs. As the nearer suburbs became completely filled up, it became necessary for the mass developers to go farther and farther out to find tracts of open land. In the meantime commutation has been deteriorating and becoming much more expensive. Particularly in the East,

* August 23, 1957 and April 20, 1959
\# June 1960

a considerable number of commuter railroad lines have been discontinued and service on many others has been curtailed. For those who drive to the city, the traffic congestion is continually becoming worse despite the new roads that are being built. Thus, those who drive have to leave a little earlier each year and get home a little later than before. In addition, the taxes in these burgeoning suburbs have been skyrocketing. In many they are already considerably higher than for the same priced dwelling in the city. And they are increasing much faster. In the meantime, the open areas around these suburbs are being filled up so completely that there is often little difference between the outskirts of the city and the suburbs twenty miles farther out. The result is that the suburbanite is spending more and more money and more and more time to get less and less of the amenities he had hoped to gain by leaving the city.

So much for preamble. Now what evidence is there that the tide is turning in New York? Our social statistics are probably less complete and less exact than those in any other major field. Even the census, which is the rock on which most of our figures are based, for any one city can be off as much as 3%. And we only have that every ten years. Yet, as Dr. Griffin's fascinating talk demonstrated, by scrounging around it is often possible to unearth assorted sets of figures which, while they may mean little by themselves, when placed together will coalesce into a convincing picture.

Here's what we have. First there is the incipient taxpayer revolt in the suburbs as witnessed by the rapidly increasing number of school budgets turned down in Nassau and Suffolk counties. About three years ago New York State passed a law providing that, if three consecutive school budgets are turned down by the voters, the state is empowered to establish a contingency budget in that school district. This contingency budget is always an austerity budget. This law was adopted so that the schools would always be able to open in the fall. Here are the number of budgets turned down in Nassau and Suffolk in the past two years, the only years for which I was able to obtain data.

	1958-9 Budgets	1959-60 Budgets
Turned Down, First Vote	6	15
Turned Down, Second Vote	2	11
Turned Down, Third Vote	2	5
Contingency Budget Established	2	5

In this two year period, only nine contingency budgets were established in the entire state, and seven of them were in Nassau and Suffolk.

Nassau and Suffolk have been the two fastest growing counties in the New York area. Here many mass developments of the Levittown type were built in the postwar years. The typical development of this type is a purely residential community peopled by couples in the same age group almost all of whom are in the process of raising families. In older communities there will be, on the average, about one child per house in the public schools at one time. In Levittown the average is three. The average cost per pupil in New York State was $587 in 1958. At that rate school taxes would come to almost $1,800 per house, if all the cost were paid locally. And these are small houses in the $15,000 range.

This, of course is an impossible situation. Across the country political pressure has forced the adoption of some sort of state or local aid programs that pay a portion of the cost for these hard-pressed communities. Even so, these people feel bitter. They don't want to sell their homes but fear they may be forced to. They voted these budgets down not because they didn't want good schools for their children but because they simply could not afford the costs entailed. It doesn't take this type of news long to get back to the people in the city who are thinking of moving out.

Second, we have some interesting migration statistics. I have classified the migrants into four groups: (1) Young adults, both male and female; (2) Older adults whose families are grown up; (3) Families where at least one child is of school age; (4) Families where no child has yet reached school age.

Though there are no good recent statistics, there is general agreement on the net movement of the first two groups. On balance, the young adults have always moved into the big city to make their fortunes. The movement of the older adults whose families have grown up is more mixed. Most of those where the husband is still working will remain where they are but there will be some net movement from the suburbs to apartments in town. Among the retired people, most will remain where they are but there will be some net movement to rural areas or warmer climates. Both of these movements, though small, have probably been increasing. The increasing movement back to town will be discussed later. Though there are no recent figures available on the movement of retired people, after such a prolonged period of prosperity as we have had since the war, it appears logical to assume that increasing numbers can now afford to move to Florida or California when they retire. Nevertheless, the net effect of these movements is still small. The great bulk of the net migration takes place in the third and fourth groups.

When we turn to the married couples with children, we have better data. The New York Public School System keeps very complete statistics on the movement of pupils in and out of the public schools because of migration. The figures are plotted on Chart 1. The five school years 1950-51 through 1954-55

witnessed the peak of the outward movement. On the average there was a net outward movement of 10,982 pupils a year. In the most recent four years, however, the chart shows how this net outward movement has dropped sharply. In the last two years it has averaged only 2,569 per year. This represents a drop of more than 75% over the previous period. While the sharp decline in net outward movement of pupils is the most striking feature of this chart, analysis of the total numbers moving in and out is also illuminating. For instance, the number moving from the city to the suburbs has declined significantly in the past four years. Also the number moving from the suburbs back to the city, while still much smaller than the number moving out, has been increasing slowly but steadily throughout the past decade.

With couples all of whose children are of pre-school age, the picture is similar. Most couples are in their thirties when they buy their first house. By that time most of them will have at least one child of school age. After the war, however, the government made it very easy for veterans to buy houses with no down payment and at very low interest rates. Buying a house cost less than renting a comparable apartment, so many veterans took advantage of the program. In the late forties many of them bought older houses in the city or new houses in the outskirts of the city. After 1950, however, the city was so built-up that increasing numbers of them had to buy outside the city limits. Since many of these veterans had been married but a short time, during this period there was undoubtedly a considerable outmigration of families all of whose children were of pre-school age.

Only an approximate measure of the size of this group is available. That is the number of requests for Veteran Administration Appraisals. These data, available only for the entire country back to 1951, are plotted on Chart 2. The sharp rise in 1954-56 and equally sharp drop thereafter is not surprising. After the Korean War veterans returned, many of them used their veterans' privileges in the next three years. Unless and until another large group of veterans is accorded the same privileges, we cannot expect to approach the totals of those years.

The number of appraisal requests are not available for the city on a historical basis. However, the population of the city was about 4.7% of that of the country in 1955 and 4.5% in 1958. Thus, we could expect an average of about 24,400 requests from veterans living in the city in each of the three peak years (1954-56) and an average of only about 9,400 requests in each of the last three years. This represents a drop of over 60% from the previous period. While a great many of these veterans may have purchased houses outside the city, undoubtedly a large number purchased homes within the city limits. In addition, some of these requests may not have resulted in sales. Thus these figures are useful only to show the upper limits of the possible migration due to this source.

While the exact level of migration represented by these figures is open to question, the marked change in trend is unmistakable. Thus, it is safe to state that, in the past three years, there has been a very sharp drop in the outward movement of the two groups who have, historically, made up the great bulk of those moving from the city to the suburbs.

Third, we have the trend of new residential construction. Since the per cent of New York City dwelling units vacant has been very low and has changed very little since 1950, these figures should prove a valuable check on the belief that the flight to the suburbs has slackened considerably in the past few years. On Chart 3 are plotted figures for both the United States and New York City. For the country the bars represent the number of non-farm dwelling unit starts compiled by the Department of Commerce. As you know, these are in large measure based on building permits issued. For New York City the figures are the number of dwelling units for which building permit applications were made.

The difference in trend between the city and country is striking. Though both started and ended high, residential construction in the nation was fairly stable throughout the decade while in New York City there was a sharp decline that lasted for most of the decade followed by an even sharper rise.

There are two drawbacks to these New York City figures. First, about 3% of building permit applications do not, on the average, result in starts. Second, since hearings are now being held on a new zoning code for the city, it is likely that some of the 1959 applications were made to get under the wire before the new zoning regulations go into effect. Accordingly, on Chart 4 I have plotted the dwelling units completed in the city in the same period. This trend is similar but dampened somewhat and delayed a year. About 70% of the units built in the decade were apartments and it takes about a year to complete a large apartment house. Therefore I have added the year ending June 1960. As you will note, completions are still rising in the city though starts have been receding in New York as well as the country, since the end of last year. For the year ending June 1960, completions were up 50% over the annual average for the years 1953-57, and still rising. For the country the increase over the same period was less than 20%.

While all five boroughs of the city have shared in the recent upsurge, in Manhattan the rise has reached boom proportions. Yet here most apartment houses are still completely rented before they are finished. Even more significant is the fact that virtually all of this recent rise in apartment construction in the

city has been in the high and medium rent categories. Since they are still renting readily, there is good basis for the belief that the flight of the better-off families to the suburbs has, at the very least, materially slackened. In support of this, several recent surveys by real estate firms show an increasing return flow from the suburbs to the new luxury apartments in Manhattan.

The migration figures kept by the New York Public Schools, the trend of Veteran Administration appraisals and the volume of new residential construction in New York City all make it appear probable that, in New York, the flight to the suburbs has slackened considerably in the past three years. Is this a permanent change or just a passing phase?

We have already noted that the comparative advantages of living in the suburbs have lessened considerably in the past decade and that the expenditures of both time and money necessary to live there have increased. As the suburbs become filled up farther and farther out from the city, there appears to be little likelihood that any of these trends will be reversed. Instead, they appear more likely to flow more strongly. The great apartment boom in New York, and especially in Manhattan, of the past year or more shows that the New York Real Estate leaders have sensed this trend and are gambling many millions that it will continue.

There are four additional factors that may accelerate the trend. First, Representative Mills has made the very cogent point that, with the present income tax law, since the war the apartment-dweller has been paying more than his share of income taxes while the home-owner has been paying less than his share. In other words, the apartment dweller has been subsidizing the home-owner.

In Congressional hearings last fall, several suggestions were made to correct this inequity. Representative Mills would like to reduce or eliminate many of the present exemptions and then to reduce the income tax rate so that the total government take would be the same. Two of the proposals were that the home-owner no longer be allowed to deduct the interest on his mortgage before computing his tax and that the home-owner be required to include the imputed rental value of his house in his income. Either of these proposals would increase the tax paid by home-owners, while reducing the tax rate would decrease the tax paid by apartment dwellers.

This is an election year, so none of these changes were made. Nevertheless, it is significant that Congressional leaders of both parties believe that the income tax law needs revision and that the present law contains inequities which favor the home-owner and penalize the apartment-dweller. Some revision appears sure in the next few years. Whenever changes are made, we may expect them to make living in an apartment in the city more attractive and living in a house in the suburbs less attractive.

Second, part of the reason for the flight to the suburbs after the war was undoubtedly because it was the accepted thing to do. As it becomes less attractive financially, it is quite possible that the fashion may change. With a dollar advantage as a lever, it is quite conceivable that it may become fashionable to live in an apartment house development which has its own swimming pool, recreational area and perhaps a marina - something like the present Levitt House development in Queens. Should this type of living replace the house in the suburbs as the accepted ideal of the upper middle class, then the return flow to the city could reach sizeable proportions.

Third, is the seldom-mentioned but very important cause of much of the out-migration from the cities, - the desire for segregated housing and schools. I have already discussed the great postwar influx of negroes to the cities of the North and West. We in New York, as do also you in the West, pride ourselves on our comparative lack of racial prejudice. But it's only comparative. Undoubtedly a great number moved to the suburbs because Negroes or Puerto Ricans were beginning to move into their neighborhoods.

To anyone interested in this subject I recommend very strongly the third volume of the recent New York Metropolitan Region study entitled "The Newcomers." It is extremely well-thought-out and well-written. While Dr. Handlin is discussing the Negro and Puerto Rican problem in New York, a great deal of what he says applies equally well to any large city which has had a recent influx of either of these groups.

As they move up the economic scale, Dr. Handlin expects both of these groups gradually to spread out through the metropolitan area. As the 1960 Census will show, this process is already well along in the New York area. As it continues, there will be less and less reason to move to the suburbs in an attempt to escape these groups. Eventually, the desire for segregation may even lead some to return to the city. It is far easier to avoid someone in the next apartment than someone whose back yard adjoins yours. Should this return movement develop, it should start sooner in New York than in most of the other large cities of the North and West. Since the Negro community in New York is comparatively old, its members have had greater opportunities to move up the economic scale.

Fourth is the changing age composition of the population. Between now and 1970, the number of persons in their thirties will decline. This is the result of the low birth-rate in the 1930's. On the other hand, the number of those in their twenties, those born in the 1940's will increase sharply. As a rule, young couples in their twenties start out in an apartment. Later, in their thirties, they are likely to buy their first house. Since most apartments are located in the cities, this unsymmetrical age distribution should prove to be a significant factor in slowing the trend to the suburbs during the next

decade. After that, it will have the opposite effect.

Now to sum up. The forces that have already led to a slackening of the flight to the suburbs should become even stronger in the future. In addition, four new forces pushing in the same direction seem likely to develop. On the other side of the ledger, only one new factor appears likely. Public pressure is almost certain to result in increased Federal and State aid for education. This will lighten the school tax burden now pressing on the suburbanite. Yet this aid, however necessary and welcome, can only be a palliative. School taxes in the fast-growing suburbs will still be very high.

So far I have been painting what looks like a very black picture for the suburbs. Let me hasten to explain that I have been looking at the suburbs exclusively from the viewpoint of one who works in the city. For those who work in the suburbs, they will still be a very attractive place to live. The suburbanite who works nearby avoids both the high commutation costs and the increasing travel time to which the suburbanite who works in the city is subjected. Since job opportunities have been expanding rapidly in the suburbs, we can expect them to continue to grow rapidly.

This brings us to the question of whether the flight to the suburbs will continue or even increase because job opportunities in the cities diminish as industry moves out to the country. This is a broad subject in itself and far beyond the scope of this paper. Fortunately, it has already been made for the New York area. In the ninth volume of the New York Metropolitan Region Study, to which I referred earlier, Dr. Vernon makes projections of employment in the city. While the book is only due to be published next December, Dr. Vernon sent me a draft copy. In it he estimates that the total number of people employed in New York City will increase by over half a million in the next quarter century. In the case of New York, therefore, we can dismiss the fear that lack of job opportunities will cause the flight to the suburbs to continue.

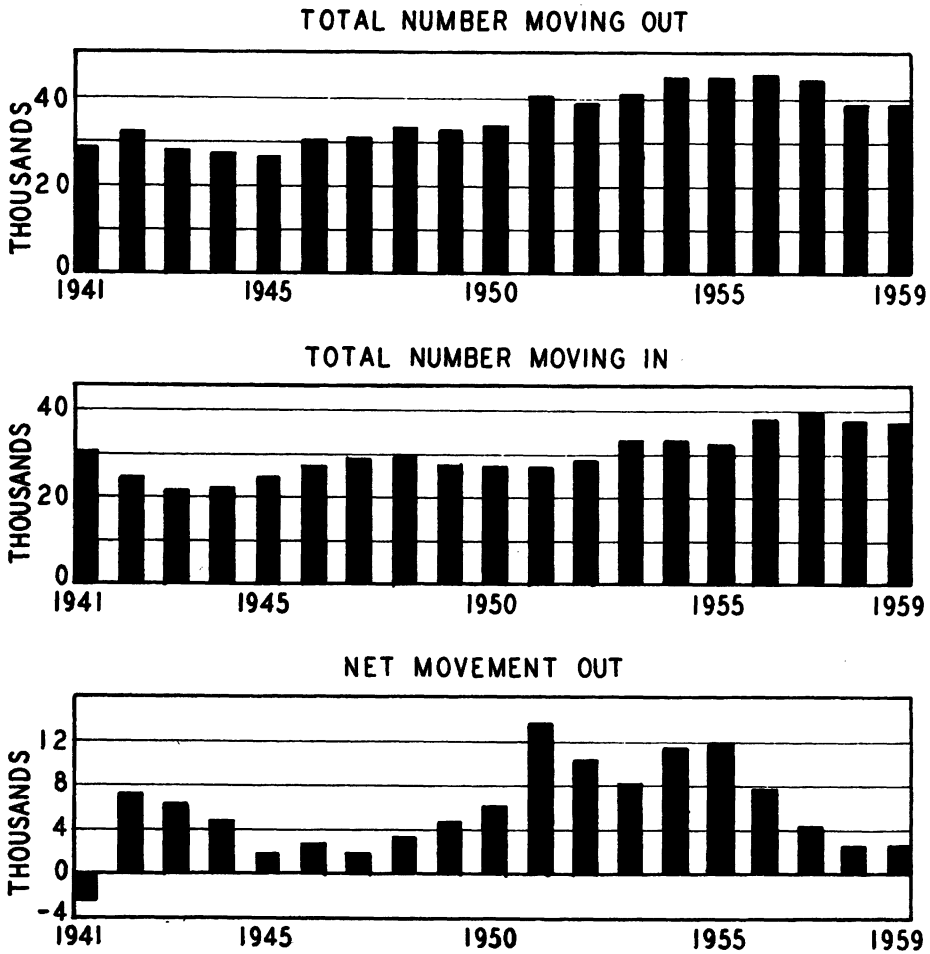
Now, what about the other large cities? Most of the developments I have covered will apply to some extent to them. And the larger and more geographically restricted the city, the more are they likely to apply. Geographical restrictions, by limiting the territory into which a city can expand or by increasing the time and cost of reaching the surrounding area, have the same effect as increasing the size of the city. San Francisco is an excellent example. Though, so far, the Census people tell me they have noticed no signs of a slackening of the flight to the suburbs, it should not be many years before some signs appear in such cities as Chicago, Los Angeles and San Francisco.

In the smaller cities, the situation is different. In many of them it is still possible to get from the middle of town to fairly open country in half an hour or less. In this automobile age, we can expect the flight from these cities to continue unabated and, perhaps, even to accelerate.

Here a caveat is in order. Though logic might indicate a slackening in the flight from the big cities to the suburbs, it may never occur unless living in the cities can be made more attractive. I have devoted considerable attention to the decline of many of the advantages which the suburbs formerly had. Nevertheless, there are a number of conditions prevalent in the big cities which were the direct cause of much of the flight to the suburbs. I will mention only the incidence of crimes of violence in the cities. Of course crimes are committed in the suburbs too, but the fact remains that, and with good reason, many people do not feel safe on the streets of our big cities after nightfall. At present, many people are disillusioned with both city and suburban living. They may still choose suburban living as the lesser of two evils. If the cities hope to benefit from the filling up of the more convenient suburbs, they must first put their own houses in order.

CHART 1

MOVEMENT OF PUPILS IN AND OUT OF THE NEW YORK CITY PUBLIC SCHOOL SYSTEM *
(YEARS ENDED JUNE 30)



* BECAUSE OF MOVING ONLY (EXCLUDING INTRA-CITY MOVES) DOES NOT INCLUDE THOSE LEAVING BECAUSE OF GOING TO WORK OR FINISHING HIGH SCHOOL.

CHART 2

NUMBER OF REQUESTS FOR VETERAN ADMINISTRATION APPRAISALS
(TOTAL UNITED STATES)

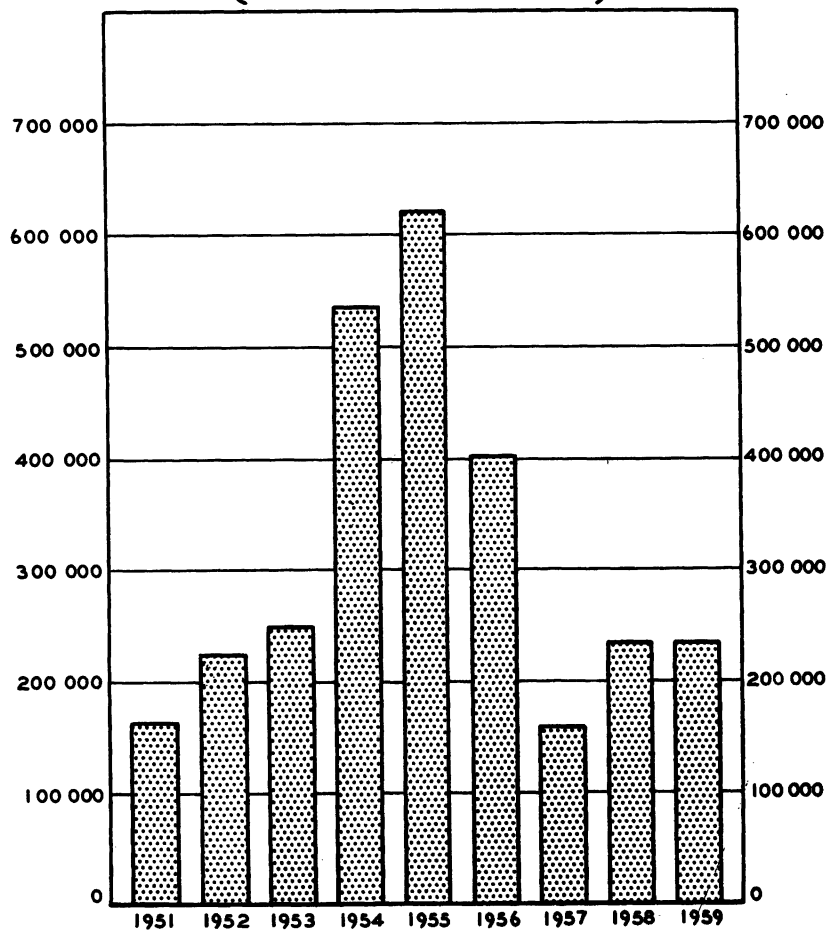


CHART 4

NUMBER OF DWELLING UNITS COMPLETED (NEW YORK CITY)

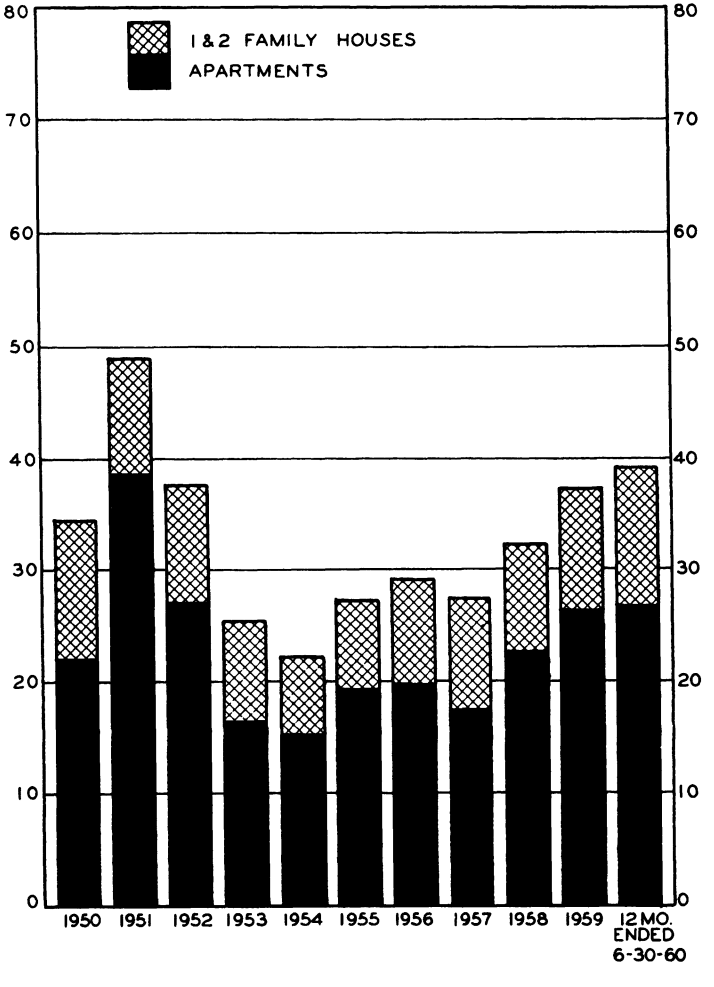
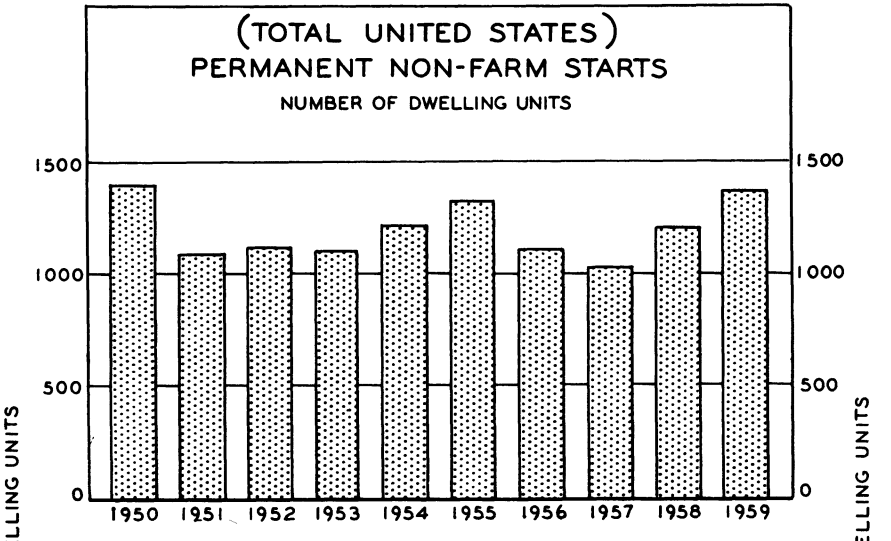


CHART 3

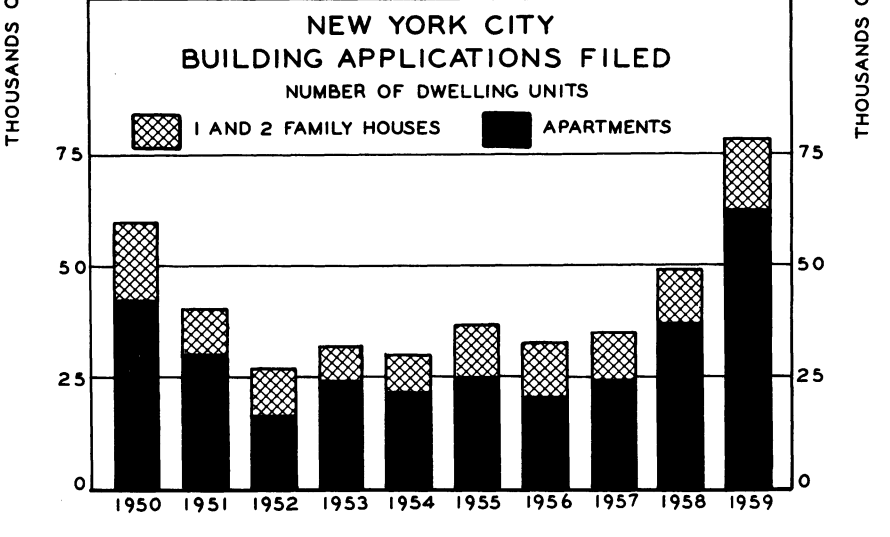
(TOTAL UNITED STATES) PERMANENT NON-FARM STARTS

NUMBER OF DWELLING UNITS



NEW YORK CITY BUILDING APPLICATIONS FILED

NUMBER OF DWELLING UNITS



XI

METROPOLITAN AREAS AND URBAN GROWTH - II

Chairman, Harry Sharp, University of Michigan

On Problems of Measuring the Distribution of Population in an Urban Area — Martin Taitel, University of Chicago

Urban Population Concentration and Occupational Structure in the United States: 1900-1950 — Georges Sabagh, Maurice D. Van Arsdol, Jr. and Hamid Zahedi, University of Southern California

Status of Research on Methods of Estimating State and Local Population — Jacob S. Siegel, Bureau of the Census

ON PROBLEMS OF MEASURING THE DISTRIBUTION OF POPULATION IN AN URBAN AREA*

Martin Taitel,** University of Chicago

The fact of human life with which we are at present concerned is this: A very large proportion of the people globulates. Human beings have not spread equally over the earth's surface (total or land). Nor have they spread unevenly with gradually and smoothly rising and falling densities. Rather they form globules or near globules of population at selected locations. These are loosely and variously referred to as cities, standard metropolitan areas, urban places, urbanized areas, and metropolises.

Quite obviously the phenomenon of population globules presents many complex facets for analysis. Variables are many--both those within globules and those as between globules. Within an individual globule, there is variation in racial composition, income, wealth, land value, density, and so on. Comparison of globules shows variation in size, income, wealth, land value, transport facilities, average density for finite areas, industrial composition, and so on. Relations--including those between the "within globules" and those between the "between globules" variables--do not appear to be simple or easy to measure.

This paper is limited to consideration of the density and distance variables of individual globules. Conceptualization of location as a point in a plane and of density as a count of persons per unit area (finite or infinitesimal) of the plane defines the variables under consideration. For those variables, this paper is concerned with methods of measuring the relation between density and location within an urban area --with methods of measuring the shape of an individual population globule.

It is convenient at this point to set forth some of the notation to be used. This is shown in Table I. It should be noted that provision is made for the distinction between "density" and "probability."

According to Winsborough (Unpublished, 1960), there is a considerable body of data of varying quality bearing on density and location within urban areas. Yet Duncan (1957) comments: "Surprisingly little systematic study has been devoted to the pattern of variation of population density from one part of the city to another." This applies to both empirical measurements of and theorizing about the shape of population globules. Since Duncan wrote in 1957, there has been some additional published work by Clark (1958) and Stewart and Warntz (1958) and within my knowledge some still unpublished work by Muth and Winsborough.

In this limited amount of study, there has developed an hypothesis expressing density as a bivariate exponential function of location. According to Duncan (1957) it was first suggested by Stewart (1947) and later by Clark (1951). In our notation, this is

TABLE I

SELECTED SYMBOLS USED

Quantity	Origin or pole O(True)	Origin or pole O'(False)
Probability	P	P
Density for persons:		
At a point	D	D
At the origin or pole	D_O	D_O'
Density for an area	δ	δ
Polar Coordinates:		
Radius	ρ	ρ'
Direction	θ	θ'
Direction OO'	θ	$*\phi$
Distance OO'	0	*C
Rectangular Coordinates:		
Abcissa	x	X
Ordinate	y	Y
X - x	0	$*C_x$
Y - y	0	$*C_y$
Parameter for exponential function	μ	μ
Size of total universe	N	N
Size of sample	n	n
Area	A	A
Frequency:		
Theoretical	f	f
Observed	f'	f'

*Parameters of location on the plane approximating the earth's surface.

*Presented before the session, Metropolitan Areas and Urban Growth--II, at the 120th annual meeting of the American Statistical Association, Stanford University, August 25, 1960.

**I wish to express my appreciation to Professor Philip M. Hauser, Director, Population Research and Training Center, University of Chicago, for making available to me the facilities of the Center in the preparation of this paper. I wish also to express my appreciation to Mr. Halliman H. Winsborough, Research Assistant at the Center, for time spent in discussion of many aspects of urban densities and of the literature relating to them and for making work sheets available to me. It seems appropriate to state also that my efforts represent a labor of love and not of duty.

I wish, in addition, to state that this paper has benefited much from the comments on an earlier draft made by Professors Philip M. Hauser, William H. Kruskal and Richard Muth, all of the University of Chicago.

$$D = \frac{1}{2\pi}\left(\frac{2}{\mu}\right)^2 e^{-\frac{2}{\mu}\rho}, \quad \left(\begin{matrix} 0 \leqq \rho \leqq \infty \\ 0 \leqq \theta \leqq 2\pi \end{matrix}\right) \qquad (1)$$

the variables of course being D, ρ, and θ. The listener is cautioned to remember that ρ and θ are polar coordinates and that (1) is not a probability function.*

Note that in formulating the hypothesis in this way ρ and θ and, consequently, D are specified as continuous variables. Further, D is a relative density at a point, not an absolute density. Absolute densities may be conceived of as given by ND. However, since N can in fact only be finite, consistency with fact must be obtained through a conceptualization of persons as divisible into infinitesimal particles.

The hypothesis (1) can be cast in terms of probability. If this is done, ρ and θ become rectangular coordinates and

$$P = \frac{1}{2\pi}\left(\frac{2}{\mu}\right)^2 \rho\, e^{-\frac{2}{\mu}\rho}, \quad \left(\begin{matrix} 0 \leqq \rho \leqq \infty \\ 0 \leqq \theta \leqq 2\pi \end{matrix}\right) \qquad (2)$$

Integrating with respect to θ gives a Gamma distribution as the univariate distribution for ρ.**

Properties of (1) and (2) are as follows:

1. μ is the only parameter (since the total volume and the origin are assumed known).
2. μ is the arithmetic mean distance from the center, 0, or the first moment of ρ.
3. $\mu/2$ is the modal distance from the center, 0, or the modal ρ.
4. The maximum density (not probability) is

$$D_0 = \frac{1}{2\pi}\left(\frac{2}{\mu}\right)^2 \qquad (3)$$

5. ρ and θ are statistically independent. Thus, distance from the center does not depend upon direction from the center and vice versa.***
6. If (1) is expressed in rectangular coordinates, then x and y are not statistically independent. This statistical dependence represents a desirable feature of the hypothesis.*

With (1) and (2) as a working hypothesis the next step is the measurement of parameters. The statement of the hypothesis assumes that the center or pole, 0, is known so that measurements are from that pole. Hence, let us proceed, in the first instance, on that basis. Let us assume, further, that sampling is of persons in the globule and is random with respect to them, i.e., we have a random sample of ρ and θ distributed as (2). And, finally, let us assume no errors of observation. Under these assumptions, the maximum liklihood estimate for μ is

$$M = \frac{\sum_{i=1}^{n} \rho_i}{n} \qquad (4)$$

It is interesting to note that sampling-wise $4nM/\mu$ is distributed as Chi-square for samples of size $4n$. Consequently, the F-distribution is applicable for testing differences between sample M's. Furthermore, the Chi-square goodness-of-fit test seems to be applicable to test the hypothesis itself.

These results are not, however, a solution to the problems of measurement which arise in practice.

For one thing, cities are not all circular in shape; there seems to be no simple general procedure for adapting the technique in cases of noncircularity.

Second, in practice, sampling can only be over a finite range for ρ. This, however, is not a serious drawback. Table II shows the

TABLE II

FRACTION OF POPULATION BEYOND RADIUS r AND RATIO OF μ TO μ_N

r/μ	Fraction of population beyond radius r	μ/μ_N*
0	1.0000	∞
1	.4060	1.8372
2	.09158	1.1923
3	.01735	1.04757
4	.003019	1.01089
5	.000499	1.002276
6	.000080	1.0004426
7	.000012	1.0000815
8	.000002	1.0000144
9	.0000003	1.0000025
10	.00000004	1.0000004

*Approximate correction factor for M_r to estimate μ.

Note: μ_N is the average distance for $0 \leq \rho \leq N$

*I am indebted to Professor William H. Kruskal for suggesting the difference between the density and probability function be called to the attention of the audience.

**It may be helpful to the audience to note that, had (1) been expressed in rectangular coordinates, the density and probability functions would be the same.

***This is also true for the normal distribution analogue, i.e.,

$$K e^{-\frac{\rho^2}{\sigma^2}} \text{ (in } \rho \text{ and } \theta\text{)}$$

*For the comparable zero-correlation bivariate normal distribution, x and y are, of course, statistically independent. This represents a very great disadvantage of the normal hypothesis.

approximate average error in M_N (the mean when sampling is only out to $\rho = N$) as an estimate of μ because it does not take into account the fact that sampling is limited to $0 \leq \rho \leq N$. If observations are over a radius five times the average distance, the error is only somewhat more than 2 parts in 1000 on the average. In practice, therefore, either observations can be obtained over a large enough range to obviate the need for adjustment or an approximation to the maximum liklihood solution as given by

$$\left[1 - \frac{2\left(\frac{N}{M}\right)^2}{e^{2\left(\frac{N}{M}\right)} - 1 - 2\frac{N}{M}}\right] M = \frac{\sum_{i=1}^{n} \rho_i}{n} \qquad (5)$$

can be used.

Third, in practical applications, not only relative densities, but also absolute densities, are sought. To obtain measurements of them either the urban area must be completely enumerated or else a supplementary estimating procedure must be used. This difficulty is not insurmountable.

Fourth, there arises the question of the validity of the random sampling assumption.* It is not easy to resolve this question, particularly since it involves the question of specifying the universe of which the complete enumeration of an urban area can be a random sample. There appears to be no forthright solution.

The fifth and last point in this discussion of the problems arising in measuring μ relates to the necessity for generalizing the measurement procedure to include measurement of the location of the center of the urban area (i.e., the origin or pole for the location coordinates). The practice has been to designate a point as the center of a city on the basis of an inspection of the data or of preconceived notions about where the center lies. For example, Blumenthal (1949) indicates the City Hall was used as the center for his data on Philadelphia. Clark (1951 and 1958), Muth (1960) and Winsborough (1960) use some point in the central business district. Reinhardt (1950) used the city hall as the center. The practice is, of course, subject to serious shortcomings as a method of measuring parameters. It should not be used particularly when more efficient procedures based upon the theory of probability are readily available.

It is convenient to demonstrate the availability of a valid procedure for measuring the parameters of location in terms of rectangular coordinates.**

*Professor O. D. Duncan, Population Research and Training Center, University of Chicago, raised the question at first sight of the author's preliminary work.

**The choice of coordinate systems depends, of course, upon the one used in obtaining the observations. Which to use in the observation process depends in turn upon considerations relating to errors of observation--not under consideration at the present time.

The maximum liklihood equations for C_x and C_y are

$$\sum_{i=1}^{n} \frac{x_i - C_x}{\sqrt{(x_i - C_x)^2 + (y_i - C_y)^2}} = 0 \qquad (6)$$

and similarly for C_y. It is interesting to note that these are the equations which locate the point from which the average distance for points on a plane is a minimum.

I see no straightforward solution of (6) for the estimates C_x and C_y. But clearly a successive approximation procedure can be used. The desired quantities are weighted arithmetic means. Unweighted arithmetic means* can be used as first approximations. On the basis of them, approximate weights can be introduced, and so on. Intuitively I believe the efficiency of such estimates to be high.

Efficient estimation of the center of an urban area is, of course, of some importance. The effect of an error in locating the center is to increase the estimate of the average distance from the center, i.e., the estimate M of μ, and to reduce the estimate of the peak density by the square of the factor involved. Unfortunately, I have been unable to perform the integrations necessary to show the relation between C and the error in M. C is, of course, the distance between true and false centers. However, it is clear that an error in M cannot exceed C and must be substantially less. Further, the square root of the second moment about the false center to the second moment about the true center is given by

$$\sqrt{1 + \frac{2}{3}\left(\frac{c}{\mu}\right)^2} \qquad (7)$$

This may serve as a guide to the order of magnitude of the error in M when computed from a false center.

In practice, various investigators could hardly err by as much as μ, assuming, of course, some validity to the basic hypothesis. And perhaps they do have sufficient insight and intuition to avoid going off-center by significant distances. Nevertheless, it is surprising to discover they had not at least raised the question of where the center of a city is. It would seem that investigators would have attempted to determine whether the center of a city is a court house, a department store, a hotel, a church, a factory, a railroad station, etc. It would seem they would have tried to determine whether the center of a city moves from time to time. And, in addition, it would seem they would have attempted to determine whether cities differed with respect to the activities located at their centers.

The procedure for measurement which I have been discussing has not been applied, though it is certainly sound and worthy of consideration. The observations required are of horizontal distance on a rectangular grid or of horizontal distance and direction on a polar coordinate grid.

*The solutions for the bivariate normal surface.

Densities as such are not observed. Nor need they be computed in the estimation process and χ^2- goodness-of-fit testing.

An alternate procedure--and the one used by other investigators--is to take D, i.e., density, as an observable variable. For the present, let us assume that there is a value D attached to each person in the population and that this value of D is the one given by his location in the (ρ, θ) or (x,y) system of coordinates. In addition, we assume, as before, no errors of observation.

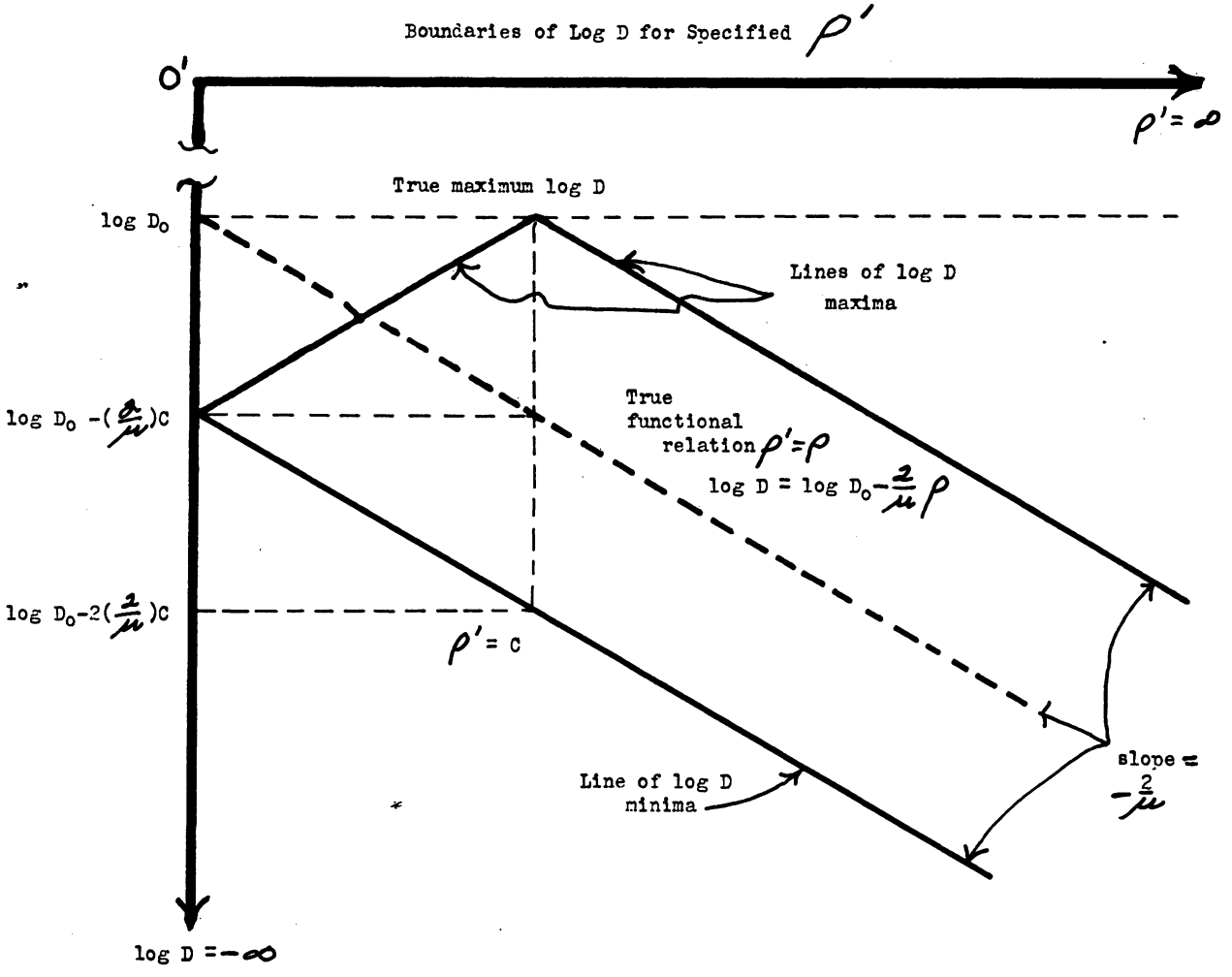
Let us first view our measurement problem on the assumption that the center is accurately known. If this assumption, in addition to those previously made, holds and the hypothesis (1) is true, then each and every observation must fall on the line

$$\log D = \log D_0 - \frac{2}{\mu}\rho \qquad (8)$$

which is shown on the chart. If one observation is not on the line, then the hypothesis must be rejected. Furthermore, it may be noted that one observation is sufficient to determine μ, if relative densities are measured; and two non-identical observations are sufficient to determine μ and N, if absolute densities are measured. Thus, if a great deal is known, the parameters can be measured without regard to the character of the sampling.

To bring our view of the measurement problems somewhat more in accord with observational conditions let us assume the center of location is not known—still assuming, of course, that (1) is true and observations are perfectly accurate. Our variables then become log D and ρ', the latter being the distance from a "false" center. And the probability function for those variables is

Chart



$$P = \frac{e^{\log D}\left(\frac{\mu}{2}\right)^2\left(\log \frac{D_0}{D}\right)}{C\sqrt{1-\left[\frac{\left(\frac{\mu}{2}\right)^2\left(\log \frac{D_0}{D}\right)^2 - \rho'^2 - C^2}{2\rho' C}\right]^2}},$$

$$0 \leq \rho' \leq \infty \qquad (9)$$

$$\log D \geqq \log D_0 - \frac{2}{\mu}C - \frac{2}{\mu}\rho'$$

$$\leqq \log D_0 - \frac{2}{\mu}C + \frac{2}{\mu}\rho', \text{ for } 0 \leq \rho' \leqq C$$

$$\leqq \log D_0 + \frac{2}{\mu}C - \frac{2}{\mu}\rho', \text{ for } C \leqq \rho'$$

where C is, of course, the distance between "true" and "false" centers. I regret that I have not had the opportunity to investigate this probability function. However, characteristics which I have been able to determine do shed some light on our measurement problem.

From a geometrical consideration of (1), the boundaries within which the universe of observations of log D and ρ' fall can be determined. These boundaries are shown on the chart and the equations for them have been written as part of (9). Actual observations, therefore, except very rarely by chance or unless deliberately selected to do so, would present a scatter. Neither the complete universe nor a sample (except extremely rarely) would appear as the functional relation (8) between density and distance. Furthermore, for the range $\rho' < C/2$, all observations would fall below the theoretical functional relation (8). Finally, it may be noted that, for noncircular urban areas, complete sections of the distribution might be missing; for example, a city with a pie slice missing might be observed as one with a much lower maxima line over the range $\rho' > C$ than the one shown on the chart.

If in the process of measurement a false center is selected and accepted as true, the hypothesis (1) would be rejected even though true. Observations would be interpreted as contradicting the hypothesis (1), first because the greatest densities would not occur near $\rho' = 0$, i.e., near the false center, and, second, because scatter would occur (except very rarely). This seems to be what other investigators have tended to do. Colin Clark (1951) states, "Then (except in the central business zone)—$y = Ae^{-bx}$." He also wrote (1958): "Density falls as we proceed outwards from the centre of a city (subject only to the qualification that at the very heart of the city most of the space will be occupied by commercial buildings, and residential density will be low)." Duncan (1957) remarks: "...the central area is essentially a non-residential district whose resident population density usually is far below that of the contiguous areas." Muth (1960) measuring log Density on distance regressions excludes central business district census tracts for theoretical reasons which presumably are in part based upon fact. Reinhardt (1950) discards the hypothesis in favor of one which is sufficiently flexible to permit a relatively flat segment or a hook in the regression of log D on distance for the shorter distances from the assumed center.

It may be that these investigators are correct and the declining exponential hypothesis is false—at least for part of the distance range. Certainly, however, an hypothesis should not be rejected without critical testing in the face of observations which might be sufficiently consistent with it to confirm it.

With respect to scatter, investigators seem to be willing to accept it as consistent with the hypothesis (1). Clark (1958) uses an average density at a given distance, saying "fitting a line to the weighted data," i.e., "density of each individual ward or census tract as a function of its distance from the centre of the city," "proved to be impracticable." Muth (1960) with correlation coefficients squared ranging from 0.022 to 0.74 for samples of 25 census tracts in each of 46 cities, apparently does not conclude this contradicts the hypothesis (for part of the distance range) or requires reconsideration of the measuring procedure. Yet, the hypothesis (1) if true, indicates correlations should be high unless the center has been falsely located. And, if not high, then the first step in further study is to test the validity of the location of the center.

If in the process of measurement the center selected is recognized as a false one, then the true center can be located—under the assumptions we have made—either (a) by a trial and error process of changing the location or (b) by the solution of a set of equations obtained from 4 non-identical observations. As before, if a great deal is known, parameters can be determined with very few observations and without regard to the character of the sampling. Of course, many observations falling on the theoretical functional relation strengthen the conclusion and, of course, one which does not so fall invalidates the hypothesis.

Thus far, density, D, has been treated as a measurable characteristic of a person. I now wish to consider density, δ, as a measurable characteristic of an area. This apparently is the definition used by at least some other investigators. Corresponding to this change in definition of density, the definition of the coordinates, (ρ, θ), or (x, y) changes to that of location for an area.

Under these definitions, the hypothesis simply becomes (1) with δ in place of D and with (ρ, θ) locating a point within the area. It is accurate for infinitisimal areas, a close approximation for small finite areas, and, perhaps, only a crude approximation for larger finite areas.*

*Of course, it is theoretically possible to select points within areas which will satisfy a known functional relation. However, under observational conditions the location points are selected without knowing the functional relation.

Unlike the case for density as a variable attached to persons, however, the theoretical functional relation for density, distance, and direction specified by the hypothesis does not entail a probability distribution for the variables. Hence, in order to proceed at all, a probability distribution must be introduced into the theory or model.* As a minimum, it is necessary to have a conditional probability function.

Other investigators seem to have made the simple blanket assumption of random normally distributed log δ deviations. This is indicated by the least squares or regression line approach with log δ as the dependent variable and ρ' as the independent variable. Further indications are provided by the failure to discuss variation in the size of areas, variation in frequencies from area to area, and effects of non-circularity of a city.

I am not, at this time, prepared to say the assumption is wrong, but I do wish to probe a bit into the problem.

Measurement of densities is, of course, essentially a counting of frequencies if the areas are assumed known. Thus we have

$$\delta_i = f_i / A_i \qquad (10)$$

Substituting in hypothesis (1), we have

$$\log f_i - \log A_i = \log \delta_0 - \frac{2}{\mu} \rho_i \qquad (11)$$

as our functional relation. If we assume the area measurements are without error, then the probability distribution we seek is that of

$$P\left(\log \frac{f_i'}{f_i} \middle| \rho_i\right) \qquad (12)$$

where, as previously noted, f_i' is the observed frequency and f_i, the theoretical frequency for the i^{th} area. Obviously, P depends upon how the f_i' are generated. But looking at (1) and recognizing that f_i' may be a binomial variable, one wonders whether the presumption of random sampling of persons is not to be used, at least as a beginning. In short, it appears that, even though we started by trying to avoid this presumption, we have come back to it.

The regression technique may provide satisfactory estimates of μ under the assumptions that f_i' is a random binomial variable, $f_i = f_j$ and any dependence between f_i' and f_j' can be neglected. Furthermore, if the f_i are not equal but each is small relative to $\sum f_i$, approximate adjustments may be possible by weighting by a simple function of A_i and $1/f_i$ (in practice $1/f_i'$).

However, the regression procedure may not provide satisfactory estimates of log δ_0. And, of course, estimation of the parameters of location is not included in the regression technique. In fact, it is precluded by the use of a conditional probability function.

The shortcomings of the regression method suggest the possibility of using the multinomial distribution--the standard for random sampling from a multi-celled universe. The mathematical expression, however, does not appear to be tractable --the term for the theoretical probability of an observation in the i^{th} cell being

$$\frac{1}{2\pi}\left(\frac{2}{\mu}\right)^2 e^{-\frac{2}{\mu}\rho_i} A_i \, k$$

where

$$\rho_i = \sqrt{(x-c_x)^2 + (y-c_y)^2} \qquad (13)$$

$$= \sqrt{\rho'^2 + c^2 - 2c\rho' \cos(\phi - \theta')}$$

and k is a factor of proportionality to obtain unity as the sum of the terms. Thus, the complete multinomial expression is obviously not easily applied to empirical data.

I have exhausted the time allotted to me. The results I have presented are, of course, incomplete and perhaps only of limited usefulness. I felt it was best to present them now, since there is no assurance I will have the opportunity to do further work. Perhaps others will and it is my hope that my work to date, such as it is, will be of use to them.

*Unless one wishes to accept least squares per se, or some similar procedure.

Selected References

Blumenfeld, Hans, "On the Concentric-Circle Theory of Urban Growth," Land Economics, Vol. 25, No. 2 (May 1949), pp 209-212.

Clark, Colin, "Urban Population Densities," Journal of the Royal Statistical Society, Series A, Vol. 114 (1951), pp. 490-496.

-----, "Urban Population Densities," Bulletin of the International Statistical Institute, Vol. 36, No. 4 (Actes de la 30e session de l'Institut International de Statistique, Stockholm, 1958), pp. 60-90.

Duncan, Otis Dudley, "Population Distribution and Community Structure," Cold Spring Harbor Symposia on Quantitative Biology, Vol. 22 (1957), pp. 357-371.

Muth, Richard, "Spatial Structure of the Housing Market," (In process, 1960).

Reinhardt, Forrest Rex, A Test of Hypotheses Specifically Related to the Cross-Sectional Distribution of Population Densities of Cities (In partial fulfillment of the Degree of Master of Arts in Sociology), University of Minnesota, September 1950, pp. 104 (typewritten).

Stewart, John Q., "Suggested Principles of 'Social Physics,'" Science, Vol. 106 (1947), pp. 179-180.

-----, and William Warntz, "Physics of Population Distribution," Journal of Regional Science, Vol. I, No. I (1958), pp. 99-123.

Winsborough, Halliman H., A Comparative Study of Urban Residential Densities, (In process, 1960).

URBAN POPULATION CONCENTRATION AND OCCUPATIONAL STRUCTURE IN THE UNITED STATES: 1900-1950

By: Georges Sabagh, Maurice D. Van Arsdol, Jr., and Hamid Zahedi,
University of Southern California[1]

In recent years, an increasing number of studies have been focused on the interrelations between urbanization, modernization, and industrialization. In much of this research one methodological problem is the derivation of measures of urbanization which provide an adequate summary of the distribution of population by community size. This paper presents data to suggest that one logical and useful solution to this problem involves the derivation of measures of urban concentration from some mathematical model of concentration. An index based on the Lorenz curve model is presented and its relationship to other population distribution measures is documented. The research further describes the associations between a Lorenz curve index and measures of economic structure. The analysis is based on data for individual states of the United States from 1900 to 1950.

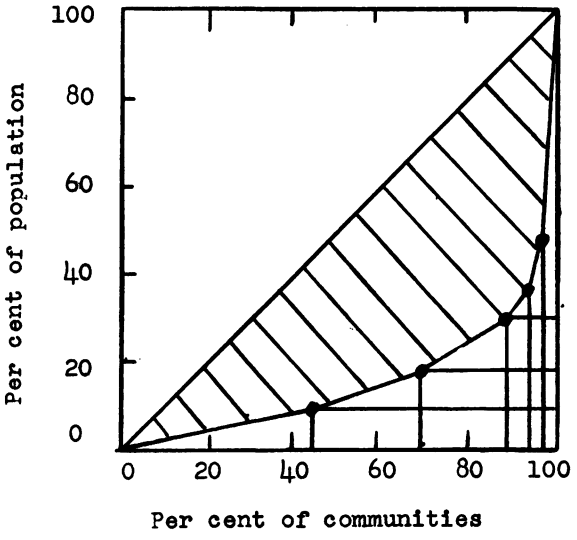
THE MEASUREMENT OF URBAN CONCENTRATION

Concentration indexes were first constructed for income size distributions. These indexes were originally developed by Pareto, Gini, Lorenz, and others, and have been applied to a wide variety of data, including size distributions of communities.[2] Two closely related expressions--Pareto's curve and the rank-size rule--have proven to be the most popular in urban research. However, analyses of the Pareto curve and rank-size rule have been more oriented toward describing empirical regularities and formulating size laws than toward deriving useful concentration indexes. Some investigators of the Pareto curve have realized that the slope of the curve can be used as an index of urban population concentration. To obtain such an index it is necessary to fit the curve to empirical data and then to measure the "goodness" and "closeness" of fit. If observed data depart from the Pareto model it is difficult to evaluate the empirical applicability of the slope as an index of urban population concentration. Concentration indexes that can be more simply derived include those based on the Lorenz curve. Few attempts have been made to measure urban population concentration on this basis, yet the method would appear to warrant further investigation.[3]

Structure of the Lorenz curve. The Lorenz curve describes the joint distributions of the cumulated proportions of communities of different sizes and of cumulated proportions of the populations of these communities. To the degree that the community and population size intervals are infinite, the Lorenz curve approaches a continuous curvilinear function. With a limited number of community and population size intervals it is discontinuous.[4] Figure 1 illustrates the Lorenz curve for community size and population for the United States as a whole in 1950. Coordinates on any of the

FIGURE 1

LORENZ CURVE FOR COMMUNITY SIZE AND POPULATION, UNITED STATES: 1950*
S=.35



*Based on "new" urban definition, and community size intervals of 2,500-4,999; 5,000-9,999; 10,000-24,999; 25,000-49,999; 50,000-99,999; and 100,000 and over.

points on the curve specify the proportions of communities and of populations below a given size level. The extent of urban concentration is proportional to the shaded area between the curve formed by the points and the diagonal. If the cumulated community and population proportions are equal, all of the observed points fall on the diagonal, all communities are of equal size, and minimum urban concentration is found. The observed points fall below the diagonal and a degree of urban concentration is obtained whenever the cumulated proportions of communities are greater than the cumulated proportions of populations. Complete concentration of population into a single city is found if the points of the coordinates delimit the total area of the triangle below the diagonal.

An index derived from the Lorenz curve. A number of indexes of urban population concentration can be obtained from the Lorenz model. The most common is one which delimits the area between the Lorenz curve and the diagonal of minimum population concentration, or equality of community size. This index, as used in the present analysis, is denoted as "S," and defined as follows:[5]

$$S = \tfrac{1}{2}\left[1 - (X_1Y_1) + (X_2-X_1)(Y_1+Y_2) + \ldots \ldots + (X_k - X_{k-1})(Y_{k-1} + Y_k)\right] = \tfrac{1}{2}\left[1 - \sum_{i-1}^{k} (X_i - X_{i-1})(Y_{i-1} + Y_i)\right]$$

where: X_i = the cumulated percentage of communities at point i; Y_i = the cumulated percentage of population at point i; and the range of S is .0 to .5. The S index value of .35 shown in Figure 1 indicates that the area between the Lorenz curve and the diagonal of minimum concentration is seventy per cent of the maximum possible area below the diagonal.

Computation of the S index for individual states. The S index was computed for each of the 48 states of the United States for the decennial census years 1900-1950. The six community size intervals used in these calculations ranged from 2,500 - 5,000 to 100,000 and over.[6] The choice of these intervals was guided by the availability of data and by the need for comparability with another analysis being prepared for various countries of the world.[7]

The S index values for individual states from 1900 to 1950 ranged from a minimum of .00 to a maximum of .43. Table 1 shows that the median values of the S index for states have increased slightly over the period 1900 to 1950 and reflect a trend toward urban population concentration documented by other investigators.

TABLE 1

MEDIAN STATE VALUES AND SEMI-INTERQUARTILE RANGE OF S, DAVIS, AND PER CENT URBAN INDEXES

UNITED STATES: 1900-1950

	S		Davis**		Per cent urban*	
Year	Mdn.	Q.	Mdn.	Q.	Mdn.	Q.
1950	.32	.08	.46	.10	.48	.09
1940	.32	.05	.47	.11	.42	.01
1930	.31	.06	.45	.11	.40	.13
1920	.30	.06	.50	.14	.36	.16
1910	.30	.06	.52	.15	.31	.15
1900	.27	.07	.59	.14	.28	.16

*Based on U.S. Bureau of the Census "old" definition.

**See footnote 9 for explanation of Davis index values.

Relations of the S index to other measures of urban population distribution. To evaluate the Lorenz curve as an independent measure of urban population distribution it was compared to two other indexes. One of them is the traditional census measure of the per cent of the population residing in incorporated communities with 2,500 or more inhabitants.[8] The other is an index comparable to one devised by Kingsley Davis to measure urban concentration.[9] This modified Davis index was obtained by taking an arithmetic average of the set of cumulated percentages of the urban population in communities of different

size, thus implicitly assigning an equal weight to each size interval. The census and the Davis indexes were calculated for each state from 1900 to 1950. The medians of these state values as well as the medians of the S indexes for states are given in Table 1. The intercorrelations of the state values of S with the Davis and percent urban indexes for 1900-1950 are shown in Table 2.

TABLE 2

ZERO ORDER CORRELATIONS OF STATE VALUES OF THE S, WITH THE DAVIS, AND PER CENT URBAN INDEXES
UNITED STATES: 1900-1950

Year	S-Davis	S-Per cent urban*
1950	-.75	.46
1940	-.51	.52
1930	-.40	.57
1920	-.40	.62
1910	-.32	.59
1900	-.21	.56

*Based on U.S. Bureau of the Census "old" definition.

The marked increase in the medians of the state values of the census index from 1900 to 1950 stands in definite contrast to the small change in S index values and a moderate increase in the values of the Davis index.

Table 2 indicates that the correlation between the S index and the per cent urban increased slightly between 1900 and 1920 and then decreased continuously until 1950. While the simple urban measure was once more closely related to the S index, the increasing amount of unexplained variance between them suggests that substitution of one measure for the other is not justified.

The marked increase in the correlation between the S and the Davis indexes from 1900 to 1950 may be explained by the fact that, with a fixed set of community-size intervals, the Davis index is more affected than the S index by the absence of communities in some of these intervals.[10] A consolidation of the upper two or three community-size intervals would have undoubtedly given a consistently higher correlation between the two indexes from 1900 to 1950. By 1950, when most states had communities in at least five of the six community-size intervals, the correlation given in Table 2 better reflects the relationship between the S and Davis indexes. The data for 1950 suggest that, in general, the S index correlates more closely with the Davis index than with the census index of per cent urban. Both the S index and the Davis index measure the extent to which the urban population of a state is concentrated in larger cities, whereas per cent urban more simply indicates the relative importance of all cities above 2,500 population. The main advantage of the S index over the Davis index is that it is based on a more general model of concentration.

URBAN CONCENTRATION AND OCCUPATIONAL STRUCTURE

The process of urbanization in Western nations has been linked closely with transformations in the basic nature of economic activity.[11] The initial shift from agriculture and home industry to factory-centered manufacturing has been accompanied by major currents in population redistribution. The next phase of economic modernization, characterized by the emergence of service and consumption oriented occupations, also witnessed significant population shifts. These economic changes can be portrayed, though somewhat crudely, by the changing importance of primary or predominantly agricultural occupations, secondary or manufacturing occupations, and service or tertiary occupations. One would expect, therefore, that there would be some association between measures of urbanization and indexes of occupational structure. In this respect, negative relationships are hypothesized between the level of urban population concentration and the proportion of the labor force in primary occupations, and positive relationships between urban concentration and the proportion of the labor force in secondary and service occupations.

Available published information

limits the scope of the description of the relationships between occupational structure and urban population concentration in the United States. By 1900, the first date for which measures of urban population concentration by states were derived, major changes in the growth and distribution of urban population had already taken place. Thus, the data to be presented are indicative of what may happen in a mature economy rather than in an economy in the initial phases of urbanization and industrialization.

Data on the occupational structure of each state of the United States in 1900, 1940, and 1950 were obtained from the University of Pennsylvania Study of Population Redistribution and Economic Growth.[12] The available occupational subclasses were combined to form primary, secondary, and service occupational classes. The proportions of gainful workers or persons in the labor force in each of the three occupational classes were computed for each state for 1900, 1940, and 1950.[13]

TABLE 3

ZERO ORDER AND MULTIPLE LINEAR CORRELATIONS BETWEEN STATE VALUES OF S INDEX AND PROPORTION OF THE LABOR FORCE IN PRIMARY, SECONDARY, AND SERVICE OCCUPATIONS
UNITED STATES: 1900, 1940, and 1950

	Years		
Variables	1900	1940	1950
S. primary	-.39	-.42	-.42
S. secondary	.23	.23	.21
S. service	.46	.34	.31
S. primary, secondary	.47	.44	.53
S. primary, service	.46	.43	.43
S. secondary, service	.49	.40	.42

Table 3 presents zero order and multiple linear correlations by states between the S index and the per cent of the labor force in primary, secondary, or service occupations for the years 1900, 1940, and 1950. While these correlations are low, they do form a pattern conforming to the hypotheses. For 1900, 1940, and 1950, the proportion of the labor force in primary occupations by states was found to be inversely related to the degree of urban concentration, without any changes over time. On the other hand, the correlations of secondary and tertiary occupations with urban concentration are positive for all years under consideration. An important historical change that is shown in Table 3 is the decreasing correlation of tertiary occupation and urban concentration over time. The same trend is noted in the multiple correlations when tertiary occupation is included as one of the independent variables.[14] In 1950 the highest multiple correlation was obtained for the combination of primary and secondary occupations.

The development of tertiary occupations may be partially a function of urban concentration which creates a need for service activities that are initially carried on within large cities. However, with the passage of time as the total society becomes more concentrated, and as techniques of transportation and communication are further developed, service activities may have become more rapidly dessiminated, allowing a deconcentration of both place of residence and place of work. This process may have been hastened by the development of white collar "dormitory" suburbs in recent years, which in the present analysis, would tend to increase the number of small communities and lower the value of the urban concentration index.

SUMMARY AND CONCLUSIONS

This paper has documented the importance of using measures of urban population concentration in the study of social and economic changes associated with urbanization. Indexes of urban population concentration derived from the Lorenz curve were described. It was shown that Lorenz curve index values tend toward independence from the traditional measure of per cent urban used by the United States Bureau of the Census. Possible uses of one of the Lorenz curve indexes in the study of changes in the economic structure of an urbanizing society were investigated. It was found that there is

a definite pattern in the relationship between urban concentration and the occupational structure characteristic of a maturing economy.

The findings of this study are tentative in nature. Further investigation is needed of the relationship of concentration indexes to other measures of urban population distribution and to more sensitive measures of economic development and structural changes in social organization. It would appear that research on urbanization needs to take into account more rigorous definitions of population distribution, and to focus on population redistribution as an important variable in the study of social change.

FOOTNOTES:

[1]This research was made possible, in part, by National Science Foundation Grant G 9452.

[2]Mary Jean Bowman, "A Graphical Analysis of Personal Income Distribution in the United States," American Economic Review, 35, 1945, pp. 607-628; and J. Aitchison and J.A.C. Brown, The Lognormal Distribution, Cambridge, Cambridge University Press, 1952. Chapter II. There is some evidence of the independent derivation of community size concentration indexes. For a discussion of early contributions in this area see: Otis Dudley Duncan, "The Measurement of Population Distributions," Population Studies, 11, 1957, pp. 28-30, and 39-42; United Nations, The Determinants and Consequences of Population Trends. New York, United Nations, 1953, pp. 175-176; and Eugenio D'Elia, Su alcuni metodi di calcolo della media aritmetica e della concentrazione nelle distribuzione multiple di frequenza. Roma, Studie e monografie della società Italiana di demografia e statistica, 3, 1949, pp. 10-18.

[3]E'Elia, op. cit., pp. 10-18, and "Methodologie statistique: la concentration," Bulletin de statistique (Belgium, Office central de statistique), 30, 1944, pp. 109-115. It should be noted that, by contrast, the Lorenz curve has been used extensively in the analysis of income distributions.

[4]It is possible to fit a curvilinear function to the observed data for the Lorenz curve, thus coming closer to a distribution that would have been observed with more detailed data. Such a procedure is not usually followed due to complications of curve fitting. Indexes derived from the actual data are approximations as they are based on a small sample of points.

[5]See: Duncan, op. cit., pp. 29-30; Aitchison and Brown, op. cit., pp. 112-113; and "Methodologie statistique: La concentration," op. cit., pp. 110-111. The notation given in the Belgian article was modified to conform with notation given by Duncan. Duncan's concentration ratio is identical with the S index, which is called "surface de concentration."

[6]See Figure 1 for a description of the size intervals used for the computations of the S index.

[7]Values of the S index may be affected by a change in the number of community size intervals used or by a modification in the lower or upper limits of the intervals selected. This problem is not given separate consideration in this paper.

[8]U.S. Bureau of the Census, U.S. Census of Population, 1950. Vol. II. Characteristics of the Population, Part I. United States Summary, U.S. Government Printing Office, Washington, D.C., 1952.

[9]Use of the Davis index is described in Kingsley Davis and Ana Casis, "Urbanization in Latin America," The Milbank Memorial Fund Quarterly, 24, 1946, pp. 186-207; Kingsley Davis, The Population of India and Pakistan, Princeton, Princeton University Press, 1951, p. 129; Jack P. Gibbs and Walter T. Martin, "Urbanization and Natural Resources: A Study in Organizational Ecology," American Sociological Review, 23, 1958, pp. 266-277.
Due to the manner in which the basic data were accumulated for the computation of the Davis index, Davis index values decrease with increasing urban population concentration and increase with decreasing urban population concentration.

[10]The number of states with no communities in at least three of the six intervals decreased from 15 in 1900, to 10 in 1910, 8 in 1920, 6 in 1930, 3 in 1940, and 1 in 1950.

[11]Colin Clark, The Conditions of Economic Progress, 3rd edition, London, Macmillan and Company, Ltd., 1957.

[12]Everett S. Lee, Ratner Miller, Carol P. Brainerd, and Richard A. Easterlin, Population Redistribution and Economic Growth, United States, 1870-1950. Vol. 1, Methodological Considerations and Reference Tables. Philadelphia, The American Philosophical Society, 1957, pp. 623-631.

[13]Ibid. Proportions are for gainful workers for 1900 and for the experienced labor force for 1940 and 1950. Primary, secondary, and tertiary gainful workers and persons in the experienced labor force are classified according to an interpretation of Colin Clark, op. cit. Agriculture, forestry, and fishing are classified as primary occupations; mining, construction, and manufacturing as secondary occupations; and transportation, trade, finance, services, and public administration as service.

[14]As the variance of a third occupational proportion is a function of the variance of the other two occupational proportions, only two occupational proportions were considered in each of the multiple regression equations.

STATUS OF RESEARCH ON METHODS OF ESTIMATING STATE AND LOCAL POPULATION*

Jacob S. Siegel, U. S. Bureau of the Census

This paper is a progress report on selected phases of the present status of research on methods of estimating State and local population, especially the work being done at the Census Bureau. This paper does not propose any new methods of estimation or present the results of any major new series of tests of methods of estimating population. At this date, it is possible to make only very preliminary comparisons on the basis of the 1960 population. Some such preliminary comparisons are described in this paper.

State Estimates

During the last decade the Census Bureau conducted systematic tests of selected methods of estimating State population using the 1950 Census counts and estimates for 1943 based on registrations for war ration books as standards of comparison. 1/ The methods tested included two mathematical methods, the Census Bureau's component Methods I and II, the vital rates method, and two variations of the composite method. (A brief outline of several methods is given below.) Component Method II gave consistently superior results in these tests, although its leads over the vital rates method, the composite methods, and averages of Method II and vital rates and of Method II and Method I were small. As you may recall, Method II calls for estimating net migration for States from the net migration of school-age children, which is, in turn, developed from data on school enrollment. In the composite methods the estimate of total population is obtained by combining estimates by age computed by using different indicators, such as school enrollment, births, and deaths.

A comparison of the preliminary census counts for States and provisional postcensal estimates for 1959 projected to April 1, 1960 provides a rough basis for measuring the adequacy of the present method of estimating State population used by the Census Bureau. Inasmuch as component Method II was used without modification for nearly all States (43), the comparison also serves as a rough basis of evaluating Method II for States. The average percent error over all States (including District of Columbia) was 2.8 percent. The estimates for only three States (Alaska, Hawaii, and West Virginia) and the District of Columbia were in error by more than 5 percent. (If the States for which the estimates were not based on Method II are omitted, the error drops to 2.2 percent.) The average error for Method II alone computed for the 48 States and the District of Columbia is 2.6 percent. These figures imply a substantial improvement over the experience of 1950, when the Census Bureau's Method II test estimates for States differed from the final census counts by an average of 3.5 percent. 2/

A program of systematic evaluation of various estimating procedures for States against the 1960 Census is planned by the Census Bureau in 1961 and 1962. The present plans call for testing the following methods and appropriate combinations of them: component Method II, vital rates method, Bogue-Duncan variation of the composite method, component Method I, the Census Bureau variation of the composite method, and the age-or grade-progression method, insofar as possible. For the most part, the previous tests provided specific guidance as to the methods which should be covered in the forthcoming tests.

For the purpose of these tests, the Bogue-Duncan variation of the composite method as originally described by the authors has been simplified by limiting the detail to only five age groups by color (except the school-age group). The Census Bureau variation of the composite method modifies the procedures and indicators used for the various age groups so as to take account of the fact that in the previous test for States, Method II showed a much smaller error for ages 18-44 than the fertility ratio procedure. In these tests Method II showed approximately the same average error as the death rate procedure at ages 45-64 and a substantially greater error at ages 65 and over. 3/ Hence, in the variation proposed for testing, Method II-type computations are to be employed at the three age groups under 45 and the death rate procedure for the two age groups over 45.

County and City Estimates

Recent changes have altered strikingly the situation and prospects with respect to the

*/ The writer wishes to acknowledge the technical assistance of Donald S. Akers, of the Bureau of the Census.

1/ J. S. Siegel, Henry S. Shryock, Jr., and Benjamin Greenberg, "Accuracy of Postcensal Estimates of Population for States and Cities, "American Sociological Review, Vol. 19, No. 4, August 1954, pp. 440-446; Henry S. Shryock, Jr., J. S. Siegel, and Benjamin Greenberg, "Current Research on Population Estimates for States and Local Areas," unpublished paper read at the 1957 Annual Meeting of the Population Association of America, Philadelphia, May 4, 1957.

2/ Siegel, Shryock, and Greenberg, "Accuracy of Postcensal Estimates of Population for States and Cities," op.cit.

3/ Shryock, Siegel, and Greenberg, "Current Research on Population Estimates for States and Local Areas," op.cit.

BRIEF OUTLINE OF METHODS BEING TESTED FOR STATES OR COUNTIES IN 1960

Method	Estimating procedure	Basic indicator
1. Census Bureau Method II	Component method: For migration, school-cohort procedure comparing expected population, based on previous census plus births, with actual population.	For migration: school data
2. Census Bureau Method I	Component method: For migration, change in local school-age population compared with change in national school-age population.	For migration: school data
3. Vital rates method	Censal ratio (birth rate and death rate)	Births and deaths
4. Composite method: Bogue-Duncan variation	Censal ratio by age	--
Age group		
0-4	Ratio of children under 5 to women 18-44	Births
5-17	School enrollment ratio	School data
18-44	Fertility ratio (births to women) and sex ratio)	Births
45-64	Death rate	Deaths
65 and over	Death rate	Deaths
5. Composite method: Census Bureau variation		
Age group		
0-4	Component Method II	School data
5-17	Component Method II	School data
18-44	Component Method II	School data
45-64	Censal ratio (death rate)	Deaths
65 and over	Censal ratio (death rate)	Deaths
6. Age or grade progression method a/	Component method: one-year school-age or grade "survival" rate for migration	For migration: school data
7. Censal ratio method using school data b/	Censal ratio	School data

a/ States only.

b/ Counties only.

availability of adequate current estimates of county population. Surveys conducted by the Census Bureau in 1955, 1957-58, and 1960 provided information on whether any State agencies prepared and published sets of estimates for the counties in their States, how frequently, and by what method. 4/ (These surveys also covered estimates of city population made by State agencies and city agencies in cities of 250,000 and over.) Over this period, there has been a substantial increase in the number of States for which county estimates are prepared, and a notable shift from the use of the less adequate methods to the more adequate ones. Now nearly all States prepare some type of county estimates (47), most of them by a component or censal ratio method. 5/

One factor in this shift is the larger role of the Census Bureau, including the publication of a report giving detailed instructions on the application of component Method II, increasing direct assistance to the States in their estimate work, the conduct of tests of the accuracy of various methods, and the conduct and publication of periodic inventories of the estimates work of State and city agencies. Other Federal agencies, in particular the National Office of Vital Statistics and the Bureau of Employment Security, have played a complementary role. These agencies have republished descriptions of recommended methods of making local population estimates;6/ and the National Office of Vital Statistics has regularly sponsored a working or study group on population estimates in its Conferences on Records and Statistics.

Another factor is the increasing recognition on the part of State governments of the need for adequate current estimates of county population and their organizing to fill this need. State population estimates committees have been organized in a number of States, among them Florida, Oklahoma, New Hampshire, and Utah. These bring together representatives of the State agencies which produce the basic data and those which are the principal consumers of estimates, as well as research technicians in the State universities, in order to exchange information, agree on the scope of the estimates program and the methodology to be used, and assign responsibility for preparing the estimates. Population studies have been given an official status in the State government in California, Oregon, Washington, and also the District of Columbia. In most of the remaining States, State Departments of Health or University Bureaus of Business Research have assumed the responsibility for preparing the county estimates. Recent legislative changes in California have given a considerable impetus to the research on, and preparation of, estimates for cities using the dwelling unit method since estimates of this kind are now acceptable in lieu of special censuses in establishing population size for the allocation of State tax funds.

No extensive test of methods of estimating county and city populations has been carried out, but several limited tests have been conducted, mostly on the basis of the 1950 Census. Among these mention can be made of the National Office of Vital Statistics' test of Method II for West Virginia counties, Schmitt's test of short-cut methods of estimating county population in Washington State, the Pennsylvania Bureau of Statistics' test of short-cut methods for Pennsylvania counties, the University of Chicago's test of the composite method in 15 urban places in Illinois, Frisén's test of the dwelling unit method in 51 California cities, and the Census Bureau's tests of several methods of estimating the population of cities of 100,000 and over and of metropolitan counties and standard metropolitan statistical areas with central cities of 250,000 or more. 7/ It was on the basis of the latter tests that the Census Bureau has been

---

4/ U.S. Bureau of the Census, Current Population Reports, Series P-25, No. 178, "Local Population Estimates Prepared by State and City Agencies: 1957-58", June 27, 1958; Current Population Reports, Series P-25, No. 116, "Current Status of Population Estimates Prepared by State Agencies," June 6, 1955; and unpublished records of U. S. Bureau of the Census.

5/ An important use made of these estimates in the late fifties was their adoption by the Federal Bureau of Investigation in computing crime rates for Uniform Crime Reports-1958 in place of 1950 Census figures. Another was their use by the Census Bureau in preparing the population estimates required for planning the field organization of the 1960 Census. For this purpose it was deemed desirable to replace the available estimates for several States, from State agencies, by other estimates. Accordingly, rough estimates prepared in the Census Bureau by the vital rates method were substituted in these cases.

6/ National Office of Vital Statistics, "A Composite Method for Estimating Postcensal Population of Small Areas by Age, Sex, and Color, Vital Statistics-Special Reports, Vol. 47, No. 6, August 24, 1959; Bureau of Employment Security, Handbook of Estimating Population of Labor Market Areas, November 1959.

7/ National Office of Vital Statistics, "Study of Population Estimates Made for Each County in West Virginia, as of April 1, 1950," processed, presented by Robert D. Grove at the Third Annual Meeting of the Public Health Conference on Records and Statistics, Washington, D. C., April 23, 1951. Robert C. Schmitt, "Short Cut Methods of Estimating County Population", Journal of the American Statistical Association, June 1952, Vol. 47, pp.232-238. Pennsylvania Bureau of Statistics, Population Statistics,

recommending the use of an average of Method II and the vital rates method for estimating city and county populations.

It is already possible to make a rough evaluation of the accuracy of the estimates for counties prepared by State agencies in recent years. Comparisons were carried out between preliminary census counts for counties and extrapolations to 1960 of the latest available estimates by State agencies, mostly for 1958 and 1959. More appropriate comparisons can be made next year or the following year when postcensal estimates for 1960 can be prepared. Although different methods are used for different areas and the present estimates are subject to substantial modification, the results are suggestive. They suggest that relatively smaller errors tend to occur with the composite and component methods than with the censal ratio methods, including the vital rates method. For New Hampshire, where school census data are combined with data from head tax counts, the average error for counties was only 4 percent. Composite estimates in Maryland and Wisconsin also had average errors of 4 percent. Average errors ranged from 4 to 8 percent for Indiana, Ohio, Oklahoma, and Florida, where an average of Method II and the vital rates method was employed. When the vital rates method was the sole basis of estimation, as in Louisiana, Mississippi, and Montana, the average errors extended from 7 percent to over 15 percent. Often the average error for a State is markedly increased because a few small counties have extremely large percentage errors. Hence, the average error weighted in relation to the population of the counties is usually smaller; and it may be considerably smaller. For example, on this basis, the error for California counties drops from 9 percent to 5 percent.

A program of evaluating methods of estimating county population has been undertaken by the Study Group on Postcensal Population Estimates, one of several working groups sponsored by the Public Health Conference on Records and Statistics, of the National Office of Vital Statistics. The work of the members of this group and of several other technicians from State agencies cooperating in this program is being coordinated by the Population Estimates and Projections Branch of the Bureau of the Census. The Study Group includes representatives of four States--Pennsylvania, Montana, Ohio, and Oregon. In addition, technicians in several other States--Oklahoma, Florida, West Virginia, and Tennessee--are carrying out all or part of the test program of the Study Group. In its several meetings over the last two years, the Study Group has developed plans with respect to the methods to be tested, the specific procedures for applying the methods, the measures of evaluation to be used, and the various factors to be taken into account in analyzing the results. The cooperating technicians receive copies of documents prepared for the Study Group and receive information and counsel from the Census Bureau by letter or personal conferences.

With certain exceptions, the methods to be tested are essentially the same as those for which the Census Bureau plans to conduct tests for States: Component Method II, the vital rates method, Bogue-Duncan composite method, Census Bureau composite method, component Method I, and the censal ratio method using school enrollment. Combinations of selected methods will be determined and carried out after the results for the separate methods have been obtained. In general, it seemed desirable to include in the test the methods which appeared most promising in previous tests for States (component Method II, Bogue-Duncan composite method, Census Bureau composite method). At the same time it seemed desirable to include certain relatively simple methods proposed in the literature which might yield satisfactory results for little investment of resources (vital rates method, component Method I, censal ratio method employing school enrollment or school census data). It may be recalled that, of the various censal ratio procedures other than the vital rates method, that employing school data has shown the most promising results. 8/ The Study Group also plans to carry out a limited test of procedures of estimating the population of cities, but these plans are rather indefinite at this time.

The members of the Study Group have been developing postcensal estimates by some of these methods for July 1, 1958 and 1959, partly to prepare for the computation of the estimates

---

(7/ continued)
Release No. P-1, "County and City Population Estimates for Pennsylvania", May 1959. University of Chicago, Population Research and Training Center, "Estimates of Population for State Economic Areas, Counties, and Cities in Illinois: 1955 to 1957," a report to the Department of Public Health, State of Illinois, dittoed, ca. 1958. Carl Frisen, Report to the League of California Cities on a Test of Population Estimating Techniques Applied to Selected California Cities, processed, California Department of Finance, March 1957. J. S. Siegel, Henry S. Shryock, Jr., and Benjamin Greenberg, "Accuracy of Postcensal Estimates of Population for States and Cities," op.cit. Henry S. Shryock, Jr., "Development of Postcensal Population Estimates for Local Areas," in National Bureau of Economic Research, Studies in Income and Wealth, Vol. 21, Regional Income, pp. 377-391, 1957.

8/ Robert C. Schmitt, op.cit; Pennsylvania Bureau of Statistics, op. cit.

for April 1, 1960, but also to observe the range of variation between methods in a particular year and the consistency of estimates by the same method from one year to another. The estimates for 1960 and the comparisons with the census figures should be completed by mid-1961. The Study Group plans to issue a report on its findings in late 1961 or early in 1962. The report would present data for each State participating in the program and for the participating States taken together. The analysis would include comparative results for the several methods and would consider the effect on estimating error of such factors as size of county, rate of growth between 1950-60, and metropolitan status (that is, whether a nonmetropolitan county, central metropolitan county, or outlying metropolitan county). On this basis, the report would make specific recommendations regarding methods to be used for estimating county population. It is quite possible, of course, that the recommended method would differ from one type of county to another. To supplement this set of tests for the counties in selected States, the Census Bureau plans to carry out a set of tests of methods of estimating the population of SMSA's, and their component counties and central cities.

In any review of the status of local estimates, it seems pertinent to mention the work of the various commercial organizations which regularly publish current sets of estimates for counties and cities. Sales Management, Inc., publishes population estimates for counties and principal cities for the previous January 1 annually in Sales Management. Standard Rate and Data Service publishes population estimates semi-annually (as of January 1 and July 1) for counties and principal cities in its radio, newspaper, and television publications. Editor and Publisher Company, Inc., annually publishes estimates as of January 1 in its Market Guide.

According to the text, the estimates in Sales Management are based on three elements: (1) "individual growth trends as established over the past decade", (2) "population growth adjusted by current sales data" (for 600 areas), and (3) "Chamber of Commerce reports". "Every chamber of commerce in the nation was solicited by mail, telegrams, and follow-up inquiries to report, for the city and county involved, changes in population since 1950, changes in the number of public utility connections since 1950, plus any unusual economic developments which might affect population growth". Standard Rate and Data Service states that it makes use of all special census counts and estimates prepared by the Census Bureau as well as estimates made by various State and local organizations, if, after analysis of the estimates, the procedures used were found acceptable. In addition, information was sought to locate geographically any unusual business activity which might cause population shifts. Editor and Publisher bases its estimates on mathematical correlation between state and county figures.

To my knowledge none of these organizations has conducted any tests of the accuracy of its estimating methods. The descriptions of the methods given in the publications do not permit independent application or testing of the methods, except possibly in the case of the estimates of Editor and Publisher Company. In order to test empirically the accuracy of some of the estimates just described, those for counties and standard metropolitan statistical areas and their central cities, given in one of the commercial compilations, were compared with the preliminary census results. It is planned to extend the test to include the other sets of commercial estimates, as resources and time permit. The most striking fact revealed by this evaluation is the considerable upward bias and the much larger error of the estimates for central cities, as compared with the estimates for SMSA's. Eighty percent of the city estimates were greater than the city counts, as compared with 48 percent for the SMSA's; the average error for the central cities was 9.3 percent, as compared with 3.4 percent for SMSA's; and 39 percent of the estimates for cities differed from the counts by 10 percent or more, as compared with 6 percent for SMSA's. The errors of the estimates for individual central cities were almost invariably far greater than the estimates for the corresponding SMSA's as a whole. In 11 cases, the city estimates and the city counts differed by more than 20 percent and in 7 of these cases, the SMSA was estimated correctly within 5 percent. This and other evidence indicates that it is possible to develop much more reliable estimates for metropolitan areas than for large cities and the considerable confidence in the available estimates for larger cities during the fifties was hardly justified.

The evaluation of the county estimates in the same commercial compilation indicated an average error of 8 percent, or nearly the same as the error for central cities. This was about the level of error in the county estimates hastily compiled from various sources by the Bureau of the Census, mentioned above, for its planning of the field organization for the 1960 Census.

Improvements in Methodology

In addition to the need for evaluating present methods and estimates, it seems desirable to give continuing consideration to their technical improvement and to the possible advantages of new types of data and methods. Although the Census Bureau developed component Method II as a demographically direct and systematic procedure for estimating the population of States and has generally adopted this method for this purpose, it has never viewed the

method as having attained an ultimate form but rather as subject to continuing reexamination and modification. Several improvements were incorporated in the method in the last decade. The most important of these was the introduction of the use of migration factors, representing the ratio of the migration rate of the total population to the migration rate of the school-age population for a given period, which factors varied with the length of the period of estimation. Derived from national sample data on gross interstate migration by broad age groups for 1-year periods from the Current Population Survey, the factors were seen to change markedly with the length of the estimating period. Specifically, they declined gradually from 1.27 for 1950-51 to 0.85 for 1950-59. This decline is due to the fact that the longer the period, the greater the exposure of younger children in the school-age cohort to migration, combined with the fact that migration rates are higher at the younger ages. 9/ Although empirical proof of the efficacy of this change, apart from theoretical considerations, is not available, the considerable reduction in the discrepancy between the national immigration figure and the sum of the computed figures on net interstate migration suggests strongly that the change constituted a considerable improvement.

A second important change in Method II has been the introduction of the use of enrollment in grades 2 to 8 and of the age range 7½ years to exact age 15½ years to match this grade range, in connection with the estimation of the population of school age. Census Bureau experimentation with various combinations of grades and ages in making estimates for States indicated substantial improvement with this combination over the combination of grades 1 to 8 with ages 6 to 13 previously used. 10/ A third change, supported by test results, is the use of adjustments to national control figures at several stages in the computation of the estimates of net civilian migration, as compared with making a single final adjustment. 10/

It seems appropriate to mention here, among the "improvements" in Method II, the preparation of a computer program for Method II for use on the IBM 650 computer by Professor James W. Tarver of Oklahoma State University. Professor Tarver has also prepared a program for the vital rates method for use on the same computer. These programs are to be published shortly.

In its future experimentation on Method II, the Census Bureau will give further consideration to the form and adjustment of the net migration rate. The rate now used represents essentially the ratio between net migration in a period and the census population plus one-half births minus net loss to Armed Forces. This definition of the migration rate is believed to be subject to criticism on at least two grounds: The selection of the categories used in the base and the use of only a portion of the births. The writer believes that the more appropriate base is the census population plus all births. However, according to a limited test, relating to several States for the period 1950-57, the effect of using all births in deriving the migration rate, and hence the amount of net migration, was small.

The possibility of using a migration factor varying from State to State on the basis of previously observed age differences in net migration rates needs to be reexamined when the 1960 Census age data on interstate migration come to hand in late 1961. At present a single factor is used by the Census Bureau for all States for a given estimating period, derived as noted earlier. Some State variation in the age pattern of net migration may be introduced by employing "color-weighted" migration ratios for each State, computed on the basis of national migration data by color to become available for 1958-59 from the Current Population Survey and for 1955-60 from the 1960 Census. This possibility also applies to county estimates. The computation of separate ratios for intrastate migrants in the preparation of county estimates, to replace or modify the ratios for interstate migrants now used, is also possible on the basis of data from the Current Population Survey. A further variation could be introduced for metropolitan counties, nonmetropolitan counties, and central cities; the basic data will become available for 1958-59 from the Current Population Survey and for 1955-60 from the 1960 Census.

In Method II more attention needs to be given to the handling of special groups which may constitute a sizeable segment of the population of certain counties. Method II already calls for the separate handling of members of the Armed Forces, but it would also be desirable to handle separately such other groups as institutional population, college students, and migratory laborers, for selected areas. Perhaps more attention needs to be given also to the inclusion of all types of schools in compiling enrollment figures, e.g., parochial schools, Federally supported schools on military bases,

9/ U. S. Bureau of the Census, Current Population Reports, Series P-25, No. 165, "Estimates of the Population of States: July 1, 1950 to 1956", November 4, 1957.

10/ Shryock, Siegel, and Greenberg, "Current Research on Population Estimates for States and Local Areas," op. cit.

Indian Bureau schools, etc., and to the definition and geographic reference of enrollment used in each State.

One type of data frequently suggested as worthy of exploration for making population estimates is income tax data. On the basis of preliminary discussions between staff members of the Census Bureau and the Internal Revenue Service, it appears that the only IRS records useful for making estimates of county population are the IBM cards located in the District Offices, used for checking the accuracy of the computations on the tax forms. The use of these data for this purpose would be quite expensive and would entail numerous estimating problems. The Bureau of the Census plans to explore this possibility further, however.

Although they are widely applied for city estimates, the various dwelling unit methods remain subject to serious hazards. In spite of the increasing availability and improvement of data on new construction, their usefulness is limited by the frequent lack of data on the other components of change in the housing inventory and, more seriously, by the problem of estimating changes in the vacancy rate and in the number of persons per occupied unit. Short of a large sample survey, there is no way to establish these changes very closely. Use of the 1950 ratios and mere mathematical extrapolation proved inadequate for the fifties. Nor have national changes, in general, been a satisfactory guide to local changes, which vary considerably in place and time. Locally sponsored estimates by the dwelling unit method for recent dates for several cities were wide of the mark, e.g., Philadelphia, New York, and District of Columbia. Characteristically, they have had an upward bias.

The only careful test of this method to date was conducted by Carl Frisén employing estimates for these cities based on building permits and the results of various special censuses of cities in California as standards. 11/ His comparisons for 32 cities outside Los Angeles County in 1955 and 1956 showed an average percent error of 7.7 percent; for 19 cities in Los Angeles County the average error was 4.0 percent. Frisén's conclusion was that if satisfactory records of change in the number of dwelling units can be maintained, as in Los Angeles County, the dwelling unit procedure can give reliable estimates. This test experience is quite limited and needs to be extended. On the whole, there seems less risk of error in the use of one of the censal ratio procedures for which there is a more satisfactory system of collection of basic data locally and for which national changes provide a more definite guide to local changes. This would seem to be true of the vital rates method or the censal ratio method based on school enrollment or the school census. The dwelling unit method calls for such broad assumptions that, like the vital rates method, it would seem to serve most effectively when used in combination with other reliable means of preparing population estimates.

The Census Bureau has explored the possibility of using special tabulations from the records of the Bureau of Old-Age and Survivors Insurance in the development of composite estimates for States. Two types of data are involved: a 1-percent sample tabulation of covered workers by State of employment and a tabulation of the interstate migration of aged beneficiaries. The latter data are now employed by the Bureau in the preparation of the annual estimates of the population 65 and over for States. They were tabulated annually from 1955 to 1958, and will be tabulated biennially henceforth. It is planned to compare the estimates of net interstate migration of aged persons for 1955-60 from this source with the corresponding census data on migration, and also the estimates of aged persons with the census counts, when the appropriate data become available. An important problem in using these data arises from differences in coverage by social insurance and from the need to make allowances for the population not covered. The covered worker data involve several additional problems. One is that they relate to place of work rather than place of residence; this could seriously affect the figures for a number of States. Another is the large sampling errors of the 1-percent sample. Furthermore, the covered worker tabulations were found to be rather expensive. Because these data have serious limitations and were costly to secure, experimentation with them has been discontinued.

Another possible source of data for making State population estimates is automobile registration. The American Association of Motor Vehicle Administrators has been encouraging in its view that the various State agencies concerned with registration could be induced to collect the data on interstate transfers, to be used in making population estimates by a component method. Plans for securing such data remain to be worked out.

In conclusion, I should like to note: (1) that substantial progress was made during the fifties in the improvement and extension of local estimates and in testing of methods, (2) that the errors of present methods of estimation are sufficiently great to warrant continuing

---

11/ Carl Frisén, op.cit.

efforts to improve them and to extend testing programs, and (3) that the outlook for the availability of reliable estimates covering the larger counties and cities of the United States by the end of this decade is moderately good.

The prospects for improving the accuracy and detail of State and local population estimates would be substantially improved if present proposals for a quinquennial census were realized.

XII

SURVEYS AND SOCIOMETRICS

Chairman, Eli Marks, National Analysts, Inc.

Estimating the Kill of Game Animals by Licensed Hunters — Lee Eberhardt and R. M. Murray Michigan Department of Conservation

Checking the Validity of the Toll Collectors' Vehicle Classifications — Kenneth C. Pearson, Massachusetts Port Authority

On Theoretical Questions Underlying the Technique of Replicated or Interpenetrating Samples — J. C. Koop, North Carolina State College

Analysis of Sociometric Structure: A Method of Successive Grouping — Jack Sawyer and Terrance A. Nosanchuk, University of Chicago

ESTIMATING THE KILL OF GAME ANIMALS BY LICENSED HUNTERS

Lee Eberhardt and R. M. Murray, Michigan Department of Conservation

INTRODUCTION

Estimation of the legal harvest of various species of game animals is an essential feature of any system wherein an attempt is made to "manage" the species as a recreational asset. While there is an increasing interest in esthetic values of wild animals, the major emphasis of agencies responsible for "game management" remains that of achieving a maximum sustained harvest of certain species by sportsmen. A number of methods have been, and are now, used to attempt measurement of game kill. Two general categories of such methods may be listed; those dependent on field contact of sportsmen during the hunting seasons, and those requiring contact after the season, either by mail or by personal interview. The former methods suffer greatly from a number of difficult sampling problems, not the least of which is the high cost per unit of game tallied. The latter methods necessarily depend on the sportsman's recollection of his hunting experiences, and the proverbial stories about fishermen and fishing illustrate the difficulty there.

Michigan game kill estimates are derived principally from sample surveys, by mail, of licensed hunters. Prior to the early 1950's, we depended on a so-called "report card" system, under which forms were issued with the hunting licenses, presumably to be returned after the end of the hunting season. A principal purpose of this paper is to describe some of the Michigan survey methods and results.

METHODS

Nearly all of our surveys have employed double-return postcards. While the size of such cards imposes a severe restriction on the amount of information that can be obtained, this drawback is considerably offset by the low costs, ease in handling, and high rate of responses. Samples of the questionnaire forms and accompanying texts are appended to this paper. As many as four reminder messages are used to insure high response rates, and each such reminder carries the original questionnaire. Different texts are used, of course, and the last two reminders are multilithed letters. We maintain a running quality check on incoming responses, and write back to the respondent (usually with a form letter) if any of several key items is omitted, or answered ambiguously.

Some six or seven separate surveys are conducted each year, covering each of several different kinds of license (i.e., big game, small game, archery, trapping), sub-populations of hunters (as those receiving a special permit to take antlerless deer), or asking for different kinds of information.

Systematic samples (with random starts) are taken from files of carbon copies of the hunting licenses. These files are maintained in a central office (Lansing) but the licenses are sold by some 4,000 issuing agents (largely merchants) throughout the state. A certain number, perhaps 5 per cent, of all licenses sold are not available for sampling, since we must begin our last mailing cycle before the last few returns from issuing agents reach Lansing. Sales of the different licenses range from a few thousand to nearly three-quarters of a million (small game), and total well over one million. The aggregate of all samples runs in the neighborhood of 25,000 individual licensees each year. In one year (1954) we used cluster sampling with individual issuing agents serving as primary sampling units. Some savings in time necessary to complete the surveys were effected, as we were able to obtain samples a good deal earlier in the year than under ordinary circumstances, but costs were considerably increased, and this and other complications made it seem advisable to return to the systematic sampling of central files.

Tabulation and estimation from the survey returns are generally relatively straightforward procedures and will not be described here, except to note that they center around the estimation of total game killed and hunting effort expended.

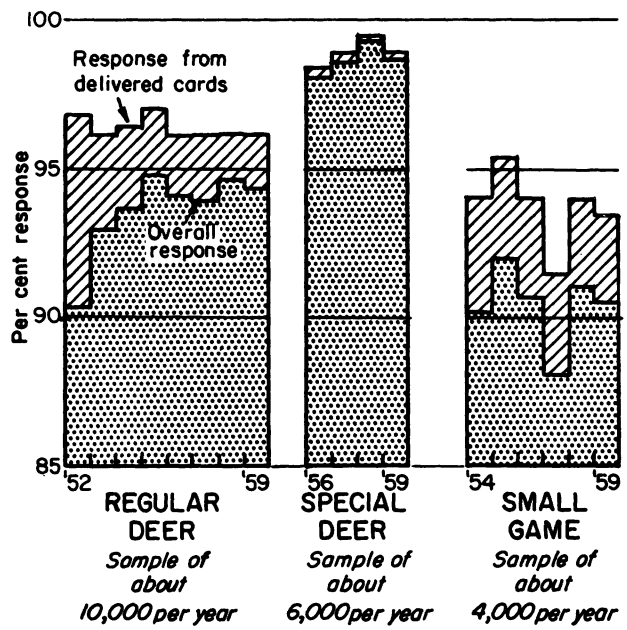


Figure 1. Survey response rates.

RESPONSE RATES

The overall response rates for our eight years of experience with these surveys (1952-1959) have, almost without exception, been over 90 per cent. Results for the three largest annual samples (Figure 1) illustrate the general situation. The two largest samples (regular and special deer) receive identical cards, but differ in that

the "special" deer survey is a sampling of persons who apply for and receive a special permit entitling them to take an antlerless deer (the basic Michigan deer hunting season is for antlered male deer only). These persons must submit an application, by mail, to Lansing, so that the return addresses are necessarily better than most such files. Also, the applications are submitted in October and are immediately processed (as part of a random drawing since quotas of permits are regularly surpassed by the applications) and returned. We thus have a complete file at the beginning of the deer hunting season (November) and are able to mail questionnaires at the close of the season, rather than one or more months later as is necessary in the other surveys. All such applicants must, however, have a regular deer hunting license, and thus constitute a subpopulation of the deer hunting licensees. In recent years, from one-quarter to one-half of all deer licensees have applied for permits.

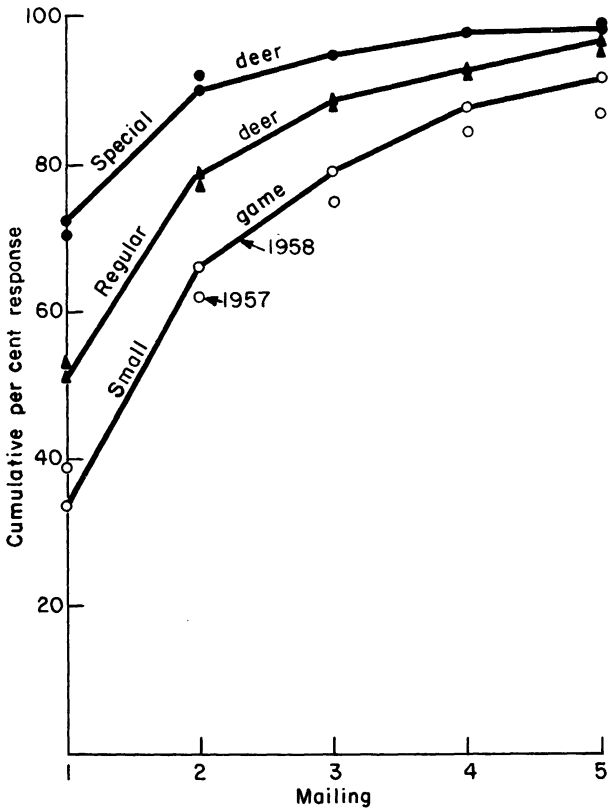


Figure 2. Cumulative response rates by mailing.

Cumulative percentage returns, when graphed by mailing (Figure 2), show necessarily much the same patterns as do the overall rates (Figure 1), but an interesting feature is the low initial return rates for the small game survey. The repeated reminders are effective, and successfully raise an initial moderate return to quite a respectable final tally. We aim at a two-weeks interval between mailings, but actually average somewhat more than that, usually 15 to 17 days. Examination of daily return rates by mailing (Figure 3), here computed on the basis of cards outstanding at the time each mailing becomes effective, shows that we probably might use a shorter interval (only the first 10 days' returns for each wave are shown in the figure). We operate with a rather small staff, however, and since we may have as many as a dozen distinct lots of cards in the mails at once, the two-weeks interval is virtually necessary to spread the work load. It seems, from Figure 3, that there is a decline in efficiency in the successive mailings, but that the general pattern is much the same. Occasionally, administrative deadlines have precluded the fifth mailing, and on one occasion, we did try a sixth mailing. This instance was the 1958 "special" deer survey, and the results were that, out of 6,707 cards initially mailed, we were able to get replies on all but 27; 8 of these were not delivered, and 19 persons failed to respond after six mailings.

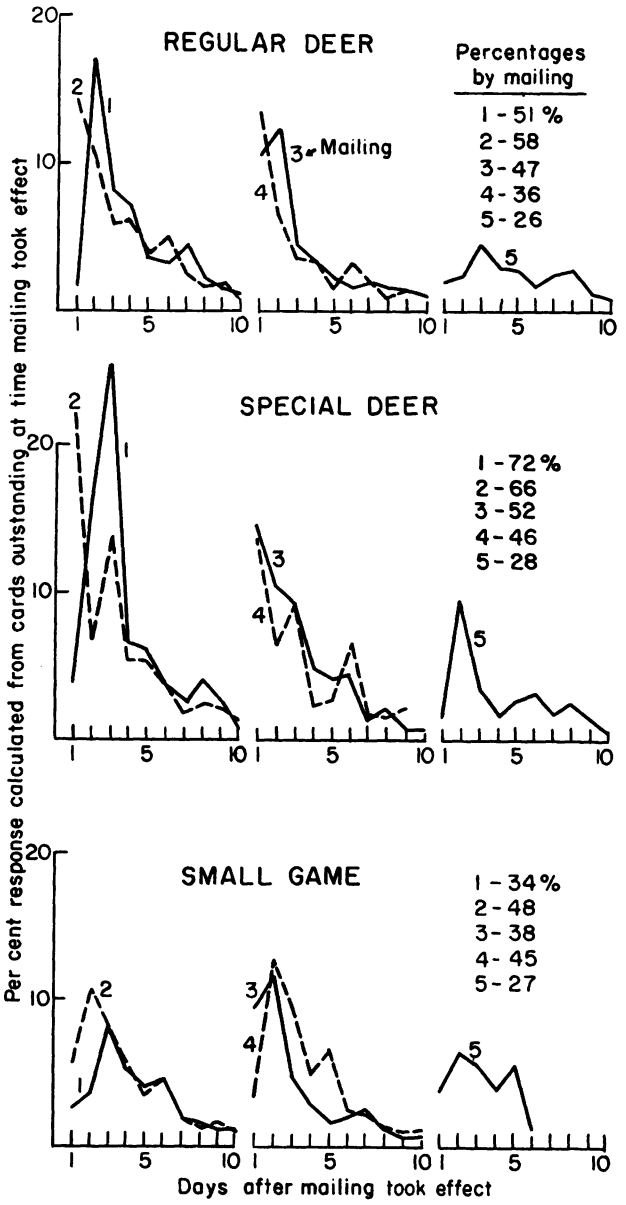


Figure 3. Response rates by mailing - 1958 surveys.

We have not been able to investigate the reasons for these high response rates in any detail. The questionnaire forms are relatively simple, requiring from a few minutes to perhaps a half an hour or more to fill out, depending on how carefully one searches one's memory for the various items requested. As mentioned above, the two deer-survey cards are identical, and a major difference is one of timing--one card comes out immediately after the hunting season, and the other from one to three months later. Michigan small game seasons range up to nearly five months duration, with many hunters doing most of their hunting in the first month or so, so that there may be a four months or more lag between the performance and the request to recall details. We suspect that this delay, and a rather more cluttered questionnaire, may be largely responsible for the lower small game survey response rates. There is also a suspicion that the average deer hunter takes his sport rather more seriously, and is more of a regular patron than is the small game hunter, who may buy a license on the spur of the moment.

Black, green, and red inks (in that order) are used on the first three cards sent out, and the last two are accompanied by a multilithed letter. Texts of the messages are kept simple, and stress the possibility that the individual may have forgotten, or mislaid his card, as well as various attempts to make the recipient feel that his reply is essential to future hunting prospects, and so on. Since the several texts, and questionnaire forms worked out so well in the beginning, we have been hesitant to make any major changes in the interim, and unhappily, have had no opportunity for any deliberate experimentation (using odd- and even-numbered cards of different format, etc.). Perhaps the only definite improvement over the first few years' operation, insofar as increasing response rates go, is an increased reliance on a file of telephone directories and city street maps for piecing out faulty addresses. Perhaps it should be parenthetically noted here, that we do have a rigid rule about substitutions--none are permitted! A very few (about 1 in 1,000) licenses are so incompletely filled out that they are discarded at the outset. Cards returned by the post office as non-deliverable are checked back to the original license-carbon and remailed, with any improvement of address that seems reasonable. At present, after one try, we relegate such cards to the "dead" file. In the past, repeated mailings to the same address would occasionally be successful, and broadcast mailings to variants on the address sometimes pay off (also, a letter apparently commands more respect in some post offices than do the cards), but our general feeling now is that these several devices do not increase the response rates enough to justify the drag they exert on our limited work-force and facilities.

We suspect that our high response rates may hinge partially on a high interest on the part of those receiving the cards. On the other hand, there is ample evidence that many license-buyers have practically never heard of the Department of Conservation, and two samplings have shown that about one-fifth of the small game hunters do not buy licenses in two successive years. Furthermore, there exists a variety of experience to show that a fair share of deer hunters, at least, are by no means wholly in accord with Conservation Department policies. As an example, we have, for some six years now, conducted a mail sampling (using postcards) of deer hunters in which we ask for various opinions. By far the most controversial issue has been that of whether or not the hunting of antlerless deer (does and fawns) should be permitted. In recent years there has been a steady sequence of legislative bills (and two Acts), public hearings, and an occasional attempt at "cease and desist" injunctions in the courts concerning this matter. Our survey results have indicated a shift from about an even split of opinion, to currently about 60 per cent of deer hunters favoring such seasons. Respone rates on surveys including questions on these seasons have, however, been very much the same as in other, less controversial inquiries.

Gray (1957, 1959), describing quite similar response rates in two surveys conducted in England, ascribes the high rates obtained there to the simplicity of questionnaire forms. We are inclined to agree that this may well be the case, but have only one instance of the use of a longer questionnaire for contrast with our postcard results. In this case, eleven questions (covering two 8½" x 11" pages), concerning both hunting experiences and opinions about types of deer hunting seasons, were asked of a sample of 1,139 deer hunters. Three mailings resulted in just over a 90 per cent return, which is about that experienced with deer hunters on the "regular" deer surveys. Sampling was, however, restricted to two particular segments of the state (of about county size) through use of a field sampling to obtain license numbers ("back-tags" carrying these numbers are required in Michigan). Probably these areas have more than the average numbers of downstate hunters.

A further point of some interest here is that a legislative act (1937) does require hunters to report the game that they bag. The act has never been enforced, and, during the years when reply cards were furnished with the licenses, returns fell off to less than 20 per cent for deer (included in Figure 6), and 5 per cent or so for small game hunters. The act was amended in 1957, and now requires hunters to report only when specifically asked to do so. We doubt that this situation now has any major influence on our response rates, and very likely only a small percentage of all hunters know that such a law ever existed.

EFFECTS OF NON-RESPONSE

There seems to be a distinct tendency for those deer hunters who are successful in bagging a deer (Michigan law permits a hunter to take only one deer in one year) to be more inclined to reply on the first mailing (Figure 4). The average bag per hunter for other species does not, however, show such a distinct difference between returns from first and subsequent mailings. In most cases, there is a definite tendency for the later responses to have somewhat lower average bags, much as shown for pheasants (Figure 4). Cottontail rabbits constitute something of a special case, as we ordinarily must begin the surveys before the end of the long

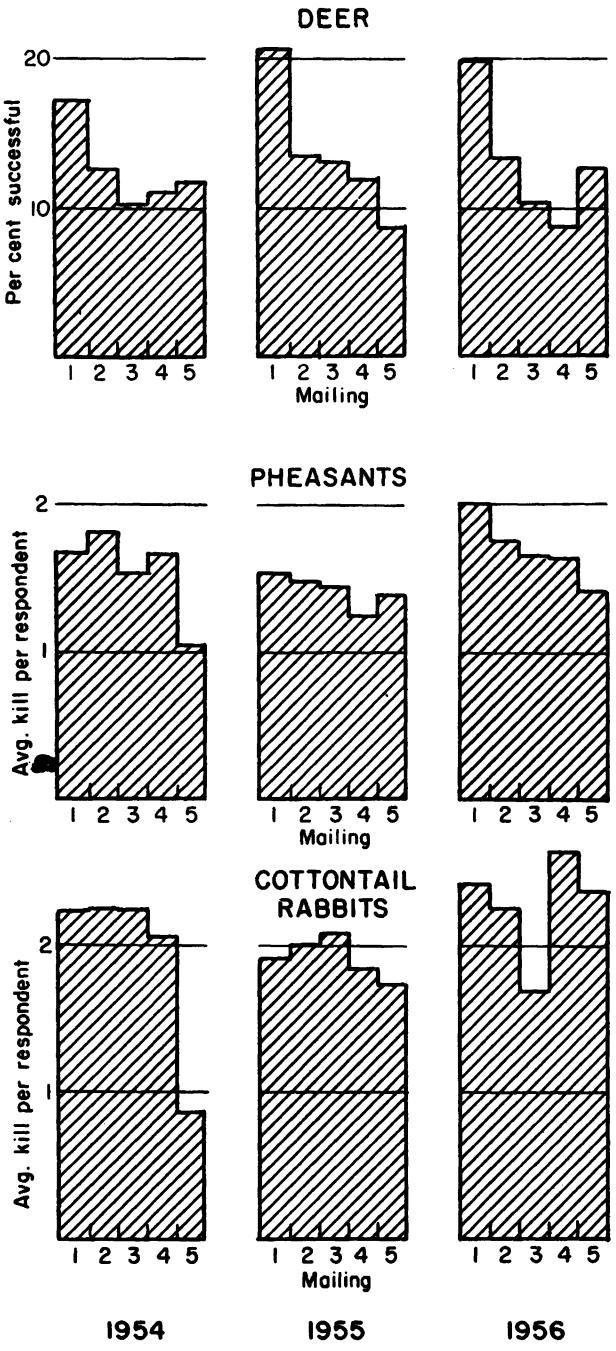


Figure 4. Reported hunting success by mailing from which response was received.

hunting season on rabbits.

In the reporting system used prior to 1952, each licensee was provided with a report form, but the only follow-ups were various press releases. Rather high initial return-rates (about 70 per cent for deer licensees, and 40 per cent for small-game licenses) soon dropped off (to about 20 per cent for deer and 10 per cent for small-game licensees). In the case of deer hunters an unusual circumstance permits approximate calculation of return-rates for successful (those killing a deer) and unsuccessful hunters.

A little more than half of those hunting for deer in Michigan's Upper Peninsula live in southern areas of the state, and virtually all of these hunters return home via the Straits of Mackinac (by car-ferry until 1957, when a bridge connecting the two peninsulas was completed). Since 1921 persons collecting crossing-tolls at the Straits have kept a tally of deer brought across the Straits. We have thus been able to compute an expected number of deer brought across the Straits and compare it with these actual tallies. This ratio of "computed to actual kills" (Figure5) increased steadily as the proportion of hunters returning their report cards dropped off.

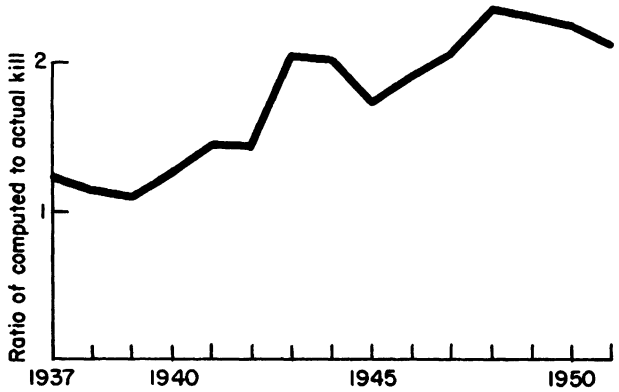


Figure 5. Ratio computed to actual kill as recorded at Straits of Mackinac.

If the bias exhibited at the Straits is considered to be that applying to state-wide estimates, the relative rates of returns shown in Figure 6 may then be computed.

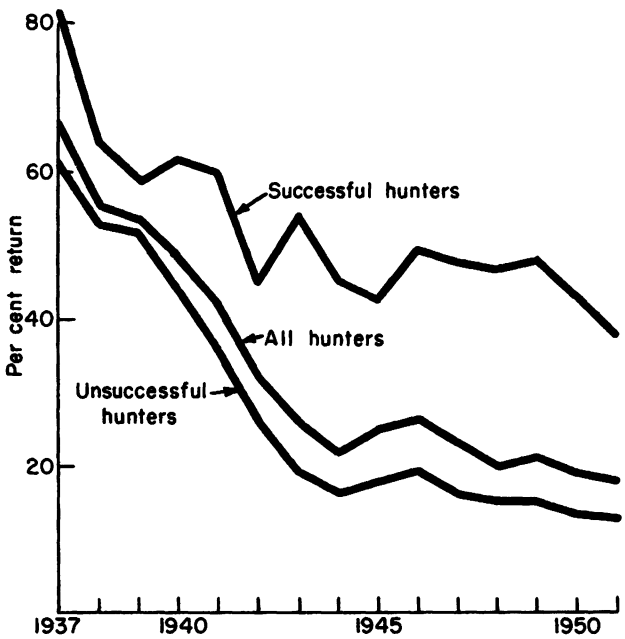


Figure 6. Return rates for report cards.

VALIDITY OF SURVEY RESULTS

The previously mentioned tallies of deer brought across the Straits of Mackinac have averaged some 15 to 20 per cent less than estimates of such crossings formed from the mail survey data. In 1957 and 1958 we made sample counts at toll booths, thus obtaining a third estimate of the number of deer brought across the Straits. Results for the two years are:

Year	Source	Estimate	Proportion of mail survey fig.
1957	Toll booth count	9,224	.771
	Sample-count est.	11,336	.948
	Mail survey est.	11,960	1.000
1958	Toll booth count	12,830	.857
	Sample-count est.	13,962	.932
	Mail survey est.	14,972	1.000

The sample counts were designed for ratio estimation, using tallies of vehicles crossing the Straits as an auxiliary variate (mechanical counts are made of vehicular traffic). Unfortunately, since precise estimates of uni-directional traffic flow were not available for the first sampling (1957), only the 1958 estimate is based on the ratio method. Variance estimates have not yet been made for the two samplings, but it seems certain that rather high precision was obtained. It appears, then, that the mail survey estimates are fairly well substantiated by the above data.

We have had several other, but less satisfactory, opportunities to check mail survey estimates of deer kill against supposedly complete tallies. These include records kept on car-ferries which service three islands (of 55 to 130 square-miles in area) and mandatory "registration" of antlerless deer shot during two hunting seasons on an area of about 1,000 square miles. In all such cases the mail survey estimates have been very nearly those of the other sources.

Hunter reports as to the kind of deer taken are an entirely different matter, however. Here we find that many hunters either cannot distinguish between adult females and juveniles (fawns) of both sexes, or do not wish to report shooting small or antlerless deer. We have consequently been forced to depend on data collected at highway checking stations to estimate sex and age composition of the deer harvest.

Precise checks on the validity of kill reports for other species are lacking. An occasional accidental duplication of returns, and a limited amount of deliberate checking for a few species, suggest that hunters do make various mistakes or report inaccurately. Atwood (1956) has reported on an extensive study of various presumed biases in reporting the kill of waterfowl. We have found a very close correlation (MacMullan, 1960) between the estimated hunting kill of pheasants and an entirely independent index to population density, suggesting that the kill estimates are at least consistent.

SURVEY COSTS

Not counting charges for office space or equipment depreciation, our costs are roughly 50 cents per completed response. Since our office facilities are multi-purpose, it is difficult to assign a specific item of cost. From 3 to 5 seasonal (6 to 9 months of the year) clerks and one full-time supervisor work on the surveys, and no special equipment beyond the usual office facilities is used, excepting IBM equipment, but this is included in the cost figures. We have attempted to include all other items of cost in the overall computation, including such things as editing and coding of responses, and preparation of routine final reports.

In our rather small operation, we find that overall costs will vary appreciably with the quality of clerical help available in a particular year. Also, some of the surveys require an appreciable amount of coding and cross-tabulation, and various collating procedures have been necessary to avoid duplication or sort out a special subpopulation in some cases. We have, however, little difficult editing, outside of the determination of locations where deer are reported killed, which does require someone rather intimately acquainted with the northern areas of the state, and this ordinarily precludes use of clerical personnel for the job. Two other "extra" items of cost perhaps should be mentioned. We find that attempts to whittle down the existing 3 per cent or so of "non-deliverable" cards by various sorts of detective work can be rather expensive, and from 5 to 10 per cent of our approximately 25,000 responses require further contacts (by letter) to attempt to get some essential item omitted from the first response.

In general, our costs range from about 30 cents per response for the simpler questionnaires requiring only straightforward coding and tabulation to as much as 60 to 70 cents for more complex questionnaires having fairly elaborate coding and tabulation (it also happens that one such survey has the lowest initial response, and thus higher costs). These are all postcard questionnaires, although the last two or three reminders may be sent out with form letters. All materials used in the survey are multilithed in a state-owned facility, and, while we have included an approximate cost figure for such services, the limited volumes used in our surveys would likely be a good deal more expensive on a commercial, job-lot basis.

LITERATURE CITED

Atwood, Earl L., 1956. Validity of mail survey data on bagged waterfowl. Jour. Wildl. Mgt., 20(1): 1-16.

Gray, Percy G., 1957. A sample survey with both a postal and an interview stage. Applied Statistics, 6: 139-153

_____, and Elizabeth A. Parr, 1959. The length of cigarett stubs. Applied Statistics, 8: 92-103.

MacMullan, Ralph A., 1960. A population survey of the ring-necked pheasant in Michigan. PhD Thesis, Mich. State Univ.. Mimeo. as Game Div. Report 2277, Mich. Dept. of Conservation, Lansing.

QUESTIONNAIRE FORM

NOV.	COUNTY HUNTED
Sun. 15	
Mon. 16	
Tue. 17	
Wed. 18	
Thu. 19	
Fri. 20	
Sat. 21	
Sun. 22	
Mon. 23	
Tue. 24	
Wed. 25	
Thu. 26	
Fri. 27	
Sat. 28	
Sun. 29	
Mon. 30	

DEER AND BEAR SURVEY-1959

1. Did you hunt this past deer season? *Yes*____ *No*____.
2. At left list **COUNTIES** you hunted in opposite dates hunted (deer and bear).
3. Did you get a deer? *Yes*____ *No*____.

IF YOU GOT A DEER:

4. On what day of the month was it killed?_____
5. Where was it killed? (Give as nearly as possible in miles and direction from nearest town. Example: Five miles east of Grayling.)_____
6. What kind was it? *Buck* ☐ *Buck Fawn* ☐ *Doe* ☐ *Doe Fawn* ☐

DID YOU GET A BEAR? *Yes*____ *No*____.

If you got a bear, where was it killed?_____

_____ Date of kill_____

DID YOU GET A PERMIT TO HUNT ANTLERLESS DEER? *Yes*_____

WHAT AREA WAS IT FOR?_____ *No*_____

TEXT OF FIRST MAILING

Mr. Deer Hunter:

You have been selected as a representative deer hunter to help us secure an accurate appraisal of the recent deer season. Your observations and experiences added to many others will give us very important information.

Please note that this survey also includes bear hunting. If you hunted for bear after killing your deer, include the counties and dates. (You do not need to indicate whether you were hunting bear or deer.)

Please fill out the card and drop it in the mail *TODAY*.

Thank you,

GAME DIVISION
Michigan Department of Conservation

TEXT OF SECOND MAILING

FIRST REMINDER

Some time ago you were asked to supply certain information to the Conservation Department's Game Division.

You probably had good intentions of mailing the return card but simply forgot it or perhaps mislaid the card. Don't forget that we work for you and that the information you give us is very necessary for the proper management of your deer herd.

We will probably continue to bother you if we don't hear from you soon. So please fill in the attached card and mail it right away.

Thanks,

GAME DIVISION
Michigan Department of Conservation

QUESTIONNAIRE FORM

SMALL GAME SURVEY 1959-1960

PLEASE CHECK YES OR NO AFTER EACH ANIMAL	DID YOU HUNT THIS ANIMAL? YES	NO	IF YOU HUNTED IN MORE THAN ONE COUNTY, USE A SEPARATE COLUMN FOR EACH COUNTY YOU HUNTED IN. COUNTY HUNTED	NO. DAYS HUNTED	NO. KILLED	COUNTY HUNTED	NO. DAYS HUNTED	NO. KILLED	COUNTY HUNTED	NO. DAYS HUNTED	NO. KILLED	COUNTY HUNTED	NO. DAYS HUNTED	NO. KILLED
PHEASANTS														
RUFFED GROUSE														
WOODCOCK														
DUCKS														
GEESE														
COOTS														
COTTONTAIL RABBITS														
SNOWSHOE HARES														
SQUIRRELS														
RACCOONS														

DID YOU BUY DEER LICENSES LAST FALL? REGULAR GUN: YES ☐ NO ☐ ; ARCHERY: YES ☐ NO ☐

TEXT OF FIRST MAILING

Dear Sir:

We are trying to get better information on small game kill this year. Will you please help us by filling out the attached post card and mailing it immediately?

It is important that you follow these directions exactly:

(1) Read over the list of game and indicate whether or not you hunted each kind by *CHECKING "YES" OR "NO" AFTER EACH ANIMAL.*

(2) After each kind hunted write in the *COUNTY* or *COUNTIES IN WHICH YOU HUNTED* and give the total number you killed in each county. Put in *ZERO* if you didn't kill any. (If you don't know the county, the name of the nearest town will do.) Also give the number of days you hunted in each county.

Thank you,

GAME DIVISION
Michigan Department of Conservation

TEXT OF SECOND MAILING

FIRST REMINDER

Dear Sir:

Some time ago we sent you a card asking about your small game hunting success last season. Your answers provide very important information, so please send in the attached card today.

It is important that you follow these directions exactly:

(1) Read over the list of game and indicate whether or not you hunted each kind by *CHECKING "YES" OR "NO" AFTER EACH ANIMAL.*

(2) After each kind hunted write in the *COUNTY* or *COUNTIES IN WHICH YOU HUNTED* and give the total number you killed in each county. Put in *ZERO* if you didn't kill any. (If you don't know the county, the name of the nearest town will do.) Also give the number of days you hunted in each county.

Thank you,

GAME DIVISION
Michigan Department of Conservation

CHECKING THE VALIDITY OF THE TOLL COLLECTORS' VEHICLE CLASSIFICATIONS

Kenneth C. Pearson, Massachusetts Port Authority

ACKNOWLEDGEMENTS

I am indebted to Raymond E. Willis, Jr. of the University of Minnesota for assistance in the formulation of the analysis procedure and to Francis X. Noonan of the Massachusetts Port Authority for assistance in checking the validity of the method.

INTRODUCTION

THE PROBLEM

At toll facilities that offer reduced rates to commuters a question which often arises is, "How do you know that some collectors are not registering non-commuter vehicles at the reduced commuter rate and then pocketing the difference"? The answer usually given is that auditing and visual controls are employed to guard against this eventuality.

A daily audit of each toll collector's deposit report is made to determine whether sufficient cash or its equivalent has been deposited to cover the transactions appearing on his register tape. This daily audit is supplemented by an occasional detailed analysis of the commutation tickets turned in by each collector. The object of the latter audit is to determine whether a collector turned in too many tickets from one book.

In both of these audits the analysis is confined to what a specific collector did without considering how the work of one collector compares with another collector when both worked under similar conditions. Questions such as, does one collector generally have more commuter vehicles than the other collectors, are not evaluated. At toll facilities that do not use commutation tickets primary control is attained by visually inspecting the toll transactions.

As the vehicles pass through the toll lanes the supervisor on duty verifies that the vehicles are being properly recorded. Theoretically, visual inspection provides for a first-hand check on the toll collectors but in actual practice certain impediments arise which nullify any systematized inspection procedure. The foremost of these is the monotony of the inspection process. Other factors are; weather conditions, heavy traffic flow, accidents and customers problems.

The purpose of this paper is to demonstrate a general method of analysis that will augment the auditing process by taking into consideration not only the toll transactions registered by the collector but also will consider the likelihood of a collector having misclassified a portion of the vehicles. The method also eliminates the necessity of depending on constant visual inspection as a contol technique.

Introduction to Statistical Control

After the auditing department has determined that all the collectors are in balance (i.e. their deposit reports balance with their register tapes) a statistical analysis is begun. The commuter and non-commuter passenger cars are entered on work sheets similar to table 1. A separate work sheet is used for each day of the week and each shift because of varying traffic characteristics.[1]

The commuter and non-commuter vehicles are not analyzed separately but are combined into a single statistic, the per cent of commuters of the total passenger cars i.e.

$$\frac{\Sigma \text{ commuters}}{\Sigma \text{ commuters} + \Sigma \text{ non-commuters}} 100.$$

This transformation is made because of daily fluctuations in each class and more importantly because of the existence of a relationship between the classes. If a toll collector deliberately registers a non-commuter as a commuter he reduces the number of non-commuters by one but he also increases the number of commuters by one. The net effect is an increase in his per cent of commuters.

[1]The discussion that follows will be restricted to passenger cars. The procedures are applicable to all types of vehicles however.

TABLE 1: COMMUTER AND NON-COMMUTER PASSENGER CARS
MONDAYS ONLY
3 P. M. - 11 P. M. - SOUTHBOUND TRAFFIC
MYSTIC RIVER BRIDGE
MARCH THRU MAY

	Lane 10			Lane 11			Lane 12			Lane 13		
Wk.	Com-muter	Non-Comm.	% Comm.	Com-muter	Non-Comm.	% Comm.	Com-muter	Non-Comm.	% Comm.	Com-muter	Non-Comm.	% Comm.
1	1061	382	73.5	1419	371	79.3	1610	401	80.1	1489	237	86.3
2	960	250	79.3	1350	316	81.0	1547	321	82.8	1203	170	87.6
3	997	363	73.3	1400	361	79.5	1648	353	82.4	1493	222	87.1
4	1056	376	73.7	1489	469	76.0	1635	342	82.8	1755	299	85.4
5	1108	407	73.1	1410	452	75.7	1691	374	81.9	1511	245	86.0
6	1178	406	74.4	1352	444	75.3	1739	360	82.8	1446	216	87.0
7	1105	409	73.0	1571	498	75.9	1669	362	82.2	1452	268	84.4
8	1179	951	55.4	1543	870	63.9	1585	612	72.1	1502	517	74.4
9	1063	565	65.3	1678	647	72.2	1561	449	77.7	1771	313	85.0
10	1292	579	69.1	1789	659	72.7	1710	491	77.7	1834	378	82.9
11	1304	607	68.2	1627	744	68.6	1763	575	75.4	1737	416	80.7
12	1164	585	66.6	1681	707	70.4	1620	575	77.0	1880	424	81.6
13	1248	616	67.0	1855	733	71.7	2020	623	76.4	1743	326	82.3

Reference to table one and the column headed "per cent of commuters" reveals two important characteristics:

1. The per cent of commuters tend to decrease with the passage of time. This is due mainly to seasonal variation.

2. When the per cent of commuters decrease or increase from week to week the same pattern generally occurs in all the lanes.[2]

With the approach of spring and improved driving conditions not only does the volume of traffic increase but the length of the average trip increases. The non-commuter vehicles consequently increase at a faster rate than the commuters and this in turn causes a decrease in the per cent of commuters. After Labor Day the reverse of this is true.

Whereas toll facilities are open 24 hours a day, seven days a week, the amount of data to be deseasonalized is extremely large. Removing the effects of seasonal variation would be very time consuming and costly. Therefore, a method which circumvents the necessity of removing seasonal is utilized.

Since the data is collected on a same day, same tour, same direction basis a good comparison of each toll collector's work is obtained if the per cent of commuters value of each lane is matched against the value that occurred in that lane last week. (To avoid the possibility of always matching the same collectors alternate assignments are used.)[3] The difference and sign of each matched pair is now determined and then these differences are ordered. For example, when week 2 is matched against

[2]In some cases where the change is very small this observation may not hold.

[3]See appendix for toll collector lane assignments.

week 1, the differences and assigned ranks are:

Lane	10	11	12	13
Difference	+05.8	+01.7	+02.7	+01.3
Rank	1	3	2	4

The largest positive difference is assigned a rank of one (1) and the smallest positive difference is assigned a rank of four (4). When ties occur, the two differences are given the average rank (i.e. if the two differences are tied for ranks 3 and 4 the average of these ranks is given to each difference)[4] Table 2 shows the differences and assigned ranks for a thirteen (13) week period.

The reader may question the use of ranks on the basis that a certain amount of information is lost. While this is true, the advantage of ranking the differences is that on an information per dollar of cost basis we come out ahead. Other advantages are:

1. The order of the differences is more important than the quantitative value of the differences. The essential point here is that in one case the size of a particular difference may be unimportant while in another case a difference of the same size may be extremely important.

2. Ranking allows persons who are unfamiliar with more technical statistical procedures the opportunity to evaluate this data.

3. Ranking lends itself to machine data processing.

Distribution of the ranks in the toll lanes

As noted earlier, the toll collector assignments are made on a prearranged basis so that the same collectors will not always be matched against each other. In an attempt to make the conditions under which the matching occurs as identical as possible and also to allow only chance variables to influence the flow of traffic, all collectors must use a standard method of collecting tolls. Each vehicle is classified before the fare is collected and each collector carries money in his hand for the purpose of making change. At this point it is of considerable interest to inquire whether the toll lanes themselves influence the rank assignments.

A review of the rank assignments in table 2 indicates that all lanes received the different ranks about an equal number of times. But since this table represents only thirteen (13) weeks, conclusions based on such a small sample could be erroneous. However, the same conclusion resulted when a random sample of two hundred matched days was drawn from a population of two years. The sum of the ranks for this sample was:

Toll lane	10	11	12	13
Sum of the ranks	483.0	523.5	490.5	503.0

Substituting these rank totals in the Friedman two-way analysis of variance by ranks formula

$$X_r^2 \quad \frac{12}{Nk\,(k+1)} \sum_{j\ 1}^{k} (Rj)^2 - 3N\,(k+1)$$

where N the number of rows
k the number of columns
Rj the sum of the ranks in each column we obtain:

$$X_r^2 \quad \frac{12}{200\,(4)\,(5)} \left[(483.0)^2 + (523.5)^2 + (490.5)^2 + (503.0)^2 \right] - \frac{3\,(200)\,(5)}{2.68}$$

Since the Friedmen statistic tends to be distributed as Chi-Square with k-1 degrees of freedom, reference to a table of Chi-Square indicates that $p > .30$ but $< .50$.[5] Since $p > .05$ (the level of significance) we can conclude that the rank assignments are independent of the toll lanes. Stated another way the mean of the rank total for each lane will be about equal. When the mean rank of each toll lane is computed in the 200 day matched sample we obtain;

Lane	10	11	12	13
Mean Rank	2.41	2.62	2.45	2.52

It will be noted that the sample means differ only slightly from the hypothetical mean of 2.5 (the value

[4] Friedman, M., "The use of ranks to avoid the assumption of normality implicit in the analysis of variance," Journal of American Statistical Assn, 32(1937)675-701

[5] Friedman, op. cit. 675-701

TABLE 2: DIFFERENCES AND RANKS OF THE PER CENT OF COMMUTERS

DATA FROM TABLE 1

	Lane 10			Lane 11			Lane 12			Lane 13		
Wk.	% Comm.	Difference	Rank	% Comm.	Difference	Rank	% Comm.	Difference	Rank	% Comm.	Difference	Rank
1	73.5			79.3			80.1			86.3		
2	79.3	05.8	1	81.0	01.7	3	82.8	02.7	2	87.6	01.3	4
3	73.3	-06.0	4	79.5	-01.5	3	82.4	-00.4	1	87.1	-00.5	2
4	73.7	00.4	1.5	76.0	-03.5	4	82.8	00.4	1.5	85.4	-01.7	3
5	73.1	-00.6	3	75.7	-00.3	2	81.9	-00.9	4	86.0	00.6	1
6	74.4	01.3	1	75.3	-00.4	4	82.8	00.9	3	87.0	01.0	2
7	73.0	-01.4	3	75.9	00.6	1	82.2	-00.6	2	84.4	-02.6	4
8	55.4	-17.6	4	63.9	-12.0	3	72.1	-10.1	2	74.4	-10.0	1
9	65.3	09.9	2	72.2	8.3	3	77.7	05.6	4	85.0	10.6	1
10	69.1	03.8	1	72.7	00.5	2	77.7	00.0	3	82.9	-02.1	4
11	68.2	-00.9	1	68.6	-04.1	4	75.4	-02.3	3	80.7	-02.2	2
12	66.6	-01.6	4	70.4	01.8	1	77.0	01.6	2	81.6	00.9	3
13	67.0	00.4	3	71.7	01.3	1	76.4	-00.6	4	82.3	00.7	2
			29			31			31			29

which would be obtained if the rank totals of all the lanes were equal). The hypothetical mean or average rank is also the value that is used as a standard to evaluate the ranks received by the individual collector. Over a period of time each collector can be expected to receive all the rank values about an equal number of times assuming only chance variations.

Evaluating the ranks of the individual collector

From a control point of view primary interest centers on those collectors who receive an abnormal number of below average ranks.[6] To illustrate this point consider the case of collector Y who worked in lane 12 during the fifth week. From table one it is noted that he registered 1,691 commuter vehicles and 374 non-commuter vehicles. Based on the differences that occurred in the other lanes, collector Y received a rank of four (4). However, suppose collector Y had deliberately misclassified 20 non-commuters as commuters. His register tape would then indicate 1,711 commuters and 354 non-commuters. The differences and ranks would then be;

Lane	10	11	12	13
Difference	-00.6	-00.3	-00.1	+00.6
Rank	4	3	2	1

The conversion of the 20 non-commuters into 20 commuters would have changed Y's rank from four (4) to two (2) and would place him below the standard on this day. The question of how many below average ranks does a collector have to receive before any action is taken will now be considered.

[6] If a collector receives the average rank several times in a sample, half of them are counted as below and half above the average rank.

Choosing a level of significance

Rather than arbitrarily selecting the level of significance at .01 or .05 consideration must be given to the consequence of committing either a type I or type II error.

In the case under discussion a type I error occurs when we reject the null hypothesis that a collector is honest when in fact he is honest. What is the cost of committing this type error? It would be the cost of conducting a visual investigation of the toll collector in question and this would be a relatively small cost. A type II error occurs when we accept the null hypothesis that the collector is honest when in fact he is dishonest. The cost of committing the type II error is the retention of the dishonest toll collector and this could be expensive over a period of time.

When the dollar costs of committing these errors are weighed, it is quite evident that we particularly want to avoid a type II error. By selecting .15 level of significance we attempt to minimize the cost of committing both types of error.

On this basis we may determine whether any collector is statistically out of control. This is accomplished by noting how many below average ranks a toll collector received in a given sample (usually the size of the sample is> 25) and then noting whether the value equals or exceeds the critical value given in the following table.[7]

Table 3. Critical number of below average ranks in a given sample.

Sample size	Critical value
26	17
27	17
28	18
29	18
30	19
31	19
32	20
33	20
34	21
35	22
36	22
37	23
38	23
39	24
40	24

(continued)

[7]These values were determined by using the normal approximation to the binomial distribution corrected for continuity, with P= Q = 1/2

A case history

In order to indicate the actual procedure which is followed in checking the validity of the toll collector's vehicle classifications, a case history is submitted.

Prior to starting his regular toll collecting duties, collector A was given the standard indoctrination. This consists of instructions on the traffic rules & regulations of the facility, plus a period of actual toll collecting under the supervision of a toll sergeant. At the completion of this training, collector A was assigned a regular tour of duty.

During his first three weeks of toll collecting he received the following ranks:

Day	Rank	Day	Rank
1	5	9	4
2	2	10	3
3	3	11	1
4	4.5	12	1
5	4	13	2
6	2.5	14	5
7	5	15	1
8	5		

A review of these ranks indicate that on eight days he was above the average rank and on seven days he was below. This is of course a perfectly normal situation.

The ranks received by collector A in the next eleven days were:

Day	Rank	Day	Rank
16	1	22	1
17	1	23	1
18	1	24	1
19	2	25	1
20	4	26	1
21	1		

When these ranks are added to the ones he previously received the results are;

above the average rank 9 days
below the average rank 17 days

Referring to table 3 we note that in a sample of 26 days, 17 below average ranks is critical. A non-statistical investigation of collector A's work was initiated at this point.

When a visual inspection of collector A's work was made it was noted that he was misclassifying non-commuter vehicles at the reduced commuter rate. On the basis of this evidence collector A became an "X" collector.

REFERENCES

Cochran, W.G., "The comparison of percentages in matched samples." Biometrika (37) 1950 256-266.

Dixon, W.J. and Wood, A.M., "The statistical sign test, "Journal of the American Statistical Association, (41) 1946 557-566.

Friedman, M., "The use of ranks to avoid the assumption of normality implicit in the analysis of variance, "Journal of the American Statistical Association, 32 (1937) 675-701.

Kendall, M.G., "The treatment of ties in ranking problems," Biometrika 33 (1945) 239-251.

Moore, G.H., and Wallis, W.A., "Time series significance tests based on signs of differnces, " Journal of the American Statistical Association, 38 (1943) 153-164.

Savage, I.R., "Bibliography of nonparametric statistics and related topics," Journal of the American Statistical Association, 48 (1953) 844-906.

Siegel, S. "Nonparametric Statistics for the Behavorial Sciences," New York: McGraw-Hill 1956.

APPENDIX

Toll collectors lane assignments, Schedules A and B.

To avoid the possibility of always comparing the same two collectors, an alternating schedule is used. Schedule A is in effect during the odd numbered weeks and B during the even numbered weeks.

SCHEDULE A

Time	Monday		Tuesday		Wednesday		Thursday		Friday		Time	Saturday		Time	Sunday	
	South	North	South	North	South	North	South	North	South	North		South	North		South	North
11-7	BB-10	DD-5	AA-10	BB-4	AA-11	EE-5	AA-10	BB-5	EE-10	DD-5	11-7	EE-11	CC-4	11-7	(6)-11	DD-5
	(1)-11	CC-4	EE-11	CC-5	BB-10	*-4	DD-11	EE-4	CC-11	AA-4		DD-10	AA-5		CC-10	BB-4
6-2	B-12	C-6	C-12	B-2	A-12	C-6	B-12	A-2	C-12	B-3						
	K-13		A-13		B-13		C-13		A-13							
											6:30-2:30	A-12				
7-3	J-10	G-3	H-9	M-3	M-11	G-3	F-10	E-3	I-10	*-4	7-3	(2)-11	H-6	7-3	I-13	E-5
	H-11	F-4	G-10	F-5	E-10	H-4	K-11	M-5	J-9	L-6		G-13			F-11	G-4
	I-9	E-5	E-11	I-4	F-9	L-5	L-9	J-4	M-11	K-5		K-9	I-5		K-10	J-6
8-4		D-2		D-6		D-2		D-6		D-2	8-4	J-10	L-4	8-4	L-12	H-3
											9-5		M-2	2-10	S-9	R-2
2-10	Q-13	O-6	O-12	Q-2	O-13	R-6	R-12	O-2	Q-13			(3)-3				
	R-12		R-13		Q-12		Q-13		O-12							
3-11	V-11	S-5	U-10	S-3	S-10	T-3	N-10	X-4	*-11	V-6	3-11	V-10		3-11	U-13	V-6
		U-4		V-5		N-5		Y-3		W-4		W-9	N-6		W-11	Y-4
	X-10	W-3	N-11	T-4	U-11	Y-4	W-11	S-5	Y-10	N-5		T-11	U-5		X-10	T-5
4-12	P-9	T-2	Z-9	P-6	P-9	Z-2	Z-9	P-6	X-9	Z-2	4-12	Y-13	X-4	4-12	(8)-12	(7)-3
													Z-2			
5-1										P-3	5-1	(5)-12	(4)-3			
Time	Monday		Tuesday		Wednesday		Thursday		Friday		Time	Saturday		Time	Sunday	

*SERGEANT WORKS THIS TOUR

SCHEDULE B

Time	Monday		Tuesday		Wednesday		Thursday		Friday		Time	Saturday		Time	Sunday	
	South	North	South	North	South	North	South	North	South	North		South	North		South	North
11-7	DD-11	BB-5	BB-11	EE-4	*-11	BB-4	EE-11	BB-5	DD-11	AA-5	11-7	CC-11	DD-5		CC-10	(6)-5
	(1)-10	CC-4	AA-10	CC-5	AA-10	EE-5	DD-10	AA-4	EE-10	CC-4		EE-10	AA-4		BB-11	DD-4
6-2	K-12	B-2	A-12	C-6	B-12	A-2	C-12	B-6	A-12	C-3						
	C-13		B-13		C-13		A-13		B-13							
											6:30-2:30	A-13				
7-3	E-10	H-3	I-9	E-3	E-9	F-3	F-10	E-5	M-9	I-5	7-3	(2)-12	G-4	7-3	E-10	I-6
	F-11	I-4	F-10	G-5	G-11	L-4	M-9	K-4	L-11	*-4		I-10	K-5		G-13	F-4
	G-9	J-5	M-11	H-4	H-10	M-5	J-11	L-3	K-10	J-2		H-11			J-11	K-5
8-4		D-6		D-2		D-6		D-2		D-6	8-4	L-9	J-6	8-4	H-12	L-3
											9-5		M-2	2-10	R-9	S-2
2-10	Q-12	R-2	Q-13	O-6	O-12	Q-2	Q-12	O-6	Q-13				(3)-3			
	O-13		R-12		R-13		R-13		O-12							
3-11	X-11	V-4	S-11	U-5	N-10	T-4	Y-11	W-4	W-10	V-5	3-11	N-10		3-11	U-11	T-4
		U-3		T-3		S-5		X-5	*-11	Y-2		T-9	V-4		V-13	W-5
	S-10	W-5	V-10	N-4	U-11	Y-3	S-10	N-3		N-4		U-13	W-6		Y-10	X-6
4-12	T-9	P-6	P-9	Z-2	Z-9	P-6	P-9	Z-2	Z-9	X-6	4-12	X-11	Y-5	4-12	(7)-12	(8)-3
													Z-2			
5-1										P-3	5-1	(4)-12	(5)-3			
Time	Monday		Tuesday		Wednesday		Thursday		Friday		Time	Saturday		Time	Sunday	

*SERGEANT WORKS THIS TOUR

ON THEORETICAL QUESTIONS UNDERLYING THE TECHNIQUE OF REPLICATED OR INTERPENETRATING SAMPLES

J. C. Koop, North Carolina State College
Raleigh, North Carolina

1. Introduction

The technique of interpenetrating samples was introduced by Mahalanobis in 1937 in jute acreage surveys in Bengal mainly for statistical control. Mention of its use in checking the work of field investigators is made in his 1939 paper where the theory behind his method of crop acreage estimation is discussed. Subsequently it was developed (Mahalanobis, 1944, 1946) and it is now used in practically all surveys conducted by the Indian Statistical Institute. The United Nations Subcommission on Statistical Sampling (1949) has recommended its use and has also suggested the alternative term "replicated sampling."

The technique consists in drawing two or more sets of samples from the same population using the same sampling procedure for each set of samples. There are two variants of the technique; one when the samples are statistically independent and the other when they are not so.

The sets of samples will be independent if each set is replaced after the entire drawing is completed. Consider for example, a multi-stage sampling procedure when k sets each containing m first-stage units, are to be drawn from a given stratum. Starting from the first set, all of the m first-stage units (as a set) must be replaced before the second set is drawn and so on; this will ensure statistical independence between the k sets in the context of any given probability system. Within a set the sampling at the first stage and each subsequent stage may be any type provided only that the units are selected according to some probability system constructed for that stage. This appears to be the most general definition of independent interpenetrating or replicated samples defined by Lahiri (1954) in a terminology slightly different from mine. It will be seen that when sampling is carried out with replacement of each first-stage unit, independent interpenetrating samples of a special type are obtained. (In this situation m = 1.)

The sets of samples will not be statistically independent if the constituent units of one sample are in some way associated or linked with those of another. Regardless of whether or not the sets of interpenetrating samples are independent, appreciable disagreement between the estimates from each set will indicate, in some sense, the discrepancies in observation.

There has been much controversial discussion (particularly in India) on the use of the technique in the assessment or control of observational, enumeration and response errors* (Panse and Sukhatme 1948, Yates 1949; Ghosh 1949, 1957; Mokashi, 1950; Sukhatme, 1952; Sukhatme and Seth, 1952; Cochrane, 1953) and it will not be discussed in this paper. Yates also discusses the question of estimating sampling errors using estimates from the independent samples. In both review articles Ghosh comments on the estimation of sampling errors and the "margin of uncertainty" (Indian National Sample Survey, 1956). His second article contains some references not cited here. Lahiri (1954) in his monograph discusses the practical and theoretical aspects of the problem of estimating sampling errors of various population estimates (totals, means, ratios) and the problem of confidence limits for them, but he gives more attention to the problems of design very relevant to Indian conditions.

Apart from the problem of non-sampling errors which will always be enmeshed in the estimate of the sampling variance, Yates, Ghosh, Cochrane and Lahiri are all of the view that the estimation of sampling errors, using estimates obtained from the independent samples, will not be precise because of the small number of degrees of freedom. This view will be examined later.

Finally, in regard to previous work related to this subject, Deming (1956) and Flores (1957) have reported on and discussed theoretically a type of sample design with very desirable properties, which in the writers view, does not strictly fall into the category of sample designs resulting from the principle of the technique, i.e., drawing two or more samples using the same sampling procedure. Jones (1956) has elaborated on Deming's ideas for this type of design.

In this paper (i) the technique of independent interpenetrating samples in its most general application will be examined from a theoretical standpoint; (ii) the problem of efficiency of linear estimates for multi-stage designs will be discussed; (iii) confidence intervals which do not depend on the distribution of the estimates will be given; and (iv) applications in which there is much theoretical uncertainty about the use of the method will also be discussed.

---

*This question has been discussed by Hansen et al. (1951) from a different point of view.

2. Theoretical basis of the technique

There is a certain probability system which shows the probabilities according to which each sampling unit is to be selected at each step (stage or phase) in sampling down to the ultimate unit which may or may not be the unit of analysis. This probability system defines the way by which each set of samples is to be drawn. It is abstract and we may say that the frame, which describes the units at each step in the sampling procedure, and the method of sampling which ensures that the units are selected with probabilities prescribed by the system, are its real counter parts.

We draw k sets of independent samples each by the same sampling procedure and with probabilities prescribed by the system and we recall that each set is replaced prior to drawing the next set. Because of the identity of the sampling procedure each set of samples will be identical in structure. This means that at a given step in sampling, a given set will have the same number of sampling units as the other sets. Hence each set of samples will have the same number of ultimate sampling units.

An example at this stage may be useful. Suppose we have a frame which shows two strata. In Stratum I there are N elements. In Stratum II there are M clusters each consisting of N_i elements (i=1, 2,...,M). We desire two sets of independent interpenetrating samples, s and s'. Each set is to consist of n elements from Stratum I and n elements each (at the second stage) from two clusters selected at the first stage from Stratum II. It is given that n is smaller than the size of the smallest cluster of Stratum II. The probability system is defined as follows:

Stratum I: Probabilities are equal and the units are not replaced after each draw.

Stratum II: Probabilities are equal at each stage and the units are not replaced after each draw and in each stage.

Consider the drawing of sample s according to the probability system just defined. In Stratum I n elements are selected; in Stratum II clusters i and j are selected at the first stage, and n elements from each of these clusters are selected at the second stage. Then the total probability of drawing the sample s (of an aggregate size 3n) is

$$P_s = \frac{1}{\binom{N}{n}} \cdot \frac{1}{\binom{M}{2}\binom{N_i}{n}\binom{N_j}{n}} . \qquad (1)$$

After this drawing, in Stratum I the n elements are replaced and in Stratum II the selected elements of clusters i and j are replaced. Another drawing is now made by the same sampling procedure to obtain sample s'. Suppose this time clusters i' and j' are obtained. Then

$$P_{s'} = \frac{1}{\binom{N}{n}} \cdot \frac{1}{\binom{M}{2}\binom{N_{i'}}{n}\binom{N_{j'}}{n}} . \qquad (2)$$

Altogether there are

$$S = \binom{N}{n}\sum_{i>j}\binom{N_i}{n}\binom{N_j}{n} \qquad (3)$$

possible samples, and therefore S possible estimates for a given characteristic. Samples s and s' are clearly statistically independent. Further

$$\sum_{s=1}^{S} P_s = 1 . \qquad (4.1)$$

This discussion illustrates the concept of the probability system (which is a construct), and the procedure of selecting the two sets of independent samples. Because of the replacement of the first set, some elements may appear again in the second set. If they appear again, they are not rejected. In the foregoing illustration the selection probabilities are equal but generally they need not be so. The probability system can be constructed to make each estimate more efficient in some sense. So also the frame can be constructed (in respect of the number of strata, the composition of the hierarchy of units incident to the stages (or steps) in sampling) to increase efficiency in some defined sense. The classical theories given in text books have discussed many of these problems.

With these preliminaries we resume where we left off at the second paragraph of this section. Let P_s be the probability of drawing a set of samples from a population or universe $\mathcal{U}$ described by a frame $\mathcal{F}$ according to a most general probability system $\mathcal{P}$. As explained above both $\mathcal{F}$ and $\mathcal{P}$ can be constructed (or chosen) so that their state (or condition) predisposes the resulting estimates to have, in some sense, some desirable property or properties.

There are k mutually independent sets of samples each of which appears with probability P_s not necessarily equal to that of the others. If we need to distinguish between the samples and their respective probabilities we may attach subscripts to s. Thus s_t will denote the $t^{\underline{th}}$ independent set of samples and P_{s_t} its corresponding probability. We will have

$$\sum_{s} P_s = 1 . \qquad (4.2)$$

Let T_s be some estimate of a population value T, computed on the basis of values revealed by sample s. T may be a mean value or a population total, in which case it is a linear function, or a ratio or some function of the values of one or more characteristics of the elements of $\mathcal{U}$. We have

$$E(T_s) = \sum_s P_s T_s, \tag{5}$$

and

$$V(T_s) = \sum_s P_s (T_s - E(T_s))^2. \tag{6}$$

Since the sampling procedure for each set is the same, each of the k estimates T_s (computed by some formula), will have an identical variance given by (6) above.

Let us consider the estimate T_s in itself which is made up of n ultimate sampling units. Suppose we are estimating some population value such as

$$T = \sum_{i\varepsilon\mathcal{U}} a_i f(x_i, y_i, \ldots) \equiv \sum_{i\varepsilon\mathcal{U}} a_i f_i \tag{7}$$

where $f(x_i, y_i, \ldots)$ is a known single-valued function of the measurable characteristics observed on the i^{th} element (ultimate sampling unit) at the last step in sampling and the a_i's are known constants specific to each i. We shall give concrete meaning to (7) shortly. For the purpose of a general discussion it is not necessary to specify the location of the ultimate unit i (e.g., its location in the hierarchy of units) in the frame $\mathcal{F}$. In this situation we can obtain an unbiased estimate of T, linear in f_i, as follows. Let

$$T_s = \sum_{i\varepsilon s} R_i f_i \tag{8}$$

where each R_i is a weight to be attached to each f_i. Now f_i is defined above and can be computed on the basis of values revealed by i. For T_s to be unbiased we must determine the R_i's such that

$$E(T_s) = E\Big(\sum_{i\varepsilon s} R_i f_i\Big) \equiv \sum_{i\varepsilon} a_i f_i \tag{9}$$

i.e.,

$$\sum_{i\varepsilon\mathcal{U}} a_i f_i \equiv \sum_s P_s \sum_{i\varepsilon s} R_i f_i$$

$$\equiv \sum_{i\varepsilon\mathcal{U}} a_i f_i \sum_{s\supset i} P_s \frac{R_i}{a_i} .$$

That is we must have

$$\sum_{s\supset i} P_s \frac{R_i}{a_i} = 1 \text{ for every } i\varepsilon\mathcal{U}. \tag{10}$$

There are as many such equations (10) as there are elements in the universe $\mathcal{U}$. But the total number of R_i's is nS where, it will be recalled, n is the number of ultimate sampling units in each set of samples and S is the number of possible samples. If $\mathcal{N}$ is the total number of elements then

$$nS > \mathcal{N},$$

and therefore the equations (10) can be solved. If S' is the number of samples which include i, then one set of solutions is obtained by choosing

$$R_i = \frac{a_i}{S' P_s} \quad \text{for all } s\supset i. \tag{11}$$

Hence we find

$$T_s = \frac{1}{S' P_s} \sum_{i\varepsilon s} a_i f_i \ . \tag{12}$$

For example if we are estimating the total characteristic

$$T = \sum_{i=1}^{N} x_i$$

of a finite population of size N with a sample of size n drawn with equal probabilities and without replacement, then $a_i = 1$, $f_i = x_i$ for all i and $S' = \binom{N-1}{n-1}$ and $P_s = 1/\binom{N}{n}$, so that we find $T_s = \frac{N}{n} \sum_{i=1}^{n} x_i$ a familiar formula.

By and large the estimating functions we may use may not turn out to be unbiased. Further it can be proved (with essentially the same approach adopted by Koop (1957)) that generally no minimum unbiased estimator to estimate $T = \sum_{i\varepsilon\mathcal{U}} a_i f_i$ exists.* Earlier this was proved by Godambe (1955) with a less general formulation of the problem.

On the basis of the k independent estimates, T_s, we wish to estimate T. The best linear combination of the k estimates will be given by

*Minimum unbiased linear estimators exist only in the special case when the units are selected with equal probabilities.

$$T' = \frac{1}{k}(T_{s_1} + \ldots + T_{s_k}). \qquad (13)$$

Clearly if each T_s is unbiased then T' will be unbiased. Regarding the variance of T' we have

$$V(T') = \frac{1}{k^2}\left[\sum_{t=1}^{k} V(T_{s_t}) + 2\sum_{t>t'} \mathrm{Cov}(T_{s_t}, T_{s_{t'}})\right]$$

We will find

$$V(T') = \frac{1}{k}\ V(T_s) \qquad (14)$$

since the estimates are statistically independent so that $\mathrm{Cov}(T_{s_t}, T_{s_{t'}}) = 0$, and each has uniform variance given by (6) above.

It is not difficult to show that $\hat{V}(T_s)$, the unbiased estimate of $V(T_s)$, is given by

$$\hat{V}(T_s) = \frac{1}{k-1}\sum_{t=1}^{k} (T_{s_t} - T')^2, \qquad (15)$$

so that

$$\hat{V}(T') = \frac{1}{k(k-1)}\sum_{t=1}^{k} (T_{s_t} - T')^2, \qquad (16)$$

which is of the same type as the formula for the sampling variance of the mean of an infinite population which was also noted by Lahiri (1954) but not proved.

Coming back to the estimate T' given by (13), had we adopted an alternative linear estimate

$$\sum_{t=1}^{k} \alpha_t\, T_{s_t}$$

where $\sum_{t=1}^{k} \alpha_t = 1$, we will find that

$$V\left(\sum_{t=1}^{k} \alpha_t\, T_{s_t}\right) = \sum_{t=1}^{k} \alpha_t^2\, V(T_{s_t}) = V.$$

By the use of the Cauchy inequality it can be shown that V attains a minimum when $\alpha_t = \frac{1}{k}$ for all t and this leads to (13) and (14).

Thus we find that the method of independent interpenetrating samples leads to an extremely simple formula (16) for the estimation of sampling variances whatever the design of the sample. As Lahiri has remarked, the calculation of estimates of variance by classical methods will involve, for a normal sample survey, very heavy computational work because of squaring numbers many hundreds of times when the classical formulas underlying the sample design are used. When two independent interpenetrating samples are used, we will find

$$\hat{V}(T') = \frac{1}{4}(T_{s_1} - T_{s_2})^2 \qquad (17)$$

which is very simple for computational purposes.

Also the standard error of T' is given by

$$(1/2)\left|T_{s_1} - T_{s_2}\right| . \qquad (18)$$

This formula is very suggestive. If we are carrying out, say, a demographic survey by the use of two independent interpenetrating samples and we desire to know the standard error of each estimate of the total characteristics, all we have to do is to tabulate the estimates of totals from each interpenetrating sample separately and compare the corresponding cells of the tables and apply (18). In the light of this discussion many of the estimates of the Indian National Sample Survey display a high degree of accuracy despite the possibility of the presence of non-sampling errors, which obviously will be enmeshed in the estimate of the variance. Elsewhere the possibility of evaluating these errors by experimental design has been mentioned (United Nations, 1949).

Finally in discussing the theoretical basis of the technique the barest assumptions about the nature of the frame $\mathcal{F}$ and the probability system $\mathcal{P}$ have been made (i.e., only their existence). If it is admitted that the universe $\mathcal{U}$ is divided up into h strata, then

$$P_s = \prod_{i=1}^{h} P_s^{(i)} \qquad (19)$$

where $P_s^{(i)}$ is the probability of obtaining one set of k independent samples for stratum i, (i=1, 2...h) obtained by the same sampling procedure as the remaining k-1, and we will have

$$\sum_{s} P_s^{(i)} = 1 \qquad (20)$$

for all possible s in stratum i. Also we will have

$$T_s = \sum_{i=1}^{h} T_s^{(i)} , \qquad (21)$$

where the $T_s^{(i)}$ are the separate independent stratum estimates so that (6) reduces to

$$V(T_s) = \sum_{i=1}^{h} V(T_s^{(i)}) \qquad (22)$$

in view of (19). Of course (22) can also be proved starting with (21). The unbiased estimate of $V(T_s)$ will then be

$$\hat{V}(T_s) = \sum_{i=1}^{h} \hat{V}(T_s^{(i)}) \qquad (23)$$

and each $\hat{V}(T_s^{(i)})$ can be estimated by formula of the same algebraic form as (16).

3. Relative efficiencies of linear estimates for multistage designs

We shall first consider a somewhat general sample design where sampling at the first stage is with unequal probabilities and with replacement and leave unspecified the probability system for the subsequent steps (stages and/or phases) in sampling.

Let the population or universe $\mathcal{U}$ consist of M first stage units which can be identified by $\mathcal{F}$. To estimate a certain total characteristic T we select k independent sets* of interpenetrating samples each containing m first-stage units each of which is replaced after drawing. In each set at each subsequent step in sampling a predetermined number of units is drawn. Let $P_i > 0$ be the probability of drawing the i^{th} first-stage unit where $\sum_{i=1}^{M} P_i = 1$. Here the probability system $\mathcal{P}$ consists of the set of probabilities $\{P_i\}$ and the remaining unspecified sets corresponding to each step in sampling.

Let X_i be the true total in unit i and X_i' the unbiased estimate of X_i calculated on the basis of estimates from the subsequent steps in sampling.

*One may wish to know why k sets of samples are taken. One practical reason is that if one or more sets are ruined(say be non-response or other unforeseen difficulties) the others will still permit valid estimation. Therefore the survey is not entirely a loss. We are playing safe.

Let s be one of the set of k interpenetrating samples. Then an unbiased estimate of T will be given by

$$T_s' = \frac{1}{m} \sum_{i \varepsilon s} X_i' \qquad (24)$$

To determine the variance of T_s' we shall apply Madow's theorem (1949). We find

$$V(T_s') = \frac{1}{m^2} \left[V\left(E\left(\sum_{i=1}^{m} X_i' \,\middle|\, i \varepsilon s\right)\right) + E\left(V\left(\sum_{i=1}^{m} X_i' \,\middle|\, i \varepsilon s\right)\right) \right] \qquad (25)$$

We find $E\left(\sum_{i=1}^{m} X_i' \,\middle|\, i \varepsilon s\right) = \sum_{i=1}^{m} X_i$ and

$$V\left(\sum_{i=1}^{m} X_i' \,\middle|\, i \varepsilon s\right) = \sum_{i=1}^{m} V(X_i') = m\, V(X_i') .$$

If we define the probability system and the procedure of sampling after the first stage, $V(X_i')$ can be determined by repeated application of the theorem. We now have

$$V(T_s') = \frac{1}{m^2} \left[m \sum_{i=1}^{M} P_i (X_i - T)^2 + mE(V(X')) \right] = \frac{1}{m} \left[\sum_{i=1}^{M} P_i (X_i - T)^2 + E(V(X_i')) \right] \qquad (26)$$

We have k sets of independent samples. The best estimate of T will be

$$T' = \frac{1}{k} (T_{s_1}' + \ldots + T_{s_k}) \qquad (27)$$

and

$$V(T') = \frac{1}{k} V(T_s') \qquad (28)$$

Had we taken a direct sample of mk first stage units and then sampled the subsequent units by the same procedures (which will mean that the number of samples taken at each step will

be the same) then the unbiased estimate of T, T'', will be given by

$$T'' = \frac{1}{mk} \sum_{i=1}^{mk} X_i' \tag{29}$$

and

$$V(T'') = \frac{1}{mk}\left[\sum_{i=1}^{M} P_i (X_i - T)^2 + E(V(X_i')) \right] \tag{30}$$

Clearly $V(T'') = V(T')$ so that when sampling is carried out with replacement at the first stage the estimate obtained by taking the arithmetic average of the separate independent estimates is as efficient as the estimate from a single sample of equivalent size.

Next consider the case when we sample the first-stage units without replacement and with equal probabilities from the same universe as above. The units at the subsequent stages are selected by some scheme which we shall leave unspecified. Suppose we take k sets of independent interpenetrating samples with m units at the first stage; then using sample, s_t, the estimate of the population total T will be given by

$$T_{s_t} = \frac{M}{m} \sum_{i \in s} \hat{X}_i \tag{31}$$

where $\hat{X}_i$ is an unbiased estimate of X_i. We will find

$$V(T_{s_t}) = M^2 \left[\frac{\sigma^2}{m} \left(1 - \frac{m}{M}\right) + \frac{1}{m} E(V(\hat{X}_i)) \right] \tag{32}$$

where σ^2 is the variance of the first-stage units defined with divisor M-1. If we take an arithmetic average of the k estimates

$$T' = \frac{1}{k} \sum_{t=1}^{k} T_{s_t}, \tag{33}$$

We will find that

$$V(T') = \frac{1}{k} V(T_{s_t}). \tag{34}$$

Had we taken a single sample of first-stage units each without replacement and employed the same sampling scheme at the subsequent steps as in the case of the separate independent samples then we will have

$$T'' = \frac{M}{mk} \sum_{i=1}^{mk} \hat{X}_i \tag{35}$$

and we will find

$$V(T'') = M^2 \left[\frac{\sigma^2}{mk} \left(1 - \frac{mk}{M}\right) + \frac{1}{mk} E(V(\hat{X}_i)) \right]. \tag{36}$$

Comparing the variances we have

$$V(T') - V(T'') = M^2 \frac{\sigma^2}{mk} \left(\frac{mk}{M} - \frac{m}{M}\right)$$

$$= M \frac{\sigma^2}{k} (k-1). \tag{37}$$

Thus the efficiency of the set of independent interpenetrating samples is less than a single sample of equivalent size. But clearly the difference in variances is the least when k=2. Therefore in this situation it is best to take only two independent interpenetrating samples because it will have the best possible efficiency but still below that of the single equivalent sample. It will be noted that this conclusion is also true for single-stage sampling. The estimate of the variance of V(T') will be given by

$$(\tfrac{1}{4})(T_{s_1} - T_{s_2})^2 \tag{38}$$

i.e., as in (17). The kind of argument usually advanced that (38) has "one degree of freedom" and therefore is a "poor estimate" does not seem quite relevant in the context of finite population theory.

If there is stratification in the universe, in both cases considered above it will be seen that the same conclusions hold.

The question of cost so far has not been mentioned in the discussion. Since the k sets of interpenetrating samples and the single sample of equivalent size have the same number of units at each step in sampling, the costs for the former cannot be greater than that for the latter in whatever way the allocation of units takes place. Indeed the k sets of interpenetrating samples are likely to have common units (since the sets are replaced) and therefore certain cost components may even be less.

4. Confidence limits for finite population estimates

Much has been said on this problem by Lahiri (1954) in his monograph and therefore the remarks which follow will be by way of a commentary and little can be added to it. Also a general theoretical discussion of this kind without consideration of this problem

would not be complete.

When the estimates from interpenetrating (or replicated) samples are linearly combined as in (13) the possibility that the resulting estimates for some or all of the several characteristics will have a normal form cannot be ignored particularly if the number of strata is large. Therefore for characteristics which are abundant, confidence limits of population values may be set on the basis of Student's distribution using the k estimates. However, we have just shown that for a given overall sample size for a multistage design without replacement at the first stage and with equal probabilities, it is best to use only two interpenetrating samples. This conclusion may or may not be true for other types of sample designs when units are drawn without replacement at the first stage and with unequal probabilities. Therefore for such designs it may be prudent not to have too many sets of interpenetrating samples. Some balance has to be struck between the opposing attractions of shorter confidence intervals (because of smaller t-values for larger number of degrees of freedom) on the one hand and on the other hand of the possibility of increased efficiency in the estimate of the standard error because of fewer replications.

The above difficulties can be remedied by certain arguments advanced by Savur (1937). He advocated the setting of confidence limits for the population median rather than the population mean. Most characteristics of finite populations are not distributed symmetrically. Now whena population of estimates (for a characteristic in question) each made up of a large number of observations is formed, most of these estimates will cluster round the true value and the median. Here it is necessary to point out that the true value and the expected value of the estimate are not necessarily identical. Suppose the difference of the cumulative probabilities up to the true value T and up to the median is δ, a very small quantity. Let the k independent estimates (which may be of any kind including ratios) be ordered as follows:

$$T_{s_1}, \cdots \quad \cdots, T_{s_k}$$

where s_1 is the least and s_k the greatest. Suppose the true value T is below the median. Then the cumulative probability up to T is $\frac{1}{2} - \delta$. Therefore the probability that all k estimates lie above T is $(\frac{1}{2} + \delta)^k$ and all lie below it is $(\frac{1}{2} - \delta)^k$. Hence the probability that some will lie below T and some above is

$$\alpha = 1 - \left\{(\tfrac{1}{2} - \delta)^k + (\tfrac{1}{2} + \delta)^k\right\} \simeq 1 - (\tfrac{1}{2})^{k-1} . \qquad (39)$$

if δ is very small as compared to its constituent binomial terms.

$$\therefore \quad P(T_{s_1} < T < T_{s_k}) \simeq 1 - (\tfrac{1}{2})^{k-1} = \alpha' . \qquad (40)$$

It will be noted that when k = 5, $\alpha' = .9375$. How small δ should be, can easily be worked out from the relevant terms in the expansion within braces in (39). However, if we cannot ignore δ, we can always make statements of inference conditional on it assuming a certain value as follows:

$$P(T_{s_1} < T < T_{s_k} \mid \delta) = 1 - \left\{(\tfrac{1}{2} - \delta)^k + (\tfrac{1}{2} + \delta)^k\right\} . \qquad (41)$$

Similar arguments have been advanced by Lahiri (1954) on the assumption that the true value and the median are for all practical purposes identical. Thompson (1936) also investigated the problem but from slightly a different point of view.

5. Interpenetrating samples which are not independent

At this stage comment on the use of such samples is appropriate. The samples may have units or elements which are linked by some rule or they may have overlapping units. Thus because the samples are not mutually independent difficulties are created in the investigation of the properties of their estimates. Even combined linear estimates may not be unbiased.

Further the derivation of estimates of variance leading to simple formulas as in (14), (15), and (16) does become possible. Also it is hard to find a way of determining confidence intervals on the same lines as in section 4 simply because the estimates are not independent.

6. Some applications

(i) Price Index Numbers. There are a large number of problems in this area which are familiar to economic statisticians. The question of sampling errors in the "weights" (in the sense of price index theory) used in the computation of index numbers does not seem to have been given attention. Kelly (1921) and Knibbs (1924) considered this problem. Recently Banerjee (1960) and Koop (1952) also considered the problem.

Because a price index number is a complicated ratio, the determination of its variance (if it has elements in its formula which can be treated as random variables) is very complicated. For example, consider an index

number of the type

$$I = \frac{\sum p_o q_o \left(\frac{p_1}{p_o}\right)}{\sum p_o q_o} . \qquad (42)$$

If the base year prices and quantities are determined by sampling methods then they will be subject to at least sampling errors the formula for which will depend on the underlying method. Now

$$V(I) = \sum \left(\frac{p_i}{p_o}\right)^2 V(w_o) \qquad (43)$$

where

$$w_o = \frac{p_o q_o}{\sum p_o q_o} \qquad (44)$$

and assuming that $\frac{p_1}{p_o}$ is determined by some reliable pricing system.

Now the expression for the variance of w_o, a ratio, will be very complicated. Indeed even if it is agreed that the methods used to find an expression for it are valid it is almost intractable.

Had we used two independent interpenetrating samples, then there would have been two independent sets of weights for the set of goods and services consumed by the population under study. With the same price information generated by some system (government) we can separately compute two index numbers which are statistically independent. Suppose they are I_1 and I_2. We will find the average, which has been shown to be the best, to be

$$I' = (\tfrac{1}{2})(I_1 + I_2) \qquad (45)$$

and

$$\hat{V}(I') = (\tfrac{1}{4})(I_1 - I_2)^2, \qquad (46)$$

so that the coefficient of variation of I' will be

$$\frac{|I_1 - I_2|}{I_1 + I_2} \qquad (47)$$

a very simple formula. In the light of the discussion in section 2 the index can be of any kind and not necessarily of a naive type used above for illustration. The sampling error will reflect the variations displayed by the population when it was surveyed. If these variations can be assumed to be conservative for at least some time then the use of $\hat{V}(I')$ will be of some help in practical situations, e.g., in wage adjustments.

Further if more than two independent interpenetrating samples, e.g., five, are used, then a confidence interval, with a high confidence coefficient,

$\alpha' = 1 - (\tfrac{1}{2})^4 = .9375$, for the true index number will be obtained.

(ii) Number of different words in a book. There has been a discussion on a class of problems of this type by Mosteller (1949) who commented on the solutions proposed by various statisticians.

Suppose we sample, by some probability system, n lines out of a total number N in a book* and count the number of different words and find it to be r, then a consistent estimate of the number of different words W will be given by

$$W' = N \frac{r}{n} . \qquad (48)$$

Now

$$V(W') = \frac{N^2}{n^2} V(r) \qquad (49)$$

and we will find that the expression for V(r) is very difficult to determine. Had we taken two independent interpenetrating samples and found the number of different words to be r_1 and r_2, then we will have

$$W' = (\tfrac{1}{2})(W_1' + W_2') \qquad (50)$$

and

$$V(W') = (\tfrac{1}{2}) \frac{N^2}{n^2} V(r). \qquad (51)$$

Now although we do not know V(r) we can estimate it as

$$\hat{V}(r) = (\tfrac{1}{2})(r_1 - r_2)^2 \qquad (52)$$

which is an unbiased estimate of V(r) and therefore we find

$$\hat{V}(W') = \tfrac{1}{4} \frac{N^2}{n^2} (r_1 - r_2)^2 \qquad (53)$$

so that we have a standard error to assess the precision of our estimate. This illustration shows another attractive property of the technique of independent interpenetrating samples. Formula (48) is essentially of the same type as Haenszel's formula. His ultimate units of analysis were elements whereas here the elements are clustered (as lines).

*We assume that the book is not a mathematical text where usually one is hard put to it to define lines.

References

Banerjee, K. S., (1960), "Calculation of sampling errors for index numbers," Sankhya, 22, 119-130.

Cochrane, W. G., (1953), Sampling techniques, New York: John Wiley and Sons, Inc., 212-214.

Deming, W. E., (1956), "On simplifications of sampling design through replication with equal probabilities and without stages," J. Amer. Statis. Ass., 51, 24-53.

Flores, A. M., (1957), "The theory of duplicated samples and its use in Mexico," Int. Statist. Inst. Bull., 36, 120-126.

Ghosh, B., (1949), "Interpenetrating (networks of) samples," Calcutta Statist. Ass. Bull., 2, 108-119.

_____, (1957), "Enumerational errors in surveys," Calcutta Statist. Ass. Bull., 7, 50-59.

Godame, V. P., (1955), "A unified theory of sampling from finite populations," J. R. Statist. Soc., B, 17, 269-278.

Hansen, M. H., Hurwitz, W. N., Marks, E. S., and Mauldin, W. P., "Response errors in surveys," J. Amer. Statist. Ass., 46, 147-190.

Jones, H. L., (1956), "Investigating the properties of a sample mean by employing random subsample means," J. Amer. Statist. Ass., 51, 54-83.

Kelly, T. L., (1921), "Certain properties of index numbers," Quarterly Publication of Amer. Statist. Ass., 17, 826-841.

Knibbs, G. H., (1924), "The nature of an unequivocal price-index and quantity-index," J. Amer. Statist. Ass., 19, 42-60.

Koop, J. C., (1952), "On the use of statistical methods in assessing the precision of index numbers," (Paper read before the first Burma Research Society Conference, August, 1952.)

_____, (1957), "Contributions to the general theory of sampling finite populations without replacement and with unequal probabilities," Ph.D. thesis, N. Carolina State College Library, and Dissertation Abstracts, 18, (7), 144.

Lahiri, D. B., (1954), "National Sample Survey Number 5. Technical paper on some aspects of the sample design," Sankhya, 14, 264-313.

Madow, W. G., (1949), "On the theory of systematic sampling II," Ann. Math Statist., 20, 335.

Mahalanobis, P. C., (1939), "A sample survey of the acreage under jute in Benegal," Sankhya, 4, 511-531.

_____, (1944), "On large-scale sample surveys," Phil. Trans. Roy. Soc., B, 231, 329-451.

_____, (1946), "Recent experiments in statistical sampling in the Indian Statistical Institute," J. R. Statist. Soc., 109, 325-378.

Mokashi, V. K., (1950), "A note on interpenetrating samples," J. Indian Soc. Ag. Statist., 2, 189-195.

Mosteller, F., (1949), "Number of different kinds in a population," American Statistician, 3, (3), 12-13.

National Sample Survey No. 7, (1956), Sankhya, 16, 230-434.

Panse, V. G., and Sukhatme, P. V., (1948), "Crop surveys in India-1," J. Ind. Soc. Agric. Statist., 1, 34-58.

Sukhatme, P. V. and Seth, G. R., (1952), "Non-sampling errors in surveys," J. Ind. Soc. Agric. Statist., 4, 5-41.

Sukhatme, P. V., (1952), "Measurement of observational errors in surveys," Int.Statist. Inst. Review, 20, 121-134.

Savur, S. R., (1937), "The use of the median in tests of significance," Proc. Ind. Academy Sc., A, 5, 564-576.

Thompson, W. R., (1936),"Confidence ranges for the median and other expectation distributions for populations of unknown form," Ann. Math. Statist, 7, 122-128.

United Nations, (1949), "Preparation of Sample Survey Reports," C (1), New York: Statistical Office of the United Nations.

Yates, F., (1949), Sampling methods for censuses and surveys. New York: Hafner Publishing Co., 44, 107, 143, 241-243.

References not cited in text

Dalenius, T., (1950), "The problem of optimum stratification," Skand. Aktuar. Tidskr., 33, 203-213.

Das, A. C., (1951), "On two-phase sampling and sampling with varying probabilities," Bull. Inter. Statist. Inst., 33, 105-112.

Durbin, J., (1953), "Some results in sampling theory when units are selected with unequal probabilities," J. R. Statist. Soc., B, 15, 262-269.

Horvitz, D. G. and Thompson, D. J., (1951), "A generalization of sampling without replacement from a finite universe," Ann. Math. Statist., 22, (2), 315.

Kitagawa, T. (1956), "Some contributions to the design of sample surveys." Sankhya, 17, (1), 1-36.

Koop, J. C. (1956), Sample survey of labour force in Rangoon: A study in methods. Rangoon: Union Government Printing and Stationery.

Midzuno, H., (1950). "An outline of the theory of sampling systems. Ann Inst. Statist. Math. (Japan), 1 (2), 149-156.

Roy, A. D. (1952), " An Exercise in errors," J. R. Statist. Soc., A, 115, 507-520.

ANALYSIS OF SOCIOMETRIC STRUCTURE: A METHOD OF SUCCESSIVE GROUPING[1]

Jack Sawyer and Terrance A. Nosanchuk

University of Chicago

The method described here was developed for defining the structure of a social group, although it can be applied to grouping problems in many other areas. Essentially, it considers a given group of individuals and assigns each to one of a mutually exclusive and exhaustive set of sub-groups.

This problem of defining sub-groups has been one of general interest since the first work of Moreno (1934) on the two-dimensional graphical "sociogram" representation. The conception of the problem was broadened by Forsyth and Katz (1946) who introduced matrices in which the ij entry is the relational measure between individuals i and j--a one for "chosen", a zero for "indifferent," and a minus one for "rejected." They proposed rearranging the rows and columns of this matrix such that the plus ones were near the diagonal and the minus ones away from the diagonal; lines could then be drawn at appropriate places to separate sub-groups. Katz (1947) suggested, as a more definite criterion, minimizing $\Sigma\Sigma e_{ij}(i-j)^2$, where e_{ij} is the element in the ith row and the jth column. Beum and Brundage (1950) provided an iterative, and sometimes lengthy, procedure for achieving the minimization. Bock and Husain (1950) also employed a matrix in grouping individuals into sub-groups, but did so on the basis of Holzinger's B coefficient.

Luce and Perry (1949) utilized the matrix formulation in a different way--to obtain, by matrix multiplication, a special kind of sub-group composed of individuals all of whom choose each other. They extended this concept to include individuals connected to each other not through direct choice, but through a third person, as in the connection of persons A and C represented by A choosing B and B choosing C. Luce subsequently (1950) further generalized the concept of the clique and provided procedures for clique identification. Harary and Ross (1957) provided a general solution for obtaining all the cliques in a given matrix.

Bock and Husain (1952) factored a matrix of choices, a technique which groups persons having similar choice patterns. A "direct" factor analysis of the score matrix (rather than of the correlations) was used by MacRae (1959) to obtain two sets of factors, one grouping persons by whom they chose, the other grouping on the basis of whom they were chosen by. Another approach has been that of mathematical graph theory, the study of directed line segments, explored by Harary and Norman (1953). A recent summary of much of the work briefly mentioned here is available in Glanzer and Glaser (1959).

[1]This research was supported in part by the National Science Foundation.

Applications in Other Areas

Although the grouping problem is particularly significant in the area of sociometrics, it also exists in many other areas. Rao (1952) employed a grouping method suggested by K. D. Tocher (1948) to determine how 12 Indian castes and tribes group together. In some cases groups are known and a multiple discriminant function may be used to classify members, as Rao illustrates (1952) in identifying the Highdown skull as Iron Age rather than Bronze Age.

A number of methods have been devised and utilized in connection with the grouping of psychological test scores, where a priori groups are not known. Holzinger and Harmon (1941) employ their method of B-coefficients to group 24 psychological tests on the basis of their inter-correlations. Cattell (1944) has summarized a number of methods of grouping psychological tests, all of which require, however, setting an arbitrary value above which two variables are considered to be related. More recently, McQuitty (1957, 1960) has proposed a number of "linkage analysis" methods.

The Grouping Methods

Any method of grouping takes as given some measure of relation among each of the pairs of individuals in the set being grouped. The present method works with any relational measure, such as strength of preference or frequency of interaction, although reference will be made to the "distance" between two points, perhaps a more general view. Such a distance measure might be determined, for example, by the method of multidimensional scaling, or by distance in a space spanned by orthogonal personality components.

Taking the matrix of relational measures as given, this method provides a completely determined, easily applied procedure for defining group structure, for any number of groups, 1, 2, ..., n-1,n, where n is the total number of points. The method proceeds by first regarding the n points as n one-point groups, and

forming the best set of n-1 groups by combining those two groups whose combination minimizes some criterion. These two groups now form a new, single group, and the procedure is repeated on the resultant set of n-1 groups. Thus the best set of n-2 groups is found, and so on; at each stage, two of the previously existing groups are combined, until finally, the last two groups are combined to form a single group of n points.

The value of the criterion measure (on the basis of which the groupings are made) after any grouping may be compared with the corresponding measure following each of the preceding groupings. A sharp jump might suggest not making that grouping, and reverting to the previous stage. There has been, however, no investigation of the stochastic properties of this measure to determine, for example, how large an increase must be to attain statistical significance.

The above method of regarding two groups, once joined, as permanently bonded, permits practical solution of the grouping problem; without this restriction there exist an overwhelming number of combinations--more than half a million different sets of groupings for only ten individuals, for example. Regarding two groups, once joined, as a single group, means that at a stage at which there are r groups, only $r(r-1)/2$ combinations need be considered.

In addition to a method for grouping, one needs a criterion on the basis of which to combine groups. The obvious general criterion is within-group homogeneity: groups should in some sense be internally homogeneous relative to other groups. Within this general conception, however, two distinctions can be made in the criteria employed. There is, first, the distinction whether in computing the "distance" between two sets of individuals (or among the individuals in a single group) one takes the sum of the individual inter-point distances, or their average. The second distinction concerns for which groups the average or total distance measure is computed.

The second distinction creates three cases; the first case minimizes the total (or average) distance between points which are in the same group. This measure starts at zero, of course, for the original n one-point groups, and grows to the sum (or average) of the distances between all the points when all the groups are finally combined into one. At each stage, one combines those two groups whose combination adds least to the intra-group distance.

In the second case, one considers each pair of groups by itself and combines those two which are "closest" in terms of the total (or average) distance between the points in the first group and the points in the second group. This case may, of course, give a different grouping from the preceding. If, for example, the two closest groups were very large, their combination would inflate the average much more than two smaller groups, even though the latter were more distant from each other. If the total distance, rather than the average distance, is employed, however, combining the two closest groups adds least to the total within-group distance over all groups, and this case becomes identical with the first.

The third case is somewhat interim, in that it considers the distance between two groups in relation to their distance from all other groups. The measure is thus a ratio, in which the numerator is the measure of the previous case, the total (or average) distance between two groups. The denominator is the total (or average) distance from the points in these two groups to all other groups. Two groups are combined, then, not simply on the basis of their absolute closeness, but in terms of their closeness relative to how far they are from all other groups. Thus two groups which were a moderate distance from each other, but exceedingly far from all the rest, might be combined prior to other pairs, which were nearer to each other but also nearer to the rest of the groups. The B-coefficient (Holzinger and Harman, 1941) is similar to this measure, although they differ in the numerator.

There thus exist five methods, as follows:

1. Total distance within groups
2. Average distance within groups
3. Average distance between two groups
4. Ratio of average distance between two groups to their average distance to points in all other groups
5. Ratio of total distance between two groups to their total distance to points in all other groups

For illustrative purposes, each of these five methods was applied to the same set of sociometric data. The data consisted of inter-person distances based upon the friendliness of each pair of 15 undergraduate fraternity members, obtained by the multidimensional scaling model (Morton, 1959). These data are presented in Table 1, columns A-O.

Computational Procedure

The balance of Table 1 illustrates the computational procedure for the first method, which minimizes the total distance between two points in the same group. The procedure is as follows.

1. Search the original nxn matrix for the smallest non-diagonal entry. Let it be d_{ij}. Record this amount, which is the increment added at this stage to the

Table 1: Total Distance Between Two Groups

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	FM	GK	CN	EL	IO	AJ	BD	CNH	AJGK	BDIO	ELFM	AJGK •CNH	BDIO• ELFM
A	317	159	357	215	209	131	204	305	(126)[6]	124	290	220	136	302	429	255	295	505	607								
B		270	(217)[7]	262	417	376	296	218	302	357	320	415	242	302	832	733	512	582	520	619							
C			327	243	280	217	167	273	164	152	296	209	(81)[3]	289	489	369											
D				323	451	401	351	347	350	388	388	437	255	355	888	789	582	711	702	707							
E					285	224	234	248	161	246	(97)[4]	271	234	333	556	470	477										
F						111	252	422	241	117	363	(70)[1]	298	396													
G							224	366	111	(71)[2]	278	113	228	338	224												
H								195	220	184	342	238	149	171	490	408	(316)[8]	576	366	424	647						
I									297	361	345	409	304	(122)[5]	831	727	577	593									
J										117	207	215	202	329	456	228	266	368	626								
K											303	146	209	359	263												
L												349	289	386	712	581	585										
M													228	408													
N														305	526	437											
O															804	697	603	719									
FM																487	1015	(1268)[11]	1635	885	1720	1505	1372	3355			
GK																	806	1051	1424	(483)[9]	1522	1214					
CN																		1062	1171	661	1094						
EL																			1312	873	1293	1638	1924	2605			
IO																				1233	(1222)[10]	1537	2657				
AJ																					1326	1085					
BD																						1741	2848				
CN•H																							(2299)[12]	3278	3143		
AJ•GK																								5505	3296		
BD•IO																									(5960)[13]	8783	
EL•FM																										6439	
AJGK•CNH																											15222

sum of the intra-group distances.

2. Create a new group, called IJ, composed of the former groups (points) I and J.

3. Determine the total distance, $d_{ij.k}$, of the new group IJ from any other group K, by adding together d_{ik} and d_{jk}. Form of these $d_{ij.k}$'s a new column IJ. Strike out rows I and J and columns I and J.

4. Repeat steps 1-3 on the resulting (n-1)x(n-1) matrix.

Thus at each stage, the corresponding entries of two groups are combined into a new column, and the rows and columns of the former groups are eliminated. In Table 1, entries corresponding to combined groups are circled and the order of combination is indicated by a superscript. The succession of particular groupings obtained by this method is further illustrated by the tree diagram of Figure 1. The total computation consists, for a group of size n, of (n-1)(n-2)/2 additions, each composed of a single pair of numbers. While the first method is by far the simplest computationally, it illustrates the general procedure of combination for all methods.

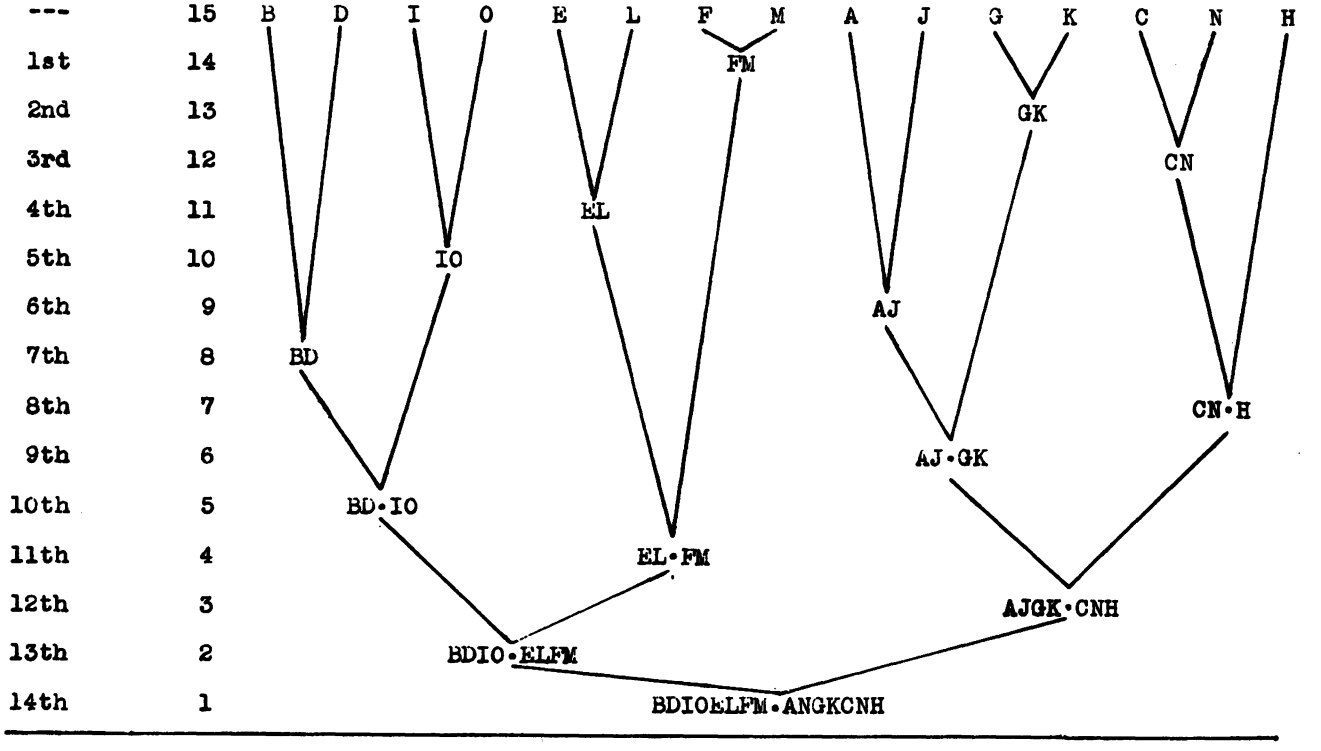
Comparison of the Five Methods

The results of applying each of the five methods are shown in Table 2. The rows indicate which two groups were combined at each stage, and the value of the criterion measure following that combination. As mentioned previously, sharp rises in the criterion may be employed as a stopping point. Thus, using the first method, one might stop after combining AJ (with nine groups) after AJ.GK (with six groups), or after AJGK.CNH (with three groups).

Comparing the methods, it is seen that FM and GK are first and second, respectively, in all five methods, and that the third and fourth groupings are always CN and EL. The agreement continues at a considerable level through the fifth and sixth groupings: IO is fifth in all but one method, where it is sixth; the combination AJ appears by the sixth grouping for three methods, and in the seventh for the other two methods. Beyond that point, there is less agreement, although in all, the 65 groupings result in only 27 different combinations (even though the more than half million

Figure 1

Successive Groupings; Total Distance Criterion



possible combinations would indicate little a priori chance of duplication). Some effects of the different methods can be observed. Use of the total distance influences toward groups of equal size, since in that way, the number of distances (and hence, likely, the total of distances) within groups is a minimum. Use of the average distance tends toward monolithic structure, one group being very large compared to the other. Other effects may be noted in further empirical evaluation. It is planned, for example, to obtain measures of interaction at successive time intervals in a newly formed group, to determine what model corresponds most closely to the process of the emerging group structure.

Aside from empirical test, certain rational advantages of the foregoing type of method may be cited, the foremost of which is that it is completely determined. Given a set of data, these methods specify precisely the groupings, for any number of groups, without use of an arbitrary level to decide when a relation is to be considered, or when an individual is to be added to a group. Thus it is amenable to computer programming, and this step is anticipated.[2] These methods have a further advantage of being able to utilize data having any values, rather than being restricted to a dichotomy or trichotomy. Thus any added sensitivity in the measurement of distance between two individuals can be employed to advantage. Computationally the first method is extremely simple, although the others are somewhat more complex.

Some open questions remain. In particular there is no provision for error in the distance measure, and the effect of this on the grouping. Further, one would like a more satisfactory way of deciding just what number of groups was most reasonable. Finally, and related to the previous questions, it is often of interest to compare the structure of two groups, or of the same group at different times, and provision is needed for this.

[2] A similar procedure, developed by David L. Wallace and Benjamin Wright, has been programmed for UNIVAC at the University of Chicago.

Table 2

Successive Groupings and Criterion Measure for Five Methods

Order of Grouping	Total Distance Within Groups		Average Distance Within Groups		Average Distance Between Two Groups		Ratio of Average Distances: Two to all Groups		Ratio of Total Distances: Two to all Groups	
1st	FM	70	FM	70	FM	70	FM	.2418	FM	.0093
2nd	GK	141	GK	71	GK	71	GK	.2990	GK	.0115
3rd	CN	222	CN	74	CN	81	EL	.3380	EL	.0130
4th	EL	319	EL	80	EL	97	CN	.3432	CN	.0132
5th	IO	441	IO	88	GK.J	114	IO	.3796	IO	.0146
6th	AJ	567	AJ	95	IO	122	FM.GK	.4212	AJ	.0214
7th	BD	784	AJ.GK	105	A.GKJ	127	AJ	.5564	BD	.0246
8th	CN.H	1100	CN.H	114	CN.H	158	AJ.FMGK	.5824	H.IO	.0337
9th	AJ.GK	1583	BD	122	AGKJ.FM	172	H.IO	.6058	FM.GK	.0383
10th	BD.IO	2805	AJGK.FM	142	AGKJFM.CNH	211	BD	.6396	AJ.CN	.0618
11th	EL.FM	4073	CNH.IO	166	BD	217	AJFMGK.CN	.6812	BD.EL	.0946
12th	AJGK.CNH	6372	BD.EL	180	AGKJFMCNH.EL	268	AJFMGKCN.EL	.8346	AJCN.FMGK	.1872
13th	BDIO.ELFM	12332	AJGKFM.CNHIO	224	BD.IO	306	AJFMGKCNEL. HIO	.9438	BDEL.HIO	.2208
14th	All	27454	All	262	All	335	All	∞	All	∞

Summary

This method takes as given a measure of relation (based, for example, upon frequency of interaction, strength of preference, or distance in a space spanned by orthogonal personality components) between all pairs of n individuals in a group. On the basis of these measures, from 1 to n groups may be formed, using mutually exclusive sub-sets of individuals. The method, at all stages, is completely determined. One proceeds by first regarding the n individual as n one-person groups, then joins those two groups whose combination minimizes some criterion. These two groups now form a new, single group, and the procedure is repeated on the resultant set of n-1 groups. As this successive grouping process is continued, the increment added to a measure of group homogeneity may be used as a criterion for stopping. Various models for behavior in the formation of groups are represented by different criteria for grouping, which vary in two respects: (1) whether the total or average relation between sets of individuals is employed as criterion, and (2) whether this criterion refers to only the two groups being joined, or to all groups. A comparison of the resultant cases is made upon friendship distances for a group of undergraduate men.

References

Beum, C. O., & Brundage, E. G. A method for analyzing the sociomatrix, Sociometry, 1950, 13, 141-145.

Bock, R. D., & Husain, Suraya Z. An adaptation of Holzinger's B-coefficients for the analysis of sociometric data. Sociometry, 1950, 13, 146-153.

Bock, R. D., & Husain, Suraya Z. Factors of the tele: a preliminary report. Sociometry, 1952, 15, 206-219.

Cattell, R. B. A note on correlation clusters and cluster search methods. Psychometrika, 1944, 9, 169-184.

Forsyth, Elaine, & Katz, L. A matrix approach to the analysis of sociometric data: preliminary report. Sociometry, 1946, 9, 340-347.

Glanzer, M., & Glaser, R. Techniques for the study of group structure and behavior: I. Analysis of structure. Psychol. Bull., 1959, 56, 317-322.

Harary, F. & Norman, R. Z. Graph theory as a mathematical model in social science. Ann Arbor: Institute for Social Research, Univer. of Michigan, 1953.

Harary, F., & Ross, I. C. A procedure for clique detection using the group matrix. Sociometry, 1957, 20-205-215.

Holzinger, K. J., & Harmon, H. H. Factor analysis: a synthesis of factorial methods. Chicago: University of Chicago, 1941.

Katz, L. On the matrix analysis of sociometric data. Sociometry, 1947, 10, 233-241.

Luce, R. D. Connectivity and generalized cliques in sociometric group structure. Psychometrika, 1950, 15, 169-190.

Luce, R. D., & Perry, A. D. A method of matrix analysis of group structure. Psychometrika, 1949, 14, 95-116.

MacRae, D., Jr. Direct factor analysis of sociometric data. Paper presented at the American Association for the Advancement of Science, Chicago, December, 1959.

McQuitty, L. L. Elementary linkage analysis for isolating orthogonal and oblique types and typal relevancies. Educ. psychol. Measmt., 1957, 17, 207-229.

McQuitty, L. L. Hierarchical linkage analysis for the isolation of types. Educ. psychol. Measmt., 1960, 29, 55-67.

Moreno, J. L. Who shall survive? Washington: Nervous and Mental Disease Monograph, No. 58, 1934.

Morton, A. S. Similarity as a determinant of friendship: A multidimensional study. Princeton: Princeton University and Educational Testing Service, 1959.

Rao, C. R. Advanced statistical methods in biometric research. New York: Wiley, 1952.

Tocher, K. D. Discussion on Mr. Rao's paper. J. roy. statist. Soc., Sec. B, 1948, 10, 198-199.