I

ADAPTING CENSUS PROCEDURES TO MEET TODAY'S NEEDS

Chairman, A. Ross Eckler, Bureau of the Census

TEAMING UP CENSUSES AND SAMPLES

By: Frederick F. Stephan, Princeton University

Census taking and sampling have been recognized as alternative methods of obtaining data about populations and economies for at least two centuries. Throughout this period they have been considered to be wholly distinct and competing methods, notably in the discussions of the International Statistical Institute at the turn of the century. Now we are beginning to understand the great advantage of using them in conjunction.

The practice of sampling to collect data quickly and economically is widespread in industry and science as well as in government statistical work. The practical problems of sampling have stimulated great progress in developing principles of sampling design and techniques for applying them. These principles now lead us to re-examine the assumptions underlying census taking and particularly the detailed planning of census operations.

Such a re-examination is inescapable when census enumeration is combined with sampling. The successive development of sampling from the 1940 Census of Population to the 1960 Census of Population and Housing serves as an excellent example of the benefits and problems inherent in the combination of the two methods.

Coincidental sampling was introduced in the 1940 Census of Population primarily to permit adding to the schedule several questions for which there was considerable pressure from users of census data without requiring the enumerators to these questions for the entire population. The sample design limited the sample questions to 5 per cent of the persons enumerated. Thus there was a 95 per cent saving in interviewing.

A similar saving was made in the subsequent operations of coding and card punching. Further savings in coding and card punching and reductions in tabulating were made by using samples of punch cards for many cross-tabulations of questions which had been completely enumerated. Sampling was also used for quality control of card punching and other processes.

Thus the result of introducing sampling in 1930 was to relieve the enumerators, increase the information obtained for census users, reduce the labor involved in some tabulations, add new tabulations, make possible the publication of some data earlier than previous schedules provided, and control quality better than previous censuses. However, the main body of questions was enumerated completely and the field work procedures were not modified significantly by the introduction of sampling. Hence the potential benefit of employing sampling was not fully realized.

The usefulness of sampling was demonstrated further by the experience of the Census Bureau with the Current Population Survey beginning soon after the 1940 censuses. By the time preparations were made for the 1950 censuses it was clear that it would be advantageous to extend sampling to some of the questions previously enumerated completely. Quite naturally there was considerable difference of opinion in the advisory committees and elsewhere on the choice of questions to be sampled. As a consequence the use of sampling was extended in the 1950 Censuses of Population and Housing somewhat less than it might have been had there been more evidence, more time for study, and less diversity of viewpoints and interests. Nonetheless, great gains were made; the sample was increased to 20 per cent; and a number of major questions were transferred to the sample.

For 1960, as Morris Hansen will tell you, the potentialities of sampling are being exploited to a still greater degree than in the two previous censuses. One of the principal benefits will be the freeing of the complete enumeration from the drag of questions, such as occupation, which are relatively difficult to enumerate and which delay the tabulations until they have gone through a time-consuming process of editing and coding. The transfer of labor force questions to the sample makes great savings of interviewing and processing costs as well as gains in the speed of publication.

Another important step is to take households instead of individual persons as the elementary sampling units. This makes possible sample tabulations which relate data for two or more members of the same household or which form aggregates of individual data, such as income, for each household. The sampling procedure is more difficult and there are other problems in the shift to a household sample but it appears on balance to be preferable to the unit and procedure used in 1950.

This, in brief, is the evolution of the incorporation of sampling into the traditional procedure of census taking. No doubt the experience of the 1960 censuses will lead to further developments and changes. It is worth while to take a broad look at what has been involved in teaming up sampling and census enumeration. After we have done that we will look at some questions and reservations which are of concern to many census users.

Clearly the primary function of the Population Census is to provide an accurate count of the populations of the States for the decennial reapportionment of Congress in fulfillment of the provisions of the Constitution. To the extent that this function is not jeopardized, additional information can be collected for the guidance of government officials and agencies in the performance of their duties and for the enlightenment of the public.

In the past, careful consideration was given to requests for the addition of questions put forth by various groups and the set of questions finally selected for enumeration constituted a compromise. One might almost call it a coalition formed out of the competing interests in obtaining information about the population. The Population Census thus acquired in addition to its Constitutional function, the function of a general-purpose statistical system.

Clearly its capacity to perform this service had some limits. As the demand for additional data increased, the difficulty of choosing the questions to be included increased sharply.

The introduction of sampling alleviated the pressure against the capacity of the system but complicated the problems of planning. Some of the problems are:

(1) Decisions about the questions to be included in the census are complicated by the necessity of deciding which of them are to be in the sample.

(2) Budgeting, scheduling, and preparatory work are complicated by the necessity of allocating and planning for the sample and of seeking an optimal relation between sample and enumeration.

(3) The effects of sampling on enumerators, respondents, and users add new problems.

(4) Field operations and the processing of data are affected in various ways.

(5) Sampling introduces problems of preparing estimates from the sample and reconciling these estimates with the results of the enumeration.

Offsetting these problems are certain advantages that ease the solution of the problems usually involved in census-taking. For example there is:

(1) Greater freedom in designing the entire data-collecting system.

(2) Opportunity to select a smaller number of better personnel to perform some of the more difficult work.

(3) A greater output of valuable information, or greater economy, or some of both with consequently better command of the allocation of resources.

Similar advantages and problems will arise in other unions of sampling and complete coverage whether they are surveys, inventories, or other canvasses. Some of them arise in the union of two sampling procedures without a complete canvass.

For statisticians, the conjunction of sampling and census-taking brings to the fore a number of technical and procedural questions.

(1) Just how should the sampling and enumeration be coupled? Should sample questions be asked at the same time as as the other questions, by the same enumerators in a separate interview, by a special corps of interviewers, or in some other way?

(2) How can the sample be designated so as to take advantage of the possibility of using the enumeration as the frame but avoiding both the biases that enumerators tend to introduce when they make the selection and the added costs that must be incurred when the sample is selected in the central offices?

(3) How should the sampling proceed in the unusual cases presented by institutions, homeless or mobile persons, and other special groups in the population?

(4) How should the enumeration be used in the preparation of estimates from the sample?

(5) How should biases and sampling errors be estimated and the accuracy of both the enumeration and sample measured?

(6) How should the interests of users in each set of data to be provided by a particular question, and in its accuracy, be given appropriate weight in the planning and processing?

Substantial progress has been made in the solution of these and other problems; interesting questions remain to be answered. The formal analysis of the sixth problem has had perhaps the least attention though the problem has been discussed at length and in great detail by advisory committees and representatives of users. The balance of this paper will sketch a general view of the problem.

We start by assuming that a well-defined purpose is served by information needed to guide certain actions and that we are concerned about the various consequences which might result from these actions - in fact that we can determine in advance the value of each of these possible consequences. Thus we may need information about the economic level of the population in a small area in order to make a decision about an investment in real estate, or the relation of education and fertility to make a decision about the expenditure of funds for further research on fertility differentials.

The consequences of action taken on information depend on the accuracy of the information.

They may be very sensitive to the accuracy of the information or they may not, i.e., moderately inaccurate information may or may not lead to action inappropriate to the objective situation and hence to losses in comparison with the consequences of action based on accurate information.

The mathematical function which expresses the value of consequences in terms of the departure from accuracy of the information may take many forms. Examples are given in Figure 1. In the case of Figure 1a if a census yields completely accurate information, it results in consequences of the highest value. Shifting the question to a sample can only result in a reduction of value.

If the result of a census is not perfectly accurate, the consequences of the action to which it leads will be less valuable than if it were a completely correct answer to the question asked. Shifting the question to a sample will reduce its value further if the sampling errors disperse the estimates under a portion of the value function which is concave downward but if, as in 1c, the census result is at the arrow, the sampling error (if they are not too great) would disperse it under a portion concave upward and would increase the value. In other cases sampling might increase or decrease the value depending on the distribution of sampling errors and the position of the consus result. We conclude that sampling may be expected to decrease the value but not always and not necessarily by a serious amount.

Consider next the lapse of time between the census date and the time census information is put to use. The accuracy of the information from an enumeration will change as the objective situation changes during the time that elapses from the census date to the date of use. Hence we may expect that the value resulting from the use of census information for a particular purpose will decrease with the passage of time somewhat as shown in Figure 2 due to its departure from perfect accuracy. If the question is shifted to a sample, the value will be changed, possibly in a manner that makes it approximately parallel to, or converging toward, the enumeration value function. In those instances in which the sample data become available sooner than they would if they were enumerated completely, the value of sample data may compare favorably to enumeration data both at the time of publication and over the ten year period between censuses. Points A and B in Figure 2 show the value at the time of publication, A' and B' the times at which the information is replaced by new results.

In the foregoing, we have assumed that perfectly accurate information leads to action, producing consequences of the highest value. Actually the utilization of information is not perfect and there is an "error of application" or "error of use" which may be constant or variable but which affects the relation of the value function to the information producing process. Also the relation of the consequences of action to the information on which it is based may be such that the greatest value results from action on information which is not perfectly accurate, i.e., that differs in a certain way and degree from the correct information sought by the questions. In some such cases, the effect of shifting from complete enumeration to sampling may actually be to make no reduction in value or even to increase the expected value of the consequences.

We need to develop definite value functions, measures of bias and error distributions before we can apply these ideas to particular cases. However, even before we obtain this implementation they warrant the following general conclusions:

(1) The ultimate effect of shifting a question to the sample is not always a reduction of value.

(2) When sample data can be published sooner than enumeration data, there may be an advantage in favor of sampling.

(3) Decisions about the choice of sample questions call for better specification of the users' value functions and determination of the departure of both enumeration and sample expected values from perfect accuracy. Judgments based on poorly defined formulations of the users' interests and on other concepts may be inappropriate or even irrelevant to the fundamental issues.

When we turn from a consideration of the effect of sampling on the utility of census statistics for an individual user to the aggregate effect for all users, actual or potential, the problem of finding a relatively precise basis for decisions appears out of reach. There is the question of how the gains and losses of the various users should be weighted in the aggregate, which government functions and which private activities should have priority, and many other considerations. Moreover, the aggregate gain or loss must now be compared with the aggregate gain or loss that would result from an alternative program of allocating questions to the complete enumeration, to the sample, or to the reject pile. These comparisons can only be made by considered judgment at the present time but a clearer understanding of the effects of sampling in individual cases can contribute to the soundness of these judgments.

In summary, the union of census and sample is a fruitful one. Statisticians do well to look for opportunities to use a similar combination of sampling with a complete canvass in other surveys and inventories. The 1960 Censuses will stand as a great demonstration of the value of joining the two methods and the progress made by the statisticians from all the major traditions of statistical work who have joined their efforts in accomplishing it.

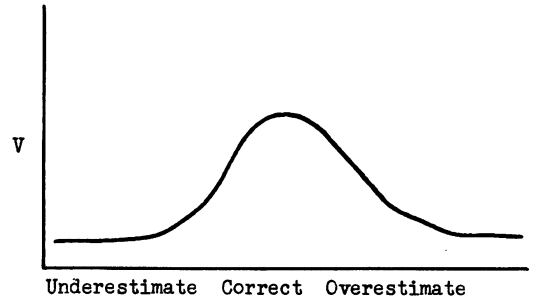


Figure 1a

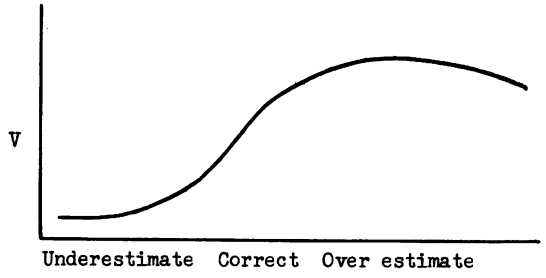


Figure 1b

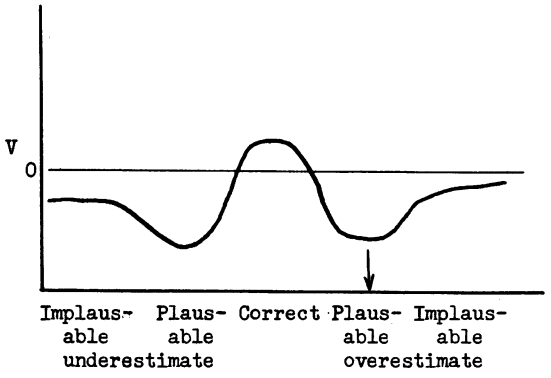


Figure 1c

Three illustrative types of relation between the expected value of the consequences of an action to the actual information on which it is based

Note: The average value of the consequences of the action taken is symmetrically related in Figure 1a, and asymmetrically related in Figure 1b, to the degree of incorrectness of the information on which the action is based. In Figure 1c, the more extreme the degree of under- or overestimate the more implausable it appears and the more it is likely induce further investigation before action is taken.

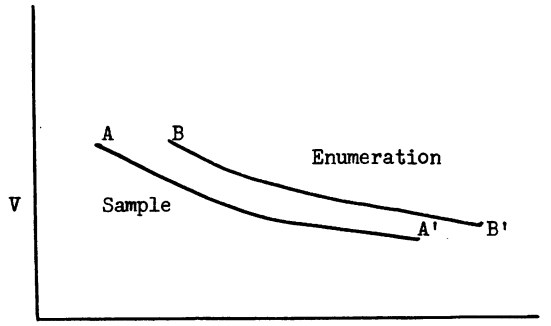


Figure 2

Relation between value of consequences of action and time elapsed since the information on which the action was based was obtained.

MEETING PRESENT DAY NEEDS FOR CENSUS DATA IN AGRICULTURE

By: Arnold J. King, National Analysts, Inc.

In this country there are two major sources of data about Agriculture - The Bureau of the Census which takes an agriculture census every five years and the United States Department of Agriculture which forecasts and estimates the current production in agriculture. These data complement each other and it is, therefore, necessary that both of them be considered in discussing the question of "Meeting Present Day Needs for Census Data in Agriculture."

The statistical program of the Department of Agriculture is based largely upon mailed inquiries. This program which has been in operation nearly a century was designed to provide accurate and timely forecasts and estimates of agriculture production to be used primarily in trading operations. To the production figures have been added, data such as taxes, prices paid and received by farmers, cost of living and production expenses, mortgage debt, grain in storage, etc. These descriptive data have come to be used extensively in government, in legislative deliberations, in business and in analytical work of the Department of Agriculture and in educational institutions. To meet the demand of the users of these data, the number of items covered, the frequency of release, and the geographical detail has increased greatly in the past three decades.

Although some kinds of information is obtained more accurately in a self-administered questionnaire, the kinds of information that can be obtained by mailed questionnaires are limited in scope and depth and the sample indications are subject to bias due to selectivity of the mailed responses. To minimize the effect of biases, the samples are expanded by correlating sample ratios of the production items and land in farms with periodic production figures shown by the census. Although this method reduces the selectivity it does not remove all the selectivity, and it introduces another error because it is assumed in the expansion factor that there is no change in land in farm between the censuses. This means that under the present system the Department of Agriculture's statistical program is dependent upon having a census of agriculture production and the accuracy of the estimates depends, in part, upon the frequency of the censuses.

The Department of Agriculture has decentralized the sampling and much of the data processing to the state level. This program is supported financially and politically by the states. This administrative set-up and the development of an information system, based upon mailed inquiries, has placed the statistical program in a methodological and an administrative strait-jacket. This strait-jacket is not likely to be broken because of self interest and political pressure at the state level. Unless an alternative system materializes it will be necessary that the agriculture census be taken at least every five years and that much of the information collected should be devoted largely to the items covered by the Department of Agriculture's statistical program.

When I try to take a broad look at the census and the United States Department of Agriculture data and, if I assume that descriptive data of production will satisfy present day needs, I would agree the information, as now being obtained, is timely, efficiently and, in its present form, essential. But, consider these data in light of the fast moving world we are now living in, the advancements made in sample design, advances in data processing through the introduction of computers, devices developed by the psychologists for pulling information from respondents in a survey interview, and analytical concepts that are now available for solving agricultural problems. It seems clear to me that the informational system developed by the United States Department of Agriculture which has not undergone a basic change in methods during the past thirty years and a census taken every five years which is confined largely to items of production, just does not meet present day needs.

Let me get more specific. There are tremendous changes taking place in agriculture. I am convinced that the five year period between census is too long a period to wait, in order to have reasonably reliable figures in agriculture. For example, the increase in the per farm man hours from 1940 through 1957 is as great as the increase in the 120 years from 1820 to 1940. In the past 10 years, farm tractors have increased almost 90 per cent. There are twice as many grain combines and milking machines, four times as many mechanical corn pickers and eleven times as many pick-up balers and forage harvesters now on farms as 10 years ago. Fertilizer consumption is about 2-$\frac{1}{2}$ times that of 1940. Changes in one year now may equal those of 5 years during a period 30 years ago.

More than 1-$\frac{1}{2}$ million farms, or about 1/4 of all farms, have disappeared since 1930. More than 1/3 of these changes occurred between 1950 and 1954 and 2/3 since 1945. Practically all decline in farm numbers has been in commercial farms. Since there has been little change in the acreage of land in farms and in total cropland, the acreage of the disappearing farms has been absorbed into existing farms. The change in the average size of farms has affected the size of commercial farms, as non-commercial farms have not changed greatly in average size. From 1950 to 1954, the average size of the commercial farm increased from 220 to 336 acres -- a 50 per cent increase.

The pressing need to balance supplies of farm production and demand, a need for a more uniform and efficient flow of agriculture

produce through the marketing system, the inroads of man made products into the markets for natural products are illustrations of a whole complex of agriculture problems in which data is needed in their solution. Agriculture is going through a technicological revolution, not only in production but in marketing. New concepts of problem analyses are being developed for agriculture production, processing and distribution of foods and fiber and in the supplying of goods and services to farmers. New and powerful analytical concepts are being devised for solving business problems, such as operations research, statistical, economic sociological models. These models will be used in solving basic agriculture problems in the future. Therefore, descriptive data of production are no longer sufficient. Whole new systems of information are needed.

A major step in developing a new system would be an annual farm visitation sample conducted by the Bureau of the Census. They have the trained field staff to conduct interviews, they have the experienced technical staff to design the samples efficiently and they have the computers and the technical staff to process the data using modern methods of analysis. However, I feel that the over-all responsibility of the agriculture statistical program should remain in the United States Department of Agriculture. They should have the responsibility for making the forecasts and estimates of production, the setting of survey objectives, and the analyses in terms of charting courses of action.

Such a sample census could be made up of three parts, each utilizing about one half of the interviewing time: - one part to provide annual estimates in such items as land use, acreage and production of crops, number of livestock, etc. (this type of information could be designed primarily to strengthen the United States Department of Agriculture's production statistical program); a second part to provide new information but on a repeat basis; and a third part to obtain new information as needed only once, which in the parlance of survey people are the "one-shot surveys." The national sample could be designed so that parts of the questionnaire could be constructed to provide regional information. If the data are needed only for the United States as a whole, it could be obtained on a sub-sample, say 1/10 of the segments and the specific questions be rotated among the questionnaires so that 10 times as many items of this kind could be obtained. Because of the smaller number of interviews and the larger and more thorough training, it would be possible to obtain accurate information from the respondent on many kinds of items which could not be obtained in a national census involving 30,000 enumerators.

An area probability sample of 180,000 farms consisting of about 1,000 segments per state (except for New England and Nevada) using revised Master Sample materials containing an average of about 4 farms per segment, supplemented by a sample of 10,000 to 20,000 large farms would provide an enormous amount of new and powerful information. A sample of this size would permit, for example, accurate estimates of the total number of farms, land in farms, land under cultivation and major livestock numbers by states. Generally, data for cross-tabulations, frequencies of occurrence could be estimated for nine types of farming areas, and for three geographic regions - North, South and West.

The sample census as indicated could, for example, provide accurate information on such important items as insurance, medical care, tenure practices, family employment, sickness, accident, farm construction, fire damage, marketing channels, transportation methods, stocks, utilization of crops, production methods, production practices, inventory of farm machines and equipment, days of use of machinery and equipment, rental agreements, debts, use of chemicals for weed control, insecticides, feeding practices, purchasing habits, cooperative marketing, use of insecticides, use of fertilizers, soil management practices, use of new and improved varieties of crops, farm population, hours of work, wage rates, etc.

In my opinion, there are many advantages of having the sample census taken in the fall rather than in the spring. If taken in the fall, October would be the ideal time for the interviewing in the rocky mountain and hard wheat areas and the first part of November for the remaining areas. If the interviewing workload averaged about 120 hours, the interviewing could be completed by the end of November. The size of the field staff would be small enough that they could be sufficiently trained to prevent the response errors from getting out of hand. If the interviewing was first completed on a sub-sample of the segments and of the large farm operators during the first week and the data processed on the computers, it would be possible that estimates for the United States could be released before January 1. States estimates from January 1 through February and cross-tabulations during March to June.

Existing or new Master Sample materials could be used to designate the sample segments. Materials prepared for the 1954 sample census of Agriculture should be satisfactory for 26 northern and eastern states. For the southern states for which census county divisions have been established, a new set of segments would be needed. Since the Ed's to be used for the last census will comprise areas larger than an MCD, the segment boundaries could be drawn so that they have distinct, identifiable boundaries and not influenced by MCD boundaries with Ed's or census divisions. A small supplementary sample of urban areas could be used to provide data for the area sample. The supplementary

sample of large farms could be taken from the last census.

The utilization of electronic equipment makes it possible to use more efficient methods of estimation than were possible in the past. For example, regression methods of estimation can be used which utilize information from the previous census in such a way that sampling variation is greatly reduced. The estimate for any item could be made by first obtaining the simple unbiased estimate, multiply the sample total by the reciprocal of the sampling rate and adding to this estimate a quantity which is the difference between the complete census total for that item, and the simple unbiased estimate of that total for the sample in the base year. This method takes into consideration the correlation between years in the identical sampling units. In most agriculture items a correlation of better than .75 can be expected which would result in substantial reduction in sampling variation.

It has been estimated that an annual sample census as outlined above would cost from $2,200,000 to $2,500,000 per survey as compared to $22,000,000 for the 1957 census. If a national sample census was taken every year and a full census every ten years, more useful information would be obtained to meet present day needs at no additional cost to the taxpayer. A vastly greater range of items could be covered and depth of information obtained, and the response error would be at a minimum because of intensive interviewer training.

With the size of the sample indicated, more accurate estimates would be forthcoming than under the present systems for the production of the major crop and livestock on a national and regional basis. Furthermore, the data from the sample would provide a more accurate basis for expanding the mailed surveys. If county figures are needed more frequently than once every ten years, I feel this is the responsibility of the states. In about one fourth of the states there has been sufficient interest in annual county figures that the states are taking annual agriculture census through the assessors. If additional states find a need for county estimates they have two courses of action open to them. They can obtain the data through the assessors or they can, no doubt, contract for this work with the Census Bureau in the same way cities contract with the Bureau for population census. This arrangement would, in my opinion, result in a better coordinated and directed statistical program between the federal and state governments.

The agriculture census, the mailed surveys of the United States Department of Agriculture and the annual sample census as outlined above would provide data that is largely descriptive and of the nose counting variety. There is wholly a different kind of survey information needed that is designed to solve specific problems. The United States Department of Agriculture has in recent years conducted a number of surveys of this type by contracting with commercial firms for the interviewing and the data processing. These surveys have been largely directed at marketing and related problems and from a survey methodological viewpoint are of high quality. Because this research has been handicapped by a lack of funds, they have too often been narrow in concept.

The psychologists have, in recent years, made great strides in developing methods for field interviewing that goes a long way in explaining human behavior by using such devices as scaling or projective methods, word-associations, multiple choice answers, and open end probing types of interviews. After all, many of the problems facing agriculture rests with people and, regardless what the facts are, people will act as they perceive the situations.

For example, the controversy continues over the role of federal agricultural agencies in establishing policies and systems having a direct effect upon levels of production and prices of agricultural commodities. A more intensive use of sample surveys to determine the motivations, perceptions, and attitudes of farmers with respect to these systems and policies is called for. The essential danger is that the directors of the agriculture agencies will attempt to implement their policies and systems while working from assumptions that are faulty estimates of the psychological characteristics of the people who must, in the final analysis, operate within their own framework. This is not to say, necessarily, that the validity of a given policy or system is to be based upon the average farmer's attitudes and opinions. The point is that the intrinsic worthwhileness of a given policy is not guarantee of its success. It can be a success only in terms of its being perceived as something designed to meet the needs and perceptions of those it is meant to effect. All too often, the barriers to success of a program rest almost solely in the attitudes and perceptions of the people who are expected to carry out the practices called for. The attitudinal and perceptual barriers cannot be changed until they are analyzed and understood in the first place. Present day market research has had a great deal of success in adapting techniques from psychological research to the study of problems such as this within the framework of sample surveys.

In keeping with this view, one can ask whether or not agriculture needs the continuing type of psychological research represented in the Federal Research Board's surveys of consumer finances and spending. In these surveys great emphasis is placed upon consumer expectations, consumer feelings of confidence in the economic

situation, consumer spending intentions, etc. The years have demonstrated that these surveys provide valuable categories of information which can be integrated into other areas of information for diagnostic appraisal of the current and near-future status of the economy. Surely, more intensive research of this type can be done upon the farmer both as a producer and a consumer.

At any given moment there is a tremendous amount of research being conducted on some aspect of the demand for agricultural commodities. And, much of this research makes use of the sample survey approach. Further, this research is generally characterized by a high degree of quality with respect to concepts and methods. The most serious criticism of all of this research effort, however, is that it is directed toward analysis of demand - problems associated with single commodities (eggs, poultry, frozen concentrated juice, etc.) or limited categories of commodities (dairy products as a class, citrus products as a class, etc.). The problem here is that the demand for these single commodities or categories of commodities does not exist in isolation. There is interaction between the demand for the specific agricultural commodities, and interaction between the demand for agricultural commodities and non-agricultural commodities. The real understanding of what is happening in terms of the demand for any one commodity must be a function, to some degree, of our understanding of these demand-interactions. As these interactions occur within people (the consumers) we must go to them in order to obtain the needed data. A few years ago we, at National Analysts, made our first real attack on this type of problem in our research on buying-decision behavior which we conducted for the Advertising Research Foundation. I believe that this kind of survey information can make a contribution to agriculture.

We sometimes seem to ignore the fact that the passage of foods and fiber through the distribution system from primary producer to ultimate consumer involves people at each of the traditional steps or stages -- processors, shippers, wholesalers, and retailers. And, whenever we have people, we must be prepared to admit that motivations, perceptions, values, information levels, etc. can be functioning in a way to prevent the operation from being a wholly rational one. Traditionally, market research, especially that using psychological concepts and methods in combination with sampling procedures, has been directed toward the consumer.

In recent years, there has been an increase in interest in applying these to problems arising within the distribution system. Several of our current projects at National Analysts are of this nature. For example, one of our projects has to do with the use of certain information sources by decision-making executives in a certain industry. It soon became apparent that we could not study this problem without going into certain motivations, perceptions and attitudes of these executives as they pertained to their decision-making function. And, of course, we encountered a most challenging problem in terms of sampling this particular universe. The main problem was that we did not want to go by title in locating our sample respondents but on the basis of actual job functioning. The studies we have conducted to date have been mainly on non-agricultural problems. We have done a few studies, of this type, though, on agricultural commodities under the sponsorship of the American Dairy Association.

It is often claimed that many of the problems in agriculture are based upon the fact that there is a drastic change occurring in the role or meaning of "farm living" in our present day society -- the values associated with the farm as a way of life are changing. To what extent do we know, in a definitive manner, just what the nature of these value-changes is? In what segments of the population are the changes occurring in the most functional sense? What are the sources of the new set of values that are relacing the old? Sociology and social psychology have matured to the extent that concepts and techniques from these fields, in combination with sampling techniques could be utilized for the study of this vitally important problem on a survey basis.

In summary, I feel that as long as the United States Department of Agriculture continue to base its forecasts and estimates of agricultural production on mailed surveys, the agriculture census should be taken every five years and devoted largely to an inventory of production so that this vital information can be continued. However, this information is limited in scope and depth.

A first step that should be taken is to break out of the administrative and methodological strait-jacket that these programs are in is to have the Bureau of Census take an annual sample census. This Bureau has the facilities to conduct the surveys and process the data efficiently. If an agriculture census was taken every ten years and a sample census taken every year, there would be an increase in the accuracy of the production data for the nation and for regions, more useful information obtained in the census, and the sample census would provide a wealth of new and powerful information at no extra cost to the taxpayer.

It should be recognized that this information is largely of the nose counting variety and there would still be a great need for sample surveys designed to solve specific agricultural problems by obtaining from farmers, processors, shippers, wholesalers, retailers and the consumers -- their perceptions, levels of information, attitudes, opinions, values, motivations, etc. -- in other words, obtaining directly from the people in the agricultural producing and marketing system information as to why they do as they do.

PROCEDURES FOR THE 1960 CENSUSES OF POPULATION AND HOUSING

By: Morris H. Hansen, Assistant Director for Statistical Standards, Bureau of the Census
U. S. Department of Commerce

Major changes are being made in the methods for taking the 1960 Censuses of Population and Housing as compared with previous censuses.[1] Some are strictly procedural and should have little or no effect on the nature of the results to be published. They will reduce the cost or speed up publication of results of the Censuses, and may be of interest to statisticians on this account as well as because of their potential application in other areas of statistical collection and compilation. Other changes in methods should improve the accuracy of results published. The changes in procedure represent an evolution from previous census methods, and the new aspects have been evaluated in various ways including tests, before, during, and subsequent to the 1950 Censuses. We have had experience in other recent censuses with some of the new procedures. The plans call for decisions and detailed planning much earlier than has been the case in the past.

Among the principal changes in methods now planned for the 1960 Censuses of Population and Housing are the following:

(1) Extension of sampling. Sampling will have a fuller role in the 1960 than in the 1950 Censuses. The complete population census will be limited to a few basic items. Information on other subjects will be collected from a 25 percent sample of households whereas in 1950 some of the subjects were handled by 100 percent enumeration and others by 20 percent and 3-1/3 percent samples. Sampling will serve a similar role in the housing census.

(2) New Equipment. All data to be tabulated in the census will be recorded on specially prepared forms by means of positioned marks (check boxes). These forms will be microfilmed, and newly developed electronic equipment will convert the data from microfilm to magnetic tape without manual punching of cards. The magnetic tape will serve as input and the returns will be edited and tabulated in the computer. The computer editing can be so controlled that it will automatically dispose of most inconsistencies and non-response but will print out for manual review those which might seriously distort subsequent tabulations. Much of the final results will be published by direct offset reproduction of the high speed printer output.

(3) Enumeration methods. Although enumerators will visit every home, much of the information will be furnished by forms which the householders themselves have filled out. There will be a two-stage operation in most areas. The 100 percent questions will be distributed in advance through the Post Office, with a request for respondent cooperation. The questionnaire, if filled out, is to be held for review and completion by the first stage enumerator. The sample questionnaire will, in a limited number of areas of low density, be filled out at the first stage, but in most areas will be left by the original enumerators at 25 percent of the households, with a request that it be filled out and mailed in. Selected enumerators who have worked on the first stage will be trained for the second-stage operation and will review and transcribe the returns to the final census schedules, and revisit the sample households as needed to complete the information. This procedure provides an opportunity for participation in the responses by all members of the household, and reduces the contribution of enumerators to response errors.

The objectives of these procedural developments are threefold. First, we expect to make substantial reductions in the time it takes to publish results; second, we hope to improve the quality of the censuses; and third, we expect to achieve the first two objectives at less cost than the 1950 cost levels adjusted for the increase in population and salary and price changes.

Extension of Use of Sampling in 1960 Censuses

In the 1950 Population Census, a 20 percent sample of persons was used for the collection of data on education, income, migration, and other subjects, and a 3-1/3 percent sample was used for fertility questions and certain other topics. For 1960 we shall, in addition, transfer citizenship, place of birth, employment status, occupation, industry, and related questions to the sample and increase the sampling fraction from 20 to 25 percent. The sample in 1960 is a sample of households. Remaining on the 100 percent population census will be the listing of the population, relationship to head, age, sex, race, and marital status. The 100 percent questionnaire will contain no questions requiring manual coding, and the questionnaires can go directly into microfilm and through processing without any delay for manual operations. Thus, a definite principle in deciding which questions to include on the sample was whether manual coding of the returns is necessary before processing.

Of course, the other and fundamental principle of deciding whether questions should be collected for 100 percent of the population or on the 25 percent sample is the requirement for data. Most of the items chosen for 100 percent coverage are basic demographic characteristics on which relatively precise statistics are needed for small as well as large areas, and for which past experience shows that relatively high accuracy can be achieved in the census. Many of the items on the sample involve concepts that are relatively difficult to measure. The use of sampling permits coverage of more items in the census, faster compilation, and lower costs.

Sample selection will be accomplished by designating every fourth household in order of visitation by the enumerator for inclusion in the sample. All members of each selected household will be included, as will data for the housing unit. Somewhat different procedures will be used for selecting the sample within institutions, large lodging houses, and other special dwelling places. In 1950 a sample of lines was predesignated on the population schedules, and the sample data were obtained from the persons enumerated on these lines.

The choice of housing items for complete coverage or sampling will follow principles similar to those described for population. However, for cities of 50 thousand or more population we plan to publish summary housing characteristics for individual blocks, and the items tabulated for blocks will be on the complete census. The items collected for all housing units in these cities will include, tenure, rent or value, number of rooms, condition, plumbing facilities and certain other questions. In the remainder of the country tenure and occupancy status, number of rooms, type of housing, condition, and plumbing will be collected 100 percent, and the remainder of the questions will be collected on a sample. For housing the sample will be somewhat more complex, however, in that some of the sample items will be collected for a 25 percent sample, and others will be from 20 and from 5 percent samples.

In both population and housing the items collected from the 25 percent (and also the housing 20 percent) samples will be tabulated in substantially as great detail as would be the case if the items were collected 100 percent. For most large area statistics, such as for states, metropolitan areas, and large cities, the use of this sample rather than 100 percent enumeration should have a barely perceptible effect in decreasing reliability of results. For small area statistics, such as census tracts, and very detailed tabulation cells for large areas, the sampling errors will be relatively large, but even here the accuracy for some items will not be substantially weakened as compared with 100 percent coverage, primarily because response errors for these items are relatively large.

We had hoped that all of the economies from the sampling would be available to finance improved quality and coverage. While some of these gains will be achieved, most of the gains from sampling will probably take the form of reductions in the total cost of the census.

A major gain from the extension of sampling is the ability to publish results much sooner (this is a joint gain from the extension of sampling and the application of electronic equipment, together with a great improvement in the advance planning of the census).

Electronic Equipment

Computers. The first Univac was delivered to the Census in 1951 and was used for compiling some of the 1950 Census. While it contributed only a small amount to that Census, the experience taught us much. During the next five years we put it into use on our current work and on many special and service projects. We needed additional capacity for the compilation of the 1954 Censuses of Business and Manufactures, and jointly with the Internal Revenue Service acquired a second Univac in January 1955.

About the time the processing of the Economic Censuses was virtually completed in 1956 we addressed our attention to the question of the computer capacity we would need to process the work load we could anticipate during the 1959 to 1962 period.

When we matched our anticipated work load against the production capacity of the Univac I's, we decided that we would need a minimum of eight Univac I's in full-time operation during the peak of our requirements. Since more advanced computers were becoming available, we embarked on a study of the comparative merits for our work of several of the new systems. The study was completed early in 1957 and resulted in a decision to replace our two Univac I's with two Univac 1105's. The 1105 system is a modification and improvement of a computer system known as the 1103A, designed primarily as a scientific computer. The modifications, which were introduced to meet the Census requirements, adapt the equipment for exceedingly effective large scale data processing work as well as for scientific computing.

The first 1105 was installed last October and the second one in December. Until about June of 1960, the 1105's will be used for current work and for processing the 1958 Censuses of Manufactures and of Retail and Wholesale Trade, and Service Establishments. Thus, we shall have considerable experience on our new computers before embarking on the tremendously large scale work of the 1960 Censuses of Population and Housing.

The two 1105 systems installed at the Bureau of the Census headquarters will be capable of delivering about 3,000 hours per quarter of productive time by operating seven days a week 24 hours a day. Approximately twice this capacity will be needed during our 1960 Census operations. To provide the additional capacity the Bureau of the Census is cooperating with the University of North Carolina and with the Armour Research Foundation of the Illinois Institute of Technology. In each of these universities an 1105 system compatible in all respects with the Census systems will be installed. In return for what will be effectively a prepayment for time on its computer, each university has agreed to make about two-thirds of the productive time on its computer available for Census use during the period of peak work load. In addition, the Census will have access to a smaller amount of time at each university for a period following the completion of the processing of the 1960 Decennial Censuses of Population and Housing.

Printers. One of our early lessons with electronic equipment was that the success of a large scale electronic data processing installation depends as much on the reliability and performance of its off-line auxilliary equipment as it does

on the central computer. A reliable high speed printer is one of the important auxilliary units, and we have two high speed printers in operation. We hope to program our output from the computers in such form that the output from the printer will be ready for final review and photographic reproduction. The ability to accomplish this through the flexibility of a computer system and a high speed printer is considerably greater than with punched card equipment. This utilization of the computer and printer will aid in speeding up and in lowering costs in the census production. We plan to do direct reproduction of printer output for the tables that are repetitive in general format for many areas or other classifications. Some of the tables, however, will be put together most effectively by manual methods, as in the past, but even in these cases the computer and printer system should simplify the manual operation to a considerable extent.

Fosdic. One additional auxilliary electronic device is particularly essential to our plans for processing the 1960 Censuses of Population and Housing. This is equipment developed for us by engineers at the National Bureau of Standards. It is known as FOSDIC, a name consisting of the initial letters of Film Optical Sensing Device for Input to Computers.

It was about 14 months after the enumeration date of April 1, 1950, that the last card was punched for a person enumerated in the 1950 Census. To improve this time schedule significantly we could plan to increase the staff of key punchers or we could search for some automatic method for accomplishing the translation from Census schedules to a tabulation medium. We described the problem to personnel of NBS and after several months of study they proposed FOSDIC.

The FOSDIC forms are designed so that answers are recorded as marks in specific locations (similar to the way data are recorded in cards by punching holes in specific locations). Almost any writing instrument can be used. The completed questionnaires are first photographed on 16 millimeter microfilm, and FOSDIC scans each frame of microfilm at a rate of about 100 to 200 a minute, depending on the amount of information recorded. It automatically reads recodes, and records on magnetic tape the intelligence that has been entered on the schedule, FOSDIC can be programmed to scan the data recorded in a microfilm frame in almost any desired sequence and has a great deal of flexibility for adjusting itself to various sizes and designs of questionnaires and for tilt, shrinkage, or expansion of the copy when it is filmed. FOSDIC has the further useful ability to identify the dominant or heaviest mark in a particular area, thereby making it possible to correct an answer by normal erasure. Experimental work with a completed prototype indicates that its reliability as an input preparation device will result in greater accuracy than has been achieved in the past with either mark sense punching or manual key punching. One of the four production models of Fosdic being built at Census has been completed and is now undergoing final checking and adjustment.

Population 100 percent Tabulations. We contemplate the use of two basic types of FOSDIC census schedules, the 100 percent (short) schedule for both population and housing, and a longer sample schedule. The basic information recorded on the 100 percent schedule will also appear on the sample schedule for these households. These two forms are used to help speed up compilation and publication. We expect to reduce the time for publication of census results by an average of about a year and a half as compared with the 1950 Census, and to compile and publish the census results obtained from all persons (the 100 percent information) within a few months after the returns are received.

It is important to recogrize that the information recorded on the 100 percent Fosdic schedules will not require coding in the office. This is in contradistinction to data on such characteristics as state or country of birth, occupation, industry, etc., which will appear on the sample schedules and which will be converted manually to number codes in the processing office before they are tabulated.

The legislation under which we conduct the Census requires us to report the total population not later than eight months after the beginning of the field enumeration. To meet this requirement we must determine only the number of persons in each state; we do not have to classify them by their age, sex, occupation, or any other characteristic. This initial or "apportionment" report has always been the result of counting by hand—no non-manual method for obtaining this count on a satisfactory time schedule has heretofore existed. However, for the 1960 Censuses our plans comtemplate that we will microfilm the 100 percent schedules, process the microfilm through FOSDIC and the resulting magnetic tape through our computers by the end of October 1960. Thus, we expect to have not only the population count from the machines, but also the tabulations of 100 percent data on the time schedule on which the hand count alone was completed in 1950.

Without electronic data processing equipment we do not believe we could work to such a tight time schedule. However, the computers alone do not bring our hopes within range of accomplishment. Only with the elimination of manual coding required for the complete census schedules, the ability rapidly to prepare our input medium which FOSDIC provides, and use of the computers does our objective become feasible. In addition, this time schedule requires advance planning beyond any level accomplished in earlier censuses, and it calls for doing the re-allocation and searching for duplicates of persons enumerated away from home as a part of the field collection operation instead of in the processing offices as has been done formerly.

The time schedule allows about four months over-all after completion of the field work for microfilming, preparing the tapes on FOSDIC, and putting them through the computers, and only about three months for any one of these operations. Three months to tabulate data for 180,000,000 persons is indeed short. Only with the most detailed

kind of advance planning can we hope for success. Our plans for this part of the work can be described as having two major characteristics. First we are organizing the job in such a manner as to reduce the number of work units to be separately identified and controlled from 180,000,000 to about 260,000 (the approximate number of enumeration districts into which the country will be subdivided). Second we expect to have "dry runs" of processing the Census between now and the time we actually begin processing the returns.

With respect to our plans for reducing the number of work units from 180,000,000 to 260,000, it was mentioned earlier that the microfilms of 100 percent content will be photographs of the schedules as prepared by our enumerators with no manual coding in the office. The only detailed pre-tabulation inspection of these schedules will be by field crew leaders and will for the most part occur during the course of the enumeration.

We recognize that at the time the schedules are microfilmed not every question will be adequately answered for every person enumerated. There will be some omissions and obvious inconsistencies. The detection and correction of such errors is what is commonly called editing, and we will direct the computer to perform the essential edits. In our judgment, this is, in many ways, a better way to edit than by manual inspection. The computers will do the work much faster. They will apply the editing rules individually to each of the 180,000,000 people, they will do it absolutely consistently, and they can be directed to keep track of how often they apply each rule. An illustration here may help to illuminate this point. The following simplified rules might be incorporated in the 1105 instruction program which would be so arranged that they were applied only when the question on the marital status was not answered.

1. If the person is under 15 years old—proceed to the next person (we do not tabulate marital status for children)
2. For a person 15 years old or older examine relationship
   a. for wife or for a male head of household followed by a wife assume married
   b. for other relationship, assign marital status by using the answer for the preceding person in this class of the same sex-age group for whom marital status was given and in addition—this is important—add "one" to a tally of the number of times an entry for marital status was imputed by the application of this rule.

In this illustration we see how the computer might proceed with the preparation of a tabulation by supplying missing information. This will seriously distort the resulting tables only if it is done too often, and the computer is keeping track for us of how often it is done for each work unit. The final editing rules may be quite different from the simple illustration above but the principle will be the same.

The technique for combining editing with tabulation leads to the plan for making 260,000 the effective number of work units to be controlled rather than 180,000,000. We expect to tabulate the 100 percent population statistics for each enumeration district (E.D.) separately. As each E.D. is completed our instructions to the computers will direct them to evaluate the resulting statistics and determine whether they are acceptable or should be subjected to further review and perhaps correction. Thus, we can instruct the computer to accept an E.D. only if there is a "reasonably small" difference between the field count of population and dwellings and the computer count, and also if the number of computer-kept tallies of the number of imputations it has made for each characteristic is "tolerable." We will have complete control of the definition of "reasonable" and "tolerable." The computers will rigorously and unerringly apply whatever standards we establish. We expect that the standards finally used will be such that about 90 percent of the ED's will be acceptable at the first run through the computers and that the remainder numbering about 25,000 ED's will be a manageable number of units which will be identified by the computer as unacceptable, and require manual investigation, correction, and rerunning.

The art of planning and computer programming is extremely intricate and requires the exercise of planning abilities of a high order and painstaking attention to detail. Also, if we are to achieve the desired time schedules, it requires that complete and final tabulation plans must be made available to technical programmers at a much earlier date than was required by the processing system previously used. With such timing the possibilities for flexibility and for improvisation of tabulation plans as you go along are extremely limited. We have for the past several months been writing, rewriting and testing some of the programs to be used on 1960 Census data during the summer and fall of 1960 and throughout 1961. This process will continue for more than another year before we are finally satisfied that we are ready for the big job.

We are certainly not sanguine about our ability to implement these plans without encountering serious difficulties. Here our plans to spend most of the next year on "dry runs" are important. During late February and March we plan to enumerate a pretest of the 1960 Census of Population. The areas we have selected for this dress rehearsal contain about 100,000 inhabitants and will allow us to anticipate many of the problems we will face in 1960. Also, we may run one more test in the summer of 1959 after the final census schedules are released for print. We plan to microfilm, process through FOSDIC, and tabulate the schedules resulting from these tests as many times as necessary between now and the summer of 1960 to acquire real skill in, and familiarity with the whole sequence of procedures involved.

Housing 100 percent tabulations. To this point our remarks concerning our electronic data processing equipment have been oriented to the 100 percent population tabulations. Our comments apply equally to the 100 percent housing tabulations except that for the Housing Census there is no such

early time requirement imposed by law. Nevertheless we hope to process the 100 percent housing data on approximately the same time schedule we have set for the 100 percent population data.

Sample data. We plan to use the same complex of microfilm, FOSDIC, and 1105 computers to process the sample data for population and housing. For the sample population schedules the manual coding of occupation, industry, place of birth, and other questions must precede microfilming. While we are processing the 100 percent schedules this coding will be started. The microfilming of the sample schedules will begin as soon as the microfilming of the 100 percent schedules is completed.

The volume of microfilm work and FOSDIC operation for the sample schedules will be greater than the work load for the 100 percent cases. The workload on the computers for the sample will be several times that for the 100 percent tabulations when measured in terms of computer time required. Whereas we hope to edit and tabulate the 100 percent data with one pass of the magnetic tape through the computer (except for rejected ED's and for summarization runs) there will be several passes necessary for the more detailed sample tabulations.

Our electronic data processing equipment has already demonstrated its utility for applying a variety of estimation formulas to sample statistics. The sample design for the 1960 Censuses is a fairly simple one. Nevertheless, the electronic processing will aid materially in obtaining estimates from the sample. We expect to use a ratio estimate, separately by age-sex groups and perhaps by size of household groups, still to be specified in detail. The ratio estimates will be applied by small areas, and fixed weights determined for each person will be carried through subsequent tabulations. Such ratio estimation would not have been practicable without the use of electronic computers. The ratio estimates will reduce the sampling variances and result in close agreement between sample and complete census tabulations.

Census Enumeration Methods

In much of the country, perhaps 85 to 90 percent, we expect to divide the collection of census data into two stages. The first stage will be the complete census canvass. In advance of this canvass, questionnaires containing the 100 percent questions will be distributed through the Post Office to give respondents an opportunity, through self-enumeration, to prepare considered answers before the enumerator calls. There will be supporting publicity. The first-stage enumerators will canvass all households in their assigned enumeration districts, as in previous censuses. They will transcribe to the FOSDIC forms from the questionnaires prepared in advance by respondents, or, as necessary, will ascertain by interview and record the information on FOSDIC schedules. In this first-stage enumeration every 4th household will be designated for the sample, and the sample questionnaires will be left with these households, to be filled out by the respondents and mailed in.

Second-stage enumerators will be selected from among the first-stage enumerators about April 15, after most of the initial canvass is completed. They will be trained on the sample questions, will review and transcribe the mailed returns to FOSDIC forms, and follow up by telephone and personal visitation as necessary to complete the sample information. In each stage the necessity of transcription by the enumerator calls to his attention any questions for which information has not been provided in the self-enumeration form, and provides an opportunity for examination of each response for acceptability. He interviews and collects the data not entered on the self-enumeration form.

In about 10 to 15 percent of the country, especially in the more sparsely populated areas, a one-visit procedure will be used, with both 100 percent and sample questions covered by the same enumerator on a single visit. In the areas where the one-visit procedure is followed, as elsewhere, the advance self-enumeration form will be distributed through the Post Office for respondents to fill in. This form will contain only the questions on coverage and content to be covered in all households. On occasion, in these areas, sample forms may be left to be mailed in, to reduce callbacks.

It has long been recognized that the census procedures place a particularly heavy burden on enumerators. They must absorb information concerning techniques of enumeration, the use of maps, administrative procedures, identification of separate living quarters, identification of residents and nonresidents and determination of who should be enumerated at a particular location, and designation of sample households. In addition, they must learn to apply the many concepts and definitions relating to the various subjects in the 100 percent and sample inquiries. Our field staff members believe that by separating the work into two stages, one primarily a canvass for census coverage and a few basic questions, and the other for the more difficult questions, we can more effectively train the enumerators for each stage separately, and can control their work more adequately, than under the one-visit approach. Also, the initial census canvass can be completed more rapidly, and thus reduce the effect of population movement on census coverage.

Under the two-visit procedure we propose to train about 160,000 enumerators on the first stage of the work and as they complete their first work stage, to choose about one-third of them to be trained for and to carry out the second stage operation.

The introduction of self-enumeration represents the first time this method has been widely used in the U. S. Censuses of Population and Housing. Experimental tests have been run on the use of self-enumeration at various times, including tests in 1948, and earlier, a test as a part of the 1950 Censuses, and tests in 1958. These tests have demonstrated widespread public cooperation, and show that the joint method of self-enumeration with enumerative follow-up is a feasible operation, and can be accomplished at about the same

cost level as with a regular direct enumeration census. Actual cost levels will depend upon the effectiveness of public cooperation.

Our plans do not assume that we shall receive substantially complete public cooperation, but we are optimistic that we shall receive very favorable public response in a national census undertaking. In the test conducted in a few counties in the 1950 Census the questionnaires were to be mailed in and about 95 percent were received through the mail. In this case an average of more than 80 percent of the individual responses on questions were acceptable and did not need further inquiry, but some questions were unacceptable on many returned forms, and follow-up on one or more items was required for a high proportion of the households. We hope to improve the self-enumeration forms and publicity and achieve even better cooperation in the 1960 Censuses.

The self-enumeration approach supplemented by follow-up enumeration as necessary is designed to take advantage of the joint contributions that can be made by respondents and enumerators. Our tests and research show that in a large army of more than 150,000 enumerators, recruited and trained in a short period, as they must be for the census, a considerable number will sometimes misinterpret the census instructions to enumerators. Particularly serious is the fact that an interpretation on the part of an enumerator may affect the results for the entire area that he covers. We have learned that such variations in interpretations consistent within the work of an enumerator but varying among enumerators, can significantly affect census statistics, especially for small areas and small cells. Self-enumeration reduces the effects of such tendencies toward consistent misinterpretations. At the same time, this approach allows more mature consideration of the responses to the questions and permits the best informed respondents to participate.

These comments on the effectiveness of self-enumeration have been directed particularly at the quality of information collected on various subjects. We hope, also, that the self-enumeration forms and the two-visit procedure will aid in improving coverage of the census. The initial short (100 percent) self-enumeration form provides for listing all persons who are members of and live in each household, whether present or not at the census date, and also all persons present in the household, whether or not they live there.

This approach, we believe, should help improve the coverage of less closely attached household members, (related members other than the wife and children of the head, and especially lodgers, guests, and persons with no usual place of residence). These groups have proved particularly difficult to cover adequately in a census. Perhaps providing for everyone on the self-enumeration form will result in a more complete listing of persons who should be listed as living in a household. This procedure should, in addition, improve our ability to pick up and reallocate people in private households who live elsewhere but have no one there to report for them.

We had hoped to arrange for the Post Office to check our census coverage, but this may not be feasible with the funds to be made available. This check would have taken the form of furnishing a card to the Post Office for each address covered in the Census. The mail carrier would sort these cards as though they were to be delivered to the indicated addresses, and then identify any addresses or households known to the postal carrier but missing from the census. Tests have shown that this would be an effective device for improving census coverage. It now appears that our resources may not permit this and certain other steps designed to improve coverage.

We do expect, nevertheless, to introduce improved quality control operations in the field collection. We have not been able previously and we do not expect to be able to impose a rigorous quality standard and then inspect, identify, and reject unsatisfactory work sufficient to insure conformance with the specified standard. But we do believe we can provide enough sample inspection to identify the weaker enumerators and the weaker spots in the census, and provide for their improvement. The difficulty, of course, is that most of the census canvass is completed in a couple of weeks, and adequate inspection in a few days covering a sample of all enumerators' work would require very substantial resources. But we do believe that real progress can be made in inspection and improvement with the limited inspection and controls that can be established with available resources. It is in this respect that a Post Office check, were it feasible, would have been especially desirable because it would make effective use of a going organization that already has an extensive acquaintance with the people and where they live.

In summary, while we shall not be able to accomplish all of the improvements in method that we would regard as desirable, we are making many important advances in census methods. These should markedly increase the timeliness of census publications, and reduce costs, and we hope also will substantially improve quality. We shall undertake to evaluate the effectiveness of these and of certain alternative methods in special evaluation and experimental work to be conducted as a part of the census.

---

[1] A detailed description of the 1950 Census methods is given in The 1950 Censuses: How They Were Taken, Procedural Studies of the 1950 Censuses No. 2, Bureau of the Census.

II

PLANNING THE 1960 CENSUSES TO MEET NATIONAL AND LOCAL NEEDS

Chairman, Robert W. Burgess, Bureau of the Census

What Will the 1960 Censuses Do?—Conrad Taeuber, Bureau of the Census

New Subjects and New Emphases in the 1960 Housing Census—Wayne F. Daugherty, Bureau of the Census

What Is New In Our Eighteenth Decennial Census of the Population?—Henry S. Shryock, Jr., Bureau of the Census

Discussion—J. T. Marshall, Dominion Bureau of Statistics

WHAT WILL THE 1960 CENSUSES DO?

By: Conrad Taeuber, Assistant Director, Bureau of the Census
U. S. Department of Commerce

The results of the 1960 Censuses will provide a basis for action and analysis to a large number of persons and organizations. Many users will be able to apply new yardsticks to the measurement of their current problems. Others who are concerned with the study of national, regional or local development will find a new benchmark to line up against those of the past. To meet these, and many other requirements for statistical information, the program of decennial censuses has been expanded and adjusted—especially during the last century. Only the first two censuses of the United States were limited to little more than a count of individuals and households. By 1810, it was already necessary to add some inquiries concerning the growing industry of the new nation. After that, the program continued to expand and change as the needs of the country changed. The Eighteenth Decennial Census includes a census of population and unemployment, one of housing (including utilities and equipment) and one of agriculture, irrigation and drainage. Censuses of industry, mineral industries and business are to be taken in 1959, covering activity during 1958.

One of the key numbers to be secured in the 1960 Censuses is the population of the various jurisdictions and areas. Size is not only a matter of civic pride; it has many direct consequences. The apportionment of representation in the Congress and in the Legislatures of most States is based on the latest official census figures. Large sums of money are distributed from the Federal Government to the States and from the States to other governmental units, with the number of people as a major element in the distribution formula. Payments of $6 to $10 per person per year are not uncommon in the allocation of State funds.

Cities and States may boast that they have outdistanced a particular rival in population size, but there are far more significant consequences from the establishment of official figures. The rights and duties of municipalities are frequently dependent on the size of their population. The number of certain officials, their salary scales, the number of licenses for particular types of business, the applicability of certain laws, and even the terms on which loans can be negotiated and bonds sold are often dependent on the number of people.

Government planning for roads, schools, hospitals, and other public services will be affected by the new figures—so will business planning for production, sales, advertising, and location of plants. With the growing interest in housing, there is an increased demand for housing data, particularly for small areas within cities. The urban renewal program especially will find the new data of great value. Federal, State, and local governmental agencies concerned with agriculture will evaluate their programs on the basis of new results, and business interests concerned with the production, transportation, processing, and sale of agricultural commodities, as well as those concerned with the sale of farm equipment or supplies, will find in the new results the materials necessary for efficient planning of their activities in the years ahead.

Many research studies will use the new data on population to analyze such topics as factors in the growth of population, the declining role of the foreign born, but the more enduring effects of differences in ethnic stocks, the geographic redistribution of the population, centralization and decentralization in metropolitan areas, the northward and urbanward movement of the Negro, the continuing growth of the West, the increase of the older population, changing patterns of family life, trends in the income distribution, rising educational levels, the relative growth of service and other white collar occupations, and many others.

Two Major Improvements Desired by Users

In preparation for the 1960 Censuses, the Bureau of the Census several years ago arranged for meetings with census users in various parts of the country. The local chapters of the American Statistical Association were particularly helpful in the conduct of these meetings. Many suggestions and criticisms were made, but two particularly stood out regardless of the part of the country in which the meetings were held. These were: (a) greater timeliness, and (b) more attention to small areas. Steps have been taken to go a long way toward meeting the suggestions, but it must be recognized that when we are dealing with 48 States, 3,000 counties, and 20,000 places, some results are going to come along earlier than others.

Steps taken to assure greater timeliness include efforts to speed up the field collection of data, a greater reliance on sampling for a considerable part of the data to be collected, and the use of new and faster equipment. This includes a document reading device which does away with the need for punching and varifying cards, high speed electronic computers to edit returns for acceptability and to tabulate them, and high speed electronic printers for copy preparation. As a result, much of the published material will appear earlier than was the case in the 1950 Census; gains of 12 to 18 months in the timetable should be possible for many of the tabulations. In the case of the population census, most items will be collected from a 25-percent sample of households. Only age, sex, color, relationship to head, and marital status will be asked of every person. This should speed up the collection of data in the field and expedite the editing and coding, since some of the items which have been placed on a sample basis require manual coding and would have slowed up processing if they were on a 100-percent basis. In the case of the housing census, items which are not needed on a block basis will be collected from a 25-percent (or smaller) sample. In the Census of Agriculture steps have been taken to reduce the

amount of information collected on a 100-percent basis, in part by moving additional items to a sample basis.

Users who are interested in small area data will take particular satisfaction in the fact that the number of census tracts has virtually doubled—from about 12,000 in 1950 to 22,000 in 1960. Nearly every city of 50,000 or more will have been tracted by the time of the 1960 Census, and for most cities of 100,000 and over, the entire standard metropolitan area will have been tracted.

Although the Bureau does not publish tabulations by enumeration districts, a number of users have found data for these small administrative work units useful for analysis where tracts are too large. In about 100 cities or counties arrangements have been made with local groups to have the enumeration districts defined in such a way that they would represent useful units for analysis should tabulations on that basis be desired.

Housing statistics will again be presented by blocks for cities of 50,000 and over. Recognizing that such statistics might be useful in some places with a smaller population, the Bureau has undertaken to develop cooperative arrangements whereby block statistics can be made available for individual places which make the necessary arrangements, including payment of added costs, in advance. Some 133 communities have entered into such arrangements with the Bureau.

The provision of data for minor civil divisions of counties in the past has been both costly and difficult. The cost arises partly from the large number of such units; the difficulty from the fact that in some States these units have little stability or standing. In cooperation with State and county officials of 17 States, stable statistical areas, known as census county divisions, have been established. These will be used for whatever data are presented for subdivisions of counties in 1960. This program is an extension of one developed for the 1950 Census in cooperation with the State of Washington.

Other Urgently Needed Improvements

The staff of the Bureau has been very much aware of the need for improvements in both coverage of the census and the quality of the data to be collected. Studies initiated by the Bureau indicate that the 1950 Census may have missed approximately 3 percent of the population, that this deficit was not evenly distributed, and that, as a consequence, some inequities may have resulted. The question of quality of the data is closely related to the question of coverage. If the census tends to miss young adult nonwhite males more frequently than some other groups, this fact affects sex ratios, age composition, household statistics, occupational statistics, and all other data in which age and sex are involved. If young couples tend to be missed more frequently than the average of the population, this may also affect the degree of coverage of infants. Analysts have long speculated on the reasons why respondents apparently failed sometimes to report such significant items as the presence of a baby. Studies following the 1950 Census, indicate that much of the underenumeration of small children was accounted for by the underenumeration of their parents.

A number of steps have been taken in the expectation that they will make for improvement in quality and coverage. The work load of enumerators in rural areas has been substantially reduced by taking the Census of Agriculture separately—in the fall of 1959. Methods are being developed to secure greater participation of each individual in providing the responses. In this way it is hoped to take advantage of the growing literacy of the population and the fact that consultation among household members may produce better replies than the off-the-cuff answers of the accessible respondents. In the last two Censuses of Agriculture significant gains were made by mailing copies of questionnaires to farmers in advance of the enumeration, so that they could have the answers ready when the enumerator called. In the Population and Housing Censuses an effort will be made to acquaint the entire population with the questions that are to be asked. A form calling for entry of the 100-percent items is to be sent to every household in advance of the enumeration with the request that it be ready when the enumerator calls. Households included in the sample will also be provided with the additional questions before they will be expected to reply to them.

Procedures and questionnaires have been subject to considerable testing in the search for improvements. Efforts are being made to simplify the enumerator's task. The information on residential finance will again be collected from a relatively small sample through a survey conducted at a time different from that of the main enumeration. A number of areas that might be especially difficult for an enumerator will be identified in advance and special arrangements made to secure the information there. There has been some testing of the possibility that the knowledge which mail carriers have about their service area might be used to help in locating households which the census enumerator may have missed. The Bureau has continued its work of developing methods of checking the quality of the work during field collection and processing. Plans are being developed to institute controls which will permit correction of errors at a time when such correction can be most effective.

Continuity of Statistics

A general purpose inquiry taken once in 10 years must always come to grips with the question of the degree to which continuity of series is to be preserved as over against the fact that conditions and needs change and that apparent comparability with the past may be less useful than forthright recognition that such comparability is not possible. This question has been in the foreground during all of the planning for 1960, as concepts and items have been reviewed to determine whether or not they should be included.

Certain items which were used in past censuses are no longer useful. No one will be able to say truthfully that the 1960 Census asked for

everything, including the kitchen sink. That particular piece of equipment is now so commonly found that no purpose would be served by asking about it. Indexes of levels of living which relied on electricity, running water, or mechanical refrigeration will have to be revised, for these items are now so nearly universal that statistics concerning their presence in the home are no longer pertinent. Certain items of farm machinery have become obsolete, and the fact that the horse is a rapidly vanishing farm animal has meant that the amount of detail concerning horses can be sharply reduced. These illustrate some of the situations which have been encountered. There are, of course, many new claimants for inclusion and some of these could be accommodated.

The spread and importance of what is known as contract farming gives this item special priority. The comparison of place of work and place of residence promises significant information for students of metropolitan area problems and this item is to be included. Dr. Shryock and Mr. Daugherty will report in more detail on these and other changes that are being made.

Where techniques for improving a particular line of information are available, the value of the improvement should be weighed against the impact on comparability. In the case of age there has long been a belief that asking for date of birth would provide data of better quality than those secured when age is asked for directly. Both devices have been used in the past but in recent years considerations of mechanical equipment were important in giving preference to the direct question on age. The new tabulating equipment makes it considerably easier to handle date of birth, and that is to be used in 1960. Some improvements are to be made in the definition of quasi- and other households and of farm residence. Efforts will be made to provide a new grouping within the rural population to identify separately the people who live in population clusters and those in the "open country."

Recognition of new needs does not necessarily mean the elimination of all measures of comparability with the past. The Current Population Survey provides the means of continuing series from the past which no longer merit consideration in the full census. It also gives the possibility of testing alternative procedures and for providing a bridge between the old and the new. Thus, the Current Population Survey is now the only source of data on literacy in the United States. The CPS late in 1958 collected statistics to help interpret the effect of the census rule which counts college students at the place where they are staying when they are attending college. The CPS will provide a bridge between the old and the new concepts of farm residence. The availability of the CPS has made it possible to reduce the pressure on the census schedule and the respondent by providing a vehicle for collecting national figures for a number of items which might otherwise have been required in the census.

Post Enumeration Survey

As an operation, a national census is unique in that it is carried out only once within a rather long time period and that it is not possible to test alternative methods under conditions which simulate the big census. The census itself offers the only real opportunity for examining certain proposed improvements which may be developed for the future. Moreover, the census operation itself is limited to a very short period of time, and once the collection of data has been started it is no longer possible to introduce improvements which may have been indicated by the experience of the first days. Therefore, it is especially important to make a systematic study of the quality of the results that are obtained. Such an evaluation of the quality of the census results is considered an essential part of the census program. The Post Enumeration Survey was the largest such effort in connection with the 1950 Census and a similar effort is planned for 1960. It is an accepted obligation on the census staff to point out the limitations of the data as they affect interpretations, comparisons or other uses that may be made of the data.

Monographs

The 1950 Census issued its findings in some 107,000 pages of reports. In addition to the basic tables, there were a number of special reports relating to individual subjects. There were also a number of analytical reports which were developed in cooperation with other agencies, particularly in the field of agriculture. A number of organizations outside the government issued special reports bringing together in readily available form the census data most needed by their clientele.

In cooperation with the Social Science Research Council a series of 13 census monographs has been issued. These publications provide an opportunity for specialists in the Bureau or in other organizations to bring together census and related materials concerning a particular topic and present them with more analysis and interpretation than is normally possible in a regular census report.

It is planned to encourage such efforts for 1960. One Federal agency has already taken steps to prepare summaries of the census data which are of particular relevance to its constituent agencies. The Social Science Research Council has established a committee to work with the Bureau in the development of a new monograph program.

Census statistics are frequently reissued by private organizations, often in combination with related statistics of interest to particular groups. Despite the fact that we occasionally have to explain why we publish such statistics when they are available in these other sources, we plan to continue to encourage such use of the census results. The widest possible dissemination of census results is desirable to give the public full value from the very large investment that is represented by a Decennial Census.

NEW SUBJECTS AND NEW EMPHASES IN THE 1960 HOUSING CENSUS

By: Wayne F. Daugherty, Chief, Housing Division, Bureau of the Census
U. S. Department of Commerce

In the past, the census programs have changed as the statistical needs of the country have changed. The 1960 Housing Census program is no exception. It will be quite different from previous censuses because it will contain many new subjects and new emphases. This reflects the changing statistical needs in current housing programs, such as urban renewal, traffic management, water supply and pollution, and urban planning in general. The change in emphasis is the outgrowth of experiences gained by governmental agencies, by university research bureaus, and by analysts concerned with markets for housing and consumer goods or with factors associated with residential financing.

Considerable effort was made to learn of these changing needs and to incorporate new subjects into the program. At the same time, it was necessary to do so without increasing the over-all scope and cost over that of the 1950 Census. To accomplish this, it was necessary to make greater use of sampling, to drop items of limited usefulness, and to restrict the types of items and the amount of detailed classification. In addition, some items which were considered fundamental to current statistical requirements had to be omitted because of budgetary limitations.

Censuses of housing traditionally have provided benchmark data on characteristics of four general types: structural, occupancy and use, facilities and equipment, and financial. Structural characteristics include such items as: number of rooms in the unit, number of units in the structure, and a description of the structural condition. Occupancy and use characteristics include such items as: tenure; number of persons in unit; and, in the case of vacant units, whether the unit is for rent, for sale, or held off the market. Equipment and facilities include some items which are descriptive of the quality of housing—toilet, bathing, and heating facilities, for example—and some items which are considered descriptive of the standards of living—air conditioning, clothes washer. Financial characteristics cover rents and value. The 1960 Census will provide not only benchmark data on characteristics, but also some data on the activity that is occurring in housing.

Of the shifts in emphasis between 1950 and 1960, two stand out as being somewhat more significant than others. The first is the inclusion in 1960 of a measure and description of the gross changes in the housing inventory. The components of change part of the 1960 Housing Census will provide information on changes between the censuses of 1950 and 1960. For two previous decades, various individuals and groups, including Bureau of the Census, attempted to estimate the magnitude of these various components, without much success. The type of information to be collected will be similar to that which was collected in the Census Bureau's 1956 National Housing Inventory. In that survey, the first of its kind, data were collected on gross changes during the period 1950 to 1956. The data consisted of the measurement and the characteristics of the major types of changes, such as, new construction, conversion, merger, demolition. In addition, the survey provided some information on units which remained essentially unchanged.

The second significant shift is the inclusion in the housing inventory of all private living accommodations. This change is an attempt to close a gap in our knowledge about the way people live—particularly in areas subject to urban renewal programs. Heretofore, the census has restricted the housing universe to living quarters which met specified criteria for classification as dwelling units. Generally, a dwelling unit consisted of quarters which were occupied as separate living quarters and which had both cooking equipment and a private entrance. The 1960 Census will include all private living quarters which are separate housing units even though the unit may have only one of the two specified criteria. Although this is an oversimplification of the comparison of the definitions, it serves to indicate the expanded coverage of the living quarters universe. The types of quarters which will be added in 1960 are, for example, rooms in converted houses which are occupied by a separate family or individual but whose occupants do not have any cooking equipment or share common kitchen facilities. In 1950, the exclusion of large rooming houses from the housing inventory resulted in the loss of many quarters which actually were separate housing units.

Because the 1960 Census will identify all separate living quarters, the identification of the unit of enumeration will be changed from "dwelling unit" to "housing unit." For purposes of comparability with 1950, housing units which have both cooking equipment and separate entrance will be identified as dwelling units.

Of the individual subject items in the 1960 Housing Census, less than one-half were included as such in the 1950 Census. These are: number of units in structure, number of rooms, year structure built, tenure, contract rent, value, water supply, bathing facilities, toilet facilities, heating equipment, heating fuel, and cooking fuel.

There are six additional items carried over from 1950 but which have been changed since then because of the necessity to improve the concept to satisfy the needs of the users. These include: condition, farm residence, classification of vacant units, gross rent, radio, and television. The 1960 plans are to classify farm residence on the basis of number of acres in the property and the value of agricultural products sold. Vacant units "held for occasional use of the owner" are to be separated from the total group held off the market. Gross rent will be obtained simply by adding the cost of gas, electricity, water, and bulk fuel to the contract rent. Estimates of rent without furniture, in the case of furnished units, will not be used. With regard to radio and television, the number of sets will be reported rather than merely

their presence in the dwelling. Another change incorporated in the 1960 Census is the measure of structural condition of housing units. In 1950, all dwelling units were classified as either "not dilapidated" or "dilapidated." In the 1960 program, the concept of "not dilapidated" has been subdivided to reflect two gradations of quality—"sound" and "deteriorating."

This leaves a list of sixteen items which are being collected for the first time. They are: type of trailer, presence of basement in structure, number of bedrooms, number of bathrooms, elevator in structure, duration of vacancy, year present household moved into unit, hot water, heating fuel, source of water supply, method of sewage disposal, number of automobiles for personal use, and five equipment items—air conditioning, home food freezer, clothes washing machine, clothes dryer, and telephone. The population items used in the tabulation of housing data—number of persons in housing unit, color of head of household, and income—are unchanged from 1950.

Many of the changes which have been made, particularly those involving concepts, are the outcome of working closely with staffs of other Federal Government agencies and with advisory committees. We have, in addition, brought in experts in various fields to supplement our staff in the research and developmental stages prior to accepting a new item or making any changes in items retained from 1950.

The aim of the Bureau of the Census to include items of high priority made it imperative to reduce the basic program in other aspects. A few 1950 items were dropped—mortgage status of owner-occupied homes, kitchen sink, type of refrigeration, and electric lighting.

A few subjects which had been proposed were discarded in favor of subjects considered to have a higher priority or considered to be less costly to enumerate. One proposal, which would add to the understanding of urban problems, was to obtain data on all structures—type, floor space, number of floors, etc. Later, it became apparent that resources were not available to develop this proposal.

Another proposal was to collect detailed information on housing costs for owner-occupied properties, e.g., taxes, utilities, fuels. Still another was to enumerate the characteristics of the previous housing for households that had moved in the two years prior to the census, e.g., where the previous house was located, rent paid, or value.

Still another proposal, which was discarded because of budgetary limitations, was the recommendation to extend the block statistics program to urban fringes and to places with fewer than 50,000 inhabitants. However, for 1960, the Bureau is contracting with these cities for block statistics. The cities will pay the incremental costs to number the blocks and tabulate the detailed data.

As has been mentioned, sampling will be used more extensively in 1960 than in 1950. Only 9 of the 31 items to be collected in cities of 50,000 or more (block cities) and 3 of 33 items to be collected elsewhere will be enumerated on a 100 percent basis. In fact, the only 100 percent items are those which are needed either for the control of the enumeration process or for providing block statistics. Of the sample items, about a third are collected on a 5 percent basis. For the 5 percent items, information would be provided for large areas such as standard metropolitan areas, counties, and cities with a population of 50,000 or more. The 20 or 25 percent samples permit statistics down to cities of at least 10,000. For smaller places, most of the items collected in the 20 and 25 percent samples will be presented, but for a very limited number of categories. For vacant dwelling units, the sample will not be adequate to provide detailed characteristics for places smaller than standard metropolitan areas and for places under 30,000 only a count of vacant units by status will be provided.

Plans for the 1960 Census, as in 1950, call for detailed statistics on residential financing. On the basis of experiences in 1950 and more recently in the 1956 National Housing Inventory, data on residential financing are expected to emphasize recent financing transactions. However, less detail than in 1950 will probably be collected on rental properties. As in 1950, we expect to obtain some of the information from the owners of residential properties, supplemented by information from lenders.

As now being developed, the enumeration on residential financing and on components of change in the housing inventory may be combined. If so, statistics for the United States as a whole and for each of 25 standard metropolitan areas can be provided for both programs. The information would be collected as of December 1959. Information on mortgaged properties will, therefore, not be an integral part of the basic 1960 Census, nor was it in 1950.

The 1960 Census was planned with recognition of the importance of including items which would be of optimum value in meeting current needs. At the same time, it was considered desirable to maintain comparability with previous censuses. For example, although the basic unit of enumeration has been revised, many users and our advisory committees stressed the need to retain comparability with the 1950 definition. Where possible, this objective is being met.

In connection with the enumeration of the housing unit as the basic unit of measure, it was the recommendation that housing statistics be published for all housing units, instead of dwelling units only. In addition, a few statistics should be published separately for housing units of the type which were omitted from the 1950 Census.

Refinement of present plans, as well as limitations of publication space, may change some of our proposals. Within this framework, it is hoped that this forthcoming census will provide data which would serve the most urgent needs for housing statistics.

Characteristics	Percent coverage in		Characteristics	Percent coverage in	
	Block cities (population of 50,000 or more)	Other urban and rural places		Block cities (population of 50,000 or more)	Other urban and rural places
Structure			**Equipment and facilities**		
Condition	100	25	Water supply	100	25
Number of units in structure	20	20	Bathtub or shower	100	25
Year structure built	25	25	Toilet	100	25
Number of rooms	100	100	Number of bathrooms	20	20
Number of bedrooms	5	5	Source of water	...	20
Type of trailer	25	25	Sewage disposal	...	20
Basement in structure	20	20			
Elevator in structure	20	...	Telephone	25	25
Occupancy			Automobile use	20	5
			Heating equipment	25	25
Tenure status	100	100	House heating fuel	5	5
Vacancy status	100	100	Cooking fuel	5	5
Duration of vacancy	25	25	Water heating fuel	5	5
Duration of occupancy	25	25			
Farm residence	...	25	Television	5	5
Financial			Radio	5	5
			Clothes washing machine	5	5
Contract rent	100	25	Clothes dryer	5	5
Gross rent	25	25	Home food freezer	5	5
Value	100	25	Air conditioning	5	5

WHAT IS NEW IN OUR EIGHTEENTH DECENNIAL CENSUS OF THE POPULATION?

By: Henry S. Shryock, Jr., Bureau of the Census
U. S. Department of Commerce

As Dr. Taeuber has described in his paper, the content of the decennial census expanded very rapidly in the first century after its modest beginnings in 1790. The peak in the number of inquiries was reached about a half century ago, however, as it was gradually realized that the burden upon the enumerator was becoming too great and that, in any case, some types of information could not be satisfactorily collected by enumerators with but little training and experience. In the process of reorganizing the Federal Government's attempts to obtain many different varieties of important statistics, certain broad fields of inquiry were dropped altogether from the decennial census but they sometimes became the objectives of other forms of statistical compilation. For example, the Federal Government cooperated in the establishment of Registration Areas for births, deaths, marriages, and divorces. Very recently and after a long hiatus, the National Health Survey has been developed as a much superior method of obtaining data on morbidity and related phenomena by means of household interviews. The separation of the field work of the Census of Agriculture by six months from that of the Census of Population and Housing is an innovation that will simplify the training and duties of the field staff and should contribute to improved quality of the respective data.

Other major changes in methods of data-gathering and office processing have already been mentioned. Probably more of these are being introduced in 1960 than have ever been made in any preceding decennial census. In contrast, the innovations in the content of the population schedule are relatively few. There will be a slight net increase in the number of items to be recorded for persons in sample households; but, because of the transfer of all but a few items on personal characteristics to the sample, the average number of items recorded per person in the population will be considerably less than in 1950.

The remaining statistical items to be obtained on a 100-percent basis will be relationship to head of household, sex, race, age, and marital status. Farm or nonfarm residence, birthplace, citizenship, the labor force items, and occupation and industry—all of which were on a complete-count basis in 1950—will be transferred to the sample in 1960. This will be a 25-percent sample of households rather than a 20-percent sample and a 3 1/3-percent sample of persons as before, thus making it possible to relate sample characteristics of different members of the household. For example, the wife's fertility can be related to the husband's education, occupation, or income. Incidentally, the number of children ever borne was among the few items on the 3 1/3-percent sample in 1950.

Probably the main new topic to be added to the schedule is that of place of work, which also happens to be the topic that the public has most frequently requested us to add. Although our experience in several pretests has indicated that it is not feasible to collect as much geographic detail on place of work as is desired for some local programs, still it is expected that the resulting data should provide a major "breakthrough" in the study of commuting since for the first time a vast amount of nationally comparable data will be available for all parts of the country, not only on the prevalence of commuting and the major streams of movement but also on the characteristics of commuters as compared with those of workers who are employed nearer to their homes. The statistics should shed light on such phenomena as the journey to work and the daytime population of various kinds of areas and, moreover, should permit further reconciliation of labor force and income data obtained from household surveys with those obtained from establishment reports.

Specifically, persons currently employed will be asked to give the county of employment and also the city, if they work in a city. Furthermore, they will be asked to give the principal means of transportation used in getting to work.

Plans for tabulations from these data are still being developed and have not yet been discussed with our advisory committees. Tentatively, we have in mind something like the following:

(1) For each county of residence, the employed workers will be classified as working in the given county, as working in other counties, or as not having a fixed county of work. The feasibility of tabulating the distribution of total workers employed in each county according to whether or not they also live there is still to be investigated.

(2) Since large-scale and long-distance commuting is essentially a metropolitan phenomenon, tabulations in greater geographic detail will probably be restricted to metropolitan areas and their environs. Within each standard metropolitan area, the central city (or cities), the balance of the county containing that city, and individual counties within the so-called metropolitan ring would be distinguished. As a long-needed aid to the Federal Government's program of delineating standard metropolitan areas, the individual counties in the next tier beyond present boundaries would also be included in this tabulation. These counties and central cities would then be used as units in the cross-classification of place of work with place of residence.

There are already indications that some local agencies would like to obtain greater detail on commuting by place of residence. This could, of course, be made available if there are sufficient resources. In fact, perhaps a supplementary tabulation in the form of a standard package could be worked out that could be purchased by interested local agencies. If this were contracted for early enough, it could be planned so as to reduce the costs and to make it available fairly promptly after the regular tabulation program is concluded.

(3) The previously mentioned cross-classifications of commuters in terms of such characteristics as size of place of residence, occupation, industry, and class of worker, and part-time or full-time status, are of considerable analytical interest. This kind of interest might be adequately served by tabulating a subsample of cases in terms of rather gross geographic categories. In so far as this sort of data is wanted for specific areas rather than for the more conceptualized types of areas, the special-package contract tabulation just described might be an appropriate vehicle. It must be appreciated that even the restricted program of regular tabulations that I have just proposed will generate an enormous amount of data for publication.

The results of two decennial censuses make it clear that the labor force questions cannot be used there as worded in the Current Population Survey but must be simplified. The general objective, however, will still be to determine labor force status in the preceding week. For those who have worked at any time during the past 10 years, a description of the current job, or of the last job, will be obtained. This description will consist not only of the occupation, industry, and class of worker, but also of the name of the company, business, organization, or other employer. The purpose of this last item is to improve the coding of industry by making it possible to use registers that will be prepared in advance of the enumeration. Moreover, it is hoped that the economic items, in general, will be among the major beneficiaries of the prospective greater use of self-enumeration since workers who are normally absent at the time of an interview would have an opportunity to supply information about their jobs, incomes, etc. A major change in the method of collecting data does run the risk of making a noticeable impact on historical comparability, of course.

Obtaining age from date of birth seems to be a procedure that is particularly well-suited to self-enumeration and should lead to greater accuracy in the reporting of age. From a question on age at last birthday, "heaping" on favored terminal digits still persists somewhat, particularly among older people, and there are directional biases for some of the functional age groups that have created problems in the use of age data.

As part of the inquiry on population mobility each person will be asked to give his place of residence five years ago. This reversion to the five-year period of the 1940 Census from the one-year period of the 1950 Census was dictated by a desire to avoid a short and unusual period of time and by a desire to secure a larger number of migrants for analysis. The 1950 Census material on migration also provided but little information on suburbanization and other forms of population movement within metropolitan areas. The 1960 questions will aim to identify movers between central cities and outlying parts of the metropolitan area, and some of the tabulations will deal particularly with this widespread contemporary form of population redistribution. Additional material on population mobility will be obtainable from a question on the year in which each person moved into his present home.

Continued interest in our mounting educational problems is recognized by the obtaining of some additional facts in this field: (1) persons attending school will be asked whether they are attending a public or a private school, (2) the number of persons who have completed at least a year of graduate work will be ascertained, and (3) the attendance question will be extended to persons 30 to 34 years old.

There has been growing concern with the measurement of the farm population by means of the simple opinion question, "Is this house on a farm (or ranch)?" Without the support of objective criteria as to what is meant by a farm, this question leads to results that are not of very high reliability. The opinions even of persons within the same household may differ here. It was probably originally assumed that a farm was a simple entity, the existence of which is obvious to any reasonably intelligent observer. Accelerating trends in agriculture and in commuting from open-country residences have made this assumption much less valid today than it was several decades ago. Our studies have shown that perhaps a quarter of the households living on so-called "farms" have little or no connection with agriculture. The current proposal is to substitute questions on the acreage and value of agricultural products concerning the tract of land on which each rural household lives. Farm residents would then become those households living on (1) tracts of 10 acres or more from which agricultural products to the value of $50 or more were sold last year and (2) tracts of lesser acreage from which the products sold had a value of $250 or more. This definition would correspond very closely to the definition of a farm to be used in the 1959 Census of Agriculture. Despite the six-months' gap between censuses, it may still be possible to prepare a matched subsample of households, homes, and farms from the three censuses so that the kind of tabulations made after 1950, cross-classifying the data of these three censuses, can be run.

A new kind of collation, not restricted to the farm population, is also being considered in the form of a person-household tape. Here the characteristics of a sample of persons would be related to those of the households and dwelling units in which they live. Thus, it would be possible to examine the living arrangements of the aged, the types of homes in which children are being raised, and the relationships between labor-force participation of mothers of young children, income of husbands, and the presence of nonworking adults and older children in the home, to name a few examples.

The electronic computer will also make it much cheaper than before to produce derived figures such as percentage distributions, ratios, means, and medians, which formerly had to be obtained by clerks using desk calculators. Therefore, it is expected that the 1960 publications will contain many more such figures, and hence the users of Census reports should be saved a great deal of time and greater use should be possible on the part of persons and agencies with limited time and means. Mechanical editing within the computer will permit more effective handling of unknowns and inconsistencies while, at the same time, a

record is kept of the number of allocations and edits made on each subject item.

Dr. Taeuber has already mentioned some of the forthcoming innovations in the field of statistical areas. Let me supplement his account by mentioning two or three others. Since some limitations of the farm residence concept have become apparent and since the new questions on this topic will be asked only of the sample, attention has been directed to supplementary ways of subdividing the rural population. In 1950, unincorporated villages of 1,000 inhabitants or more were first identified and given separate treatment. Under consideration is an effort to extend this program to smaller unincorporated places so that the population in villages and hamlets can be separated from a residual that might be called the "open country" population. The size and characteristics of these two major segments of the rural population could then be presented by States, counties, and economic areas. Such data should be of particular interest to rural sociologists and others who have been concerned with the heterogeneity of the rural-nonfarm population.

A different approach to reducing this heterogeneity was begun in the 1950 Census when urbanized areas were first delineated. The closely built-up suburban fringes of these areas were transferred from rural to urban territory, thus resulting in a major improvement in our urban-rural classification. In 1960, methods of bounding urbanized areas will be simplified somewhat, and an up-to-the-moment picture of the extent of suburban development around all cities of 50,000 or more will be available. Unfortunately, it does not now seem feasible to extend the delineation of urbanized areas to smaller cities.

One of the most prominent trends in municipal development during the 1950's has been the greatly increased number of annexations to incorporated places. No account of urbanization during the period could be complete that overlooks this process, and measurement of the population involved seems to be called for. We propose to publish, for every urban place that has had one or more annexations during the decade, the 1960 population in the total annexed area. Such statistics should be particularly useful to the many persons who estimate the intercensal net migration for individual cities and for urban and rural areas. Annexation has been an uncontrolled factor in the study of the components of population change to which most analysts have been able to give merely qualitative recognition.

Finally, this will be the first census in half a century when we shall be faced with the problem of how to handle an addition to the area covered by the States. The statistical integration of Alaska will be accomplished as far as the 1960 data are concerned; and probably to a limited extent, some alternative historical revisions can be presented. We hope that a question mark will not be poised over the status of Hawaii at the time we are having to conclude our publication plans.

INQUIRIES FOR 1960 CENSUS OF POPULATION

[NOTE: Inquiries 1 - 7 will be made of all persons

Inquiries 8 - 16 will be made of all persons (as applicable) in every fourth household

Inquiries 17 - 25 will be made only of persons 14 years old or older in every fourth household]

1. Name

   Last name, first name, middle initial

2. Address

   Street, avenue, or road
   House (and apartment) number

3. Relationship to head of household

   What is the relationship of this person to the head of this unit?

   Head
   Wife of head
   Child of head
   Other relative
   Non relative
   Inmate

4. Sex

   Male
   Female

5. Color or race

   White
   Negro
   American Indian
   Chinese
   Japanese
   Filipino
   Other—give specific color or race

6. Month and year of birth

   What is the month and year of birth?

7. Marital status

   Is he now:

   Married? Separated?
   Widowed? Never married?
   Divorced?

8. Farm Residence

   a. Are there 10 or more acres in this place?
   b. (If less than 10 acres) Did agricultural products sold have a value of $250 or more?
   c. (If 10 acres or more) Did agricultural products sold have a value of $50 or more?

9. Place of birth

   In what State or foreign country (or territory or possession) was he born?

10. Citizenship

    If born in foreign country is he a citizen of the United States?

    Naturalized citizen?
    Alien?
    Born abroad of American parents?

11. Country of birth of parents

    a. Where was his father born?

       U. S. or
       Write name of foreign country, or U. S. territory or possession.

    b. Where was his mother born?

       U. S. or
       Write name of foreign country, or U. S. territory or possession.

12. Length of residence and migration

a. In what year did he (last) move into this house (or apartment)?

1959 or 1960
1958
1955 to 1957
1950 to 1954
1949 or before

b. Where did he live on April 1, 1955?

Same house (or apartment)
or
Enter city, county, and State or foreign country: and specify whether "in city limits" or "not in city limits"

13. Educational attainment

a. What is the highest grade (or year) of regular school he has ever attended?

Never attended school
Kindergarten
Elementary school—1, 2, 3, 4, 5, 6, 7, 8 years
High school—1, 2, 3, 4 years
College—1, 2, 3, 4, 5, 6 or more years

b. Did he finish this grade (or year)?

14. School attendance

a. Has he attended regular school at any time since February 1, 1960?

Yes, regular school
No (did not attend school or attended special school only)

Born before April 1925 (35 years old or older)

b. If he attended school since February 1—is it a public school or private school?

Public school
Private school

15. Times married and date of marriage

If the person has ever been married—Has he been married more than once?

Yes: When did he get married for the first time?

No: When did he get married?

16. Fertility

If this is a woman who has ever been married—How many babies has she ever had (not counting stillbirths)?

None, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more

17. Employment status and hours worked

a. Did he work at any time last week? (Include part-time work such as a Saturday job, helping on a farm, or delivering papers. Do not count own housework)

b. If Yes in "a"—How many hours did he work last week?

1-14 hours
15-29 hours
30-34 hours
35-39 hours
40 hours
41-48 hours
48-59 hours
60 hours or more

c. If No in "a"—Was he looking for work or on layoff from a job?

d. If No in "c"—Even though he did not work last week, does he now have a job he usually works at?

18. Date last worked

When did he last work at all, even for a few days?

1960
1959
1955-1958
1950-1954
Before 1950
Never worked

19. Occupation, industry and class of worker

If he worked in 1950 or after: Describe job or business held last week, if any, and give name of employer. If no job or business last week, give the information for last job or business.

a. For whom did he work?

Write name of company, business, organization, or other employer or
If on active duty with Armed Forces, skip parts b, c, and d

b. What kind of business or industry was this?

(Examples: poultry hatchery, county junior high school, auto assembly plant, radio and TV service, retail supermarket, highway construction)

c. What kind of work was he doing?

(Examples: truck driver, 8th grade English teacher, paint sprayer, repairs TV sets, grocery checker, civil engineer)

d. Class of worker?

For private company, business or individual for wages or salary
With a Federal, State, or local government
In his own business, professional practice or farm
Without pay in a family business or farm

20. Place of work

If he worked last week: What city or county did he work in last week?

Did not work
No fixed place
Worked at: City, county, and State: and specify whether "in city limits" or "not in city limits"

21. Means of transportation

If he worked last week—How did he get to work? (Mark principal means of transportation used last week)

Railroad
Subway, elevated
Bus, streetcar
Taxicab
Private automobile
Walk only
Worked at home
Other means, specify

22. Weeks worked

a. Last year (1959) did he work at all, even for a few days?

b. If yes—How many weeks did he work in 1959, either full-time or part-time? (Include paid vacation, paid sick leave, and military service)

13 weeks or less
14-26 weeks
27-39 weeks
40-47 weeks
48-49 weeks
50-52 weeks

23. Earnings in 1959

If worked in 1959—

a. How much did he earn in 1959 in wages, salary, commission, or tips from all jobs? (Before taxes, bond deductions, etc.)

Enter amount or none

b. How much did he earn in 1959 working in his own business, professional practice, partnership, or farm? (Net income after business expenses)

Enter amount or none

24. Other income in 1959

Last year (1959) did he receive any income from:

Social security, pensions or veteran's payments:

Rent, interest or dividends:

Unemployment insurance or welfare payments:

Any other source not already reported

Enter amount or none

25. Veteran status

If a male—Did he ever serve in the Army, Navy, or other Armed Forces of the United States during-

a. Korean War (June 1950 to January 1955)

b. World War II (September 1940 to July 1947)

c. World War I (April 1917 to November 1918)

d. Any other time, including present service

USCOMM—DC

Panel on

Planning the 1960 Censuses to Meet National and Local Needs

Discussion by J. T. Marshall - Canada

I am sure, Mr. Chairman, we have all listened with interest and envy to the many challenging and stimulating plans which our colleagues from Suitland have presented for our instruction this morning. Conrad Taeuber has given us a broad and timely insight into the proposals for the 1960 Population Census of the United States which introduces many new and intriguing approaches in the application of the enumeration, tabulation and publication of the Census results. Wayne Daugherty has outlined to us the plans for Housing and here, too, we find new approaches being brought into play. Also Henry Shryock has sketched the new items that will be available in your 18th Census in so far as it measures the human resources. I am sure you will agree, Ladies and Gentlemen, that they have given us a very comprehensive picture of what would happen in the Census, the sort of things we may expect out of it and, in fact, have covered the "Census Waterfront" very thoroughly indeed.

The plan for Agriculture sets out, at least to the Canadian census takers, one of the most challenging departures in the census proposals. The decentralization of the central collection, tabulation and analysis processes to the Town of Parsons in the State of Kansas will, I am sure, produce many new problems and stimulate many new techniques for meeting and solving them. I am sure everyone here wishes the Bureau of the Census a full measure of success in these undertakings and that their reward will be in keeping with all the hard work that has gone into their efforts.

We Canadians face many of the same problems confronting our United States friends, though on a smaller scale. While the scale of our operation is far smaller, with a population of 18-19 million expected in 1961 compared to the 190 million for the United States, we have the added problem of a small population scattered over a large area. Thus, although we are aware of the advantages of separate enumeration for each census, we feel that we must live for a time yet with a simultaneous enumeration for our four censi - Agriculture - Housing - Merchandising and Services and Population. Our decision is as much due to physical, as it is financial considerations. It is rather obvious, Mr. Chairman, that large or small, so many of the problems which underlie these proposals are of such great interest to all census takers - and to those who use the product of their efforts - it is hard to know where to start this discussion. It has been suggested that I tell you something about the Canadian proposals in so far as the technicians have developed them, but I must point out that the Cabinet has not yet given its approval to any census proposal. One thing is almost certain and that is because of the constitutional requirement, Census Day in Canada will be June 1, 1961.

Of special interest to me personally is the question of timeliness. Our speakers have emphasized the need for making census material available as quickly as possible and have pointed out how they propose to use sampling procedures and mechanical equipment to reduce production times for the results by 12 to 18 months. We in Canada also propose to reduce our enumeration time in 1961 by utilizing sampling techniques. In our population census, for instance, there will be some 34 questions on the 1961 schedules of which 12 will be taken on a sample basis. This compares to a total of 30 questions asked in 1951, all of which were on a hundred per cent basis. It is hoped that such devices will be of use in relieving pressure on the enumerators and assuring a speedier return of the material from the field.

From the point of view of saving time by the use of electronic equipment, we were a little bit proud of ourselves because in this respect we accomplished a good deal in 1951. At that time we introduced the "mark sense" reading devices and cut sharply from the several release dates - two to four years. As we all realize - one of the "devils" of volume statistics is the long delay in the preparation of material for processing by high speed computers and tabulators. The trick of getting the material into the machines is the "defile" through which all our armies of data must pass. By the use of electronic "mark sense" readers we were able - in 1951 and 1956 - to cut short this operation and the resulting increase in timeliness over the previous Census releases was astonishing. We trust that the U.S. Bureau of the Census will achieve even greater results from their planning for 1959 and 1960 than we did with ours nearly a decade ago, because even the gains of 1950-51 - over 1940-41, which we still hope to maintain, - were not enough to meet the demands for more and more complicated analyses and faster release of the results.

As a further step along the "census road of romance" - that Bill Madow referred to at the session on Saturday - in respect to machines we expect to use an electronic computer in the 1961 Census of Canada, but we know full well that further time reductions comparable to those of 1951, cannot be expected again. The added saving in time will be largely a problem of getting the material into the computer faster and new methods of achieving this - we are now confident - will be worked out in time for the Canadian test in June, 1959. The advantages of using "brains" or "memories" in metallic form for processing would seem to be cheaper operations and the possible improvements in quality and variety of the data, as so ably pointed out by Mr. Shryock. Of course, in the use of electronic computers, the Bureau of the Census people are the "Grandaddies" of all census takers - so once again we are looking forward to

the valuable lessons that we shall learn - in computer application - from their efforts and their experiences.

We are all, I am sure, exceedingly interested in the heroic efforts being planned by all countries taking censuses to improve coverage and quality. I am afraid some of us may feel that these efforts will prove somewhat costly, and in Canada we have had to ask ourselves "whether to us the gain will compensate for the increase in cost". In our case the 1951 Census of Population and Housing was checked by the independent and continuing monthly sample, and though we may be "kidding" ourselves, the enumeration in terms of coverage did not seem to be inaccurate by more than 1.5% when the checks were completed. While we may have been more lucky than accurage, I am reminded that in the "days of my youth in census learning" it was Virgil Reed who on my first visit to Washington - in the early thirties - impressed all who "sat at his feet" with "it matters not too much that a census misses this 2 per cent of the population but it is important that the loss can be reasonably measured - sometimes the cost of picking up the remaining 2 per cent is more than the actual counts are worth". Thus, while we cannot afford to incorporate many of the proposals we would like to in our planning, we do hope to achieve in the Canadian plans an even better coverage than we had in 1951 and I expect those plans will not be too divergent from those we have heard outlined at this meeting.

In the agricultural field the pressures to produce additional statistics have been acute in our country, as I guess they have been in yours from the remarks of Mr. Taeuber. We have been asked, by private sources and by Government Departments engaged in administration of agricultural and forest products programmes, to supply additional information on a variety of items including farm mechanization, contract farming and farm classification. To some extent, however, we have been able to resist pressure towards greatly enlarging our agriculture census, largely due to the fact that under the Canadian Statistical system we prepare a co-ordinated system of annual and monthly agriculture statistics. Thus, while many people would like us to ask a great number of questions about the machinery on farms, we are able to supply up-to-date and continuing production figures on farm machinery from other sources. From these figures we can prepare reasonably accurate estimates of machinery stocks and therefore feel quite justified in holding down the number of questions on our census schedules. At this time we are taking a very extensive survey of farm revenue and expenditure which will yield a great deal of information on farm household activities, thus eliminating the need for extensive census enquiries along this line.

By sharpening up their definitions of a farm, our census "ag" boys hope to improve the coverage and consequently the quality of their statistics. Recognizing that continuity must be maintained, they plan to ask about 225 questions on agriculture compared to 338 for 1951 and 124 in 1956. One of the real areas of difficulty has been caused by the pressure from all the forestry agencies who are interested in "forestry products taken off farms". But this appears to be a pretty genuine pressure in our country at least, because these are data which cannot be collected and measured successfully by means other than a census.

In the 1961 Canadian Census of Housing, we are planning a number of changes and it is proposed to drop about nine questions which were carried in 1951, to make room for a slightly larger group of "new" questions. Of the so-called "new" questions, several were carried prior to 1951, but for various reasons were dropped at that time. Demand has been such as to warrant their inclusion again. However, the overall size of the 1961 Housing Census will be little changed to that of 1951. The chief change will be in emphasis - the proposed inquiries in 1961 being weighted more heavily toward items directly related to the dwelling (such as age of dwelling, value, and number of bedrooms) at the expense of items relating more to the household's standard of living (such as washing machines, vacuum cleaners, telephones and radios). Statistics relating to these latter are already available from the annual "Household Facilities" survey.

While we in Canada have been afforded an intimate insight into all the planning and discussion that has gone on - we must say that we are grateful that in 1960 the U.S. will once again provide a full-scale test for what DBS has to do in 1961 and that once again we may, by close observation, benefit from the successes they will achieve in their census operations. We shall approach our problems with the satisfaction that during the next two or three years it will be the privilege of both Bureaux to live together in the stress and in the strain, as well as in the accomplishment, because since the Canada-United States joint census committees were established in June, 1953, one of the good things that has come out of all this has been the mutual trust and understanding that has prevailed in planning - together - the censuses for both countries.

III

SMOKING AND CANCER

Chairman, Thomas Parran, Avalon Foundation

THE MORTALITY OF SMOKERS AND NONSMOKERS 1,2/

By: Harold F. Dorn, National Institutes of Health

The question of the effect of the use of tobacco upon health has been vigorously debated since tobacco was introduced into Europe approximately four centuries ago. A member of the royal house of England, King James I, who obviously was opposed to the use of tobacco, wrote a pamphlet, "A Counter Blaste to Tobacco," published anonymously in 1604. Typical of the more extreme assertions concerning the harmful effects of tobacco is the following quotation from "Doctor of Physick in Bathe" published by a Dr. Venner in 1650:

> "I will summarily rehearse the hurts that Tobacco inferreth, if it be used contrary to the order and way I have set downe. It drieth the brain, dimmeth the sight, vitiateth the smell, dulleth and dejecteth both the appetite and stomach, destroyeth the decoction, disturbeth the humours and spirits, corrupteth the breath, induceth a trembling of the limbs, exsiceateth the windpipe, lungs and liver, annoyeth the milt, scorcheth the heart ..."

No substantial quantitative evidence was produced at that time to support this assertion nor indeed was any produced for nearly 300 years thereafter. Although numerous studies of the acute effects of nicotine subsequently were made, investigations of the possible relationship between the onset of specific diseases and the long continued use of tobacco date mainly from the past 30 years, the majority having been published during the past decade. References to the probable relationship between the use of tobacco and certain diseases can be found in the medical literature of the past century but these were based almost entirely upon clinical impressions and the experience of practicing physicians.

A rapid increase in the number of diagnosed cases of lung cancer began to attract widespread medical attention during the decade of the 1920's. At first, many believed that the observed increase in the number of diagnosed cases was due almost entirely to more accurate methods of diagnosis and more effective case finding. This opinion has gradually changed; nearly all who have carefully examined the evidence now agree that the observed increase cannot be accounted for solely by more thorough case finding and improved methods of diagnosis. However, existing data are insufficient to permit a reliable estimate of the proportion of the increase that may be attributed to an increase in the risk of developing lung cancer.

The dust from macadam roads, air pollution from the combustion of coal and petroleum products, special occupational hazards, and the use of tobacco, especially cigarette smoking, are the principal agents that have been mentioned as possible causes of the increase in lung cancer. Most of the early investigators selected the use of tobacco as the first to be studied. By the end of 1950, eight retrospective studies of the use of tobacco by patients with lung cancer and by persons without lung cancer had been published, four in the United States and four in Europe.

Plan of Study

The present study was planned in the summer of 1952 as a prospective investigation of a defined population with the general objective of studying the relationship of the use of tobacco, residence, and occupation to mortality.

With the cooperation of the Veterans Administration, policyholders of U.S. Government Life Insurance were selected for study. This insurance was available to persons who served in the armed forces of the United States from 1917 to 1940. Most of the policyholders were veterans of World War I; the remainder first served after that date. Over 99 percent of the policyholders were men.

All persons with an active policy at the end of 1953 were included except for a few special groups such as persons with total and permanent disability. Beginning in January 1954 a questionnaire requesting information concerning the use of tobacco, usual occupation and industry was mailed to 291,800 policyholders. Usable replies were received from 198,926 persons or 68 percent of those included. A second questionnaire was mailed to the nonrespondents beginning in January 1957. Usable replies were received from an additional 50,000 policyholders, making a total of 249,000 or 85 percent of those included in the study for whom information concerning the use of tobacco, occupation, industry, and residence was available. The nonrespondents have been kept in the study and the same medical information is available for them as for the respondents.

Whenever a claim is filed for the payment of a policy, a copy of the death notice, usually a copy of the official death certificate, is sent to the Public Health Service. Additional medical information including verification of the causes of death entered on the death certificate, the procedures used to establish these diagnoses, whether the deceased had cancer even though it was not considered to be an underlying or contributory cause of death, and the histological type of cancer, is requested from the physician who signed the death certificate or

---

1/ This study would not have been possible without the cooperation of the Veterans Administration. However, the Veterans Administration is not responsible for, nor does it necessarily endorse, any of the findings or conclusions of this report.

2/ This study was carried out in cooperation with Field Investigations and Demonstrations Branch, National Cancer Institute, Dr. R. F. Kaiser, Chief. Dr. W. S. Baum, Division of Indian Health, Public Health Service, assisted in the planning and initiation of the study.

from the hospital where the death occurred.

Verification of the cause of death is not requested if the death occurs outside the United States, is due to an accident, or is certified by a coroner. Replies have been received to more than 99 percent of the letters of inquiry.

Basis of this Report

This report is based on the mortality experience during the two and one half year period, July 1954-December 1956, of policyholders for whom a tobacco-use history was obtained prior to July 1954. The experience during the first six months of 1954 has been excluded since the original questionnaire was mailed during this period. Persons who died before they received the questionnaire obviously had to be assigned to the nonrespondent group. It is reasonable to suppose that a larger proportion of persons who were seriously ill at the time the questionnaire was mailed failed to reply than actually did reply. This also tended to increase the death rate of the nonrespondent group.

As a result, the death rate per 1,000 of the nonrespondents during January-June 1954 was 23.2 per annum and during July-December 1954 it was 17.9 per annum. During 1955 and 1956 the annual death rate of this group was 18.3 and 18.8 respectively. Although the death rate of the entire group of policyholders was not affected, the experience of the first six months of 1954 has not been used in this report since the mortality rates of the nonrespondents were higher than normal while those of the respondents were lower than normal during this period.

The following data are based upon 478,952 person years of exposure, of which 89,774 were contributed by persons who had never smoked and 389,178 by persons who had smoked tobacco during their lifetime but not necessarily during 1954 to 1956. The number of person years exposure of persons with different types of smoking history is shown in table 1.

Characteristics of the Policyholders

Almost all of the policyholders were white males. Less than one half of one percent were females and only a negligible number of the remainder were nonwhite males. In age, they ranged from 30 years to more than 80 years. The vast majority--84 percent--were between 50 and 70 years of age; only two percent were more than 70 years old.

Eighty-two percent of the policyholders were white collar or skilled workers, 7 percent were semi-skilled or unskilled workers, and 6 percent were farmers or farm laborers. Roughly comparable percentages for the white male population aged 20 to 54 years with work experience from the 1950 Census of Population are 50, 35, and 13 respectively (2). It is clear that compared to the white male working population of the United States in 1950 a much smaller proportion of the persons in this study were semi-skilled and unskilled workers. This is not surprising since most of the policyholders were persons who had held whole life or endowment insurance policies for many years.

Since the policyholders were mainly from the middle and upper socio-economic classes, their death rate from all causes during the study period was less than that for the white male population of the United States. Their average death rate for the three years, 1954-1956, was about 70 percent of that for the total white male population in 1955.

Smoking History

No attempt was made to obtain a precise estimate of the amount of tobacco used. Instead, each person was asked to place himself in one of six categories by amount of each form of tobacco used--cigarette, cigar, pipe, and chewing and snuff--both with respect to current use and maximum past use. The categories were comparable to those used by Hammond and Horn in the study carried out by the American Cancer Society and to those used by the National Cancer Institute and the Bureau of the Census in their study of the smoking habits of the population of the United States (3,4).

Finkner and associates of the Institute of Statistics, North Carolina State College investigated alternative methods of measuring the current daily number of cigarettes smoked by an individual (5). Among the methods compared were a lighter-counter, the questionnaire used in the National Cancer Institute-Bureau of the Census survey, and a questionnaire developed at the Institute of Statistics. With respect to the classification of individuals by categories of the average daily number of cigarettes smoked, they found that each of the two questionnaires gave results that differed from those obtained by using the lighter-counter. There was no way of determining which of the results was closest to the true classification. However the differences in classification were not sufficiently great to support the belief that the results reported here would have been materially changed if one of the alternative methods of classifying individuals by the average daily amount smoked had been used.

Another way of measuring the reliability of the replies to a smoking history questionnaire is to request a number of persons to complete the same questionnaire on two occasions sufficiently far apart to reduce the effect of memory carry-over to a negligible level. This was inadvertently done at the time of the original mailing in 1954.

Some policyholders held more than one policy. The addressograph plates were arranged by policy number so that duplicate policies could not be readily identified when the questionnaires were addressed. At the time of the original mailing which extended over a period of nearly four months, instructions were given to identify and destroy the questionnaires for duplicate policies prior to mailing. By error, some duplicate questionnaires were mailed and completed copies were returned by 1714 policyholders.

The percentage agreement in the replies to selected items on the pairs of questionnaires are shown in table 2. Although some individuals wrote "duplicate" on one of the questionnaires, others gave no indication of sequence so that it

was necessary to collate the punch cards for all coded questionnaires in order to identify the duplicates. The two questionnaires of each pair were coded independently.

About 90 percent of the 1714 persons were classified identically on the basis of replies to two questionnaires with respect to the amount of tobacco currently smoked, the number of years the current amount had been smoked, age started smoking each form of tobacco, and a summary grouping based on a comparison of current amount smoked and the maximum amount smoked. The agreement of replies to the question on the total number of years of use of tobacco in any form was 89.7 percent based on a classification into 11 groups.

No evidence of systematic differences in the replies to the two questionnaires was found except for the question on the current amount smoked. The matching of punched cards for duplicate questionnaires was not done until after the tabulations used in this report had been completed. One of each pair of punched cards including those for questionnaires marked "duplicate" was excluded from the tabulation.

When the two sets of punched cards were matched, the number of policyholders classified as never having used tobacco was slightly larger based on the set of punched cards used in tabulation than it was based on the duplicate cards. The percentage differences were 3 for pipe, 6 for cigar, and 9 for cigarette. In other words these data suggest that some of the persons included in the group of nonsmokers probably had smoked, at least occasionally, during part of their lifetime. No information is available to indicate whether those who completed the duplicate questionnaires were a representative sample, with respect to smoking history, of the entire group who replied. If they were, the above difference would have the effect of probably decreasing the mortality ratio so that the true difference in death rates between smokers and nonsmokers would be slightly greater than that shown.

Table 3 presents a comparison by smoking history of USGLI policyholders and the male population of the United States. The lifetime smoking histories of the two groups are very similar with respect to the percentages who have never smoked, smoked occasionally, and smoked regularly. The proportion reporting the use of multiple forms of tobacco is definitely greater among the policyholders. This is to be expected since each policyholder reported for himself whereas the smoking history of 42 percent of the men included in the National Cancer Institute-Bureau of the Census survey was obtained from another member of the family. Twenty-nine percent of the self-respondents reported that they had never smoked; the corresponding percentage for the total sample was 32. A self-respondent probably is more likely to report the use of a second form of tobacco, which could account for the difference in the relative number of users of two or more forms of tobacco in the two populations. Although there are differences in detail, the lifetime smoking histories of the two populations, taken as a whole, are very similar.

Measures of Mortality

The primary purpose of this study is to determine whether the death rate of persons who have used tobacco is greater or less than that for persons who have never used tobacco. A direct way of measuring the difference in mortality is to compute the ratio of the death rate of smokers to that of nonsmokers. In practice, this mortality ratio is calculated by dividing the number of observed deaths for any group of smokers by the number of expected deaths computed on the assumption that the death rate of that group is the same as the rate for those who have never smoked. The ratio of 1.32 shown in table 1 for persons who have used tobacco means that the death rate for these persons is 32 percent greater than the rate for persons who have never smoked.

The mortality ratio as used here is a measure of the relative excess mortality of smokers compared to nonsmokers. It provides no basis for judging the absolute magnitude of the numbers of deaths on which it is based. An alternative measure that reflects the absolute numerical difference in mortality is the excess number of deaths computed as the difference between the observed and the expected number of deaths. These two measures emphasize different aspects of the variation in mortality between two groups. In the following discussion the mortality ratio will be used when it is desired to call attention to relative differences in mortality while the excess number of deaths will be used to indicate the relative rank of any cause of death in comparison to other causes of death.

Mortality by Smoking History

The death rate of persons who have smoked regularly is 36 percent greater than that for persons who have never smoked (table 1). The largest increase in mortality--58 percent--is for those who have regularly smoked only cigarettes. (fig. 1). Persons who have smoked regularly only cigars, a pipe, or both die at a rate only slightly greater than that of nonsmokers. Individuals with a history of regularly smoking cigars and/or a pipe in addition to cigarettes have a mortality rate 29 percent greater than that for nonsmokers. Their excess mortality is about one half that of persons who have smoked only cigarettes. Occasional smoking, irrespective of the form of tobacco used, does not increase the total death rate.

The higher mortality of regular smokers exists throughout the entire range of age--30 years and over--for which data are available (fig. 2). For all classes of smokers except those who have smoked only cigarettes, the excess mortality either decreases or disappears altogether after age 70 but this cannot be regarded as firmly established due to the small number of persons who were more than 70 years old at the start of the study (table 4).

Mortality by Amount Smoked

Persons who had ever smoked were requested to indicate which of five classes of amounts smoked most accurately described their smoking

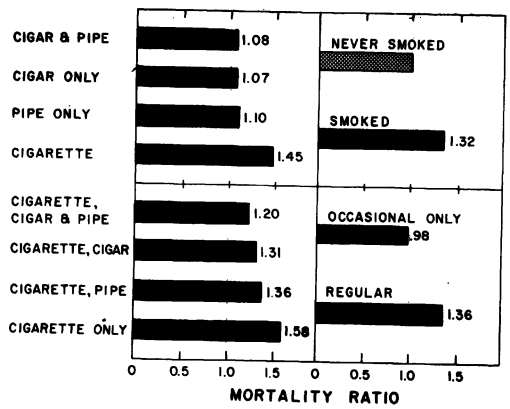


Figure 1. Mortality of smokers and nonsmokers; ratio of observed to expected number of deaths; death rate of nonsmokers = 1.00.

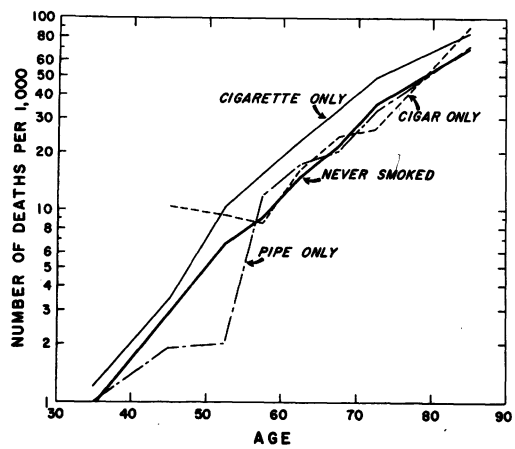


Figure 2. Mortality of regular smokers and of nonsmokers by age and type of tobacco used.

habits with respect to each type of tobacco (a) at the time the questionnaire was filled out during the spring of 1954, and (b) when their use of tobacco was a maximum. Table 5 presents mortality ratios for persons currently smoking specified quantities of tobacco in 1954. All of these persons had smoked regularly at some time, even those who were smoking only occasionally at the time they completed the questionnaire.

The excess mortality of cigarette smokers is directly related to the average daily number of cigarettes smoked (fig. 3). Those who smoke two

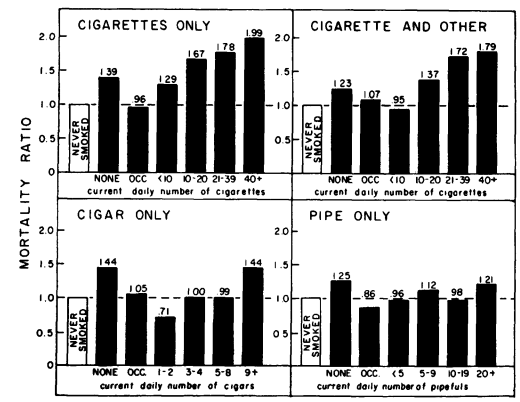


Figure 3. Mortality of regular smokers by current amount smoked in 1954 and type of tobacco; ratio of observed to expected number of deaths; death rate of nonsmokers = 1.00.

packs or more per day have a death rate nearly twice that for nonsmokers. The data in table 5 suggest that the higher mortality ratio for cigarette-only-smokers than for cigarette-and-other smokers is due in part to the fact that the latter smoke fewer cigarettes. After adjustment for differences in age and number of cigarettes smoked, the death rate for cigarette-only-smokers is 14 percent greater than the rate for persons who smoke cigars and/or a pipe as well as cigarettes. Without adjustment for amount smoked the relative difference in death rates is 20 percent.

Only very heavy cigar or pipe smokers experience a higher mortality than nonsmokers. The death rates for the heaviest cigar and pipe smokers, those currently smoking 9 or more cigars or 20 or more pipefuls per day, are slightly less than the rate for persons who smoke from one half to one pack of cigarettes per day. The mortality ratio is significantly greater than one only for persons who regularly smoke cigars only and for the combined total of cigar and pipe smokers.

Essentially the same relationship between mortality rates and amount smoked is shown when smokers are classified by the maximum amount smoked. Mortality ratios for persons who were smoking during the spring of 1954, classified both by

current and maximum smoked, are shown in table 6.

For cigarette smokers, mortality ratios based on the maximum amount smoked are slightly smaller than those based on the current amount smoked but the increase of death rates with an increase in the amount of tobacco used remains unchanged. For cigar and pipe users mortality rates also increase slightly with an increase in the amount of tobacco smoked. The trend is consistent for each of the three groups shown in table 6 but the relative increase is small and is not statistically significant for any group by itself although it is for the total of the three groups.

The above analysis has been based on the experience of policyholders who were smoking one or more forms of tobacco when the questionnaire was completed early in 1954. There was an additional group of policyholders who formerly had smoked but who had stopped prior to 1954. The percentage who had stopped of the total number who had ever smoked regularly was as follows:

Cigarette only	26 percent
Cigarette and other	36 percent
Cigar only	24 percent
Pipe only	25 percent
Cigar and pipe	23 percent

These percentages are based upon the number of persons who at one time had regularly used each form of tobacco shown and who had stopped smoking that form before 1954. However, they may have been smoking another form of tobacco. For example, 26 percent of the persons who had ever regularly smoked only cigarettes had stopped smoking cigarettes but some of them may have been smoking cigars or a pipe occasionally. Similarly, 36 percent of the persons who had ever regularly smoked cigarettes and cigars or a pipe or both had stopped smoking cigarettes but some of them may have been smoking cigars and/or a pipe either regularly or occasionally.

Mortality ratios by maximum amount smoked for ex-smokers as just defined are presented in table 7. For cigarette smokers the ratios are somewhat less than those for persons who were still smoking but, with one exception, the ratios are all greater than unity. The difference between the death rate of persons who had smoked one pack or less per day and those who had smoked more than one pack per day is less than the corresponding difference for persons who were still smoking but the heaviest smokers had the highest death rate in both groups.

In contrast to cigar and pipe smokers who were still smoking these forms of tobacco, those who had stopped smoking cigars or a pipe had a higher death rate than persons who had never smoked, although for two groups, pipe smokers who had smoked less than 10 pipefuls per day and cigar and pipe smokers who had smoked 5 or more cigars per day, the difference is not statistically significant (fig. 4). Furthermore, for cigar smokers the mortality ratio for persons who had smoked less than 5 cigars per day is slightly, but not significantly, greater than the ratio for persons who had smoked 5 or more cigars

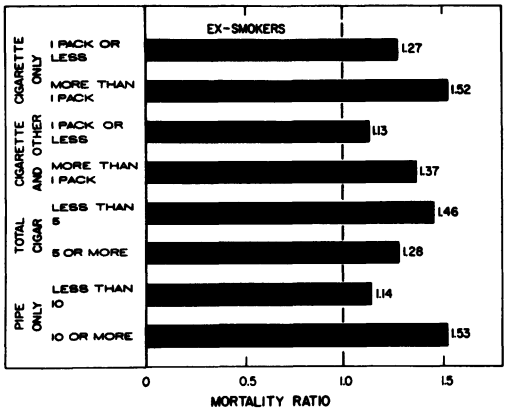


Figure 4. Mortality of ex-smokers by maximum amount ever smoked and type of tobacco; ratio of observed to expected number of deaths; death rate of nonsmokers = 1.00.

per day. These data suggest that many cigar and pipe smokers may have stopped smoking because of ill health but it is not obvious why this should be true for cigar and pipe smokers and not for cigarette smokers.

To summarize, the death rates for cigarette smokers are definitely greater than those for nonsmokers and this difference increases with the average daily number of cigarettes smoked for current smokers and for ex-smokers. Moreover the higher mortality exists whether the current amount smoked or the maximum amount smoked is used as a measure of amount.

The picture is not as clear for cigar and pipe smokers. The death rate for persons who were continuing to smoke was not appreciably higher than that for nonsmokers except possibly for the very heaviest smokers. By contrast, cigar smokers who had stopped smoking experienced mortality rates definitely greater than those for nonsmokers but these rates did not increase with the average daily number of cigars smoked. The death rates for ex-pipe smokers were highest for the heaviest smokers and were greater than those for nonsmokers.

Mortality of persons who had reduced the amount smoked

In addition to former smokers who had discontinued smoking when the histories were collected in 1954, some of the current smokers were using less tobacco than in the past. The percentage of current smokers who were smoking less than in the past was 11.7 for cigarette only, 19.9 for cigar only, and 7.2 for pipe users. These percentages refer to persons who were continuing to smoke the same form of tobacco but at a reduced rate; persons who had switched from one form of tobacco to another, for example, from cigarettes to pipe, are not included.

Policyholders who had ever smoked regularly

were grouped into three classes with respect to changes in smoking habits, (a) those who were currently smoking as much as ever, (b) those who were currently smoking less than in the past, and (c) those who had stopped smoking. Mortality ratios by maximum amount smoked for these three groups are shown in table 8. For those who were still smoking, the maximum amount was the current amount smoked but for those who had cut down or stopped smoking, the maximum amount refers to previous smoking habits. Mortality ratios for ex-smokers and all current smokers are shown in table 1.

Looking first at the summary figures in table 1, it is evident that regular cigarette smokers who stopped smoking prior to 1954 have a lower death rate than those who continued to smoke. Nevertheless their death rate is 30 percent greater than that for nonsmokers. The greater mortality of men who regularly smoke only cigarettes is emphasized by the fact that the mortality ratio for those who had stopped smoking, 1.39, is slightly higher than the ratio, 1.35, for persons who were continuing to smoke cigarettes in combination with cigars or a pipe.

In contrast to ex-cigarette smokers, ex-smokers of cigars and a pipe have a higher death rate than persons who continue to smoke these forms of tobacco. The death rate of those who were still regularly smoking cigars or a pipe at the beginning of 1954 does not differ significantly from that of persons who had never smoked; but the corresponding rate for ex-smokers ranged from 21 to 44 percent higher.

Turning to table 8 which shows mortality ratios for persons who had reduced the amount of smoking as well as for persons who had stopped, we find that, for regular cigarette smokers, those who had stopped smoking cigarettes have lower death rates than either those who continued to smoke as ever or who continued to smoke but at a lower rate than in the past (fig. 5). The latter group experienced the highest mortality, especially those who formerly had smoked more than a pack of cigarettes per day. However, irrespective of whether they had stopped smoking cigarettes, cut down on the number smoked, or continued to smoke as many as ever, the death rate of cigarette smokers was higher than that of nonsmokers.

Among policyholders who had regularly smoked only a pipe, the death rate is not significantly different for those who had stopped and those who continued to smoke at a reduced rate. The mortality ratios for both groups are significantly greater than unity for persons who had smoked 10 or more pipefuls per day but not for persons who had smoked less than this amount. Pipe smokers who were smoking as much as ever had a death rate no greater than that for nonsmokers. No difference in mortality rates exists for the two amount-of-use groups shown in table 8.

The mortality experience of cigar smokers is generally similar to that of pipe smokers. The highest death rates are found among those who have either cut down on the number of cigars smoked or have stopped smoking cigars. Those who

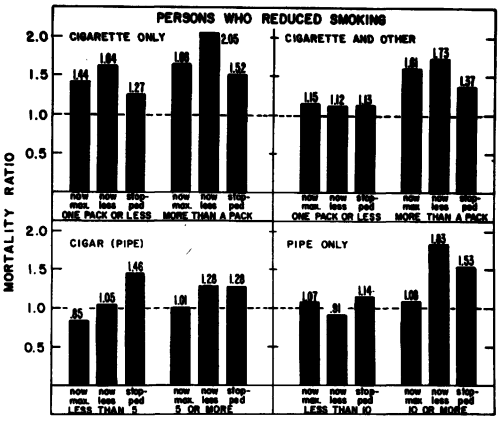


Figure 5. Mortality of persons who had reduced the amount of tobacco used by type of tobacco, amount of tobacco used, and amount of reduction; rate of observed to expected number of deaths; death rate of nonsmokers = 1.00.

continued to smoke as much as ever have a death rate no greater than nonsmokers.

Duration of Smoking History

It is worthwhile investigating whether a difference in the duration of smoking history could account for the higher mortality of ex-cigar and ex-pipe smokers than that of persons who were still smoking at the beginning of 1954. If the action of tobacco is like that of many other agents known to induce specific diseases, its effect upon health should become more marked the longer it is used. At least there should be a minimum period of use necessary before any effect is noticeable and possibly a maximum period of use after which no increase in effect is perceptible.

Mortality ratios by the number of years of use of different types of tobacco separately for ex-smokers and for current smokers are shown in table 9. The death rate of cigarette smokers who have smoked for less than 25 years is not significantly higher than that for nonsmokers (fig. 6). This holds true both for ex-smokers and for current smokers (fig. 7). The same absence of any effect upon mortality is equally true for persons who have smoked more than a pack per day and for persons who have smoked a pack or less per day (table 10). In other words, among men who have smoked cigarettes for less than 25 years, the number of cigarettes smoked has no noticeable effect upon the death rate. This observation might not hold true if there were a sufficient number of deaths to subdivide the data for those smoking more than a pack per day.

Mortality ratios for men currently smoking more than a pack of cigarettes per day are greater than unity for durations of less than 25 years but the number of deaths is not large enough to ensure that this excess mortality is statistically significant. For ex-smokers of cigarettes, there is no evidence of a relationship between the maximum number of cigarettes formerly smoked and

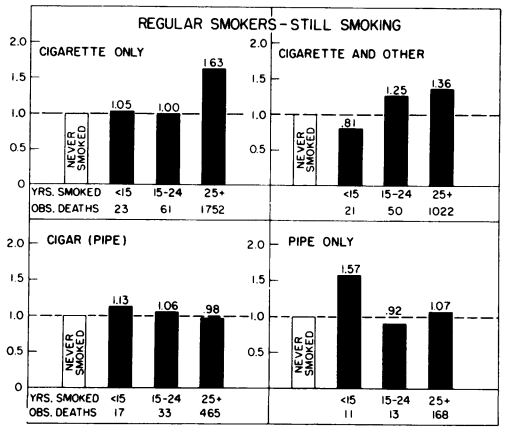


Figure 6. Mortality of regular smokers who were currently smoking in 1954 by type of tobacco and number of years smoked; ratio of observed to expected number of deaths; death rate of nonsmokers = 1.00.

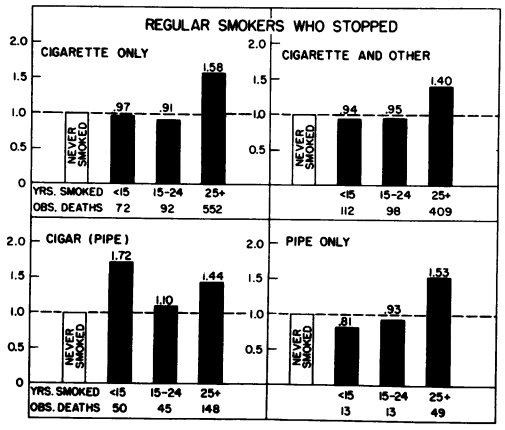


Figure 7. Mortality of regular ex-smokers by type of tobacco and number of years smoked; ratio of observed to expected number of deaths; death rate of nonsmokers = 1.00.

mortality for durations of less than 25 years.

Unfortunately when the tabulations were planned, the data for durations of 25 or more years were combined. Until another tabulation can be made it will not be possible to determine whether death rates increase with an increase in duration of smoking beyond 25 years.

The death rate for regular pipe smokers who were currently smoking is no greater than that for nonsmokers, even for men whose current rate was 10 or more pipefuls per day and who had been smoking a pipe for 25 or more years. The only group of ex-pipe-smokers with a mortality ratio significantly greater than unity were men who had smoked a pipe for 25 or more years before stopping.

The death rates for regular cigar users who were currently smoking are similar to those for current pipe smokers. The mortality ratio for men who were smoking 5 or more cigars per day and who had been smoking cigars for 25 or more years is on the borderline of significance but even so it reflects a death rate only 11 percent greater than that for nonsmokers.

The effect of amount of smoking and duration of smoking history upon health is more marked for ex-cigar smokers than it is for pipe smokers. All of the mortality ratios, except that for men whose maximum rate of use was 5 or more cigars per day and who had smoked from 15-24 years prior to stopping are greater than unity although some of the differences are not statistically significant due to the small number of deaths. The most puzzling aspect of the mortality of ex-cigar smokers is that the death rates are as high for men who have smoked less than 25 years as they are for men who have smoked 25 or more years. This reinforces the supposition that many men who stop smoking cigars may do so because of ill health.

Age Started

Another way of examining the effect of duration of smoking upon mortality rates is to classify persons by the age they began to smoke. Mortality ratios for regular ex-smokers and for regular current smokers by age started smoking are shown in table 11. For this purpose, the policyholders were classified by the age they started to smoke each form of tobacco; those who had smoked both cigarettes and cigars or a pipe were classified by the age they started to smoke cigarettes, those who had smoked cigars and a pipe were classified by the age they started to smoke cigars.

For regular cigarette smokers, irrespective of whether they had stopped smoking or were still smoking when the study was begun, the mortality ratios are greater than one and decrease with an increase in the age at which smoking was started (fig. 8). In other words, men who started to smoke cigarettes in their late teens or early twenties have a higher death rate than those who did not start to smoke until after age 25 or 30. Except for men who were continuing to smoke only cigarettes, the death rates for those who started to smoke after age 25 are not significantly greater than the death rate of nonsmokers.

The same relationship between age started

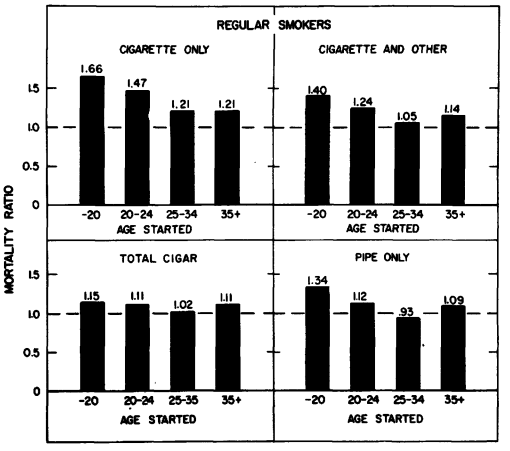


Figure 8. Mortality of regular smokers by age started and type of tobacco; ratio of observed to expected number of deaths; death rate of nonsmokers = 1.00.

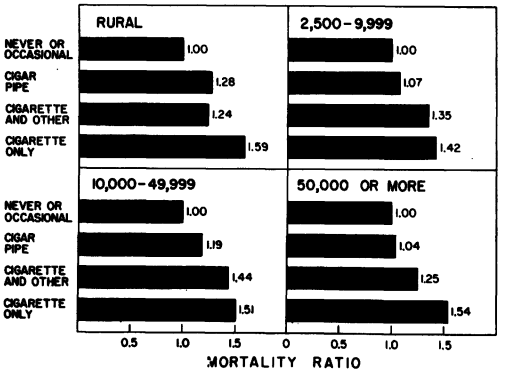


Figure 9. Mortality of smokers and nonsmokers by size of community and type of tobacco; ratio of observed to expected number of deaths; death rate of nonsmokers and occasionally-only-smokers for each region = 1.00.

smoking and mortality rates is found when current cigarette smokers are classified by the amount smoked. The heaviest smokers have the highest death rates for each age-of-starting group except for mixed tobacco users who started to smoke cigarettes after 25 years of age.

The death rates of ex-cigar and ex-pipe smokers exhibit the same trend by age started as do those for cigarette smokers but no consistent relationship between age started and mortality exists for men who were currently smoking cigars or a pipe even when they are classified by amount smoked.

Size of Community

The long established difference in mortality of rural and urban residents has gradually narrowed during recent years. For the United States as a whole the age-adjusted death rate for white males during 1949-1951 was only 10 percent greater for residents of central city metropolitan counties than for residents of non-metropolitan counties. A similar small differential exists for the males included in this study.

On the basis of their residence in 1954, policyholders were classified as living in places of 50,000 or more population including the urbanized areas around each city as defined by the Bureau of the Census, 10,000-49,999 population, 2,500-9,999 population, or in rural areas. The death rate for persons who have never smoked or at most have smoked only occasionally is from 7 to 9 percent greater in urban than in rural areas. The rural-urban difference in death rates is smaller for regular smokers than for nonsmokers. However the absolute value of the urban excess in both instances is so small that residence can be neglected as an important factor in studying total mortality.

The relationship between smoking and mortality is similar for residents of each of the four types of communities (fig. 9). The death rate is higher for men who have regularly smoked than for nonsmokers. Among regular smokers those who have smoked only cigarettes have the highest death rate while those who have smoked cigars and/or a pipe have the lowest rate. No consistent relationship between amount smoked and mortality exists for cigar or pipe smokers but men who have smoked more than a pack of cigarettes per day experience a higher death rate than those who have smoked one pack or less per day.

Geographic Region

Mortality ratios by region of residence in 1954 are presented in table 12. Inspection of these data show that the relationship between smoking history and mortality described above for the entire group of policyholders and for those living in rural and urban areas separately also exists in each geographic region of the United States.

Mortality By Broad Groups of Causes

Causes of death were classified according to the rules of the International Statistical Classification of Diseases, Injuries, and Causes of Death published by the World Health Organization. One underlying and a maximum of two contributory causes of death were coded. The selection of the underlying cause of death was based on the opinion of the deceased's physician except in instances where it was obvious the physician had misunderstood the intent of the question and had selected a terminal condition such as pulmonary failure or edema as the underlying cause of death. In addition, a separate code was provided for cases with cancer which the physician stated was not a contributory or underlying cause of death. Examples of these are cases with skin cancer or clinically quiescent cancer of the prostate discovered at autopsy, or cases with cancer who died suddenly from coronary artery disease.

The choice of one of several coexisting

diseases as the underlying cause of death frequently must be done arbitrarily and may not accurately reflect the complex of diseases that caused the death. For the purpose of this study, the important question is whether a given disease had ever occurred, and not whether this disease may have been chosen by some method as the underlying cause of death.

For comparability with the data from other published studies and with official vital statistics some of the following tables present mortality rates computed from underlying causes only. Each death was assigned to only one cause irrespective of the number of coexisting diseases. In other tables, the mortality rates are based upon both underlying and contributory causes. For example, if a policyholder had both diabetes and cancer of the lung at the time of death, he was counted twice, once in the group with diabetes and once in the group with cancer of the lung. Consequently the sum of the number of deaths by cause in these tables is greater than the number of persons who died.

The expected number of deaths by cause were computed from the death rates for persons who had never smoked or who had smoked only occasionally in order to have a statistically more stable basis of comparison. The death rates for these two groups did not differ significantly either for all causes combined or for any individual cause. The crude death rates per 100,000 from lung cancer were 11 and 25 respectively but this difference is not statistically significant. For simplicity, the combined group will be referred to as having never smoked.

By far the greatest increase in the risk of developing a disease for smokers is that for lung cancer (table 13). Among the entire group of men who had ever smoked, there were 312 observed deaths from lung cancer compared with 52 expected if they had been no more likely than nonsmokers to develop this disease, a mortality ratio of 6.0 based on underlying causes only.

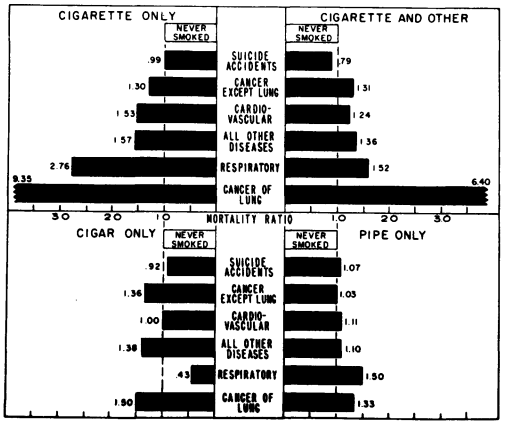


Figure 10. Mortality of smokers from broad groups of causes of death by type of tobacco; ratio of observed to expected number of deaths; death rates of nonsmokers and occasionally-only-smokers = 1.00.

However, the size of the increased risk of dying from lung cancer varies widely among men who smoke tobacco in different forms, ranging from an excess of 33 percent for pipe smokers to 835 percent for cigarette-only-smokers (fig. 10). The mortality ratio for none of the three cigar and/or pipe smoking groups differs significantly from one; the ratio for the three combined is barely significant at the 5 percent level.

A sharp difference exists between the death rates from lung cancer for men who have smoked cigarettes and for men who have smoked cigars or a pipe. The chances of developing lung cancer are 9.35 times greater for regular cigarette smokers than for nonsmokers compared to an average risk of 1.60 times for regular users of cigars and/or a pipe.

For no other disease does the excess mortality among smokers approach that for lung cancer. The next highest mortality ratio is for a group of respiratory diseases including pulmonary tuberculosis, asthma, bronchitis, emphysema, pneumonia, and pleurisy. Although these diseases are important causes of morbidity, they are numerically unimportant as primary causes of death; only 118 deaths among smokers and 24 deaths among nonsmokers were due directly to this group of disease.

An increased death rate from the respiratory diseases is found only among regular cigarette smokers for whom the mortality ratio is 2.24. Similarly to lung cancer, the highest ratio, 2.76, is for persons who smoke cigarettes only. The death rate for cigar or pipe smokers is no higher than that for nonsmokers.

Nearly two-thirds of the deaths of persons who had used tobacco were attributed to diseases of the cardiovascular-renal system, including arteriosclerosis, rheumatic heart disease, coronary heart disease, chronic endocarditis, hypertension, and chronic nephritis. The risk of dying from one or more of these diseases is 31 percent greater for regular smokers than for nonsmokers. Again the risk is greater for regular cigarette users, especially those who have smoked only cigarettes, than it is for users of other forms of tobacco. There is no indication that regular cigar or pipe smokers have a higher death rate than nonsmokers.

The mortality ratios for cancer, exclusive of lung cancer, are similar to those for cardiovascular disease except that the ratios for cigar and pipe smokers are as high as those for cigarette smokers. The mortality from the individual forms of cancer will be discussed below.

Mortality from Specific Causes of Death

The data in tables 14, 16, 17, and 18 are based upon underlying and contributory causes of death and for cancer also include cases with this disease even though it was not considered to be a primary or contributory cause of death.

For regular cigarette smokers the diseases with a mortality ratio greater than 2.0, a signifying a death rate more than double that for nonsmokers are bronchitis, emphysema, and allied respiratory diseases, cirrhosis of the liver, ulcer of the stomach or duodenum, cancer of the

prostate, and cancer of the esophagus and buccal cavity (table 14). Several studies have reported that heavy smokers also tend to drink alcoholic liquors excessively so that the increased death rate from cirrhosis of the liver may reflect the effect of alcohol rather than that of cigarettes. An increased mortality of cigarette smokers from the other diseases mentioned also have been reported by other studies. An explanation of the high mortality ratio for cancer of the prostate is not apparent.

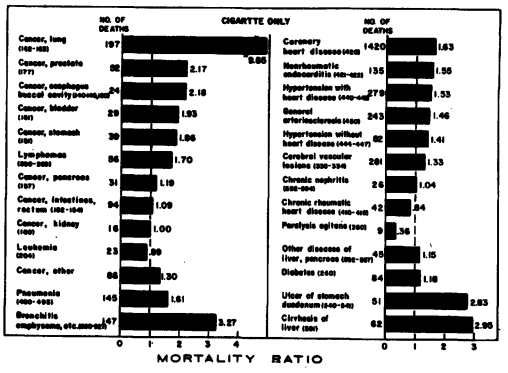


Figure 11. Mortality of regular smokers of cigarettes only from specific diseases; ratio of observed to expected number of deaths; death rates of nonsmokers and occasionally-only-smokers = 1.00.

The death rates from the principal cardiovascular diseases, including coronary heart disease, are from 33 percent to 63 percent greater for regular cigarette-only-smokers than for nonsmokers (fig. 11). No increase in mortality exists for chronic rheumatic heart disease, whose cause is known, or for chronic nephritis.

Mortality ratios for men who were regularly smoking cigarettes only by current amount smoked are shown in table 16. For several diseases the number of deaths is so small that the mortality ratios do not differ significantly by amount smoked. But for cancer of the lung and cancer of the buccal cavity and esophagus the increase in the death rate with an increase in the average daily number of cigarettes smoked is very marked (fig. 12). The death rate for men who regularly were smoking more than a pack of cigarettes per day is nearly 16 times the death rate for nonsmokers. A similar although numerically smaller increase in the mortality ratio with an increase in the number of cigarettes smoked is also found for cancer of the buccal cavity - lip, mouth, tongue, pharynx and esophagus.

The leading cause of death of the policyholders included in this study is coronary heart disease. The death rate from this disease is significantly greater than that for nonsmokers for each of the amount-of-smoking groups shown in table 16 but the difference is considerably less than that for cancer of the lung. Moreover the mortality ratio for men who were smoking more than a pack of cigarettes per day is no greater than the ratio for men smoking from half a pack to one pack per day although these ratios

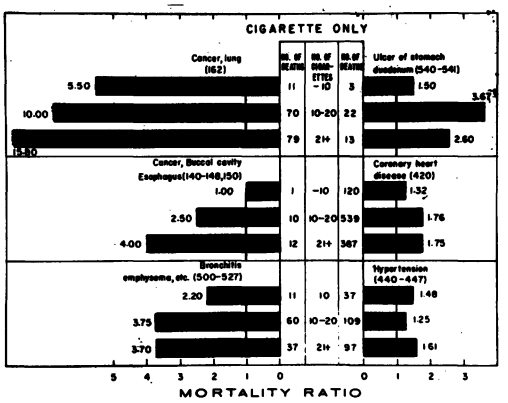


Figure 12. Mortality of regular smokers of cigarettes only from specific diseases by current amount smoked in 1954; ratio of observed to expected number of deaths; death rates of nonsmokers and occasionally-only-smokers = 1.00.

are greater than that for light smokers - those smoking less than half a pack per day.

None of the mortality ratios for specific diseases for men who have regularly smoked only a pipe differ significantly from one except those for bronchitis and emphysema, 1.88, and coronary heart disease, 1.16 (table 17). For both diseases, the death rate for pipe smokers is less than that for cigarette smokers.

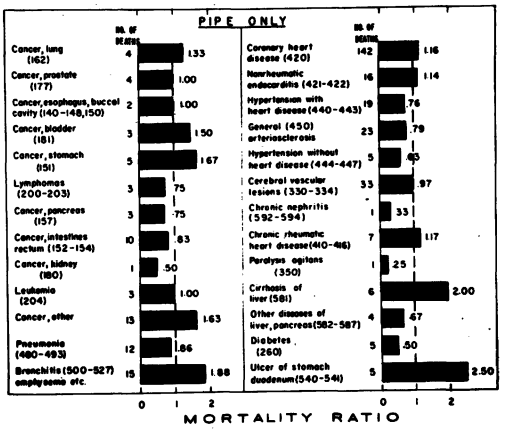


Figure 13. Mortality of regular smokers of a pipe only from specific diseases; ratio of observed to expected number of deaths; death rates of nonsmokers and occasionally-only-smokers = 1.00.

In general, the death rates of regular cigar smokers are less than those for regular cigarette smokers not only for all causes of death combined but also for most of the specific diseases shown in table 18. However, cigar smokers are more likely to die from diabetes, cancer of the prostate, and cancer of the intestinal tract than are cigarette smokers; for cirrhosis of the liver and cancer of the buccal cavity and esophagus the death rates for the two groups of smokers are practically equal. The mortality ratios for these diseases are significantly greater than one (fig. 14).

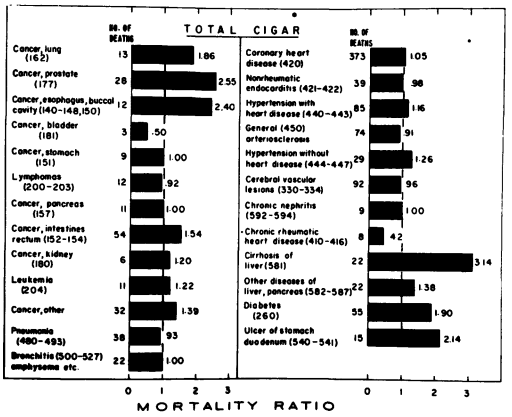


Figure 14. Mortality of regular cigar smokers from specific diseases; ratio of observed to expected number of deaths; death rates of nonsmokers and occasionally-only-smokers = 1.00.

The death rate from cancer of the lung for cigar smokers falls between those for pipe and cigarette smokers. Although it is significantly greater than the rate for nonsmokers, the excess is very much less than that for men who have smoked only cigarettes, the mortality ratios being 1.86 and 9.85 respectively.

Relative Importance of Causes of Excess Deaths

Up to this point the effect of the use of tobacco upon health has been measured by the ratio of the death rate of smokers to that of nonsmokers. The relative importance of a disease as a cause of death cannot be determined solely from the magnitude of the mortality ratio; for this purpose some index of the absolute number of deaths due to each disease is necessary.

The numerical value of this index depends upon the rules for selecting the underlying cause of death and the way specific diseases are grouped into cause-of-death categories. In practice, some cause-of-death categories usually include a single disease while others may include several diseases. Moreover, the relative rank of the cause-of-death categories varies with age at death so that the ranking for two populations will not be the same unless both the death rates and the percentage age distribution are the same. These limitations upon the use and interpretation of a ranking of causes of death should be borne in mind during the following discussion.

As the data in table 15 show there is no necessary relationship between the mortality ratio and the relative rank of a disease as a cause of death. Cancer of the lung with a mortality ratio of 9.35 caused 6.8 percent of the deaths of men who had regularly smoked only cigarettes; coronary heart disease caused 43.7 percent of the deaths but had a mortality ratio of 1.58. The death rate from bronchitis, emphysema, and allied conditions among cigarette-only-smokers was nearly three times that for nonsmokers but these diseases are relatively unimportant as primary causes of death, accounting for only 2.5 percent of the total.

Another way of ranking diseases as a cause of death is by percentage of the total number of excess deaths assigned to each disease. The number of excess deaths in table 15 is the difference between the number of observed deaths and the number of expected deaths based on the death rate from each cause of death for persons who had never smoked or who had smoked occasionally only. This ranking is strongly affected by the mortality ratio.

The relative importance of coronary heart disease is the same for both methods of ranking; it is by far the leading cause of death and also the leading cause of the excess deaths. The change in relative importance is most striking for cancer of the lung which caused 6.8 percent of the total number of observed deaths but 16.8 percent of the number of excess deaths of men who regularly smoked only cigarettes. Lung cancer caused about one half as many deaths as all other forms of cancer combined but it caused twice as large a proportion of the number of excess deaths as all other forms of cancer.

Summary

1. The death rate from all causes of men who have smoked tobacco is 32 percent greater than that for persons who have never smoked.

2. Men who have regularly smoked only cigarettes have the highest death rate of all groups of smokers - 58 percent greater than the rate for nonsmokers.

3. The death rate from all causes of men who have regularly smoked cigars and/or a pipe is not appreciably higher than that of nonsmokers.

4. The excess mortality of regular cigarette smokers increases with the average number of cigarettes smoked per day. The death rate for cigarette-only-smokers who regularly smoke two or more packs per day is twice the rate for nonsmokers. Only the heaviest users of cigar and pipe tobacco experience a significant increase in total death rate and this increase is less than that for cigarette smokers.

5. Regular cigarette smokers who had stopped smoking cigarettes have a lower mortality rate than those who continued to smoke. Reducing the number of cigarettes smoked per day but still continuing to smoke regularly does not result in a reduction in the death rate.

6. The death rate of regular cigarette smokers who had smoked for less than 25 years is not significantly greater than that for nonsmokers except for those who were continuing to smoke more than a pack of cigarettes per day.

7. The greatest increase for smokers in the risk of developing a disease is for cancer of the lung. Regular smokers of cigarettes only have a death rate from lung cancer nearly 10 times that of nonsmokers. The death rate for men smoking more than a pack of cigarettes per day is 16 times that of nonsmokers.

8. Regular cigarette smokers also are subject to an increased risk of dying from cardiovascular diseases, from bronchitis, emphysema, and allied respiratory diseases, from ulcers of the stomach and duodenum, and from cirrhosis of the liver.

9. The death rate from coronary heart disease among regular users of cigarettes only is 63 percent higher than the rate for nonsmokers.

10. The leading cause of death among these men is coronary heart disease. Cancer of the lung caused 6.8 percent of the deaths of men who smoked only cigarettes but accounted for 16.8 percent of the number of excess deaths.

References

1. Cited in Annals of Eugenics II, 246, 1927.
2. Moriyama, I. M. and Guralnick, L., "Occupational and Social Class Differences in Mortality", in Trends and Differentials in Mortality pp. 61-73 published by Milbank Memorial Fund 1956.
3. Hammond, E. Cuyler and Horn, Daniel, "Smoking and Death Rates - Report on Forty-Four Months of Follow-Up of 187,783 Men", J. Am. Med. Assoc. 166,1159-1172, 1294-1308, 1958.
4. Haenszel, William, Shimkin, Michael B. and Miller, Herman P., "Tobacco Smoking Patterns in the United States", U. S. Public Health Service, Public Health Monograph No. 45, 1956.
5. Finkner, A. L., Horwitz, D. G., Foradori, G. T., Fleischer, Jack and Monroe, John, "An Investigation on the Measurement of Current Smoking by Individuals", North Carolina State College, Institute of Statistics, Mimeo Series No. 177, 1957.

Appendix

Classification by Smoking History

Persons were classified by smoking history in accordance with the following definitions.

Used tobacco: Persons who had smoked at least 5 to 10 packs of cigarettes or 50 to 75 cigars or 3 to 5 packages of pipe tobacco.

Smoked occasionally only: Persons who had never regularly smoked any form of tobacco but who had occasionally smoked one or more forms. Also included here are persons with unknown amount used either currently or in the past provided that the maximum amount of known use was occasional only.

Regular smoker: Persons who at sometime during their lifetime had regularly smoked cigars, cigarettes, or pipe tobacco. These were further classified by the form of tobacco used and whether or not they were smoking at the start of the study in 1954.

Regular smoker, cigarettes only: Persons who had regularly smoked only cigarettes. They may have occasionally smoked cigars or a pipe; they had never smoked either of these regularly.

Regular smoker, cigarette and cigar: Persons who had regularly smoked both cigarettes and cigars. They may have smoked a pipe occasionally but never regularly. Similar definitions were used for regular smokers of other combinations of tobacco.

Amount unknown: Persons who stated they had smoked more than the minimum amount to qualify as a user of tobacco but who did not report the amount used either currently or in the past with sufficient accuracy to permit assignment to one of the groups of regular or occasional smokers.

Amount used: In this report a classification by amount of tobacco used is based on the current amount used at the time the questionnaire was filled out in 1954. Regular smokers of (a) cigarettes and cigars, (b) cigarettes and pipe, and (c) cigarettes, cigars, and pipe were classified by the current number of cigarettes smoked. Regular users of cigars and pipes were classified by the current number of cigars smoked.

Classification by Cause of Death

ISC number	Title
162,163	Cancer of lung
140-205, except 162,163	Cancer except lung
	Respiratory diseases
001-008	Respiratory tuberculosis
241	Asthma
480-493	Influenza and pneumonia
500-502	Bronchitis
527.1	Emphysema without bronchitis
470-475; 527.0; 527.2; 510-526	Other respiratory diseases
	Accidents
810-835	Motor vehicle accidents
800-802; 840-962; 980-991	Other accidents
970-979	Suicide
	Cardiovascular
330-334	Cerebral vascular lesions
400-402	Rheumatic fever
410-416	Chronic rheumatic heart disease
420	Arteriosclerotic heart disease
421-422	Nonrheumatic chronic endocarditis
430-434	Other heart disease
440-443	Hypertension with mention of heart disease
444-447	Hypertension without mention of heart disease
450	General arteriosclerosis
451-468	Other circulatory system disease
592-594	Chronic nephritis
Remaining categories	Other diseases

TABLE 1. MORTALITY OF SMOKERS AND NONSMOKERS

Ratio of observed to expected number of deaths (all causes) by type of smoking history and current use, July 1954-December 1956

Smoking History	No. of person years exposure	Current Use		
		Total	Smokes	Does not smoke
Never smoked	89,774	1.00	. . .	. . .
Used tobacco	389,178	1.32	1.37	1.24
Occasionally only	28,144	.98	.91	1.05
Regular smoker	339,903	1.36	1.40	1.30
Cigarette total	271,757	1.45	1.54	1.30
Cigarette only	161,172	1.58	1.65	1.39
Cigarette and other	110,585	1.29	1.35	1.21
Cigarette and cigar	21,188	1.31	1.34	1.27
Cigarette and pipe	53,168	1.36	1.41	1.26
Cigarette, cigar, pipe	36,229	1.20	1.26	1.11
Cigar only	28,422	1.07	.94	1.44
Cigar and pipe	21,944	1.08	1.04	1.21
Pipe only	17,780	1.10	1.05	1.25
Amount unknown	21,131	1.06	1.43	1.05

Expected number of deaths - computed by multiplying the number of person years exposure in each age group for each smoking history category by the age-specific death rates of persons who had never smoked.

TABLE 2. RELIABILITY OF REPORTED SMOKING HISTORY

Percentage change in classification by selected items of smoking history between the original and a duplicate questionnaire received from 1714 policyholders

Item	No. of Code groups	Percentage Agreement		
		Cigarette	Cigar	Pipe
Amount currently used	8	90.2	89.7	93.8
Years used current amount	12	90.7	90.2	92.8
Age started	8	91.0	90.8	94.5
Summary of use	7	92.8	91.0	93.7

TABLE 3.--LIFETIME SMOKING HISTORY
Percentage who have ever smoked, U.S. males, 1955
and USGLI policyholders, 1954

Smoking history	USGLI	U.S. [1]
Total	100.0	100.0
Unknown[2]	4.5	6.5
Never smoked	18.7	20.3
Smoked occasionally only	5.9	5.2
Smoked regularly	70.9	68.0
Cigarette only	33.6	43.8
Cigarette and cigar	4.5	3.1
Cigarette and pipe	11.1	8.3
Cigarette, cigar, pipe	7.5	2.3
Cigar only	6.0	3.6
Cigar and pipe	4.5	2.2
Pipe	3.7	4.7

1/ Tobacco Smoking Patterns in the United States, Public Health Monograph No. 45, page 57; data for males 35 or more years of age.

2/-USGLI policyholders--unknown amount smoked; U.S. males--not reported.

TABLE 4. MORTALITY BY AGE AMONG SMOKERS AND NON-SMOKERS
Number of deaths per 1,000 per year by smoking history and age, July 1954 - December 1956

Smoking History	Number of deaths	Death rate per 1,000								
		All ages	30-39	40-49	50-54	55-59	60-64	65-69	70-74	75 and over
Never smoked	1179	13.1	0.7	2.9	6.6	9.0	14.8	21.6	35.8	70.0
Used tobacco	6203	16.0	1.3	3.2	9.3	13.1	19.0	28.1	38.6	73.3
Occasionally only	345	12.3	2.0	2.7	3.2	9.5	12.9	23.6	30.6	85.2
Regular smoker	5564	16.4	1.3	3.4	9.8	13.6	19.8	28.9	39.6	72.6
Cigarette total	4513	16.6	1.4	3.3	10.4	14.4	20.9	31.2	43.5	72.8
Cigarette only	2771	17.2	1.2	3.4	10.4	15.3	22.9	33.4	49.8	84.7
Cigarette and other	1742	15.8	1.8	3.1	10.4	13.0	18.1	28.8	37.4	62.9
Cigarette and cigar	363	17.1	3.0	7.4	10.2	10.5	17.8	37.0	43.6	48.2
Cigarette and pipe	805	15.1	1.4	2.0	10.1	14.1	19.3	28.3	34.9	73.5
Cigarette,cigar,pipe	574	15.8	2.2	4.5	11.0	12.8	16.7	24.5	36.3	62.5
Cigar only	433	15.2	0.0	10.4	9.3	8.5	16.3	24.1	26.7	89.6
Cigar and pipe	342	15.6	0.0	2.3	7.3	11.1	15.8	23.5	34.6	56.5
Pipe only	276	15.5	1.0	1.9	2.0	11.8	17.2	20.6	32.4	71.7
Amount unknown	294	13.9	0.7	0.0	7.8	10.2	15.6	22.9	33.8	67.8

TABLE 5. MORTALITY OF REGULAR SMOKERS BY CURRENT AMOUNT SMOKED

Ratio of observed to expected number of deaths (all causes) by type of smoking history and current amount smoked, July 1954-December 1956

Smoking History	Ratio of observed to expected					Number of observed deaths				
	Current number of cigarettes smoked per day									
	occas'l	less than 10	10-20	21-39	40 or more	occas'l	less than 10	10-20	21-39	40 or more
Cigarette only	.96	1.29	1.66	1.77	1.99	25	205	1019	663	137
Cigarette and other	1.08	.95	1.37	1.72	1.79	72	159	492	319	59
Cigarette and cigar	1.08	.90	1.30	1.75	2.71	13	35	91	56	19
Cigarette and pipe	1.00	1.03	1.36	1.88	1.40	27	71	252	180	21
Cigarette,cigar,pipe	1.14	.88	1.41	1.46	1.73	32	53	149	83	19
	Current number of cigars smoked per day									
	occas'l	1-2	3-4	5-8	9 or more	occas'l	1-2	3-4	5-8	9 or more
Cigar only	1.05	.71	1.00	.99	1.44	20	62	97	79	26
Cigar and pipe	.93	1.16	.99	.87	1.33	40	101	67	32	8
	Current number of pipefuls smoked per day									
	occas'l	less than 5	5-9	10-19	20 or more	occas'l	less than 5	5-9	10-19	20 or more
Pipe only	.86	.96	1.12	.98	1.21	6	44	73	45	29

Expected number of deaths - computed by multiplying the number of person years exposure in each age group for each smoking history category by the age-specific death rates of persons who had never smoked.

TABLE 6. --MORTALITY OF REGULAR SMOKERS BY AMOUNT SMOKED
Ratio of observed to expected number of deaths (all causes) by smoking history, current and maximum amount smoked for persons smoking in 1954.

Smoking History	Current amount smoked	Maximum amount smoked
Cigarette only		
one pack or less	1.59	1.48
more than a pack	1.81	1.75
Cigarette and cigar		
one pack or less	1.16	1.12
more than a pack	1.92	1.76
Cigarette and pipe		
one pack or less	1.27	1.15
more than a pack	1.81	1.72
Cigarette, cigar, pipe		
one pack or less	1.22	1.12
more than a pack	1.50	1.54
Cigar only		
less than 5	.86	.83
5 or more	1.07	1.11
Cigar and pipe		
less than 5 cigars	1.08	1.01
5 or more cigars	.93	1.14
Pipe only		
less than 10	1.05	1.06
10 or more	1.06	1.11

Expected number of deaths based on the death rates of persons who had never smoked.

TABLE 7. --MORTALITY OF EX-SMOKERS BY MAXIMUM AMOUNT EVER SMOKED
Ratio of observed to expected number of deaths (all causes) by smoking history and maximum amount smoked.

Smoking History	Mortality ratio	Number of observed deaths
Cigarette only		
one pack or less	1.27	417
more than a pack	1.52	304
Cigarette and cigar		
one pack or less	1.26	107
more than a pack	1.27	42
Cigarette and pipe		
one pack or less	1.21	173
more than a pack	1.31	81
Cigarette, cigar, pipe		
one pack or less	.97	144
more than a pack	1.50	81
Cigar only		
less than 5	1.59	94
5 or more	1.33	56
Cigar and pipe		
less than 5 cigars	1.29	62
5 or more cigars	1.19	31
Pipe only		
less than 10	1.14	50
10 or more	1.53	29

Expected number of deaths based on the death rates of persons who had never smoked.

TABLE 8. --PERSONS WHO REDUCED AMOUNT OF SMOKING
Ratio of observed to expected number of deaths by type of smoking history and maximum amount smoked

Smoking history and maximum amount smoked	Ratio of observed to expected deaths			Number of observed deaths		
	Current use is maximum	Smokes less than in past	Stopped	Current use is maximum	Smokes less than in past	Stopped
Cigarette only	1.53	1.90	1.37	1502	285	721
one pack or less	1.44	1.64	1.27	847	92	417
more than a pack	1.66	2.05	1.52	655	193	304
Cigarette and cigar	1.40	1.18	1.26	155	40	149
one pack or less	1.20	1.06	1.26	92	19	107
more than a pack	1.85	1.31	1.27	63	21	42
Cigarette and pipe	1.34	1.43	1.24	385	129	254
one pack or less	1.16	1.07	1.21	216	47	173
more than a pack	1.69	1.78	1.31	169	82	81
Cigarette, cigar, pipe	1.19	1.50	1.11	201	120	225
one pack or less	1.11	1.19	0.97	122	50	144
more than a pack	1.34	1.84	1.50	79	70	81
Cigar only	0.86	1.06	1.48	159	69	150
less than 5	0.76	0.93	1.59	84	27	94
5 or more	1.03	1.17	1.33	75	42	56
Cigar and pipe	0.97	1.24	1.26	114	108	93
less than 5 cigars	0.97	1.12	1.29	84	54	62
5 or more cigars	0.97	1.38	1.19	30	54	31
Pipe only	1.07	1.24	1.27	146	21	79
less than 10 pipefuls	1.07	0.91	1.14	80	10	50
10 or more pipefuls	1.08	1.83	1.53	66	11	29

Expected number of deaths were computed from death rates for persons who had never smoked.
Amount is maximum use for persons who had stopped or who smoked less than in the past and current use for those whose current use is maximum.

The ratios for those who had stopped differ slightly from the ratios for ex-smokers in Table 1 due to the fact that broader age groups were used for this table.

TABLE 9. MORTALITY BY DURATION OF SMOKING HISTORY
Ratio of observed to expected number of deaths (all causes) by smoking history, duration of use, and current use

Smoking History	Mortality Ratio			Number of Observed Deaths		
	Number			Years		
	- 15	15-24	25+	-15	15-24	25+
	Regular smokers - Currently smoking					
Cigarette only	1.05	1.00	1.63	23	61	1952
Cigarette and cigar	1.29	1.00	1.36	9	9	196
Cigarette and pipe	.50	1.22	1.39	5	22	517
Cigarette, cigar, pipe	.78	1.46	1.30	7	19	309
Cigar only	1.44	.95	.92	13	18	243
Cigar and pipe	.67	1.25	1.06	4	15	222
Pipe only	1.57	.93	1.07	11	13	168
	Regular - Ex-Smokers					
Cigarette only	.97	.91	1.58	72	92	552
Cigarette and cigar	1.03	1.08	1.47	30	27	91
Cigarette and pipe	.84	.62	1.53	27	24	202
Cigarette, cigar, pipe	.95	1.21	1.17	55	47	116
Cigar only	2.00	1.29	1.45	32	31	87
Cigar and pipe	1.39	.82	1.42	18	14	61
Pipe only	.81	.93	1.53	13	13	49

Expected number of deaths based on death rates of persons who had never smoked.

Number of years is the number of years each specific form of tobacco had been used; for users of two or more forms duration refers to the first mentioned.

TABLE 10. MORTALITY BY AMOUNT AND DURATION OF SMOKING HISTORY
Ratio of observed to expected number of deaths (all causes) by smoking history, amount used, duration of use, and current use

Smoking History Amount Used	Mortality Ratio			Number of Observed Deaths		
	Number of Years					
	-15	15-24	25+	-15	15-24	25+
			Regular Smokers - Currently Smoking			
Cigarette only						
one pack or less	1.00	.92	1.52	17	36	1038
more than a pack	1.20	1.14	1.78	6	25	914
Cigarette and other						
one pack or less	.55	1.23	1.16	11	32	541
more than a pack	2.00	1.29	1.68	10	18	481
Cigar (including cigar and pipe)						
less than 5 cigars	1.00	1.00	0.89	12	21	258
5 or more cigars	1.67	1.33	1.11	5	12	207
Pipe only						
less than 10 pipefuls	1.50	1.00	1.03	9	9	92
10 or more pipefuls	1.00	.80	1.12	2	4	76
			Regular Ex-smokers			
Cigarette only						
one pack or less	.97	.92	1.48	60	61	292
more than a pack	1.00	.91	1.70	12	31	260
Cigarette and other						
one pack or less	.91	.88	1.36	93	66	259
more than a pack	1.12	1.14	1.47	19	32	150
Cigar (including cigar and pipe)						
less than 5 cigars	1.65	1.23	1.48	38	32	86
5 or more cigars	2.00	.87	1.38	12	13	62
Pipe only						
less than 10 pipefuls	.67	.90	1.56	10	9	28
10 or more pipefuls	.75	1.33	1.50	3	4	21

Expected number of deaths based on death rates of persons who had never smoked.

Number of years is the number each specific form of tobacco had been used; for users of two or more forms, duration refers to the first mentioned.

TABLE 11. MORTALITY OF REGULAR SMOKERS BY AGE STARTED
Ratio of observed to expected number of deaths (all causes) by smoking history, current use, age started, and current amount used.

Smoking History Amount Used	Mortality Ratio				Number of Observed Deaths			
	Age Started Smoking							
	-20	20-24	25-34	35+	-20	20-24	25-34	35+
				Regular Ex-Smokers				
Cigarette only	1.51	1.29	1.08	1.20	393	207	85	30
Cigarette & other	1.34	1.07	1.05	1.16	324	162	89	43
Cigar (including cigar and pipe)	1.59	1.47	1.26	1.00	65	103	59	15
Pipe only	1.71	1.23	1.00	1.00	29	27	14	8
				Regular Smokers - Currently Smoking				
Cigarette only	1.72	1.54	1.27	1.21	1144	603	214	75
one pack or less	1.66	1.51	1.20	1.18	644	391	147	58
more than a pack	1.80	1.61	1.46	1.31	500	212	67	17
Cigarette & other	1.43	1.35	1.06	1.13	591	307	125	70
one pack or less	1.28	1.21	1.10	1.04	358	208	100	53
more than a pack	1.77	1.77	0.93	1.42	233	99	25	17
Cigar (including cigar and pipe)	1.01	0.98	0.94	1.14	129	197	133	56
less than 5 cigars	1.12	0.96	0.86	0.98	99	142	91	40
5 or more cigars	0.75	1.04	1.20	1.78	30	55	42	16
Pipe only	1.21	1.08	.91	1.12	57	65	42	28
Less than 10 pipefuls	1.16	1.05	0.97	1.33	29	39	28	24
10 or more pipefuls	1.27	1.08	0.82	0.57	28	26	14	4

Expected number of deaths based on death rates of persons who had never smoked.
Age started is the age of starting to smoke each form of tobacco; for users of two or more forms, this refers to the first mentioned.

TABLE 12. MORTALITY BY SMOKING HISTORY AND GEOGRAPHIC REGION
Ratio of observed to expected number of deaths (all causes) by smoking history, geographic region and maximum amount used.

Smoking History Maximum Amount Used	Mortality Ratio				Observed Number of Deaths			
	North-east	North Central	South	West	North-east	North Central	South	West
Cigarette only	1.47	1.60	1.53	1.55	794	864	636	468
one pack or less	1.28	1.55	1.39	1.50	408	515	311	280
more than a pack	1.73	1.69	1.70	1.62	386	349	325	188
Cigarette & other	1.19	1.39	1.29	1.23	498	587	368	285
one pack or less	1.03	1.21	1.14	1.17	295	342	204	178
more than a pack	1.53	1.77	1.55	1.37	203	242	164	107
Cigar only	.93	1.15	1.16	1.14	119	152	104	57
less than 5 cigars	.96	1.14	.88	1.16	68	90	44	36
5 or more cigars	.91	1.17	1.50	1.11	51	62	66	21
Cigar and pipe	1.02	1.15	1.23	1.04	108	106	81	47
less than 5 cigars	1.00	1.13	1.21	1.00	70	72	51	30
5 or more cigars	1.06	1.21	1.25	1.13	38	34	30	17
Pipe only	1.07	1.19	1.14	1.02	87	87	57	42
less than 10 pipefuls	1.08	1.02	1.18	1.04	54	48	33	27
10 or more pipefuls	1.07	1.50	1.09	1.00	33	39	24	15

Expected number of deaths based upon the death rates of those who had never used tobacco or had used tobacco only occasionally for each region separately.

TABLE 13. MORTALITY OF SMOKERS FROM BROAD GROUPS OF CAUSES

Ratio of observed to expected number of deaths by type of smoking history and cause [1/]

Smoking History	Ratio of observed to expected deaths2						Number of observed deaths					
	Cancer of lung	Cancer except lung	Respiratory diseases	Cardiovascular	Suicide, accidents	Other diseases	Cancer of lung	Cancer except lung	Respiratory diseases	Cardiovascular	Suicide, Accidents	Other Diseases
Never smoked or smoked occasionally only	1.00	1.00	1.00	1.00	1.00	1.00	17	248	24	1017	97	121
Used tobacco	6.00	1.26	1.66	1.27	.92	1.38	312	981	118	3983	294	515
Regular smoker	6.64	1.30	1.80	1.31	.94	1.41	299	877	112	3556	261	458
Cigarette total	8.32	1.30	2.24	1.40	.91	1.48	283	675	101	2887	203	363
Cigarette and others	6.40	1.31	1.52	1.24	.79	1.36	96	290	32	1107	72	145
Cigarette only	9.35	1.30	2.76	1.53	.99	1.57	187	385	69	1780	131	218
Cigarette and cigar	7.00	1.42	1.00	1.23	.94	1.41	21	64	4	227	16	31
Cigarette and pipe	6.29	1.22	2.25	1.32	.75	1.49	44	122	18	518	33	70
Cigarette, cigar, pipe	6.20	1.35	1.25	1.14	.79	1.16	31	104	10	362	23	44
Cigar only	1.50	1.36	0.43	1.00	.92	1.38	6	87	3	271	22	44
Cigar and pipe	2.00	1.48	0.40	1.00	1.11	1.12	6	74	2	211	20	29
Pipe only	1.33	1.03	1.50	1.11	1.07	1.10	4	41	6	187	16	22
Amount unknown	2.00	1.00	0.75	1.06	.71	1.45	6	45	3	196	12	32

1/ Includes underlying causes only.

2/ Expected number of deaths - computed by multiplying the number of person years exposure in each age group for each smoking history category by the age-specific death rates from each cause of death of persons who had never smoked or who had used tobacco only occasionally.

TABLE 14. MORTALITY OF REGULAR CIGARETTE SMOKERS FROM SPECIFIC DISEASES

Ratio of observed to expected number of deaths of persons who regularly smoked cigarettes only.

Disease	No. of Observed Deaths[1]	Ratio to Expected Deaths[2]
Cancer of lung (162,163)	197	9.85
Cancer of prostate (177)	52	2.17
Cancer of mouth,pharynx esophagus (140-148,150)	24	2.18
Cancer of bladder (181)	29	1.93
Cancer of stomach (151)	39	1.86
Malignant lymphomas (200-203)	56	1.70
Cancer, other forms	86	1.30
Cancer of pancreas (157)	31	1.19
Cancer of intestines and rectum (152-154)	94	1.09
Cancer of kidney (180)	16	1.00
Leukemia (204)	23	.89
Pneumonia (480-493)	145	1.61
Bronchitis, emphysema, etc. (500-527)	147	3.27
Arteriosclerotic (coronary heart disease (420)	1,420	1.63
Nonrheumatic chronic endocarditis (421-422)	135	1.55
Hypertension with heart disease (440-443)	279	1.53
General Arteriosclerosis (450)	243	1.46
Hypertension without heart disease (444-447)	82	1.41
Cerebral vascular lesions (330-334)	281	1.33
Chronic nephritis (592-594)	26	1.04
Chronic rheumatic heart disease (410-416)	42	.84
Paralysis agitans (350)	9	.36
Other diseases of liver, gallbladder and pancreas (582-587)	45	1.15
Diabetes (260)	84	1.18
Ulcer of stomach and duodenum (540,541)	51	2.83
Cirrhosis of liver (581)	62	2.95

1/ Includes underlying and contributory causes of death.

2/ Expected number of deaths - computed by multiplying the number of person years exposure in each age group of regular smokers of cigarettes only by the age-specific death rates from each cause of death (including underlying and contributory causes) of persons who had never smoked or who had used tobacco only occasionally.

TABLE 15. RELATIVE IMPORTANCE OF DISEASES CAUSING EXCESS MORTALITY AMONG REGULAR CIGARETTE-ONLY SMOKERS

Cause of Death[1]	Observed Deaths		Observed Minus Expected[2]		Mortality Ratio
	Number	Percent	Number	Percent	
Cancer of Lung	187	6.8	167	16.8	9.35
Respiratory Diseases	69	2.5	44	4.4	2.76
Coronary Heart Disease	1212	43.7	446	44.8	1.58
Hypertensive Cardiovascular Disease	148	5.3	53	5.3	1.56
Other Cardiovascular Diseases	420	15.2	119	11.9	1.40
Cancer except Lung	385	13.9	88	8.8	1.30
All Other Diseases	349	12.6	79	7.9	1.29
Total	2770	100.0	996	99.9	

1/ Underlying causes only.

2/ Expected deaths based on death rates of persons who had never smoked or who had smoked occasionally only.

TABLE 16. MORTALITY OF REGULAR CIGARETTE SMOKERS FROM SPECIFIC CAUSES BY AMOUNT SMOKED

Ratio of observed to expected number of deaths of persons who had regularly smoked cigarettes only by current amount smoked, July 1954-December 1956.

Disease	Mortality Ratio 2/			Number of observed deaths 1/		
	Current amount smoked					
	-10	10-20	21+	-10	10-20	21+
Cancer of lung (162,163)	5.50	10.00	15.80	11	70	79
Cancer of prostate (177)	1.67	2.00	2.33	5	16	14
Cancer of mouth, pharynx esophagus (140-148,150)	1.00	2.50	4.00	1	10	12
Cancer of bladder (181)	1.00	1.83	2.75	2	11	11
Cancer of stomach (151)	4.50	2.00	1.40	9	14	7
Malignant lymphomas (200-203)	0.67	1.91	1.89	2	21	17
Cancer, other forms	1.00	1.39	1.39	7	32	25
Cancer of pancreas (157)	0.67	1.00	2.00	2	9	14
Cancer of intestines and rectum (152-154)	1.22	1.00	1.14	11	30	25
Cancer of kidney (180)	1.50	1.17	0.75	3	7	3
Leukemia (204)	0.33	0.78	1.43	1	7	10
Pneumonia (480-493)	1.70	1.78	1.82	17	57	40
Bronchitis, emphysema, etc. (500-527)	2.20	3.75	3.70	11	60	37
Arteriosclerotic (coronary) heart disease (420)	1.32	1.76	1.75	120	539	387
Nonrheumatic chronic endocarditis (421-422)	1.30	1.68	1.62	13	52	34
Hypertension with heart disease (440-443)	1.32	1.34	1.63	25	87	75
General arteriosclerosis (450)	0.84	1.62	1.46	16	97	57
Hypertension without heart disease (444-447)	2.00	1.10	1.57	12	22	22
Cerebral vascular lesions (330-334)	1.54	1.27	1.46	37	95	73
Chronic nephritis (592-594)	-	1.00	1.14	-	9	8
Chronic rheumatic heart disease (410-416)	0.80	0.94	0.77	4	16	10
Paralysis agitans (350)	0.33	0.11	0.17	1	1	1
Other diseases of liver, gall-bladder and pancreas (582-587)	1.00	0.64	1.50	4	9	15
Diabetes (260)	0.62	0.96	1.39	5	24	25
Ulcer of stomach and duodenum (540,541)	1.50	3.67	2.60	3	22	13
Cirrhosis of liver (581)	3.00	3.14	4.17	6	22	25

1/ Includes underlying and contributory causes of death.

2/ Expected number of deaths based on the death rates from each cause of death (including underlying and contributory causes) of persons who had never smoked or who had used tobacco only occasionally.

TABLE 17. MORTALITY OF REGULAR PIPE SMOKERS FROM SPECIFIC CAUSES

Ratio of observed to expected number of deaths of persons who had regularly smoked a pipe only (lifetime history) July 1954-December 1956.

Disease	No. of Observed Deaths 1/	Ratio of Expected Deaths 2/
Cancer of lung (162,163)	4	1.33
Cancer of prostate (177)	4	1.00
Cancer of mouth, pharynx esophagus (140-148,150)	2	1.00
Cancer of bladder (181)	3	1.50
Cancer of stomach (151)	5	1.67
Malignant lymphomas (200-203)	3	0.75
Cancer, other forms	13	1.62
Cancer of pancreas (157)	3	0.75
Cancer of intestines and rectum (152-154)	10	0.83
Cancer of kidney (180)	1	0.50
Leukemia (204)	3	1.00
Pneumonia (480-493)	12	0.86
Bronchitis, emphysema, etc. (500-527)	15	1.88
Arteriosclerotic (coronary) heart disease (420)	142	1.16
Nonrheumatic chronic endocarditis (421-422)	16	1.14
Hypertension with heart disease (440-443)	19	0.76
General arteriosclerosis (450)	23	0.79
Hypertension without heart disease (444-447)	5	0.62
Cerebral vascular lesions (330-334)	33	0.97
Chronic nephritis (592-594)	1	0.33
Chronic rheumatic heart disease (410-416)	7	1.17
Paralysis agitans (350)	1	0.25
Other diseases of liver, gallbladder and pancreas (582-587)	4	0.67
Diabetes (260)	5	0.50
Ulcer of stomach and duodenum (540,541)	5	2.50
Cirrhosis of liver (581)	6	2.00

1/ Includes underlying and contributory causes of death.

2/ Expected number of deaths based on the death rates from each cause of death (including underlying and contributory causes) of persons who had never smoked or who had used tobacco only occasionally.

TABLE 18. MORTALITY OF REGULAR CIGAR SMOKERS FROM SPECIFIC CAUSES

Ratio of observed to expected number of deaths of persons who regularly smoked only cigars or cigars and a pipe, lifetime history, July 1954-December 1956.

Disease	No. of Observed Deaths 1/	Ratio of Expected Deaths 2/
Cancer of lung (162,163)	13	1.86
Cancer of prostate (177)	28	2.55
Cancer of mouth, pharynx esophagus (140-148,150)	12	2.40
Cancer of bladder (181)	3	0.50
Cancer of stomach (151)	9	1.00
Malignant lymphomas (200-203)	12	0.92
Cancer, other forms	32	1.39
Cancer of pancreas (157)	11	1.00
Cancer of intestines and rectum (152-154)	54	1.54
Cancer of kidney (180)	6	1.20
Leukemia (204)	11	1.22
Pneumonia (480-493)	38	0.93
Bronchitis, emphysema, etc. (500-527)	22	1.00
Arteriosclerotic (coronary) heart disease (420)	373	1.05
Nonrheumatic chronic endocarditis (421-422)	39	0.98
Hypertension with heart disease (440-443)	85	1.16
General arteriosclerosis (450)	74	0.91
Hypertension without heart disease (444-447)	29	1.26
Cerebral vascular lesions (330-334)	92	0.96
Chronic nephritis (592-594)	9	1.00
Chronic rheumatic heart disease (410-416)	8	0.42
Paralysis agitans (350)	-	-
Other diseases of liver, gallbladder and pancreas (582-587)	22	1.38
Diabetes (260)	55	1.90
Ulcer of stomach and duodenum (540,541)	15	2.14
Cirrhosis of liver (581)	22	3.14

1/ Includes underlying and contributory causes of death.

2/ Expected number of deaths based on the death rates from each cause of death (including underlying and contributory causes) of persons who had never smoked or who had used tobacco only occasionally.

DISCUSSION

By: Carl E. Hopkins, University of Oregon Medical School

Statisticians and public health workers alike are in debt to Dr. Dorn and the National Cancer Institute for this painstaking contribution to our knowledge of mortality and its relation to smoking. It is truly remarkable that so little is still known about causes and effects of such a widespread human activity as tobacco use. Physicians, philosophers and poets have for centuries speculated, eulogized and moralized on the subject, without adducing substantial evidence, until Lombard, and separately, Pearl, made in the late 1930's some epidemiologic studies of the effect of smoking on mortality and longevity. The results were not shocking, and it was not until the 1950's when a sharp rise in lung cancer deaths became noticeable here and in England that intensive exploration of smoking as a possible health hazard began. In retrospect it is easy to see that just as in much exploratory experimentation in the laboratory, a main effect had been completely obscured by excessive dilution until a sufficient concentration was reached to make the main effect stand out above the noise level.

Now that we have a main effect standing out clearly, it is also remarkable, tho not unprecedented, that rather unusual standards of credibility are being required for the conclusions of such epidemiologic investigations. While arising out of diverse motivations and confused understandings of the statistician's role, these unusual requirements of "proof" have been a stimulating challenge to the professional statistician and produced some worthwhile introspection on his role and his method.

If we view his role as that of making some sort of rational decision on whether to adopt or reject some incompletely specified but far-reaching public health program, then it is quite clear that he does not yet have values for all of the parameters of the pertinent decision functions. For that matter it is not clear that he even has the necessary tools. But if we view his role more narrowly as that of evaluating the credibility of inference drawn from observations, then it would appear that his usual tools of probability and sampling theory as applied to statistical epidemiology may be presumed to suffice here as they have in so many other similar problems where direct experimental test seemed impossible. I prefer to speak from this latter viewpoint, and will therefore discuss Dr. Dorn's contribution only as to the inference he draws from his observations, viewing wider implications, however important they may be to the public health, as irrelevant in a statistical discussion.

The National Cancer Institute study was ingeniously conceived to eliminate certain ambiguities remaining in earlier studies due mainly to their retrospective design, their susceptibility to bias in selection of smokers and non-smokers, and their size limitations. While being motivated by interest in the lung cancer problem, the study was wisely broadened to cover mortality of all kinds in relation to smoking. Veterans provided a large, defined and apparently suitable population for following in a prospective study. The study as reported shows every evidence of care and great technical skill in planning, execution and interpretation. Few statisticians could really wish to undertake such a task, but most could certainly feel proud of the job reported here. For the first time in all the welter of shaky claims and counter-claims the relations of smoking to overall risk of dying and to risk of dying of particular disease stand out clearly as facts against which hypotheses may be tested.

It should be added quickly that the study says nothing about causes or mechanisms for explaining these facts. This must be done eventually by the epidemiologist and the pathologist who must integrate these and all other known facts into a plausible and preferably testable explanatory hypothesis. Several of these are on the market now, and facts from this study should help in selecting from among them.

Yet there are limitations to these facts. Some of the old bugaboos that hobbled earlier studies are still with us.

1. Retrospective aspects. Altho prospective since 1954, the crucial separation into experimental and control groups by smoking history is in some senses retrospective. The subjects of the study are U.S. uncolored males, who (a) became veterans of the armed forces before 1940, (b) continued their life insurance in force until 1954, (c) survived until 1954, and (d) responded to a smoking history questionnaire in 1954.

The first 3 of these selections are obviously influenced by such host factors as health status, including mental health and personality, social and occupational class, etc.; all of which could be associated with smoking and non-smoking. Since we end up with a reference population only about 16% of which report as non-smokers, a little arithmetic will show that rather moderate biases in the selection of smokers and non-smokers into these successive stages of the pre-study history could result in a disproportionate overloading of extra-healthy persons into the non-smoking part of the 1954 study population, thus giving non-smokers a superior survivorship due not to non-smoking but to health and social factors that made them non-smokers. Until we know more about the social, psychological and physical factors in the individual person that select him into the non-smoking, smoking and heavy-smoking classes, we cannot be certain that these mortality data do not at least partly reflect the mortality prognosis of the groups prior to their selection into the smoking classes. These selective effects may be presumed to have been washed out by the regression phenomenon in the interim from beginning of smoking until entry into this mortality study, but we have no way to know how much to allow for this.

2. The memory problem. Here, as in Kinsey's studies, the question is not whether the questionnaire obtained reliable, repeatable responses, but whether the responses were <u>valid</u> in the sense of corresponding with

objective fact. Some skepticism is required on the accuracy of the memory instrument in a subject as charged with deep psychological content as smoking habits.

3. What of the non-response bias? It is presumed that the final report will explain what effect the 16% non-respondent group could have had on the outcome.
4. What of experimental artifacts? Did the policy holders know they were queried thru their life insurance company? Could this have influenced their responses? Could they have feared to lose their policies or to have premium rates changed as a result of their replies?
5. To what larger population can these results be generalized? They were obtained from U.S. uncolored males, somewhat upper class, with 40% better than average mortality experience. What can the epidemiologist do with this?

These are not detractions. But they are questions and limitations that must be kept in mind in weighing and using the evidence of this study. If the study were to be viewed in isolation one would have to be suspicious of dangerous biases possibly lurking in the background. Its credibility tonite rests not only on the evident craftsmanship of the work itself, but also and perhaps mainly on agreement at crucial points with numerous other independent studies subject to <u>different</u> <u>sets</u> of biases.

In sum, we are indebted to Dr. Dorn and his coworkers for their significant contribution in confirming previous but less well established facts, and bringing out new detail on kind, amount and length of smoking history, as related to specific cause of death. We are another long step ahead with facts which will eventually resolve controversy.

DISCUSSION

By: E. Cuyler Hammond, Sc. D., American Cancer Society

I would first like to congratulate Dr. Dorn on an excellent study well presented. My discussion will be confined to his statistical findings rather than to an interpretation in terms of possible cause and effect relationships.

The two major variables in the study were: 1) status of subjects in successive periods of time (i.e., alive or dead), and 2) past and current smoking habits as ascertained at the start of the study. There was little chance of recording a living man as dead since deaths were verified by death certificate. Some under-reporting of deaths may have occurred due to failure of beneficiaries to make a claim. This would result in an artificially low apparent death rate but there is no reason to suppose that it would bias the findings in respect to smoking. Men who dropped their life insurance policies after the start of the study may be somewhat different from those who retained them, and it is not stated whether all such persons were traced. However, it seems unlikely that this could have had much influence on the findings.

It is more difficult to estimate the amount of error in reporting of smoking habits. The most serious error would be confusion between smokers and non-smokers. However, the subjects certainly knew the answer to this simple question and the only problem is how many of them deliberately or inadvertently gave a false answer. Conceivably, a few smokers claimed to be non-smokers because they were afraid that it might affect their insurance status. Unless confined to the healthiest subjects, an error in this direction would tend to reduce the apparent relationship between smoking habits and death rates. The possibility of errors in the other direction seems less likely considering the detailed questions asked of smokers. Incidental to our prospective study,[1] we questioned 45,000 subjects about their smoking habits in 1952 and again in 1954. Of those who in 1952 said that they had never smoked, less than 1% said that they were smoking cigarettes regularly when questioned in 1955. This suggests that it is rare for a non-smoker to record himself as a smoker.

Finkner[2] and his associates have studied the accuracy of reporting daily amount of cigarette smoking, making use of lighters with counters attached and cans in which individuals were asked to deposit their cigarette butts. An analysis of the basic data they present for 76 regular cigarette smokers among workers in a research laboratory showed the following: 80% were classified in the same amount category by lighter counts as by questionnaire answers and all the remainder were in adjacent categories when classified independently by the two methods. A similar study of 52 regular cigarette smokers among office workers showed essentially the same thing.

Todd and Laws[3] made a study for the Tobacco Manufacturers' Standing Committee of Great Britain on the accuracy of information on current smoking habits obtained by questionnaires. They came to the following conclusion:

> ".....since in most tables in this paper informants are classified into broad groups consisting of non-smokers of cigarettes and of smokers of 0-4, 5-14, 15-24, and 25 or more cigarettes a day, the percentage of smokers, particularly in view of the proportion smoking 10 or 20 cigarettes a day, likely to be classified wrongly even with a 15% error is negligible."

Considering this evidence, it seems unlikely that errors in either of the two major variables could have introduced much error in the apparent association between current smoking habits and total death rates reported by Dr. Dorn. Estimates of past smoking habits are probably less reliable than are estimates of current smoking habits.

The population chosen for study was well defined, namely the 291,800 veterans who held United States Government Life Insurance at the end of 1953. Almost all of them were traced but 31.8% of them failed to answer the questionnaire on smoking habits in 1954. About half of the non-responders of 1954 who were still living in 1957 answered a smoking questionnaire at that time.

The responders of 1954 had a death rate far below that of the general white male population of the United States. This was due to two factors: 1) The insured veterans as a whole (responders and non-responders together) had a death rate only about 70% of that of the general white male population, and 2) the responders were self-selected in such a way that their death rate was considerably less than that of the non-responders. Dr. Dorn reasonably postulates that the difference in death rates between responders and non-responders was due to failure of critically ill persons to answer his questionnaire. The rapid drop in the death rate of the non-responders during the subsequent few months gives strong support to this explanation. However, there is no way of knowing for sure whether this selective factor operated differentially on smokers and non-smokers. If such a differential did exist, it would bias the findings at least temporarily. The bias would have been maximum during the first few months of the study and then would have diminished rather rapidly thereafter. I imagine that this was one of the reasons that Dr. Dorn omitted the first six months of experience from the data he presented today. The trend in mortality ratios during the first several years of the study will provide a good indication of the degree and direction of this

bias if indeed any such bias existed. I should add that there is no reason to suspect that a serious bias of this type occurred.

From the standpoint of possible selective bias, Dr. Dorn's analysis of death rates in relation to duration of cigarette smoking and age at which smoking began presents an interesting problem. Candidates for the armed services are screened by medical examinations which eliminate those with impairments. Some of these impairments (e.g., cardio-vascular fitness,[4] chronic severe cough,[5] etc.) may be associated with smoking habits. Subjects who started smoking at an early age were smoking prior to the time of medical screening, while most of those who started smoking later in life were non-smokers at the time of medical screening. This raises a question as to the validity of the comparison of later death rates in these two groups. In all probability, the selective bias would have worn off long before the start of Dr. Dorn's study. However, some evidence on this point would be most interesting.

It is of interest to compare Dr. Dorn's findings with the findings in two other prospective studies on smoking in relation to death rates.

Doll and Hill[6] mailed questionnaires to all physicians registered in Great Britain and 68% of them replied. The subjects have been traced ever since 1951. A direct comparison cannot be made with Dr. Dorn's findings because Doll and Hill used a different method of classifying smoking habits. Nevertheless, the findings in relation to smoking and death rates were essentially the same in all major aspects, taking sampling variation into consideration. That is, they found that death rates increased with amount of smoking. This was true not only for lung cancer but for many other diseases including coronary artery disease, lung diseases other than lung cancer, peptic ulcers, and cancer of several sites.

Hammond and Horn[1] studied death rates in a sample of 187,783 white American males who were in age group 50 to 69 at the time of selection in 1952, and were followed for 44 months. The subjects were friends and relatives of volunteer workers of the American Cancer Society. Obviously ill men were deliberately excluded. We have reason to believe that very few men (probably less than 3%) refused to fill out a questionnaire when asked to do so. Thus, there was very little self-selection on the part of the subjects. The death rate in the last year of the study was about 81% of that of the general white male population of the United States as compared with 70% for the total group of veterans and less than 70% for the responders studied by Dr. Dorn. The very lowest socio-economic groups were somewhat under-represented and the institutionalized segment of the general population was almost entirely excluded.

Table 1 shows mortality ratios by type of smoking history for Dorn's study and for the Hammond and Horn study. I think that it is fair to say that the two sets of figures are in essential agreement. The major difference is that the mortality ratios tend to be a bit lower in Dorn's study than in our study. It would have indeed been surprising if they had agreed any more closely considering that the subjects were selected in different ways, the age distributions were not the same, and the subjects were followed for different lengths of time. The fact that all veterans were originally selected on the basis of medical examinations which screened out those with serious impairments and the factor of self-selection of the subjects may have had some influence on Dr. Dorn's findings. The deliberate exclusion of obviously ill persons probably had some influence on our findings.

Table 2 shows mortality ratios by current amount of cigarette smoking. Again the figures from the two studies are in essential agreement although not identical.

Table 1

MORTALITY RATIOS OF SMOKERS AND NON-SMOKERS
DORN STUDY COMPARED WITH HAMMOND AND HORN STUDY

Smoking History	Current Use					
	Total		Smokes		Does Not Smoke	
	Dorn	H.& H.	Dorn	H.& H.	Dorn	H.& H.
Never Smoked	1.00	1.00	-	-	-	-
Occasional Only	.98	1.09	-	-	-	-
Cigarette Only	1.58	1.68	1.65	1.74	1.39	1.43
Cigarette & Other	1.29	1.43	1.35	1.50	1.21	1.32
Cigar Only	1.07	1.22	.94	1.09	1.44	1.68
Pipe Only	1.10	1.12	1.05	1.09	1.25	1.32

Table 2

MORTALITY RATIOS OF CIGARETTE SMOKERS BY CURRENT AMOUNT SMOKED
DORN STUDY COMPARED WITH HAMMOND AND HORN STUDY

Current Cigarettes Per Day	History of Cigarettes Only		History of Cigarettes and Other	
	Dorn	H.& H.	Dorn	H.& H.
1-9	1.29	1.34	.95	1.27
10-20	1.66	1.70	1.37	1.49
21-39	1.77	1.96	1.72	1.70
40+	1.99	2.23	1.79	1.83

Now let us turn to a consideration of death rates by cause of death as reported by physicians. Table 3 shows mortality ratios for men with a history of regular cigarette smoking by broad categories of underlying causes of death. In other words, each death shown here was classified according to the cause which in the opinion of the physician was the major factor leading to death. The figures from the two studies are in fairly good agreement. There appears to be no association between cigarette smoking and death rates from violence, accidents, and suicide. In each of the other categories, the mortality ratios are appreciably greater than 1.00 in both studies.

Table 4 shows mortality ratios for men with a history of regular cigarette smoking only for a number of specific diseases. While doctors are required to name one disease or injury as the underlying cause of each death, additional diseases are often specified as contributing to death. Dr. Dorn chose to classify each death according to every disease mentioned, thereby putting many deaths into two, three, or more categories. In our original study, we chose to classify each death by underlying cause only. These figures are shown in the last column of this table. For this presentation, I had our cases reclassified according to the procedure used by Dr. Dorn and the figures are shown in the middle column. Except for pneumonia (which frequently occurs in the terminal stages of other diseases) and peptic ulcer (which is a common chronic disease with a low case fatality rate), the two methods of classification yielded about the same mortality ratios.

The finding which is likely to attract the greatest attention is the very high mortality ratio for lung cancer. This has now been found in such a large number of independent studies and has been discussed so often that it is pointless for me to say anything more about it at this time.

In my opinion, much more attention should be given to the findings in relation to other diseases. Mortality ratios, as shown here, only tell a part of the story. A table showing differences in observed and expected number of deaths would give a very different impression of the findings. The association between cigarette smoking and lung cancer accounts for only a small part of the excess deaths among cigarette smokers as compared with non-smokers. In our study, 52% of the excess deaths associated with cigarette smoking were

Table 3

MORTALITY RATIOS OF MEN WITH A HISTORY OF REGULAR CIGARETTE SMOKING
DORN STUDY COMPARED WITH HAMMOND AND HORN STUDY

Underlying Cause of Death	Dorn	Hammond and Horn
Cancer of Lung	8.32	10.73
Other Cancer	1.30	1.51
Respiratory	2.24	2.85
Cardiovascular	1.40	1.57
Accidents, Violence, Suicide	.91	.94
Other	1.48	1.29

Table 4

MORTALITY RATIOS OF MEN WITH A HISTORY OF REGULAR CIGARETTE SMOKING ONLY
DORN STUDY COMPARED WITH HAMMOND AND HORN STUDY

Disease	Underlying & Contributory		Underlying Only
	Dorn	Hammond & Horn	Hammond & Horn
Cancer of Lung	9.85	12.75	12.45
Cancer of Mouth, etc.	2.18	5.00	4.76
Cancer of Prostate	2.17	1.85	1.73
Cancer of Bladder	1.93	2.33	2.37
Cancer of Stomach	1.86	2.19	2.19
Cancer of Rectum, Colon	1.09	0.74	0.70
Coronary	1.63	1.84	1.83
Rheumatic Heart	0.84	1.13	0.98
Cerebral Vascular	1.33	1.38	1.36
Bronchitis, Emphysema	3.27	3.27	3.25
Pneumonia	1.61	2.76	3.77
Peptic Ulcer	2.83	3.94	4.64
Cirrhosis of Liver	2.95	1.97	2.21
Diabetes	1.18	0.84	0.84

accounted for by cases in which coronary artery disease was specified as the underlying cause of death by the certifying physicians. The corresponding figure for Dr. Dorn's study is 45%.

To me, the most striking finding in all three of the prospective studies is that death rates attributed to a number of diverse diseases were found to be higher among cigarette smokers than among non-smokers. The question is whether this can be attributed to errors of some sort which occurred in roughly the same way in all three studies. Faulty diagnosis of cause of death appears to be the most likely source of serious error.

If we accept the evidence that total death rates are higher among smokers than among non-smokers, then it follows that death rates from at least one specific disease must be higher among smokers than among non-smokers. Considering the number of cases involved, lung cancer cannot possibly account for all of the difference, so one or more other diseases must be involved. If, because of errors in diagnosis, Dr. Dorn's figures give an over-estimate of the mortality ratio for certain diseases, it follows that they give an under-estimate for some other diseases.

The diagnosis of cancer is seldom wrong since it is microscopically proved in 80% or more of the cases so reported. An unknown number of cases are missed. The greatest chance for error here is erroneous diagnosis of primary site of the disease which is sometimes little more than a guess. Conceivably, lung cancer occurring in cigarette smokers is often diagnosed as primary cancer of some other site. If so, it is possible that cigarette smoking is not associated with cancer of primary sites other than the lungs. In that event, the figures presented are an under-estimate of the degree of association between cigarette smoking and lung cancer. More evidence will be required to test this hypothesis.

Pneumonia can occur secondary to lung cancer and the symptoms of lung cancer can be confused with symptoms produced by bronchitis and other lung diseases. If confusion of this type accounts for the apparent association between cigarette smoking and pneumonia, bronchitis, and emphysema, it follows that the death rate from lung cancer is badly under-reported in general mortality statistics and that the association between cigarette smoking and lung cancer is greater than shown by Dr. Dorn's figures.

Evidence from a well controlled experiment on human subjects[7] as well as evidence from animal experiments [8] supports Dr. Dorn's findings of an association between cigarette smoking and death rates from peptic ulcer.

Perhaps the most interesting finding is the apparent association between cigarette smoking and death rates from coronary artery disease. While the mortality ratio is only moderately high, the absolute difference in death rates is very large. If correct, this finding is extremely important since coronary artery disease is reported to account for about 37% of all deaths among white males in the United States over the age of 50.

Many people who have studied the problem are under the impression that recording of coronary artery disease is one of the least reliable diagnoses reported on death certificates in the United States today. It is often put down in cases of sudden death when the certifying physician has had little or no opportunity to make a proper diagnosis. If

erroneous diagnosis entirely accounts for the apparent association between cigarette smoking and this disease, then the association between cigarette smoking and some other disease or diseases is grossly under-estimated in the figures presented here. This is necessarily so because of the large number of cases involved.

In closing, I would like to make a comment on the use of the mortality ratio as an index of association in studies of this type.

The probability of a person dying from a particular disease at a particular moment depends upon a multiplicity of factors operating at various times during the entire life span of the individual. There are two reasons for believing this to be true. First, there is reason to suppose that many causes of death are selective as to type of individual; so a high death rate from any one cause may alter the composition of the surviving population. Second, and aside from this, a number of factors (including susceptibility, exposure, treatment, etc.) influence the probability of death from any particular disease. At least this appears to be true of almost every disease so far studied.

Now let us assume that a particular factor can have an influence upon death rates from a particular disease. In the exposed population, the death rate from that disease is dependent not only upon that factor but upon other factors as well. The same is true of the unexposed population (provided the factor in question is not necessary for death from the disease). If this be true, then the mortality ratio for a particular disease and a particular factor depends upon the total set of conditions and will vary under other sets of conditions. For example, it is virtually certain that long exposure to a heavy concentration of uranium dust results in lung cancer in a very large proportion of people so exposed whether or not they smoke. Therefore, if a study such as Dr. Dorn's were to be carried out among uranium miners, it is virtually certain that the mortality ratio (smokers compared with non-smokers) would be far smaller than he has reported.

Assuming that a causal relationship actually exists, it seems to me that the mortality ratio is a reasonably good index of the contribution of a particular factor to the variance in death rates under a particular set of conditions. If the mortality ratio is small, it simply means that the factor contributes little to the variance under that set of conditions; it might make a much larger contribution under some other set of conditions. In other words, a mortality ratio of any size greater than 1.00 may reflect a causal relationship.

Now suppose that it is unknown whether a causal relationship exists. A mortality ratio above 1.00 suggests the possibility of a causal relationship but it is necessary to rule out other possible explanations before arriving at that conclusion. It seems to me that if the mortality ratio is very large, then it should be relatively easy, with appropriate studies, to determine whether some other associated factor or selective bias accounts for the relationship. This is more difficult if the mortality ratio is small.

In the case of smoking and lung cancer, we have evidence of an extremely high mortality ratio but a relatively small absolute difference in death rates (cigarette smokers compared with non-smokers). In the case of coronary artery disease we have the appearance of a much smaller mortality ratio but a very large difference in death rates (cigarette smokers compared with non-smokers). The latter may be as truly a reflection of a causal relationship as the former. Furthermore, if we had more accurate information on diagnosis and greater knowledge of the total situation, we might find that in a particular segment of the population the mortality ratio for coronary artery disease (cigarette smokers compared with non-smokers) is as high as the mortality ratio for lung cancer. I do not assert that this is so, but considering the magnitude of the difference in death rates I suggest that it is an important field for further investigation.

REFERENCES

1. Hammond, E. C. and Horn, D.: "Smoking and Death Rates--Report on Forty-Four Months of Follow-Up of 187,783 Men." Journal of the American Medical Association, 166:1159-1172 & 1294-1308 (Mar. 8 & 15, 1958).
2. Finkner, A. L.; Horvitz, D. G.; Foradori, G. T.; Fleischer, J.; and Monroe, J.: "An Investigation on the Measurement of Current Smoking by Individuals." Institute of Statistics Mimeo Series #177, Univer- of North Carolina (July 1957).
3. Todd, G. F. and Laws, J. T.: "The Reliability of Statements about Smoking Habits." Research Papers #2, Tobacco Manufacturers' Standing Committee, 6-10 Bruton Street, London W. 1, England (1958).
4. Diehl, H. S.: "The Physique of Smokers as Compared to Non-Smokers." Minnesota Medicine, 12:424-427 (July 1929).
5. Hammond, E. C.: "Smoking in Relation to Lung Cancer and Other Respiratory Diseases." To be published.
6. Doll, R. and Hill, A. B.: "Lung Cancer and Other Causes of Death in Relation to Smoking." British Medical Journal, 2:1071-1081 (November 10, 1956).
7. Doll, R.; Jones, F. A.; and Pygott, F.: "Effects of Smoking on the Production and Maintenance of Gastric and Duodenal Ulcers." Lancet, 1:657 (March 29, 1958).
8. Toon, R. W.; Cross, F. S.; and Wangensteen, O. H.: "Effect of Inhaled Cigarette Smoke on the Production of Peptic Ulcer in the Dog." Proceedings Experimental Biological Medicine, 77:866-969 (August 1951).

DISCUSSION

Robert C. Hockett, Ph.D.
Associate Scientific Director
Tobacco Industry Research Committee

I want to state at once and emphatically that I am not a statistician. Hence, I am not at all competent to comment upon Dr. Dorn's paper on the technical level of the professional statistician. This fact was pointed out to your committee at their first approach, and I also took the liberty of suggesting at least five other persons whom I considered far better qualified for the present assignment. However, since your Chairman was quite insistent that I should make this appearance, I warned him that my comments would necessarily be those of an amateur and non-specialist.

My personal orientation is essentially that of the experimentalist and my reflections upon these figures that Dr. Dorn has presented therefore tend naturally to revolve around their meaning and challenge to the man in the laboratory.

It is, of course, not fair and proper to dwell unduly upon the relatively minor indications of these figures where the small numbers of cases available for analysis reduce the reliability of the correlations. Nevertheless, there are a good many individual points where the figures seem to suggest contradictory, irrational or difficultly credible conclusions. Dr. Dorn has called attention to some of these but a few others may perhaps be mentioned in passing. For example, the data seem to indicate that cigarette smokers, as a group, die of stomach cancer more often than non-smokers. Yet, the light smokers show a higher mortality from this disease than medium smokers, and medium smokers a higher mortality than heavy ones. One wonders whether an extrapolation of the trend line would indicate some level of very heavy cigarette smoking at which the mortality from stomach cancer would equal or fall below that for non-smokers. Data from other studies have already shown that the mortality from stomach cancer in general has been declining for a number of years during a period in which cigarette smoking was increasing sharply. In view of the fact that a great deal of the extraneous material inhaled into the lungs is eventually swept out by the flow of mucous, received into the mouth and swallowed, the figures for stomach cancer seem to present an anomalous picture.

A similar situation appears to exist in the case of cancer of the kidney.

Also, the higher mortality of cigarette smokers who reduced their levels of smoking as compared to those who continued at their former rates, appears anomalous, unless the reduction was indeed due to poor health as postulated.

Another puzzling point is that so large a number of cigars, eight or more, must be smoked daily before any effect upon the mortality ratios is found.

However, as I have stated, comment should properly be concentrated upon the largest and most pronounced trends and relationships that have been reported. The most remarkable observation in this study as well as in its predecessors is the great difference between cigarette smokers on the one hand and cigar or pipe smokers on the other, with respect both to total mortality rates and the mortality from a number of individual diseases. Our knowledge of the chemical composition of pipe and cigar smoke as compared to that of smoke from cigarettes is still quite inadequate for the kinds of comparisons that are indicated by these striking statistical differences among various types of smokers. We do know that pipe and cigar tobaccos tend to run higher in nicotine content than the usual blends of cigarette tobaccos.(1) It is reasonable to suppose that nicotine absorption is of the same order for pipe and cigar smokers as for smokers of cigarettes. Studies on this point are under way which may help in the interpretation of the statistical comparisons.

Very deficient also is our present information on the relative biological activity of these several different types of tobacco smoke. A few studies have reported activity for cigar smoke condensates comparable to that of cigarette smoke condensates as measured by the skin reactions of sensitive mouse strains.(2) Comparisons of this kind, however, have been delayed and handicapped by a lack of reliable and particularly of rapid bio-assay methods sufficiently standardized to be capable of interpretation. A great deal of work is currently under way in the effort to provide better assay techniques for use in the resolution of such questions.(3) There is also a great dearth of information about the mechanics of smoking by users of pipes and cigars, as well as about the kinds of differences that exist among cigarette

smokers in their methods of smoking. The greatest question attaches to the problem of inhalation since this would presumably affect actual physiological dosage of smoke.(4) While it is easy to assume that cigarette smokers inhale more often and more deeply than the users of pipes and cigars, and to assume that inhalation differences account for the statistical differences in disease and mortality statistics, we actually have very little information on this point.

Dr. Dorn's figures tend strongly to reinforce the conviction that fuller knowledge of the comparative chemistry of these several smokes, better methods of bio-assay, and improved knowledge of smoking mechanics, particularly with respect to smoke inhalation, may provide the answers to some key questions. I would be doubtful whether inhalation practices can ever be studied adequately by questionnaire methods. Better information seems likely to be provided by direct mechanical measurements of mouth and chest movements, and perhaps by assay of body fluids for metabolites of nicotine or other smoke ingredients that may serve as indicators of the degree of overall smoke contact with the tissues.

All the problems mentioned above are either included in the present program of the Scientific Advisory Board to the Tobacco Industry Research Committee or are in some stage of planning.(3) Solutions to these problems should provide tools eventually capable of showing whether the statistical comparisons of pipe and cigar smokers to cigarette smokers present a truly anomalous picture or whether they can be given a rational and coherent interpretation. It would not seem to me that further statistical observations alone could carry us much closer to a final solution.

The question has been raised many times before whether and to what extent the psychological and physiological make-up of individual persons may determine whether they remain non-smokers or whether they adopt the use of pipes, cigars or cigarettes. In a society where the opportunities and inducements to use these products are virtually universal, the determinants of use or non-use, and the choice of form might be expected to reflect prominently the family and group mores as well as such psycho-physiological factors. The latter might be expected also to play a role in determining the frequency and manner of smoking within the several categories of tobacco use.

Evidence of personality differences among heavy smokers and non-smokers has been found by Heath in an analysis of data collected, largely by him, over a period of many years on a group of Harvard students.(5) Some indications of sociological factors in smoking patterns have been provided by McArthur in a further analysis of those data.(6) Such data suggest that the various patterns of smoking are not distributed among the population at random but that they tend to select various types from among the general population and thus to separate groups that may have differing inherent life and disease expectancies. Similar evidence of selection has been provided by Sir Ronald Fisher through study of identical and non-identical twins. He found that identical twins are far more often similar in their smoking habits than non-identical twins even though reared apart in different environments.(7)

Further evidence along such lines has been published by Dr. Caroline Bedell Thomas from a study of the medical students at the Johns Hopkins University. After first establishing to her satisfaction that certain cardiovascular conditions show a familial tendency, she has shown that young men with a family history of such conditions are more often smokers than those lacking such a history.(8) Although these young men do not now have any disease of the heart or arteries, it is presumed that they constitute a selected group that may eventually show a high incidence of disease in this category. Of course, if that time arrives, the relatively heavy cigarette smoking of this group will show a correlation with the disease incidence. I do not suggest, naturally, that this relationship furnishes the complete explanation for the observed correlations.

Dr. Dorn's data on the association of smoking with ulcer of the stomach and duodenum are especially interesting to me on account of the experimental studies that have recently been conducted in this field. He shows a ratio of actual to expected deaths from these diseases among regular cigarette smokers of 2.83, which reflects one of the relatively strong associations. While the etiology of ulcer cannot be said to be fully understood, it is widely considered that gastric hypersecretion and hyperacidity are important contributory factors especially since direct control of these conditions is often effective in palliation. The effects of cigarette smoking on gastric secretion have been studied rather extensively in several experimental studies of human subjects

in projects sponsored by our Scientific Advisory Board.(9) In general, no significant differences have been found between the responses of persons with an ulcer history and those lacking such a history. Indeed the smoking of cigarettes by subjects in either category produced no significant changes in the several gastric functions that were measured including volume, density, viscosity, acidity or pepsinogen content of the secretions or gastric motility. The absence of significant response to nicotine absorption, therefore, fails to sustain any hypothesis that smoking contributes directly through physiological action to the etiology of ulcers. The observations tend rather to recall the prevalent concept that ulcer is an anxiety or stress disease containing a strong psychosomatic component. If this is truly the case, one wonders whether the statistical association of smoking and ulcer may not prove to be due to the fact that the candidates for ulcer are selectively inclined to seek solace from smoking to a greater degree than the average population.

According to Dr. Dorn's statistics, cirrhosis of the liver shows an even stronger association with cigarette and cigar smoking than ulcer. This is a disease now widely, if not generally, believed to be related to nutritional deficiency, particularly to the lack of nutrients providing labile methyl groups, in the presence of relatively high calorie intake. Cirrhosis has long been associated with alcoholism and the hypothesis that it might be caused primarily by the direct action of some alcohol metabolite was formerly in vogue. At present the concensus appears to be that the consumption of alcohol has no direct causal relation to liver cirrhosis. The association is regarded to be indirect and to reflect the fact that alcoholics often neglect to consume diets adequate in choline and other methyl donors to balance the calorie content of their alcohol consumption.

Since an association between heavy smoking and heavy alcohol consumption has also been reported by several investigators, it would appear likely that the association of smoking with liver cirrhosis may be removed by a still further step from a direct causal relation to this disease. Discussion of cirrhosis recalls the recent study of Trieger and his collaborators on cancer of the tongue.(10) In a series of about one hundred cases, the factors of malnutrition, syphilitic infection, local irritation, alcohol consumption and smoking were studied. The salient observation of this team was that a combination of three or more of these factors was found in a large majority of the tongue cancer cases. In only a small minority of cases was a single factor such as smoking found to be present alone. If tobacco use had been the only factor studied, this research would have appeared as still another in which an association between tobacco use and a disease was reported. In actuality the study placed some emphasis upon the prevalence of incipient cirrhosis as a factor in mouth cancer and implied that smoking might function as a non-specific minor contributing influence through local irritation.

I have dwelt somewhat upon these several diseases because they all illustrate cases where the initial statistical association has received some degree of elucidation through experimental study or additional statistical data. Several of them illustrate the usual complexity of etiological pictures and in some the association with tobacco use has been shown to be indirect, incidental or accidental. The moral is that statistical association alone is not able to indicate whether a specific factor is actually a part of the causal complex or to distinguish between a direct and major factor and one that is involved in an indirect, incidental or accidental manner. The distinction will usually have to come, in the end, through direct experimental investigation that can trace out the etiological picture step by step.

This brings us to the consideration of lung cancer, which though it accounts for far less mortality than the cardiovascular diseases, stands out as the disease which, according to Dr. Dorn's ratios, shows by far the greatest difference in incidence among the smokers and non-smokers of cigarettes. Nevertheless, a striking feature of the statistics is that the great majority even of the heaviest smokers do not develop the disease, whereas some non-smokers do. These facts alone appear to me to constitute quite incontrovertible evidence of the complex nature of the etiology of this disease. Evidently some unknown series of factors, either extrinsic, intrinsic, or both, must be combined in the proper manner to permit development of the condition designated as primary bronchogenic carcinoma.

These several factors may be likened to the elevations and depressions on the key to a modern lock. A whole series of such elevations and depressions must be aligned in the proper order and must have the proper relative dimensions if the key is to turn in the lock. So it is with the numerous etiologic factors that enter into the web

of causation of any disease. It is unlikely that many of the elevations and depressions in the key to lung cancer have as yet been identified or evaluated, and it is certainly reasonable to suppose that interaction of factors is involved. Influences which are incapable of producing the disease when acting singly, may well do so when operating synergistically.

There has been much rather pointless semantic debate over the relationship between statistical associations and the concept of causation. The practical issues, however, seem to me to be actually rather clear. The universal objective is to find the simplest, most effective, and also cheapest and least disturbing method of interrupting the etiologic chain and obtaining effective prevention of lung cancer or reduction in incidence.

Malaria has been associated for centuries with the proximity of stagnant water. Even when the disease was erroneously considered to be "caused" by inhalation of poisonous miasma (whence the name), some degree of control was attained occasionally by draining a swamp or by removal to higher ground. It is conceivable that malaria might even have been eradicated by heroic measures along such lines, but this proved hardly practical.

When the role of the female Anopheles mosquito was explained as a vector, more effective control became available through spraying stagnant water with oil and later with insecticides to kill the mosquito larvae. Complete understanding of the role played by the plasmodium paved the way for better chemotherapy with consequent reduction in the reservoir of human sources for infection of the mosquitoes. Still later, the more detailed study of mosquito behavior revealed their habit of alighting on a nearby surface for a siesta immediately after every meal of human blood. This knowledge revealed the efficacy of spraying all house walls in endemic areas, with insecticides to destroy potentially infectious mosquitoes before they can leave the premises. This method of control promises to be the one which may finally bring the total eradication of malaria within the realm of practical possibility for the first time in history.(11)

It is obviously pointless to debate the relative degree to which swamps, mosquitoes or plasmodia should be designated as the "causes" of malaria. The point is that as the total web of causation has been elucidated step by step, the control measures have correspondingly improved progressively, in simplicity, practicality and efficacy.

It is well known that pellagra was long ago associated with the consumption of corn, that is maize, as a staple food. For a long time debate continued as to whether this deficiency disease was due to an infection or to a poison in the grain. It is reported that the prevalence of pellagra in certain parts of France actually led to a ban on the growing of maize in that country. This measure is reputed to have been successful in the sense that pellagra diminished or disappeared. In the light of present knowledge, however, the story has a tragic aspect, since a very valuable economic crop was lost to the farmers and a potentially valuable food was sacrificed on account merely of an inadequate understanding of the etiology of pellagra. Now we know that a simple dietary supplement to supply niacin or tryptophane would easily have eradicated the disease and preserved the values of maize cultivation. The moral of these stories should be obvious.

When the statistical relation between cigarette smoking and lung cancer was first reported, it was reasonable to assume as a primary working hypothesis, that some one or more of the various known chemical carcinogens or close relatives of these might be present in cigarette smoke at levels of biological consequence. The search for such agents has now been continued so long in the hands of so many able investigators and with such meager results that many scientists no longer believe it likely that tobacco smoke exerts any significant effect as a direct or specific carcinogen for human tissues. The universal failure to produce lung cancers in animals by simple smoke inhalation reinforces this point of view. Many investigators now expect to find that if tobacco smoke exerts any significant effect in the complex of cancer etiology it will prove to be of a non-specific, indirect, accidental or synergistic character in combination with many others. In the elucidation of such effects and in the search for a satisfactory point d'appui for control, the laboratory and clinic must certainly be primary arenas of activity henceforth.

It seems to me very unfortunate that some of the popular magazines nevertheless have seen fit to publicize the "carcinogenic substance" theory so prematurely, vigorously and sensationally, as to create a strong public demand for supposedly remedial measures which are actually of a wholly uncertain efficacy.

In my opinion, the clue hunt has only begun, and certainly needs to be continued and intensified. Through the Veterans Administration hospitals, the national health agencies have access to a relatively

large number of lung cancer cases. I hope that opportunity will be found to search the medical and personal histories of these patients very exhaustively for any common elements that may be revealed. Dr. Walter Finke(12) and Dr. Sheldon Sommers(13) have already made fruitful approaches along such lines. The discovery of common elements among lung cancer victims should provide additional clues for test and verification in the clinic and laboratory as possible factors in the causal complex. Identification of such contributory factors may well provide a simple and effective method of control or prevention.

Bibliography

1. Garner, W.W. The Production of Tobacco. Philadelphia, 1951. p. 430ff.

2. Croninger, A.B., Graham, E.A., and Wynder, E.L. Experimental Production of Carcinoma with Tobacco Products V. Carcinoma Induction in Mice with Cigar, Pipe and All-Tobacco Cigarette Tar. Cancer Research, 18, 1263 (1958).

3. Little, C.C. 1958 Report of the Scientific Director. Tobacco Industry Research Committee. New York (1958).

4. Schwartz, D., and Denoix, P.F. L'enquête française sur l'étiologie du cancer broncho-pulmonaire. Rôle du tabac. Sem. Hôp. Paris 33, 424 (1957).

   Todd, G.F. Tobacco Manufacturers' Standing Committee, Research Papers, No. 1. Statistics of Smoking. London (1958).

5. Heath, C.W. Differences Between Smokers and Nonsmokers. A.M.A. Archives of Internal Medicine, 101, 377 (1958).

6. McArthur, C., Waldron, E. and Dickinson, J. The Psychology of Smoking. J. Abnormal and Social Psychology. 56, 267 (1958).

7. Fisher, R.A. Lung Cancer and Cigarettes? Nature 182, 108 (1958).

   Fisher, R.A. Cancer and Smoking. Nature. 182, 596 (1958).

8. Thomas, C.B. Familial and Epidemiologic Aspects of Coronary Disease and Hypertension. J. Chronic Dis. 7, 198 (1958).

9. Cooper, P. and Knight, J.B., Jr. Effect of Cigarette Smoking on Gastric Secretions of Patients with Duodenal Ulcer. N.E.J. Med. 255, 17 (1956).

   Cooper, P., Stein, H.L., Moore, G.F. and Harrower, H.W. Effect of Cigarette Smoking on Excretion of Uropepsin and Concentration of Plasma Pepsinogen. R.I. Med. J. 40, 215 (1957).

   Cooper, P., Saltz, M., Harrower, H.W. and Burke, D. Effect of Cigarette Smoking on Dissolved Gastric Mucins and Viscosity of Gastric Juice. Gastroenterology 33, 459 (1957).

   Cooper, P., Harrower, H.W., Stein, H.L. and Moore, Goldwyn F. The Effect of Cigarette Smoking on Intragastric Balloon Pressure and Temperature of Patients with Duodenal Ulcer. Gastroenterology 35, 176 (1958).

10. Trieger, N., Weisberger, D., Ship, I.I. and Taylor, G.W. Cirrhosis and Other Predisposing Factors in Carcinoma of the Tongue. Cancer, 11, 357 (1958).

11. World Health Organization. Malaria Eradication. Geneva (1958).

12. Finke, W. Chronic Pulmonary Diseases in Patients with Lung Cancer. New York State J. Medicine 58, 3783 (1958).

13. Sommers, S.C. Host Factors in Fatal Human Lung Cancer. A.M.A. Arch. Path. 65, 104 (1958).

Discussion

Paul Meier, University of Chicago

The Present Situation

There now exist reports of a number of large scale studies dealing with the association of smoking and mortality. Starting with the observation that the reported death rate for lung cancer had increased markedly in recent years, Doll and Hill conducted a retrospective study which seemed to indicate a relationship between smoking and Cancer of the Lung. The special difficulties in the interpretation of the results of retrospective studies led Doll and Hill, and also Hammond and Horn, to do separate prospective studies, and again smoking seemed to be implicated.

However, serious questions were raised about the interpretation of these prospective studies also. First was the question of sampling technique. In the Hammond-Horn study the subjects were recruited by volunteers and the population sampled is difficult to define or, at any rate, to study. In the Doll-Hill study the population--all British physicians--was well defined, but the proportion of nonrespondents was about as large as the proportion of nonsmokers and, since nothing was known about the nonrespondents, there appeared to be appreciable possibilities for biased selection here also.

Second was the question of possible biases and inaccuracies in responses to the mailed questionnaire and in the diagnoses of cause of death.

Third, as Berkson has emphasized, the results of both prospective studies seemed to indicate that only a minor portion of the increase in death rate among smokers is attributed to lung cancer. It seems that almost all causes of death have elevated rates, and almost two-thirds of the increase is attributed to coronary artery disease, rather than to lung cancer. Berkson suggests that this phenomenon might be taken as evidence that there is something fishy in the methodology or, at least, that the effect of smoking is a general one and not specifically that of a cancer producer. Others, on the contrary, see no incompatibility between the general and the carcinogenic effect. In any event, however, we do seem to be getting more than we had originally bargained for.

At this point it may be well to expose my own prejudices. I suspect that smoking really is bad for one's health, and that people not already slaves to the habit ought to be advised not to take it up. On the other hand, I think that some of the questions raised about the interpretation of the Doll-Hill and Hammond-Horn studies are nontrivial. I don't consider the question settled beyond a reasonable doubt, and I would like to see more conclusive evidence on the subject, if it can be had.

The Problem

Now it seems to me that we can profitably divide the problem of interpretation into two parts.

First, is the association between smoking and mortality (lung cancer and over-all) found in the study populations simply a sampling artifact due to nonrandom selection?

Second, supposing that the observed association in the sample from the study population is not a sampling artifact, does it indicate causation? It might be, for example, that factory workers are subject to industrial hazards which cause cancer and other illnesses and, at the same time, but for quite independent reasons, they tend to be heavy smokers. A population consisting of factory workers and others might then show an association between smoking and mortality which does not represent causation. Indeed, R. A. Fisher urges us not to overlook the possibility that cancer "causes" smoking, in the sense that people in the early and possibly unrecognized stages of their illness might seek the mild narcotic effect produced by smoking. Finally, so far as lung cancer alone is concerned, by now one may have serious doubts about the independence of a lung cancer diagnosis from smoking history. A lung cancer occurring in a smoker may have an appreciably better chance of being diagnosed than one occurring in a nonsmoker.

Contribution of the V.A. Study

So far as the sampling problem is concerned, the V.A. study population has some outstanding advantages over the populations investigated in the previous prospective studies. The population is capable of precise definition, and many characteristics are ascertainable for it without having to contact each individual. Furthermore, the mechanism for learning about the event of death seems almost foolproof and, although the nonrespondents have unknown smoking histories, their deaths are as well followed up as are those of the respondents. If the advantages conferred by use of this

population are fully exploited, our confidence in the assertion that the association between smoking and mortality is not purely a sampling artifact may be greatly strengthened.

In some other respects, however, although the findings in this study agree well with those of Hammond and Horn, the interpretation of them seems open to much the same kind of question. The smoking histories were obtained by mailed questionnaire and, although the physician signing the death certificate was queried, the possibility that a diagnosis of lung cancer may have been directly influenced by the smoking history still remains. Furthermore, since the group as a whole was of higher socio-economic status and had a much more favorable survival experience than does the white male population generally, one may wonder whether the smokers are not economically worse off than the nonsmokers and have a higher mortality for reasons not related to smoking.

For the V.A. population it seems possible to settle the question of whether smokers do or do not have a higher over-all death rate than nonsmokers. Whether this is true for lung cancer as a specific cause is a separate question, and a different type of study might be helpful in resolving it. There seems to be a tendency at present in favor of large studies which depend on mailed questionnaires and routinely reported deaths. However, the effects of smoking appear to be so large that they should be distinguishable in a much smaller study in which smoking histories could be taken by trained interviewers, and subjects might be closely followed in the hope of getting more definitive diagnoses of cause of death.

Some Further Comments

on Dr. Dorn's Report

The study reported here represents a substantial addition to our knowledge of the association of smoking and mortality and was well worth the considerable effort required to carry it out. For this we are much indebted to Dr. Dorn and his colleagues. There are, however, some points which Dr. Dorn's paper does not cover and which I hope he will take up in his final report.

a) I'm not quite clear on just why he chose to define the study population as he did. I infer that the population consists of all subjects with government insurance, excluding certain categories, who were alive or for whom no claim had been filed by January 1, 1954. Those who died in the mailing period, January to June, were carried as nonrespondents. Would it not be more useful to deal with the population surviving at some date after the mailing had been completed?

b) I see that the number of nonrespondents is not negligible--it is about twice the number of nonsmokers. However, although medical information was obtained for them and, I imagine, their ages were known, they are not carried through the analysis as an additional smoking category. Their average death rate, even in the last year, is higher than that of all smokers. This indicates a bias of some kind--perhaps only that the nonrespondents are somewhat older--but this is a point that I should like to see discussed.

c) I hope that we may expect to see many more detailed tables of death rates in which the reader himself may hunt for interesting leads, and that we will find a greatly expanded discussion of the possible or probable magnitude of the inaccuracy and bias which may be present.

I realize that there is no end to the tabulations and discussion which a curious reader might demand, but I do hope that Dr. Dorn will go further than he has in meeting this demand.

In closing let me again congratulate Dr. Dorn for having conceived and carried out this extremely valuable study.

IV

TRANQUILIZING DRUGS - WHAT IS BEING LEARNED

Chairman, Morton Kramer, National Institute of Mental Health

The Social Effects Of Tranquilizing Drugs—Dean J. Clyde, National Institute of Mental Health

Community Adjustment Of Former Mental Patients and Needed Steps For Their Assistance—Else B. Kris, New York State Department of Mental Hygiene

What The Tranquilizing Drugs Are Doing To The Population In Mental Hospitals—Robert E. Patton, New York State Department of Mental Hygiene

Discussion—James C. Munch, Hahnemann Medical School, and Pharmacodynamics, Inc.

THE SOCIAL EFFECTS OF TRANQUILIZING DRUGS

By: Dean J. Clyde, National Institute of Mental Health

With the introduction of tranquilizing drugs a few years ago, we entered a new era in the chemical modification of human behavior. Of course, the venerable tranquilizer alcohol had been known for a long time, but it possessed many disadvantages and had not been found useful for treating severe mental illness. The most dramatic effects of the new drugs were seen in social behavior: patients who were emotionally disturbed and combative calmed down and it was possible for others to associate with them as human beings again.

Medical scientists were puzzled as to how to evaluate these new drugs objectively. Traditional measuring techniques employed in pharmacology and medicine did not adequately embrace the changes seen as a result of tranquilizers. Some psychological techniques seemed closer to the mark, but even they had been designed for other purposes and when they were tried in drug studies often proved uninformative.

Many scientists felt that new techniques were needed for measuring the effects of drugs on mood and social behavior. We wished to study the effects of the drugs from the viewpoint of the psychiatric patient and his family, as well as from the viewpoint of physicians, nurses, and other professional observers.

In order to pinpoint a drug's specific effects on social behavior, a new and simple rating procedure has been developed at the National Institute of Mental Health. It can be utilized by untrained subjects of no more than average intelligence, as well as by professionally-trained experts. The descriptive terms have been aimed squarely at the changes in mood and social behavior which were apparently being produced by the new tranquilizers and stimulants.

The rating procedure consists of a deck of specially-printed, prepunched IBM cards. Each card has on it an adjective which may be relevant to drug effects. For example, some of the adjectives are "friendly," "impulsive," "suspicious," and "amused." The person making the rating sorts these cards into four piles to show to what extent they are descriptive of the patient. The patient may sort the cards himself, a member of the family may sort them to describe the patient, or a professional observer may sort the cards to describe the patient's behavior.

After a deck of cards is sorted, it is picked up and fed directly into a computer which summarizes the results. No laborious hand tabulations are necessary. The computer punches numerical scores for various aspects of mood and behavior, compares the observations of different raters to see if they agree, and shows differences before and after a drug.

Eight investigators throughout the country are now using the new rating procedure and sending the decks of cards back to us so that their results may be compared. Each investigator is studying not only a group of subjects on a new drug, but also one or more control groups at the same time. This of course makes it possible for us to tell whether a new drug is having an effect over and above those changes which may be produced by the passage of time, by the suggestive effects of taking a pill, and what not.

The new rating procedure is already proving to be sensitive to rather subtle drug effects. Let me cite two examples. Each of these investigators will be publishing his results in full.

Dr. Leon J. Warshaw is the director of a large employee health service in New York. He asked a number of office workers in his company if they would like to participate in an experiment, and almost without exception they agreed. He assured them that only standard drugs would be used in small, safe doses. Each of the subjects sorted a deck of our cards before and two hours after swallowing an unidentified pill. One of the pills was meprobamate, otherwise known as Miltown or Equanil, and another was an inert placebo which should have no discernible effect on behavior.

We picked out two extremes among his subjects, based upon their self-ratings before they took the drug. One group was very jittery and tense at the beginning of the experiment, and the other group was unusually calm and relaxed. When we scrutinized the data from the tense group, we found that their ratings showed a slight, but statistically-significant, difference between the meprobamate and the placebo. After taking meprobamate, they said they felt less sluggish and more amused. The group which was calm at the beginning of the experiment felt no difference between the drug and the placebo.

These findings of Dr. Warshaw's may well mean that in evaluating new psychiatric drugs, we must be careful to try them on people who need them. Giving a tranquilizer to a perfectly healthy, relaxed person may not yield a fair indication of its effectiveness for someone else.

Another investigator who is trying our new rating scale is Alberto DiMascio at the Massachusetts Mental Health Center in Boston.

He gave a number of drugs to medical students who volunteered for an experiment. In addition to our deck of cards, they

described their reactions by means of other rating techniques, and a psychiatrist also observed them and rated them on a check list. DiMascio gave the drugs at two dosage levels, a low and a high. All of the rating techniques differentiated between the drugs and the placebo at the high dose, but only the ratings made by the subjects with our deck of cards differentiated at the low dose. DiMascio concluded that the deck of cards was more sensitive to small changes than the other rating procedures.

We are hoping that it has become possible to describe the effects of tranquilizers objectively, and that we will be better able to assess their uses and limitations.

COMMUNITY ADJUSTMENT OF FORMER MENTAL PATIENTS AND NEEDED STEPS FOR THEIR ASSISTANCE

By: Else B. Kris, M. D.

New York State Department of Mental Hygiene

During the past fifty years, different therapeutic approaches in the field of psychiatry have been advoated sporadically. However, it is only recently that the psychiatric profession at large became alert to new concepts in the areas of intra- and extramural psychiatric treatment and the profession at large became aware of the vast psychosociological responses evoked byoour new therapeutic measures. The entire concept of aftercare and follow-up, although known for sometime, has only recently begun to receive greater attention. This all started when new pharmacological compounds, generally known as tranquilizing drugs, were found to be effective in the treatment of psychotic patients. It is true that chemical agents have been used before in the care of mental patients, but it has to be stressed that never before has any drug been used on such overwhelming numbers of individuals. Dr. Paul H. Hoch, New York State Commissioner of Mental Hygiene, recently stated that today some 40,000 of the State's mental hospital patients, or 45% of the total mental hospital population, are receiving drug therapy as compared to other treatment. As a result, large numbers of patients have been released from our mental hospitals and returned to their communities. Since July 1, 1955 when it reached the peak of an ever-ascending spiral, the population of the New York State Mental Hospitals has dropped by some 3,500 patients. This reduction in population was achieved in the face of rising admissions and in spite of an economic recession, the latter factor being known to generally bring about an increase in the state hospital population.

The use of these new drugs has resulted, not only in a significant increase of discharges, but in a tremendous lessening of disturbed behavior in all psychiatric hospital wards where these drugs are used. There was no outstanding increase of either personnel, time, or in the expenditure of fiscal money appropriated. The only appreciable change was in the fact that during 1955-56, 30,000 patients, and lately, as was mentioned before, 40,000 patients received drug therapy in New York State alone, an increase large enough to produce a material effect or release, if the treatment is therapeutically active.

It is interesting to note that the highest discharge rate from our mental hospitals came from those categories of patients who received the most intensive drug therapy. It is also noteworthy that the greatest improvement in rate of dishharges occurred among patients with a long hospital residence, and was less among the newly admitted cases. It was found that during the past four years the state hospital population in New York State decreased by a total of 3 to 40000 patients, instead of an expected increase of about 8,000 patients.

The question has been posed, why this form of therapy has had this result? There are possibly a number of reasons why this was possible, but one of the most outstanding reasons seems to be that for the first time an agent became available which could easily provide continued treatment for a prolonged period of time, on a maintenance basis, thus making it possible for patients to be kept on a level of satisfactory mental functioning for a prlonged period of time. This is due to the fact that for the first time therapy can be carried out outside of the hospital by just administering medication, whereas before, complicated forms of therapy, such as, electroshock, were necessary in order to maintain a certain level of mental functioning. Also, for the first time, we were able to come near our therapeutic goal; that is, to achieve for our patients optimal psychological, social, and vocational capacities in the community.

There is another factor of importance to be added here. While before the advent of our modern drugs, the return rate of patients, who after release from the hospital required renewed hospitalization, amounted to about 33 to 35%, it now, for the first time, became possible to keep this return rate down to between 10 and 20%. However, for the first time, also, the assistance of the general practitioner has become imperative, because extensive drug therapy requires supervision, and this frequently can only be provided by the general practitioner.

In considering the needs for continued care of the former mental patient in the community, several outstanding questions are generally asked:

1. Which patients do require maintenance therapy? After several years of investigation, it can now be stated that most all chronic patients, that is, those who have had extensive hospitalization, or several hospital admissions, must be kept on maintenance therapy, if recurrence of symptoms are to be prevented. Some cases, however, although of a more acute nature, do require maintenance therapy, if the stress situations in the environment are considerable. It appears, for instance, that the excitement of coming home and adjusting to the outside world if better tolerated if maintenance therapy is given for at least the first few weeks, after which it can be again discontinued.

2. How long should maintenance therapy be continued? Chronic cases seem to require infinite

continuation of maintenance therapy. If such therapy, however, is started in order to bridge the period of readjustment, this can be safely discontinued after the patient is comfortable and establishes a new routine of life. Occasionally, when a patient starts employment and the tension interfers with his sleep, maintenance therapy over the first few weeks has proved to be helpful in assistaing the individual in work adjustment.

3. How high should maintenance dosage of the mostly used drugs be? Generally, most of our patients released from mental hospitals are given information regarding the type of drug used effectively inside the hospital, as well as the amount required, so that they can pass this information on to their family doctor, and, not infrequently, the family physician is contacted directly by the hospital doctor prior to the patients' release. These recommendations usually cover existing needs to maintain the level of mental improvement achieved inside the hospital prior to patients' release. If, however, symptoms, such as anxiety, insomnia, hallucinations, delusions, or vague somatic complaints appear, the dosage has to be increased for several weeks With the disappearance of symptoms, the dose can be lowered again to the previous amount. In this way, it has been possible to control cases of impending serious relapse. A single daily dose at bedtime seems in the majority of cases to maintain the leveloof improvement without causing drowsiness that might interfere with work. But it appears imperative for these patients to be under regular observation of a psychiatrist or a general practitioner for control of dosage and of possible complications, and in order to vary the dosage according to individual needs. Another important reason for seeing these patients at regular intervals, is the need to determine whether the drug is actually being taken. Sometimes, patients coming to the office show considerable irritability and tension, and, upon questioning, it will be noted that these patients have not taken their medication for several days. Moreover, stress situations requiring change of dosage can be discovered only if these patients are seen frequently enough.

4. Another question frequently posed is: What is the incidents of complications? The question of greatest concern, particularly in connection with the aftercare of mental patients requiring maintenance therapy outside of the hospital, centered aroudn this possibility of complications resulting from prolonged drug administration. Thus far, however, it seems that there are far less untoward side effects caused by any of these phenothiazine derivatives, even when taken for a prolonged period of time, than might have been anticipated The only side effects observed on a large number of patients kept on maintenance therapy in the community over a period of years were: mild skin rashes, constipation, and occasionally, drowsiness.

It should be mentioned here, that generally the literature on psychiatric complications arising during pregnancy and childbirth reveals that almost all authors agree that rapid and dramatic changes in the course of normal life processes frequently are accompanied by emotional reverberations. Thus, many authors considered pregnancy and childbirth to be immediate precipitants of schizophrenic reaction types. There seems to be general agreement that persons with a history of earlier emotional and personality disorders are liable to become overtly psychotic in reaction to the stress of child bearing and childbirth. In my own study, I have had the opportunity to observe twenty-four women who had been hospitalized between two and ten years, how, after returning to the community, became pregnant again. They went through pregnancy, childbirth, and the post-delivery period on maintenance therapy withone of the psychopharmacological compounds. As a result, the relapse of the mothers into a psychotic condition, was prevented, and, up-to-date, these women continue to function successfully in the community and are able to take care of their infants. These infants, up-to-date, have shown no ill effects and are developing normally. The oldest of these children are now about two and one-half years old. In this group, only two women who had refused to continue with drug therapy, relapsed at the end of their pregnancy and had to be returned to the hospital.

It actually can be stated now that the greatest problem encountered in this form of therapy lies not in any ill effects of such therapy, but in the fact that the patients themselves are frequently reluctant to continue with drug therapy. To overcome this, better indoctrination of families, as well as of the patients themselves, is absolutely necessary.

In addition to these factors, we have to consider what returning our patients to the community really means psychologically and socially. One cannot disregard the fact that, frequently, families show a great deal of anxieity when faced with the necessity of accepting relatives back into their circles sometimes after years of separation which that member had spent in a mental hospitall One has to consider that institutionalization does effect an individual's habits, and that these habits and modes of life, at times, are difficult to be tolereated by the environment. It thus will frequently be important for the family physician to discuss existing problems with all members of the family, as well as with the patient, and to attempt to assist the in straightening out such existing problems.

Among other factors, for instance, families frequently wonder why the patients returned home cannot start working immediately. Although they are willing to admit that after any physical illness an individual requires some time for recuperation and readjustment, they are unwilling to allow the same thing after prolonged hospitalization for reasons of mental illness. Others, again, are rather over-protective, interfering with the patients' attempts to do things on their own; to work, to start any kind of social

life, frequently being afraid of what neighbors might think or say.

In the unfolding of a special study of the social factors involved in the general rehabilitation of mental patients, it became apparent that a greater number of individuals, about 32%, were able to find gainful employment on their own initiative than had been anticipated. These findings are particularly interesting, as many of the patients in this group were hospitalized from over two to thirteen years. While, presently, a number of patients are unemployed, only about 10% are considered not employable at all. About 47% need some form of vocational rehabilitation, either to learn new skills, to brush up on old skills, or to develop work tolerance.

While several of the men and women returned to their former jobs, others found employment through contacting private or public agencies; others, through contacts of one or the other member of the family, or through the daily newspapers. While the employers of those patients who returned to their former jobs knew about their employee's former illness, most of the other patients withhelf this information after having had experienced on two to three occasions when applying for a job, that they were not hired after having told the truth. When asked about recent employment thereafter, these patients stated that they had been physically ill for some time, or gave some other personal reason for their recent period of unemployment.

In a group of three hundred and fifty patients under observation for the past three years, 65% of the men and women gainfully employed, required maintenance pharmacotherapy. In none of the cases did this interfere with their work capacity or work performance, nor was their any accident proneness observed. But it was noticed, that when these patients were exposed to too much pressure on the job, symptoms were quickly reactivated; they disappeared as soon as pressure eased off.

Several patients were able to learn new skills and to go through some type of formal education, as for instance, working toward a high school diploma. Pharmacotherapy did not interfere with these undertakings.

In the majority of cases, a single dose of any of the tranquilizing drugs given at bedtime, was all they needed to help maintain these patients' functioning in their community.

In concluding, it can be stated, that new psychopharmacological compounds are of great value in helping to maintain former hospitalized mental patients in the community. No drug, however, can change the social and economic pressures which prove to be the underlying cause of many of the relapses encountered. Our new drugs can help insulate the former mental patient from the stresses caused by ignorance and prejudice, but the stresses are still there, and every attempt possible should be made to gain better knowledge about their nature and to find means by which they can be eliminated, or, at least alleviated.

To achieve this, we will have to provide sufficient extra-mural services, making available for these patients all forms of psychatric treatment.

Some of the other existing needs which became apparent are:

1. Better education of employers and the public in general to counteract the still existing prejudices.

2. More sheltered workshop facilities where patients could develop work tolerance, as well as be provided with the opportunity to refresh previous skills.

3. Supervised residences of the boarding house type for patients who have neither a family, or where realignment of the family circle has proven to provide unfavorable circumstances for the returning patient.

4. More and better organized family care programs, particularly for younger patients.

5. Social clubs of the A.A. type for these patients to counteract the so frequently existing isolation.

6. Community clinics to provide, where necessary, supervision of pharmacotherapy, and particularly, to provide the so frequently needed supportive psychotherapy.

In taking these needs into consideration, we will be better able to achieve the goal we are striving for- that is, full social, economic and vocational rehabilitation of our mental patients returned to the community.

WHAT THE TRANQUILIZING DRUGS ARE DOING TO THE POPULATION IN MENTAL HOSPITALS

By: Robert E. Patton, New York State Department of Mental Hygiene

For fifty years the trend of the patient population in the New York State Mental Hospitals had been steadily upward. There was a slight drop during the height of World War II, but the rising trend quickly made up for the momentary hesitation. Then after this almost perfectly steady increase lasting for decades, in July 1955 something happened. The trend reversed. The population in the hospitals started to go down.[1] Today, three and one-half years later, the trend is still downward. It has been as steadily downward in the last three and a half years as it was steadily upward in the preceding fifty years. What caused this shift? Can we on the basis of statistical data assign a cause to this dramatic and long hoped for effect?

The patient population in a mental hospital is the net balance resulting from patients entering and patients leaving. It is a function primarily of how long patients stay in the hospital. In any analysis of a change in mental hospital population it is necessary to look first at each factor which affects the size of the population and see how it has changed. The major factors are: first, number of patients admitted; second, number of patients dying; third, number of patients released alive and; fourth, number of released patients returning to the hospital. If the change in patient population can be shown to be the result of a single one of these four factors then the search for a cause can be quickly narrowed to the things which can affect that factor.

Let us look first at the number of patients admitted. Over the years the number of admissions has been steadily increasing. Since 1910 the number of admissions per year to the New York State Mental Hospitals has increased from 7,066 to a high of 23,286 in 1958. From 1950 to 1957 the number remained relatively stationary fluctuating between 20,140 and 21,828. Hence we may say quite categorically that the decrease in resident patients did not occur as the result of a decrease in the number of patients admitted to the hospitals. This does not rule out the possibility that there might have been changes in the characteristics of the patients admitted which could have some effect on their length of stay, but it does mean that the changes which occurred did happen to patients who did enter the hospitals. Hospital data should therefore be able to shed light on the nature of these changes.

In regard to the possibility that the characteristics of the patients admitted have changed, we will discuss this in regard to age and diagnosis groups a little later. The other important characteristic of admissions affecting release is the legal type of admission. In 1955, 74% of the admissions were court certifications. In 1958 only 58% were court certifications. The biggest increase occurred in admissions by Physicians Certificate which increased from 1 to 11 per cent. This method of admission is used primarily for elderly patients who do not object to hospitalization and whose families desire it. These patients would previously have been court certified. This shift should have no effect on releases. The other type of admission that increased was the voluntary certificate. The per cent of patients admitted on this form increased from 7 to 15 per cent. This increase may have had some effect on releases. In absolute numbers the increase was from 1,534 to 3,372 over the four year period.

Table 1

Selected Statistics of Patients in New York Civil State Mental Hospitals

Fiscal Year Ending in	Resident Patients End of Year	Admissions	Deaths
1958	91,191	23,286	9,421
1957	92,409	21,828	8,555
1956	92,862	21,454	8,345
1955	93,314	21,459	8,078
1954	90,893	21,577	8,056
1953	88,868	21,309	8,120
1952	86,298	20,140	7,680
1951	84,608	20,420	7,629
1950	82,906	20,903	7,432
1945	72,700	16,502	6,779
1940	71,160	16,614	5,959
1935	59,828	14,540	5,164
1930	47,330	11,504	4,195
1925	40,281	9,436	3,726
1920	35,848	8,511	3,679
1915	33,155	7,934	3,036
1910	30,445	7,066	2,536

Table 2

Selected Statistics on Movement of Patients

New York Civil State Mental Hospitals by Age Group

Fiscal Year Ending March 31	Resident Patients End of Year	Admissions	Deaths	Releases* Alive	Returns from Convalescent Care	Net Releases Alive
			All Patients			
1958.....	91,191	23,286	9,421	19,334	5,120	14,214
1957.....	92,409	21,828	8,555	18,011	4,896	13,115
1956.....	92,862	21,454	8,345	17,060	4,280	12,780
1955.....	93,314	21,459	8,078	14,362	3,968	10,394
			Patients Less Than 65 Years of Age			
1958.....	62,049	16,334	1,892	17,591	4,664	12,927
1957.....	63,299	15,124	1,686	16,373	4,404	11,969
1956.....	64,211	14,646	1,757	15,530	3,809	11,721
1955.....	65,703	14,847	1,758	13,040	3,534	9,506
			Patients 65 Years of Age and Older			
1958.....	29,142	6,952	7,529	1,743	456	1,287
1957.....	29,110	6,704	6,869	1,638	492	1,146
1956.....	28,651	6,808	6,588	1,530	471	1,059
1955.....	27,611	6,612	6,320	1,322	434	888

*Direct discharges plus placements on convalescent care.

Many of these would have been court certified if attitudes and policies had not changed. It doesn't seem reasonable to expect that a change in legal status from court certification to voluntary would cause an earlier release in all cases. It might in some cases. In others it might work the other way. All in all changes in legal status are not large enough to be a conclusive cause of the change.

Next let us look at the number of patients dying in the state hospitals. This number has also increased over the years. It has increased from 2,536 in 1910 to 9,421 in 1958. The ratio of deaths to admissions has stayed remarkably constant over these years. It has only varied from 36 to 43 per hundred over this 48 year period. This is a meaningful relationship since many deaths occur among elderly patients shortly after their admission. Since deaths have increased we must admit that they have played a factor in the decrease in resident patients. This is probably best put as a negative factor. If deaths had not increased there would have been a smaller decrease in resident patients.

We will have to examine this increase in deaths in more detail to see how it has affected the resident population in the last 4 years. Perhaps the first thing is to examine the age distribution of the deaths that have occurred. The increase in the number of deaths has occurred in the 65 years of age and older group. Interestingly enough the decrease in the resident patient population has occurred in the less than 65 years of age group. The decrease in this age group cannot be attributed to excess deaths. If the deaths in the over 65 group had not increased it would only have meant a larger increase in the number of patients 65 years of age and older. We can then conclude that the increase in deaths is not the factor in the decrease in resident patients less than 65 years of age.

The third factor to look at is the change in the number of patients released alive. Both the psychiatric condition of the patient and legal considerations affect releases from mental hospitals. There are several ways patients may be released. Patients may be completely discharged directly from the hospital. More commonly patients are released from the hospital on convalescent care status. This status allows them to return if necessary to the hospital with no difficulty during the period of convalescence. If, however, the patient adjusts satisfactorily outside the institution, he is then discharged.

We will call the sum of the releases by these two principal methods(direct discharges and placements on convalescent care) releases alive. The figures for releases alive in the last four years show an amazing increase from 1955 to 1958. The number of releases alive has increased from 14,362 to 19,334. This is an increase of almost 5,000 in the annual number of patients released alive in four years. Over 4,500 of this increase is in the less than 65 years of age group. This then would seem to be the factor that caused the decrease in the resident patients. However, before we conclude that an increase in releases alive was the principal cause of the drop in population, we must look at the fourth factor, returns to the hospital If these additional released patients returned to the hospital almost immediately, then they could not be the cause of the population decrease.

From 1955 to 1958 the number of patients

returned from convalescent care did increase. In 1955 there were 3,968 such returns and in 1958 there were 5,120. This is an increase of about 1,150. Since there was an increase of 5,000 in releases alive over this same period, then the net increase in releases alive is really about 3,800. On the basis of these data, I think we can conclude that the decrease in resident patients was due primarily to an increase in patients released alive.

We now can get to the major point of interest in this paper. What factors are associated with this increase in releases? If the use of tranquilizing drugs is one of these factors, how important is it?

In Table 3 the distribution of releases alive by age group for each of the last four years is shown. These data show that the increase in releases alive has occurred among all age groups in a rather steady pattern over the four years. Expressed on a rate basis as the number per 100 patients under treatment the same steady pattern of increase appears. When put in index number form using the 1955 age specific release rates as bases we can see that the release index for all ages reached 130 in 1958. At the same time the indices for the age groups from 15 to 54 had increased even more. The youngest age group (less than 15) and the age groups above 54 had increased less. The greatest increases have occurred in the age groups from 15 to 54 years.

Now let us look at these same releases alive distributed by diagnostic category in Table 4. The releases alive of schizophrenic patients increased from 6,427 in 1955 to 8,970 in 1958. On a rate basis this represents an increase from 10.7 per hundred under treatment to 15.0. For patients with alcoholic psychoses the rate increased from 29.6 to 31.6. For patients with senile psychoses or psychoses with cerebral arteriosclerosis the rate increased from 4.7 in 1955 to 5.3 in 1958. For patients with all other diagnoses the rate of releases alive increased from 20.6 to 26.1 in the

Table 3

Releases Alive* from New York Civil State Mental Hospitals by Age

Age Group	Fiscal Year Ending March 31			
	1958	1957	1956	1955
Total	19,334	18,011	17,060	14,362
Less than 15	353	250	220	130
15-24	2,246	2,009	1,938	1,705
25-34	4,298	4,036	3,936	3,293
35-44	4,521	4,238	3,991	3,318
45-54	3,773	3,494	3,327	2,748
55-64	2,400	2,346	2,118	1,847
65-74	1,281	1,187	1,108	957
75 and over	462	451	422	364
Release Rates (per 100 under treatment)				
Total	16.7	15.7	14.9	12.8
Less than 15	23.5	21.2	24.4	18.9
15-24	44.4	43.0	41.0	31.9
25-34	36.4	33.5	31.4	24.3
35-44	26.3	24.2	21.8	17.7
45-54	17.4	16.1	15.0	12.9
55-64	10.7	10.6	9.7	8.9
65-74	6.4	6.1	5.8	5.4
75 and over	2.9	2.8	2.8	2.6
Index Number (1955 Base)				
Total	130	123	116	100
Less than 15	124	112	129	100
15-24	139	135	129	100
25-34	150	138	129	100
35-44	149	137	123	100
45-54	135	125	116	100
55-64	120	119	109	100
65-74	119	113	107	100
75 and over	112	108	108	100

*Direct discharges plus placements on convalescent care.

Table 4

Releases Alive* from New York Civil State Mental Hospitals by Diagnosis Group

Diagnosis Group	Fiscal Year Ending March 31			
	1958	1957	1956	1955
Total	19,334	18,011	17,060	14,362
Schizophrenia	8,970	8,478	7,863	6,427
Alcoholic Psychoses	1,747	1,643	1,561	1,456
CASSP/	1,164	1,096	1,052	970
All Other	7,453	6,794	6,584	5,509
Release Rates (per 100 under treatment)				
Total	16.7	15.7	14.9	12.8
Schizophrenia	15.0	14.2	13.0	10.7
Alcoholic Psychoses	31.6	31.3	30.4	29.6
CASSP/	5.3	5.1	4.9	4.7
All Other	26.1	24.3	23.6	20.6
Index Number (1955 Base)				
Total	130	123	116	100
Schizophrenia	140	133	121	100
Alcoholic Psychoses	107	106	103	100
CASSP/	113	109	104	100
All Other	127	118	115	100

*Direct discharges plus placements on convalescent care.
/Psychosis with cerebral arteriosclerosis plus senile psychoses.

same period.

On an index number base we see that the greatest improvement, from 100 to 140 occurred in the schizophrenic group. The smallest improvements occurred among the patients with alcoholic psychoses and with the psychoses associated with old age.

Can we conclude anything from these improvements in the rate of release alive? We must first recognize that there has been some improvement in all age groups and in all major diagnosis groups. The improvement has been general. It has increased most in the relatively young age groups and in the schizophrenic group, but it has affected all classes of patients. The cause or causes therefore must be general. They must affect all groups of patients. What things have occurred in the hospitals which meet these criteria? To the best of my knowledge there have been three major developments within the New York State Mental Hospital system that could be possible causes of the observed changes. They are all interrelated and it is very difficult to separate them neatly. They are: first, the introduction of the tranquilizing drugs; second, the intensive treatment program; and third, the open hospital policy.

Experimentation with the tranquilizing drugs began in the New York State Mental Hospitals in 1953. In February 1955 after a six months large scale trial, general use of tranquilizing drugs began. On March 31, 1955, 7,000 patients were on tranquilizing drugs. The figure had increased to 20,000 on March 31, 1956, 28,000 on March 31,1957 and 42,000 on March 31, 1958. This last figure represented 46% of the patients in the hospitals on that date. The drugs have been used on all categories of patients. However, they have been used most extensively in the younger age groups (except for the youngest) and in the non-organic psychoses (principally schizophrenia). These are the categories that have shown the largest improvement in release rates.

The intensive treatment program provided additional treatment personnel in the reception services of the hospitals to insure that every patient admitted received the maximum benefits from known treatment modalities in the crucial initial period after admission. This program began in 4 hospitals in 1956. It was extended to an additional 4 in 1957 and finally in 1958 it was started in the remaining 10 hospitals. This treatment program makes extensive use of the tranquilizing drugs, hence it seems almost impossible to separate the effects of the two programs. However, the intensive treatment program did not begin until 1956 and then only in four hospitals of the 18 in the department. The overall increase in releases cannot be ascribed solely to this program It almost certainly has had an important effect

but it can not have been the decisive one.

The open hospital policy is not new. In the nineteenth century many American mental hospitals were operated as open hospitals. It was only in the late nineteenth century and the early part of this century that closed or locked hospitals became the rule in this country. The New York State hospitals have always had some open wards. These were usually either so-called convalescent wards or workers wards. In the convalescent wards were patients almost ready to be released, in the workers wards were patients working in the hospital industries. Other patients living on closed wards were given honor cards which allowed them the freedom of the grounds under specified conditions. The experience of the British mental hospitals in recent years with unlocked wards led to a resurgence of interest in this policy as a therapeutic technique. In April 1956 about 6% of the 93,000 patients in New York State mental hospitals were in open wards or on ground parole. Late in 1956 deliberate attempts to apply the open door policy in the hospitals were begun. In April 1957, 8% of the patients were on open wards By October 15% were on open wards. In April 1958 the proportion had risen to 43% and in October 1958 had reached 58%. The tranquilizing drugs have played an important part in the success of this program. They have helped reduce the amount of violent behavior and helped make possible a tremendous drop in the use of restraint and seclusion. The use of these has decreased 90% from 1954 when about 25 patients per 1,000 average daily resident patients were in some form of restraint or seclusion. Now the rate is less than 3 per 1,000.

These three programs have brought about a revolution in the care of the mentally ill. They all involve the use of the tranquilizing drugs. In our analysis of population trends we must admit that we cannot separate out the effects of these drugs alone. We have to consider the effect of these drugs when used as part of an expanded and enlightened treatment plan.

Since admissions have remained steady,deaths increased only slightly, releases alive increased markedly and since returns have increased only slightly we can, I believe, safely conclude that the effect of the tranquilizing drugs when used as part of an expanded treatment plan has been to reduce the patient population in the New York Civil State Hospitals. What then of the future? Can we assume that the population will continue to decline indefinitely until there is no more need for mental hospitals? Or is the population in the hospitals merely undergoing a readjustment, a clearing up of a backlog, which will be followed by a renewed increase? Or will it hit a plateau at some lower or higher level and stay there? All of these are possibilities. The forecasting problem in this area is an unsolved one. The evidence so far is scanty. However it is possible to note some straws in the wind.

To do this it is necessary to look first at the present age and duration distribution of the patient population in the hospitals. Of the 91,191 patients in the New York State mental hospitals 23,283 have been there for 20 or more years. Another 19,500 have been there from 10 to 19 years. Thus 42,783 or 47% of the patients have been in residence at least 10 years. Until some major research breakthrough occurs, it is not likely that any great number of these patients who have been in the hospital for these periods of time will be released. Most of them may be in residence for the rest of their life. Every year some patients enter the 10 or more years duration category as they reach the 10th anniversary of their admission to the hospital. As long as more patients enter this category by staying in the hospital than leave it by death or release the size of the chronic patient load will continue to increase. This phenomenon is still going on.

In 1955, 41,829 were in this category. In 1956 the number rose to 42,347 and it continued to increase till it reached 42,783 in 1958. The growth of the long term chronic patient load has slowed down but it hasn't stopped yet.

Since the total population is decreasing while the population in the category of "duration of residence more than 10 years" is increasing,it means that the category of patients in residence less than 10 years must be decreasing very rapidly. This indicates that in the future the number of patients entering the over 10 years duration group must decrease.

To check this let us look in Table 5 at the distribution of patients less than 65 years of age by time since admission. The total number of these has declined from 65,707 in 1955 to 62,049 in 1958. At the same time those in residence 20 or more years has increased from 12,677 to 13,792. This is an increase of 9%. At the same time the number of patients under 65 in the 10-19 year distribution group dropped from 16,646 to 15,289; in the 5-9 year duration group dropped from 11,590 to 11,206; and in the 1-4 year duration group from 15,656 to 12,859. As time goes on the number of patients entering the longer duration groups must decrease and eventually the number in the longest duration group in the under 65 years of age group must decrease.

In the over 65 age group the situation is a little different. Patients move into the longer duration groups of this age group in two ways. They can age in by passing their 65th birthday. They can do this into any duration group except the shortest. They can also enter a longer duration group by virtue of time spent in the hospital beyond the age of 65. Thus the increase of the 20 years and over duration group from 8,394 in 1955 to 9,491 in 1958 is the result of patients over 65 years of age in the 10-19 duration group moving to it as well as patients under 65 years of age in the 20 and over duration group passing their 65th birthday. As a result of these two types of flow all of the duration groups except the shortest have shown increases since 1955.

What can we expect in the future. First, for the patients less than 65 years of age, I think we can expect continued decreases in each duration

Table 5

Resident Patients in New York Civil State Mental Hospitals

by Age Group and Time Since Admission

Fiscal Year Ended March 31	Time Since Admission (years)					
	Total	Less than 1	1-4	5-9	10-19	20+
1958.....	91,191	12,674	20,304	15,430	19,500	23,283
1957.....	92,409	12,483	21,436	15,735	19,927	22,828
1956.....	92,862	12,361	22,310	15,844	20,375	21,972
1955.....	93,314	13,051	23,089	15,345	20,758	21,071
Less than 65 Years of Age						
1958.....	62,049	8,903	12,859	11,206	15,289	13,792
1957.....	63,299	8,575	13,782	11,566	15,803	13,573
1956.....	64,211	7,937	14,887	11,835	16,318	13,234
1955.....	65,707	9,138	15,656	11,590	16,646	12,677
65 Years of Age and Older						
1958.....	29,142	3,771	7,445	4,224	4,211	9,491
1957.....	29,110	3,908	7,654	4,169	4,124	9,255
1956.....	28,651	4,424	7,423	4,009	4,057	8,738
1955.....	27,607	3,913	7,433	3,755	4,112	8,394

group except for the less than 1 year group. This latter group should increase as a result of increased admissions. There should be an increase for a few years of the 20 and more years duration group. Then it should start to drop along with the others.

This means a substantial drop in resident patients less than 65 years of age in the future. What this means for these patients is that the average stay has been and will continue to be drastically reduced. There will be no substantial number of them entering the hospitals in their 20's and staying the rest of their lives. There will no doubt always be some patients in this category, but the vast majority will be able to return to the community. They may have to return to the hospitals for short stays several times during their lives but they should not have to spend their entire lives there.

Now let us consider the patients 65 years of age and older. In the next 10 years I believe this number will continue to increase. If we assume that admissions will continue at the same rate, I would assume about 5,000 in the less than 1 year duration group; 8,000 in the 1-4 year group; 6,000 in the 5-9 year group; 6,000 in the 10-19 year group and 14,000 in the 20+ duration group. This would mean a total of 39,000 in the over 65 years age group. This would be an increase of 10,000.

These estimates are based on the assumptions of a continued increase in the number of patients in the older age groups admitted to the hospitals as well as on a continued improvement in the treatment procedures for these elderly patients. It must be remembered, however, that the main result of improved treatment for elderly patients may be longer hospitalization. In many cases the patient's condition may be improved, death can be postponed for long periods but the patient may not be well enough to leave the hospital. Thus the average stay in the hospitals for older patients admitted may well be increased.

In addition many long term patients now in the hospital will continue to be there 10 years from now. By then nearly all of them will be more than 65 years of age. All of these factors will tend to increase the number of older patients in the hospitals for at least 10 years.

At some time in the future, perhaps 20 years from now, we can look to the day when the backlog of long term mental hospital patients has been cleared up. Then we can expect really substantial reductions in the mental hospital population. However, two factors working to increase the patient load should not be overlooked. One is the increase in the general population. Particularly the increase in the number of aged in the population. Unless there is a tremendous expansion in the field of community psychiatry or in the provision of alternative facilities which will eventually reduce the admission rates, we may have so many patients admitted that even with better treatment and shorter stays the mental hospital population will start to rise again.

Second is the possible expansion of the classes of patients treated in mental hospitals. If narcotic addicts, non-psychotic alcoholics, and other types of patients different from those now received are admitted to the mental hospitals, then of course major decreases in population are most unlikely.

A note on methodology should be inserted in this paper. The proper way to study the population of an institution such as a mental hospital

is through cohort analysis.[2,3] That is, one should follow each homogeneous group of admissions throughout the time any of its members remain in the hospital. We are now doing this for all patients admitted to New York State Hospitals since 1954. We have not found it possible to do this for patients admitted prior to this date because of gaps in the record system. On the basis of preliminary data from cohorts of patients admitted since 1954, the conclusions concerning length of stay and release derived in this paper seem to be fully confirmed.

Summary

For fifty years until 1955 the patient population of the New York State mental hospitals had climbed steadily from 28,000 to 93,000. Early in 1955 the large-scale use of tranquilizing drugs began in these hospitals. In the three years since then the patient population has decreased to 89,000 in spite of an increase in admissions. The issue is whether the relationship between the use of tranquilizing drugs and the population drop is coincidental or causal. Evidence is given to show that the population decrease is due primarily to an increase in the number of patients released and that this increase in releases is probably related to the use of these drugs.

The future patient population is estimated on the basis of present trends. The characteristics of the resident patient population will continue to change rapidly. In general, patients will stay for much shorter periods of time. There will be an increased proportion of aged patients in the hospitals. The use of tranquilizing drugs is being accompanied by an intensified treatment program and by an expansion of the open hospital policy. In summary, the mental hospitals will continue to shift their emphasis from custodial care to active treatment.

References

(1) Brill, H. and Patton, R. E.: Analysis of 1955-1956 Population Fall in New York State Mental Hospitals in First Year of Large-scale Use of Tranquilizing Drugs. Am.J.of Psychiatry 114: 509-516, December, 1957.

(2) Kramer, M. et al: A Historical Study of the Disposition of First Admissions to a State Mental Hospital. Public Health Monograph No.32, November 1955.

(3) Pollack, E. et al: Patterns of Retention, Release, and Death of First Admissions to State Mental Hospitals in 1954. Public Health Monograph No.__, (in press)

Discussion:
James C. Munch, Hahnemann Medical School, and
Pharmacodynamics, Inc.

It is a definite contribution to scientific progress to develop any type of test which is more objective in nature and permits recording results numerically. Dr. Clyde has done this by preparing a set of 3 x 7 inch cards carrying perforation for machine classification. The 133 traits listed include " angry, anxious, absent-minded" etc. Each subject classifies himself with one set; the examiners with a different set. Results fall in 4 categories :(1) Not at all; (2) A little; (3) Quite a bit; (4) Extremely. With a weight of 1 for each card, this permits compuation of an overall score at any time. Changes in score produced by drugs or other treatment may be expressed quantitatively.

This is a decided improvement over scoring many psychodiagnostic tests in use. Other questions may be added to the present list; validation of present attributes confirmed; the troublesome question of relative weights to be assigned various questions resolved; and the applicability of these questions to measuring psychological behaviour changes. Such studies are desired in various types of mental disturbances, as well as in "normal" individuals.

Mental patients comprise half of the total hospital population in the United States. Dr. Kris points out that an increasing number of these patients are recieving "tranquilizers" and that there is an increasing rate of discharge from mental hospitals to the community. Also, the remaining patients tend to be more tractable. Our experience with one of these tranquilizers at Eastern State Penitentiary suggests that one in four inmates showed significant improvement in behaviour pattern during administration, not shown while on a placebo.

Tranquilizers differ pharmacodynamically, and are not interchangable. Individual patients may be expected to show better response to some than to others. The effective dose will vary between patients as well as within any patient from time to time, possibly as a result of such factors as tolerance,metabolic or pathological alterations, and differences in intensity of external stresses. Dr. Kris emphasizes the growing need to indoctrinate the general practitioner in adequate therapy for furloughed and/or discharged mental patients, which will further decrease readmission rates. The development of side effects such as jaundice or blood dyscrasias should be anticipated during continuing therapy. There is very little data on the effect of chronic administration of these drugs. Dr. Kris makes 6 very pertinent suggestions, perhaps the most important being need for cooperation by possible employers.

Mr. Patton presents convincing evidence that there is an association between the decrease in the number of mental patients in the New York State Hospitals, and the increase in therapy with "tranquilizing drugs." Further explorations will be made to determine if this is a true correlation. This will be done by cohort analysis, which was started in 1954. It is hoped that an attempt can be made to determine the particular "tranquilizer" used, and the optimal dose for each, since these drugs are not interchangable. The development of any toxic effects with prolonged use should be carefully observed. Other factors such as intensive treatment ond open hospital policy made only a minor contribution to this decrease in hospital stay, which is real and in spite of increase in the number of general population. Any possible difference in mortality rate on mental patients in hospitals as compared with the same age groups outside hospitals, would be interesting to determine.

Patients between 1 and 65 years of ages and who have been in mental hospitals less than 10 years show greatest improvement, and half of them may expect discharge after a comparatively short hospital stay. The greatest benefits may be expected in schizophrenia, and less in alcoholic psychoses.Mental hospitals will be needed for custodial, rather than active treatment. Increasing knowledge by the general practitioner regarding therapy of such patients may be expected to materially decrease re-commitments.

V

RECENT SURVEYS IN HEALTH

Chairman, Theodore D. Woolsey, Department of Health, Education, and Welfare

Survey Research and Medical Care: Strategy and Tactics of a Research Program—Odin W. Anderson, Health Information Foundation

A Sample of Developments From Sampling Projects of the U. S. National Health Survey—Walt R. Simmons, National Health Survey

What Americans Think About Their Medical Care—Jacob J. Feldman, National Opinion Research Center

Discussion—Sam Shapiro, Health Insurance Plan of Greater New York

SURVEY RESEARCH AND MEDICAL CARE: STRATEGY AND TACTICS OF A RESEARCH PROGRAM

By: Odin W. Anderson, Ph.D., Health Information Foundation

I. Background

Without the development of survey research methodology over the past 20 years, the research program of Health Information Foundation would not have been possible. It is the use of survey research methods in the Foundation's research program which may be of interest to the members of the American Statistical Association. In fact, unless survey research methodology had been relatively well developed when the Foundation was established in 1950 it is unlikely that the Foundation could have gotten off the ground. This paper is devoted to a description of survey research in medical care and how it was used in the strategy and tactics of the development of a research program in the social and economic aspects of personal health services directed to current problems and issues in the health field.

In 1948 several leaders in some of the large pharmaceutical, drug, and chemical firms were considering how the industry could best be of assistance to the private or non-governmental sector of paying for and providing personal health services in this country. There had been a great deal of ferment in the health field -- and this is continuing -- little of which was understood by the consumer, the insurance agencies, or the providers of service.

In due course, the industry decided that the public interest could best be served with simultaneous creation of goodwill for the industry by establishing a research foundation chartered as a non-profit, tax-exempt research agency which may not engage in propaganda nor influence legislation. This was done in 1949. In 1950 a nucleus staff was in operation which was trying to determine research policy and course of action. During the first two years three community studies were launched to determine how local areas went about surveying their health problems and solving them, or as the case may be, not solving them.1/ It was hoped that "successful" communities could be used as models of self-determination by other communities wishing to do the same. Useful as these community studies were, they could not yield definitive data or information on social process and they contributed to research on communities in a long-term sense, but not to research for immediate application. It was at this point late in 1952 that I became research director and was charged with developing a research program. There was no tangible research policy and no staff, but there was money for research and a real desire to have a program developed which would be in the public interest.

By 1952 various forms of voluntary health insurance had reached a point where over 50 per cent of the population of the country was covered by some type of health insurance. In its rapid growth since the late Thirties it has become the dominant force to help families pay for personal health services, particularly hospital care and surgery. It caused and continues to cause endless discussion and debate regarding problems of enrolling more and more people and expanding the range of benefits covered to include services provided outside of the hospital, particularly physicians' home and office calls. In 1952 no systematic review had been made of the extent of voluntary health insurance -- only gross figures by state were known -- and what such insurance was doing to help families pay for costs of personal health services of all types. The consumer side of the costs of personal health services and the extent to which prevailing benefit patterns of insurance help families pay for services then became the focal point for the Foundation's research policy, and one which the sponsors could support in the public interest and in the framework of their general social philosophy. No other agency, governmental or private, was at that time (or since) sponsoring research of any magnitude or depth directed at the forces which would help to shape the financing and eventually the organization of our personal health services for years to come. I will give the rationale for choice of research problems, and explain the methods by which they were formulated in the execution of a research program in the sections to follow.

II. Defining the Dimensions

The decision to focus on the financing and organization of personal health services and more specifically prevailing voluntary health insurance provided a point of reference for planning the other necessary aspects of the operation of the Foundation. Since the Foundation was completely new, dependent for funds from its business sponsors on an annual basis and without an endowment, and entering an area of great importance to public policy formulation, it was necessary -- as with any new organization in society -- to work toward the creation of a public image. The image desired was that of an independent research agency, selecting problems of research directly relevant to current problems and issues in the financing of personal health services, publishing and disseminating the research findings freely through reputable publishing houses and setting high standards of methodology and interpretation of data.

The voluntary health insurance field appeared to have developed to a point where there was enough consensus among contending parties that insurance was here to stay -- so the question was how to evaluate it and continue to make it work and expand. At this stage there was a favorable climate for research in which to flourish to help determine the scope and

components of the problems and provide a basis of facts for intelligent policy formulation. Such I believe was the intuitive timing of the Foundation. The next step was to take advantage of this timing so that research could influence and move in the midstream of continuing developments. The creation of an image -- unlike that of creating the image for a certain product or service -- was not difficult. No publicity as such was necessary. What was necessary was the formulation of a research program based on the criteria mentioned earlier and gradually let the results and the record speak for themselves.

From the first we became an "operating" rather than a "granting" Foundation, i.e., the research program was formulated by the staff and ways were sought to carry it out in terms of projects. In instances where the Foundation did not carry out its project directly with my own staff, grants were made to research agencies for specific purposes agreed on between the Foundation and the research agency. Eventually, three ways of carrying out research evolved, all three flowing from policy formulated by the Foundation staff. They are: (1) a project could be carried out under me with a project director on the staff and conducted from the Foundation office; (2) a grant could be made to a research person or research agency interested in the problems the Foundation was committed to, the Foundation, in effect, buying a manuscript, but the manuscript always being under the control of the agency conducting the project; or (3) the Foundation would work cooperatively and jointly with a research agency right through publication. The three methods described permit great flexibility in operation and entail only a small core staff at the Foundation headquarters. My main responsibility is then that of selecting research problems, outlining general context, and working with people qualified to carry out the technical and methodological requirements. Occasionally a member of my staff or I also write up the results, but such a decision is made beforehand depending on the arrangement.

Within the framework of our sponsors' intent as articulated by the staff, the Foundation conducts research mainly within the prevailing structure of personal health services and is so far engaged in describing and evaluating the problems in voluntary health insurance from the standpoint of the consumer. Thus our emphasis to date has been mainly on the consumer problems of paying for and using health services so that voluntary health insurance has more benchmarks by which to judge its efforts in helping families pay for personal health services. Almost automatically our research projects have implications for public policy formulation and are selected with that purpose in mind.

The very immediacy of the problems and issues we wish to cast more light on dictates that we engage in short-term projects of two years duration or so, although we can plan long-term policy. This again necessitates our drawing largely on relatively well-tested and perfected social research and statistical techniques and methods which can be brought to bear on an applied field like the social and economic aspects of personal health services. In view of the Foundation's focus on consumer problems, survey research methods provide an ideal tool for large-scale household surveys on various problems to be described later.

Seemingly out of nowhere several things converged -- sponsors representing an important and relatively neutral part of the health field wishing to finance a research foundation in the public interest, the emergence of voluntary health insurance as a problem of prime importance in public policy and debate, the concurrent perfection of adequate social research methodology which is directly applicable to household surveys, and a climate conducive to research in that contending parties had reached a consensus that health insurance was here to stay and we now needed more information on how to apply it effectively.

III. Development of the Research Program

With the financing and organization of personal health services as the general area for research it was decided that in 1953 the most important area for public policy at that time was a comprehension of the distribution of costs of personal health services among families throughout the country, the extent of health insurance, and the degree to which such insurance was helping families to pay for services. The costs of personal health services among families had been done before but not since health insurance had become an important method of paying for services. The research was planned to examine at least two aspects: (1) the extent to which insurance paid for covered services, mainly hospital care and surgery, and (2) the extent to which the costs of other services outside of the hospital, such as home and office calls, drugs, and dental care, not usually covered, were difficult for families to meet without some insurance mechanism. These were crucial questions, answers to which were not known. Some basic data would immediately provide benchmarks of accomplishments and expose problems still in need of solutions.

This study entailed a survey of a sample of households in the United States and represented a formidable undertaking. In view of survey experience in medical care going back to the Thirties and further development and refinement of survey research since World War II, a survey of this nature was entirely feasible. Sampling science had been refined to a point where a relatively small sample of households was necessary to provide estimates within narrow margins of error. This, of course, reduced the cost and made surveys much more feasible financially than even a decade earlier.

The survey research agency which the Foundation worked with on this survey was the National Opinion Research Center. In formulating the design it was decided that a sufficiently

large sample would be about 3,000 households comprising 9,000 individuals. The interviews were conducted in the households and there was only one interview. Costs and utilization of services were recalled for 12 months prior to the interview. There was concern with respondents' ability to remember adequately for our purposes the costs and utilization of services for a 12 month period. Pretests were made in several areas of the country to compare people's responses with what was recorded in physicians' offices with sufficiently high congruence to go ahead on a one-interview basis. For hospital care, hospital records were examined in the actual survey to compare respondents' statements with the hospital records. The congruence in this survey was extremely good. It should be remembered that we were after nationwide estimates for costs and utilization by age, sex, residence, and insurance status to show the general patterns.

Highlights of data are of interest both for their implications for public policy and for their influence on plans for future research. As was to be expected, we found that some families incurred no costs for services and a small minority incurred extremely high costs.2/ Again, as was to be expected, families which carried some type of health insurance were likely to live in cities, be employed in large industries, and have relatively good incomes. We also showed -- to no one's surprise -- that individuals with insurance used more hospital and surgical services (the prevailing benefit pattern) and incurred higher costs than families without insurance. At this point we left generally anticipated results and produced data which were more or less novel. We showed precisely what portion of all costs of personal health services for insured families health insurance was paying at the time, namely, 19 per cent; and the proportions of costs for hospital care and surgery were medians of 89 per cent and 76 per cent respectively.

Concurrent surveys but in a separate project were conducted in Birmingham, Alabama, and Boston, Massachusetts, of subscriber-households in Blue Cross-Blue Shield plans in these two cities and employed groups covered by Aetna Life Insurance Co. in Boston. This project was supplementary to the nationwide survey in order to pinpoint what prevailing benefit patterns as represented by the plans mentioned did to help families pay for personal health services. It was felt that in the nationwide survey individual insurance companies and Blue Cross-Blue Shield plans could not see themselves, although the figures given above showed in general what the prevailing pattern of insurance benefits did for families. By selecting specific insurance plans and relating a definite known benefit pattern to the subscriber-households total costs of services in a year, such a survey of households would seem more real to people who were actively concerned with health insurance problems. This project could also serve as a model for plans that wished to test the adequacy of their benefits.

Samples of subscribers were drawn from the files of persons currently insured by the agencies studied for both group and individual contracts. The contracts must have been in effect for one year in order to provide retrospective costs and utilization for that period. The size of samples in each of the three plans rounded are: Birmingham, 950 families, 2,900 individuals; Boston 1,200 families, 3,900 individuals; and Aetna 350 families, 975 individuals.3/ Except for changes necessitated by local conditions the questionnaires were the same as for the nationwide survey. Comparability was then assured.

This survey showed what plans that might be called "typical" did to help families pay for costs of personal health services. They certainly represented the prevailing benefit pattern of voluntary health insurance in the country, i.e., hospital care and physicians' services in the hospital. The results were not grossly different from the nationwide survey, although the three plans were in general better than seemed to be true for the national picture. In both surveys -- the nationwide and the two cities -- it was clearly brought out that even including insured services the patient pays from 70 to 80 per cent of the total charges and the costs of services outside of the hospital are also of financial consequence to families. Prevailing insurance benefit patterns had been formulated on the assumption that hospital care and physicians' services in the hospital were the services that had the greatest impact on families. In these surveys it was shown that over a year just as many families incurred high costs for home and office calls as for hospital care, for example, and some families incurred higher costs for home and office calls than for surgery. A reorientation of current insurance thinking was then indicated.

The nationwide survey brought out another fact of importance to a consideration of the influence of insurance on the level of utilization. It had been known in hospital care, for example, that when a group of people acquired hospital insurance the hospital utilization rate generally increased compared to the time when the same people had no hospital insurance. The reduction or elimination of the financial barrier presumably made it less burdensome for people to be hospitalized. In comparing insured and uninsured individuals nationally it was found that 14 per cent of the insured individuals were hospitalized in a year compared with 9 per cent for uninsured individuals. Similarly it was found that 9 per cent of the insured individuals had a surgical operation as compared with 5 per cent of the uninsured individuals. Concurrently, however, it was found that insured individuals were much more likely to use services like dental care, home and office calls, and drugs which are normally not covered than uninsured individuals. This indicated that a simple explanation of the presence of insurance influencing hospital care or surgery was not an adequate one. In addition, there were data on utilization and costs by

family income which could not be explained by economics alone. There were attitudinal factors involved, perceptions of health services and health; life values which dictated priorities in a prosperous economy like ours.

On the basis of the nationwide and two city surveys three problems were defined which called for further research and projects were planned accordingly. First, having shown that physicians' costs outside of the hospital are of financial consequence to families, compared with hospital care and surgery, the Foundation needed to study methods of providing all physicians' services under prepayment insurance and assist in the expansion of the range of benefits. Second, having shown that insured individuals use more of all kinds of personal health services, whether or not they were covered by insurance, the Foundation saw the value of making a study of adults in the general population, both insured and uninsured, to learn if there was a difference between them in how they perceived and appreciated health services, their concept of health, their images of the doctor, hospital, dentist and so on, and general level of health knowledge. Third, since a large portion of the uninsured population were not in employed groups, the usual mechanism of insurance enrollment, what was the problem of enrolling the so-called non-group portion of the population and what methods were in use at that time.

In due course projects directed to these three problems were set up, one outside the Foundation with the University of Michigan and University of Washington, one jointly with the National Opinion Research Center, and one under my immediate direction and supervision. The problem studied by research teams at the University of Michigan and the University of Washington related to insuring the full range of physicians' services, including home and office calls. Physicians' service plans operating in Windsor, Ontario, and the State of Washington sponsored and controlled by county medical societies in the two areas had for 20 years or more provided comprehensive physicians' services within the traditional and desired structure of private practice, fee-for-service and physicians practicing in their own offices. The medical profession was generally fearful of expanding in-hospital physicians' service plans to home and office calls because of expectations of ever-rising utilization and costs eventually entailing a reorganization of medical practice along lines that were resisted. Examples are the group practice closed panel salaried plans in several parts of the country.

The Foundation felt that an intensive examination of the plans as described in Windsor and Washington would be useful as an answer to the possible expansion of the range of physicians' services in insurance which would still be within the traditional structure of medical practice. Accordingly in both areas samples of subscriber-households (and in Windsor households without insurance as well) were drawn to learn how the plans were used, the attitudes toward the doctors, costs incurred out-of-pocket, stability of the physician-patient relationship and other matters. Heads of households were interviewed in their homes. Representative samples were selected of participating physicians and interviewed (300 out of 900 in Seattle, and virtually the total universe of 216 in Windsor) to learn physicians' reactions to medical insurance effect on their incomes, patient load and patient pressures, and so on. In addition some analysis was made of internal operation regarding costs of administration. Time and space do not permit even highlights of these two studies. One has just been published, and the other is being prepared for publication.4/

The second problem, public attitudes and perceptions, was carried out by the National Opinion Research Center on an area probability sample of 2,600 adults 21 years of age and over. An intensive and structured interview was carried out on a one-interview basis. Samples were also drawn of about 500 physicians the public mentioned as having attended them or to whom they would go if needed and 500 pharmacists who had filled prescriptions for them the previous year. Thus there were interlocking samples of the public, physicians, and pharmacists. The physicians were queried as to their attitudes towards patients, insurance, hospitals, and similar matters. Parallel questions were put to pharmacists in retail pharmacies. The manuscript for this study is in preparation, but several short reports have been published treating topics more intensively than will be true in the main report.5/

The third problem, non-group enrollment, was carried out by a staff member working in close cooperation with all the 88 Blue Cross hospital plans in the country. In addition, further case studies were made of five plans which had attempted various methods of enrolling individuals outside the normal mechanism of employed groups. Because of the great interest shown by the Blue Cross executive director and enrollment directors and certain personal relationships in the field, a long mailed questionnaire yielded a return of 100 per cent.6/

Studies of voluntary health insurance led to the consideration of another problem area undergoing intensive discussion today -- the people 65 years of age and over. The Foundation felt that there was no perspective on the components of this area since problems of the aged were defined mainly by physicians, social workers, hospital administrators and others who saw the segment of this population group with acute problems. Accordingly the Foundation again worked with the National Opinion Research Center to formulate a household survey of people 65 years of age and over (later reduced to 60 years of age and over) to learn directly from the older persons themselves how they viewed their life situations and problems with particular reference to health and health services. In addition, a sample of people mentioned by the older persons as being next in

line of responsibility if something should happen to them was interviewed. This enabled us to get some idea of the sense of mutual responsibility in American families today. Further, a relatively small sample of the general adult population was interviewed to learn their attitude toward the aged and aging and what preparation was being made in middle life for retirement, decrease in activity and related matters. The manuscript for this survey is in preparation.7/

Another survey research project which can be mentioned deals with an emerging problem in the health insurance field. Since the development of voluntary health insurance in this country two general types of arrangements have existed simultaneously characterized by those which are payment mechanisms grafted on, as it were, to the existing structure of medical practice and those which set up group practice panel plans with physicians usually on a salary and providing the full range of physicians' services. The first type potentially permitted choice of any physician in the community and the group practice panel type restricted choice to physicians on the panel. Labor unions had been and are quite active in supporting the group practice type, but during the last few years in both New York and California individual members wished to select a so-called free-choice fee-for-service type of plan or a group practice panel type. This resulted in the procedure known as dual-choice in which members of a labor union could select one type or the other within the same union.

In New York several local labor unions voted whether to join Group Health Insurance (G.H.I.), the free-choice fee-for-service type of plan, or Health Insurance Plan of Greater New York (H.I.P.), a group practice panel type. G.H.I. and H.I.P. asked the Foundation if it would finance and conduct an opinion survey of several unions to determine why union members chose one plan or the other within the same union. This involved perceptions of change and choice of doctor, attitudes toward health and insurance, and so on. The Foundation was interested providing the survey could include a study of the extent to which these two plans helped subscriber-families pay for costs of personal health services. It will be recalled that both G.H.I. and H.I.P. offer the full range of physicians' services in different settings. Also, both have arrangements with the Blue Cross hospital plan in New York for the same hospital benefits.

This survey presented involved relationships and a complicated sampling problem. Three local unions were selected as the survey units and the leadership in each union had to be seen personally and the objective of the survey explained. Eventually letters expressing support were obtained. In the formulation and final design of the project it was necessary to keep both G.H.I. and H.I.P. informed. The sampling problem then involved three unions and matching samples within each union to reflect G.H.I. and H.I.P. members and the total of the two parts in each union to be strictly comparable as to the usual demographic characteristics. About 450 households were drawn for each group. Our problem was to compare the experience of union members enrolled in G.H.I. and H.I.P. necessitating matched samples which, naturally, did not represent the G.H.I. and H.I.P. enrollees in the unions. This survey is now being written up.8/

IV. The Planning and Staging of Research and Dissemination of Results.

I have been giving you the general rationale for selecting research projects -- timeliness, relevance, and so on -- and the more obvious mechanics of operation. There are other projects but time does not permit describing them. Obviously, in order to promote and stage research touching on matters of public policy in the health field it is simple common sense to stage the projects so that they will have maximum acceptance and impact when results are publicized.

Our first nationwide family survey carried out in 1953 and described earlier was a large undertaking implicating many interests -- medicine, hospitals, dentistry, drug industry and pharmacists, Blue Cross and Blue Shield plans, and insurance companies. Each of these was briefed as to our intentions and a broad advisory committee was set up as a sounding board and sanctioning group for the project. Each interviewer from NORC was armed with letters from the American Medical Association, the American Hospital Association, and the Blue Cross and Blue Shield Commissions, as well as a letter from the Foundation President. As for myself, more directly, I had the services of three personal consultants whom I selected as my technical sounding board, as it were.

After the research findings were available, a press conference was held aimed at Sunday release. Our press coverage was exceedingly good. This signalized that the Foundation was on its way and now began the image of the Foundation -- area of research, scope, objectivity, and complete publication of results -- which was desired.

After this large initial venture the staging of our activities has been much less formal, because relationships outside and inside the Foundation have become stabilized as all parties have become familiar with our objectives and methods of operating. This means, of course, that the staff assumes and is given considerable responsibility and the staff must maintain good and continuous relationships with the Foundation Board and the many interests and agencies in the field that are implicated by our research. This is done on appropriate levels by the President of the Foundation, myself and other staff members. Since we are a relatively small organization -- 15 people with 8 in research -- it is necessary to personalize the Foundation systematically and constantly in our dealings with research agencies and

other agencies mentioned. This requires travel and personal contacts in establishing the Foundation and stabilizing its desired role.

Considering the fact that we are an operating research foundation directing our research primarily to current problems and issues as described and desiring it to have continuous and regular impact on the health field, to this end we have three chief methods of disseminating information. One is the intermittent publication of the findings of major studies in book form and through our research series in pamphlet form when projects are completed. A second is a regular monthly bulletin 10 months a year called Progress in Health Services, each issue standing on its own right as a research document. In our monthly bulletin we wish to give the health field factual data on many topics relevant to an evaluation of the field and at the same time provide for the Foundation a monthly medium by which to sustain on a regular basis the image it is creating. The books and research series are sporadic and intermittent and offer no apparent continuity to those outside the Foundation. The Foundation staff collects and organizes a great deal of data from vital statistics reports and financial reports, issued by public and private agencies, which usually do not see the light of day in a form readily comprehended. We also present highlights of some of our research findings before publication in full-scale reports. Some examples of the monthly bulletin are:

Declining Crude Mortality in the United States: 1900-1954.

Changes in the Leading Causes of Death.

What Americans Spend for Personal Health Services.

The Growth of American Hospital Facilities.

Health Insurance Benefits and the American Family.

Control of the Communicable Diseases.

Accidents in the United States.

The Increased Use of Medical Care.

Our monthly mailing list is in excess of 60,000 and the bulletin goes to over 3,000 newspapers and magazines, members of boards of trustees of voluntary hospitals, legislators, public health departments, medical societies, hospital associations, dental associations, medical schools, and many related agencies and individuals. You can see we are aiming at audiences which shape the scope, content, and direction of our health services.

The third method is an annual inventory of research in progress in the social and economic aspects of the health field by means of a questionnaire to agencies that normally are carrying on research in this area. This provides the staff and me with rather detailed knowledge of what is going on to help the Foundation in its planning, as well as to give research people interested in this field a systematic source book. The audience for the annual inventory is research personnel and it is hoped that the inventory will diminish duplication of projects and stimulate investigation in new areas or add to research already in progress. The 1958 survey just published contains almost 800 projects.

V. Conclusion

We are now probably at a point where there will be less emphasis on the consumer's problems, as such, because the broad and in many ways the detailed dimensions of their problems are being spelled out. The next line of development will be directed at the problems of applying the principle of insurance, which leads us into studies of how hospitals are used, methods and problems of administration, measurement of need and demand, and many others. Survey methodology is being more and more perfected and adapted to particular purposes. Although the Foundation draws on relatively well tested and perfected survey research methods to apply to health problems, it is certain that the research we have sponsored is feeding back into research methodology so that not only are we building a body of knowledge in the health field but we are also helping to improve methodology.

This is then a case presentation of the establishment of a research foundation in the public interest and the development of a research program and methods of dissemination of information. The image desired was that of an independent research agency, selecting problems of research directly relevant to current problems and issues in the financing and providing of personal health services, publishing and disseminating without fear or favor research findings through reputable publishing houses, and following the highest accepted standards of methodology and interpretation of data.

REFERENCES

1. Christopher Sower, John Holland, Kenneth Tiedke, and Walter Freeman. Community Involvement: The Webs of Formal and Informal Ties That Make for Action. Glencoe, Illinois, Free Press, 1957; Floyd Hunter, Ruth Connor Schaffer, and Cecil G. Sheps. Community Organization: Action and Inaction. Chapel Hill, N.C., University of North Carolina Press, 1956; Solon T. Kimball and Marion Pearsall. The Talladega Story: A Study in Community Process. University, Ala., University of Alabama Press, 1954.

2. Odin W. Anderson and Jacob J. Feldman. Family Medical Costs and Voluntary Health Insurance: A Nationwide Survey. New York, McGraw-Hill, 1956.

3. Odin W. Anderson and the National Opinion Research Center. Voluntary Health Insurance

in Two Cities: A Survey of Subscriber-Households. Cambridge, Mass., Harvard University Press, 1957.

4. Benjamin J. Darsky, Nathan Sinai and Solomon J. Axelrod. Comprehensive Medical Services under Voluntary Health Insurance: A Study of Windsor Medical Services. Cambridge, Massachusetts: Harvard University Press, 1958.
George A. Shipman, manuscript in preparation.

5. Wallace Croatman and Paul B. Sheatsley, The Prescription Pharmacist Today. Health Information Foundation, Research Series 3, 1958.
Eliot Freidson and Jacob J. Feldman, The Public Looks at Hospitals. Health Information Foundation, Research Series 4, 1958.
Eliot Freidson and Jacob J. Feldman, Public Attitudes Toward Health Insurance. Health Information Foundation, Research Series 5, 1958.
Eliot Freidson and Jacob J. Feldman, The Public Looks at Dental Care. Health Information Foundation, Research Series 6, 1958.

6. Sol Levine, Odin W. Anderson and Gerald Gordon. Non-Group Enrollment for Health Insurance. Cambridge, Massachusetts: Harvard University Press, 1957.

7. Ethel Shanas, National Opinion Research Center.

8. Odin W. Anderson and Paul B. Sheatsley -- title not yet determined.

A SAMPLE OF DEVELOPMENTS FROM SAMPLING PROJECTS OF THE U. S. NATIONAL HEALTH SURVEY

By: Walt R. Simmons, National Health Survey

It was just a little over two years ago that the National Health Survey Act became a part of our laws. The act authorized the Public Health Service to develop and carry out a continuing program of statistics on the illnesses, injuries, impairments, and related characteristics of the population. Need for this legislation had been long felt, and carefully studied. The history of experimentation and study in this area, and delineation of many of the purposes and objectives of the new law have been reported elsewhere.[1] They will not be reviewed here except to note that the health of a Nation's people is possibly its greatest asset, and that measurement of levels and trends in health matters is essential to the efficient promoting of better health.

Within the Public Health Service the Surgeon General established the U. S. National Health Survey as the organization immediately responsible for conducting the program. Enabling legislation specifically authorized a three-phase program to accomplish its general intent. The first of three phases is a continuing survey of the general civilian population to provide statistics on the incidence and prevalence of acute and chronic illnesses; on numbers of accidental injuries and impairments; on volume and kind of disabilities; on use of medical, dental, and hospital facilities; and on a variety of allied topics.

The second phase consists of special studies, designed to complement the continuing survey with selected determinations which are narrower in scope than the more general undertaking. The special studies may relate to a particular aspect of specified health statistics, to a type of measurement which cannot be made in the more general survey, or to other features of the health status of people.

The third phase of the program is of particular note to members of a statistical association. It is of course common that any substantial ongoing activity will devote a part of its energy to improvement in operation. In the case of the Health Survey, however, the Congress directed that methods and survey techniques for securing statistical information on health be studied, with a view toward their continuing improvement.[2]

As had been intended, approximately the first 12 months following the passage of the Health Survey Act were devoted to establishing the nucleus of an organization, securing staff, initial planning, testing,[3] and as urged by the Act, to making contracts with other agencies whose facilities and capacities would be especially useful in fulfilling objectives of the program.

Now the National Health Survey has completed recently its first full year as an operating organization. It is the purpose of this paper to offer a sampling of that first year's activity. This is sampling somewhat in the nature of a cook sampling the soup with a spoon; or the commuter's survey of current American and world affairs by scanning his favorite morning paper. The report is not intended as a complete summary of the year, nor yet a scientific random sample of such a summary. It is hoped that identification of a few developments in each of the three phases of the program, with brief comments on those developments, will give some impression of the course the National Health Survey has been taking.

Findings From the Continuing Survey of the Population

It was decided early that a principal part of the continuing survey should be an area-type probability sample of households. In accordance with requirements and specifications set by the Public Health Service, arrangements were made for the Census Bureau to design the sample and to plan and conduct the household interviewing.

The household survey follows a highly stratified multistage probability design which permits a continuous sampling of the civilian population of the United States. The first stage of the initial design consists of an area sample of 372 from among 1,900 geographically defined Primary Sampling Units (PSU's) into which the continental United States has been divided. A PSU is a county, a group of contiguous counties, or a Standard Metropolitan Area. Subsequent stages of selection result in the choice of ultimate units of compact segments containing an expected six households each.

The central design relates to a sample of the population to be interviewed over the span of one year. In the year just passed, approximately 115,000 persons in 36,000 households over the Nation came into the sample. A striking and distinguishing feature of the survey is its continuous nature. About 700 households are interviewed each week. This block of households is itself a representative sample of the civilian population of the United States. Thus the design permits a continuous measure of characteristics of high incidence or prevalence, and through consolidation of, say, 13 or 52 weekly samples, more detailed analysis and treatment of less common characteristics and smaller categories.

The interview is rigidly structured, consisting in its first year of 40 items for identification of persons and households and their socio-economic characteristics, and of 54 questions concerning health conditions of respondents.

Both manual and electronic editing and tabulating are performed at the Census Bureau under Public Health Service specifications. The formal estimation process is a two-stage ratio method, sample results being controlled by residence, age, sex, and color of the population as determined by the 1950 decennial census and more recent independent estimates.

Five principal reports[4-8] of substantive findings from the household interviews have been released. They cover physician visits, dental visits, persons injured, disability, and a set of subjects termed "Selected Survey Topics." The first four of these are provisional, being based on either 13 or 26 weeks of interviewing. The fifth is in the nature of a partial summary of data collected during the first full year's interviews. A sixth report on acute conditions has been completed and will be released in a few days. Reports on hospitalization and on impairments are scheduled for publication before spring.

Copies of these reports are available. Here we shall note only a very few findings, chosen principally to illustrate types of information being released.

In a preliminary report on volume of physician visits it was observed that during the third calendar quarter of 1957 residents of the United States visited the doctor at an average rate equivalent to five times a year. Two thirds of all visits involved diagnosis or treatment, with one third representing preventive care and miscellaneous purposes. Over-all global figures of this type were issued initially to help sketch the general dimensions of various health problems, even though it was recognized that they rested on rather small samples, might have seasonal biases—since they were based on experience in but one calendar quarter—and would have other shortcomings. In some measure all findings from the survey will be considered tentative and provisional, pending subsequent evidence, although clearly reports based on longer periods of sampling and those treating phenomena of high frequency and relative stability will be more nearly definitive than other types of reports. Preliminary reports will be superseded with more reliable data as time and resources permit. For example, that first estimate of number of physician visits per person per year already has been replaced with an estimate based on 52 weeks of interviewing. The estimated rate was revised from a little under five visits per person per year to a little over five visits.

One figure picked up by the press was another provisional estimate from one quarter's interviewing. It indicated that there are 22 million edentulous persons in the United States. Reporters phrased this finding, "By Gum: 22 Million with Nary a Single Tooth."

A beginning has been made in assembling information on injuries and accidents—their types and distribution. Data from the first year's interviewing show nearly 50 million persons being injured annually. Injury here refers to bodily hurt or damage sustained from external causes either accidental or purposeful. Survey reports show some 40 percent of all injuries occurring in the home, about 15 percent at work outside the home, another 10 percent involving motor vehicles, and the remainder otherwise classified.

First publications on injuries point up a feature which will be a part of many Survey statistics. This is the severity of the condition. It will be recognized that ailments come in many degrees, and that the statistics on illnesses or injuries are the resultant not only of chosen definitions but also of the entire measurement procedure. Without going into this matter at length, it is noted that Survey publications include what it is hoped are sufficiently detailed descriptions of statistics to permit the reader to acquire a reasonably clear understanding of just what it is that has been measured. In the case of the above numbers on injuries, for example, a condition which might be a candidate for recognition as an injury is not so counted in survey publications unless it involved medical attendance or at least one full day of restricted activity of the person hurt.

Incidentally only about 50 percent of the possible-injury conditions picked up by interviewers have met the criteria of medical attendance or one day of restricted activity. A further sidelight might be noted in this connection. Statistics on injuries grow out of a sequence of questions which locate the injury as occurring in the first week prior to interview or in the second week prior to interview. If there be no differential bias of recall for the two weeks, one would usually expect about 50 percent of reported injuries to occur in each of the two weeks. Sample checks show experience is consistent with this hypothesis. For lesser injuries not meeting the severity test, however, more than 75 percent occur in the first previous week, thus suggesting a rapid falling off in recall for these minor hurts as the period of recall lengthens.

To most of us it is the human and personal aspects of illnesses and impairments that are their dominating characteristics. But survey reports underline also the social and economic impact of ill health. Data from 52 weeks of interviewing produced an estimate of $3\,^{1}/_{3}$ billion man-days of restricted activity including $1\,^{1}/_{3}$ billion bed-days for the civilian noninstitutional population for the year ending last June. This is an average of 20 days of restricted activity and almost 8 bed-days per person per year. Another significant figure is 600 million days lost from work by persons 17 years old and over.

Technical Developments in the Continuing Survey

As implied earlier, it is a policy of the National Health Survey to prepare fairly extensive and detailed descriptions of its procedures and statistics. The first reason for this policy is one of self-discipline—it might even be called a selfish reason: it is simply that the exercise of reducing concepts and procedures to written form promotes a more precise understanding by the Survey staff of what is being attempted. But there are two other reasons for the policy. The first is that it permits consumers to know the data much better, and consequently to make better use of the data. The second is the belief that a widespread dissemination of information concerning Survey methods is very likely to lead to critical review, re-evaluation, and improvement of Survey products.

In line with this policy, each substantive publication issued thus far by the Survey has carried appendices covering such matters as definitions, sample structure, data collection, response rates, and sampling variances. More comprehensive treatment of these and other matters is being offered in a separate series of technical publications,[1,9,10] three of which have been issued thus far. The first of these papers dealt with the purposes, objectives, origin, and program of the Health Survey. The second is a detailed description of the statistical design of the household survey. The third is an exposition of the concepts and definitions employed in the household survey.

The initial household-survey design gave unusual emphasis to the securing of separate estimates for a large number of geographical sectors of the country. This course was taken in order to provide data which would detect the highlights of geographical differentials, even though it was apparent that the scale of operation was such that sampling errors for individual geographic sectors would be very substantial. Three factors have led to a design modification which is to be introduced the first week in January. These factors are (1) trends in consumer desires, and especially as expressed by advisory bodies, toward greater emphasis on national figures including somewhat finer detail in subject matter; (2) after 18 months of collection, the existence of data which can be tabulated by relatively small geographic sectors and thus perhaps permit a partial realization of the earlier demand for such information; and (3) the experience in operation of the survey which has pointed the way toward certain efficiencies. At a later date descriptions of the design modifications—which do not change the household-interview survey in any of its fundamental structural features—will be prepared. There are noted here just three leading features of the changes.

1. In line with the trend of consumer interest, resources have been reallocated to give greater emphasis to statistics of national totals based on 52 weeks of interviewing. But estimates still can be prepared for 12 geographic regions and for 4 different degrees of concentration of population.
2. The number of Primary Sampling Units (PSU's) has been changed from 372 to 500 in accord with the evidence of data on unit costs. Some PSU's will appear in the sample each week; some only once a year; the average number of appearances in a year for a single PSU is six. Each week's sample will contain some 50-60 PSU's and each quarter's sample approximately 400 PSU's. Each week's sample continues to be a representative sample of the population.
3. These two changes and others of lesser impact will yield a 25-percent reduction in variance for the new design over the old, with respect to a typical National Health Survey national statistic based on a full year's interviewing.

All that has been said thus far, on both methodological and substantive matters has given primary attention to results based on interviewing over a year or perhaps a calendar quarter. And indeed that emphasis is in accord with planning and policy of the Health Survey. In the very first weeks of operation of the household survey, an event occurred which, coupled with basic survey design, has led to experimental work in a somewhat different area.

The event was the epidemic of Asian influenza. The design feature was the fact that each week's interviewing in the National Health Survey is a representative sample of the population. Since there was urgent demand for information on the course of the epidemic, many people felt the new Survey should attempt to trace that course on a weekly basis.

It should be recalled that there are only about 700 households in each weekly sample. The Health Survey had never expected to attempt a weekly report on any topic. But Asian influenza was thought to be a phenomenon of very high incidence and desire for statistics on the topic was insistent. Accordingly, there was some hurried investigation of the possibilities. Since the investigation was sufficiently promising, the Public Health Service decided to issue weekly provisional releases on number of new cases of acute upper respiratory diseases (including pneumonia and influenza), and average number of persons in bed each day from such diseases. These reports were issued throughout the epidemic and then were extended to provide weekly data for the period from July 1957 to

May 1958.[11] At the peak of the epidemic there were nearly 12,000,000 new cases a week, and more than 6,000,000 persons were in bed.

The experience with what in shop talk came to be known as the Flu Data has led to exploration of other possible useful features of weekly data from the Health Survey. These explorations have not yet resulted in definite conclusions, although they may be characterized perhaps as encouraging. It is likely that any weekly estimates which may be produced would be useful on a retroactive, rather than a current basis.

The statistical features of this investigation and particularly the problems of estimating population parameters and sampling variances are intriguing. The present paper will not be diverted to those problems. One or two aspects may be noted briefly, however, not only for their intrinsic interest, but because they suggest features of the Household Survey which have implications outside the matter of estimates based on one week's interviewing.

For several types of conditions persons each week are asked in the Health Survey for onset of illness: (1) "Last week or the week before," and (2) separately for "Last week" and for "week before." Thus if α_i is the true measure of incidence for week i for a particular type of illness, two primary estimates of α_i are

α'_i, the estimate for "last week" obtained from interviewing in week (i+1), and

β'_i, the estimate for "week before" obtained from interviewing in week (i+2).

A number of other estimates may be derived from these primary estimates. Among those to which special attention has been given in our investigations are derived estimates u'_i and x'_i defined by the following equations.

(1) $u'_i = \frac{1}{2}(\beta'_{i-1} + \alpha'_i)$, and

(2) $x'_i = \frac{1}{2}(u'_i + u'_{i+1})$.

The latter quantity x'_i has several attractive features as an estimator of α'_i. It is a weighted average of four primary weekly estimates, two of which referred originally to week i, one to week (i-1) and one to week (i+1). The estimate x'_i comes from <u>two</u> weeks of interviewing in weeks (i+1) and (i+2), rather than from a single week.

These characteristics give x'_i a relative stability as compared with several other possible estimates. Specifically, the variance of x'_i is for all items less than ½ that of α'_i and for some items is a much smaller fraction of the variance of α'_i.

The statistic x'_i is a biased estimate. Its expected value being $\alpha_i + 1/4(\alpha_{i-1} + \alpha_{i+1} - 2\alpha_i)$. It is apparent, however, that the form of the bias is such that bias will be small for most statistics in most periods, except possibly for the trend turning point of a very rapidly changing phenomenon. Even in such instances, the bias is not a cumulating one, and is almost certain to be minor over a period of several consecutive weeks. Figure 1 reflects the interviewing and estimation scheme just described. Experimentation with weekly data continues. It is expected that conclusions on technical characteristics of these data will be reached shortly. Attention will turn then to more specific consideration of whether and how they can make an informational contribution to the program.

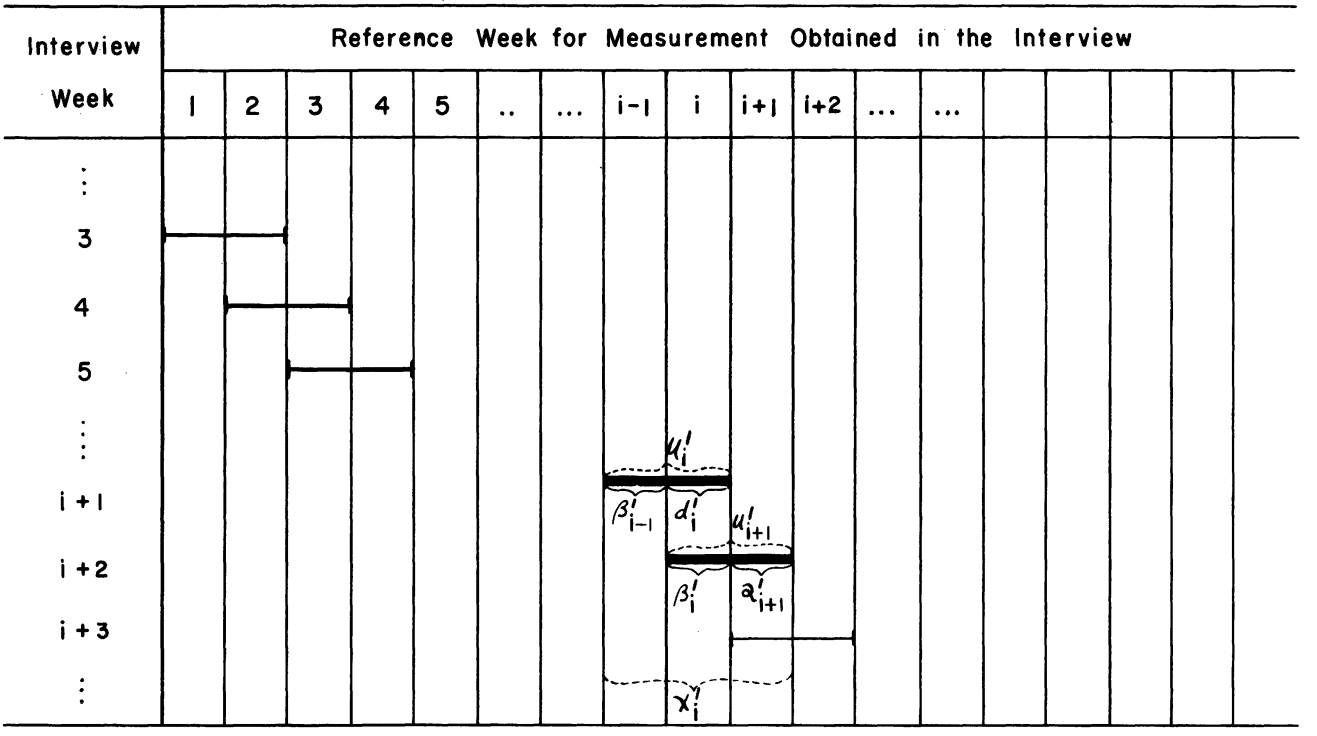
Other Technical Developments

During the past year most technical statistical work outside that connected immediately with the continuing household interviews has been related to testing or exploring certain aspects of possible alternative schemes of measuring health phenomena. This work is being done largely by other organizations under specifications and contracts prepared by the Health Survey. It includes such activities as:

A. Efforts to develop an effective standard questionnaire for taking medical histories as part of a procedure for a health examination survey, i.e., a survey in which medical and dental examinations are administered to a sample of persons. This work is being done by the Survey Research Center, University of Michigan.
B. Analysis of differences between information from a household interview and medical records on diagnoses of chronic disease, and on the fact of hospitalization, for a selected group of members of the Health Insurance Plan of Greater New York City.
C. Pilot study of hospitalization of decedents, since the household interview covers the living population only. This investigation is being conducted by the National Office of Vital Statistics.
D. Study of attitudes toward participation in a health examination survey, being conducted by the National Opinion Research Center.

These projects and others are currently in progress. A preliminary report on the attitude study

Figure 1. Relationship Between Interview and Reference Week for Selected Questions in the Health Household-Interview Survey. *And* Pattern of Estimates d_i' β_i' u_i' & χ_i'



The estimate $u_i' = \frac{1}{2}(d_i' + \beta_{i-1}')$

The estimate $\chi_i' = \frac{1}{2}(u_i' + u_{i+1}')$

just mentioned has been given by Borsky and Sagen[12] in a paper presented at the American Public Health Association meeting in St. Louis. The National Health Survey expects to institute a health examination survey at a later date. This was the reason for the study on motivation, since both evidence and opinion have suggested that nonresponse will be a major problem for the proposed health examinations. Borsky and Sagen concluded tentatively that favorable response is correlated with favorable attitudes toward doctors, medical research, and the government's role in health matters. They noted too, a close relationship between acceptance of the examination and the individual's concern over his own health. The next step is to see if these findings can be used in reducing the nonresponse rate.

The Health Examination Survey

In point of time the continuing household survey is the first major NHS activity. The second major NHS undertaking is expected to be a Health Examination Survey. As has been evident from previous remarks, this survey is in its planning and testing stages. The precise form which it will take has not been finally determined. The general plan contemplates the selection of a small probability sample of the population, and the administering of a limited single visit medical and dental examination to persons selected. The objective is to obtain medical statistics that cannot be gathered by other types of surveys, and in particular, to complement information obtained in the household survey.

Theoretical and experimental work in this survey have included in the past year—in addition to those items already mentioned—such matters as general survey and sample designs, proper content of examination, methods of conducting examination, and a pilot study of the project conducted in the Washington, D. C., metropolitan area. The pilot study applied to a sample of 180 persons, ages 18-64, who had been given the usual household interview and who had been offered the health examination. Examinations were actually given to 119 of the 180 persons. Experience in this pilot study is being analyzed while two more pilot studies are being planned, one in a smaller city, and one in a rural area. It is hoped that these

three studies will provide sufficient information that they may be followed by initiation of the Health Examination Survey itself.

Evolutionary Pattern of the National Health Survey

One final item should be added to this sampling of developments in the National Health Survey. This is an idea which was present at the start of the program and which has become one of the guiding principles of the Survey. It is that the National Health Survey is no single specific procedure, but rather is a coordinated set of undertakings intended collectively to provide intelligence on the health conditions of the population—and further that the Survey is an evolutionary program which must develop in response to need and experience.

REFERENCES

1. U. S. National Health Survey. Origin and Program of the U. S. National Health Survey. Health Statistics. Series A-1. PHS Publication No. 584-A1. Public Health Service. Washington, D. C., May 1958.

2. Public Law 652 - 84th Congress, Chapter 510-2d Session - S. 3076.

3. One aspect of this testing, concerning a pilot study of household interviewing, was reported on at the 1957 annual meeting of the American Statistical Association, by Harold Nisselson and Theodore D. Woolsey.

4. U. S. National Health Survey. Preliminary Report on Volume of Physician Visits, United States, July-September 1957. Health Statistics. Series B-1. PHS Publication No. 584-B1. Public Health Service. Washington, D. C., February 1958.

REFERENCES—Continued

5. U. S. National Health Survey. Preliminary Report on Volume of Dental Care, United States, July-September 1957. Health Statistics. Series B-2. PHS Publication No. 584-B2. Public Health Service. Washington, D. C., March 1958.

6. U. S. National Health Survey. Preliminary Report on Number of Persons Injured, United States, July-December 1957. Health Statistics. Series B-3. PHS Publication No. 584-B3. Public Health Service. Washington, D. C., May 1958.

7. U. S. National Health Survey. Preliminary Report on Disability, United States, July-September 1957. Health Statistics. Series B-4. PHS Publication No. 584-B4. Public Health Service. Washington, D. C., June 1958.

8. U. S. National Health Survey. Selected Survey Topics, United States, July 1957-June 1958. Health Statistics. Series B-5. PHS Publication No. 584-B5. Public Health Service. Washington, D. C., November 1958.

9. U. S. National Health Survey. The Statistical Design of the Health Household-Interview Survey. Health Statistics. Series A-2. PHS Publication No. 584-A2. Public Health Service. Washington, D. C., July 1958.

10. U. S. National Health Survey. Concepts and Definitions in the Health Household-Interview Survey. Health Statistics. Series A-3. PHS Publication No. 584-A3. Public Health Service. Washington, D. C., September 1958.

11. Provisional Tabulation From the U. S. National Health Survey. Current Statistics on Respiratory Diseases. Numbers 1-29, November 7, 1957-May 22, 1958.

12. Borsky, P. N., (NORC) and Sagen, O. K. (NHS): Motivations Toward Public Health Examinations. Paper presented at the American Public Health Association, St. Louis, Mo., October 1958.

WHAT AMERICANS THINK ABOUT THEIR MEDICAL CARE*

By: Jacob J. Feldman, National Opinion Research Center

In recent years considerable interest has been expressed concerning the public's image of the medical establishment. In part, this interest stems from public relations concerns. The spokesmen for medical organizations assume that favorable public opinion will serve to protect the current situation of the profession. It is feared that popular sentiments of distrust and hostility might be translated into unfriendly legislation or into some other undesirable consequences. Thus, there appears to be a great deal of anxiety among medical leaders concerning what people think of doctors.

There is also a second reason for concern about the public's sentiments on this issue. Attitudes are assumed to be major determinants of the extent to which medical facilities are utilized. It is held that antipathy towards doctors can act as a barrier to a person's seeking medical care. Consequently, we see that both health educators and leaders of organized medicine are quite concerned with how satisfied or dissatisfied the public is with the medical care which is available to them.

A number of journalistic accounts of public thinking on this subject have appeared in the mass media in recent years. They have been based primarily on hearsay rather than on systematic investigations, and so only the views of the more vocal advocates or critics of the medical profession are brought forward. One gets the impression, though, from most of these accounts, that the public is undergoing a widespread disenchantment with medicine. Ostensibly, a substantial segment of the public yearns for the "good old days" of the "horse-and-buggy doctor" and no longer holds the medical profession in high regard. Let me paraphrase a few sentences from a recent magazine article on the subject:

> According to the snap answers of most people you meet, a substantial majority of the population firmly believes that the U.S. medical profession is formed almost entirely of men who are avaricious, self-seeking, inhuman, lackadaisical, arrogant, and hypocritical. It is only recently that the medical profession has had to face this reaction of widespread and intense suspicion. Even thirty years ago the U.S. physician was generally held in pretty high esteem, even though his patients got a good deal less scientific diagnosis and care -- and died a lot earlier -- than is the rule today. For all that, the doctor was a leader in his community, was looked on not just with respect, but with affection, too.
>
> Now all this has changed. Affection has just about gone, and respect has dwindled. Even those who swear by their own physician's decency and honesty "know" that most other doctors are as rich as Croesus, drive nothing but Cadillacs, persistently cheat on their taxes, and are always boosting their already sky-high fees.

The article which I have just paraphrased goes on to explain why the medical profession is held in such disrepute by the public. Actually, this particular article happens to be a rather well-balanced and intelligent evaluation of the current position and problems of the medical profession. It is, in part, a defense of the profession against unjust criticisms and stereotypes ostensibly held by broad segments of the public. But the general soundness of this particular article does not concern us here. The present issue is the accuracy of the author's image of what the general public thinks of doctors and medical care.

In discussing this question I will draw primarily on data collected by the National Opinion Research Center during the summer of 1955. At that time we conducted interviews with an area-probability sample of some 2,400 adults. We also conducted interviews with random samples of some 450 physicians and 450 pharmacists whom members of the general public sample had designated as their regular physicians and pharmacists. The general public interviews averaged over two hours in length and covered a wide range of issues pertinent to the utilization of medical facilities. There were, for instance, batteries of questions tapping the individual's level of medical knowledge, his conception of the need for medical care in various situations, his attitudes toward doctors and hospitals, his satisfaction with the care he and his family had recently received, his health status during the preceding year, his own and his family's medical utilization and expenditures during the preceding year, and many other topics. Obviously, none of these areas could be probed very intensively in an omnibus survey of this type, but the wide range of issues touched upon enables us to view our data within a rather broad context.

The study was financed and sponsored by the Health Information Foundation. Paul Sheatsley of NORC and I are currently in the throes of preparing a comprehensive interpretive report of the results. Today I shall touch on only a few of these results.

Our respondents were given ample opportunity to express the negative sentiments to which the previously discussed magazine articles alluded. For instance, a series of statements, involving

*The data upon which this paper is based were derived primarily from a survey financed by the Health Information Foundation. A portion of the costs of preparing this particular paper was borne by a general grant to NORC from the Behavioral Sciences Division of the Ford Foundation.

supposedly common criticisms of doctors, was submitted to the respondent and he was asked whether or not he felt each statement was true of most doctors. Such an approach should, if anything, lead to an overestimate of the prevalence of dissatisfaction. It is extremely easy, with such a question structure, for the respondent to reflect back the negative stereotypes with which he has unquestionably come in contact in the past. Yet we generally found a majority of our sample rejecting these stereotypes. For instance, seventy per cent of the respondents rejected the validity with respect to majority of doctors of the following statement: "They don't give you a chance to tell them exactly what your trouble is." Sixty-three per cent rejected: "They don't take enough personal interest in you." Sixty-six per cent disagreed with: "Doctors like to give you medicine, even if you don't need it." Admittedly, some of the other critical statements elicited greater agreement. Half the respondents considered as applicable to most doctors the statement: "They don't tell you enough about your condition; they don't explain just what the trouble is," while almost half accepted: "Doctors make you wait entirely too long when you try to see them in their office." Thus, certain complaints attributed to the public are actually fairly widespread. Still, by and large, the medical profession seems to come off quite well if one considers the devastating character of the line of questioning employed in this study.

This stereotyped-statement device has also been used by several other agencies in connection with studies of attitudes toward doctors. Their results have been quite similar to ours. For instance, a study conducted by Ben Gaffin and Associates for the AMA examined, by this device, the extent to which greed and arrogance are attributed to doctors.[1] Gaffin found a much lower prevalence of the ascription of such undesirable traits than would have been predicted on the basis of the aforementioned journalistic assessments. In general, he found that the vast majority of the population hold their own doctors in remarkably high esteem. Sentiments toward the profession as a whole are far less favorable than those toward one's own doctor. Still, one could hardly help concluding that most people attribute relatively few faults even to doctors as a group.

Doctors are obviously not blindly worshipped. There is clearly quite a bit of dissatisfaction with the amount and frankness of the explanations which doctors give concerning what ails their patients. There are also rather common complaints about the amount of time allotted to each patient and the amount of time a patient has to spend in the waiting-room before he can see his doctor. But there is some evidence that these problems are viewed as being somewhat beyond the doctor's control. They are felt to be situationally induced rather than manifestations of flaws in the doctor's character. For the vast majority of the population, the good qualities of doctors seem to outweigh their faults by a rather substantial margin.

The pattern of intercorrelation among the agree-disagree-type items from the NORC survey has also been examined. While there is clearly a common esteem-antipathy factor underlying all these items, this factor can explain only a relatively small portion of the variance for the bulk of the items. It seems clear that attitudes toward various facets of the doctor-patient relation are reasonably differentiated. The "halo effect" appears to operate less strongly in this area than in many others because practically everyone has either personal or vicarious experiences with doctors rather frequently. These experiences tend to be somewhat idiosyncratic and thus preclude a rigid patterning of the attitudes toward different aspects of the relationship. One person may be led to complain about the fact that one has to waste a lot of time in the waiting-room, but he may be satisfied with the thoroughness with which the doctor explains one's ailments. Another person, with different experiences, may complain about the lack of thoroughness of doctors' explanations but not about having to spend excessive time waiting to be seen. This may imply that the conceptions which many people hold about doctors are quite solidly grounded in experience rather than being mere expressions of more general attitudes.

We also have a substantial amount of other data from our 1955 survey which confirm the generally favorable light in which physicians are seen. I have time to touch on this additional evidence only briefly. For instance, we found that a vast majority of the population considers the social standing of physicians to be extremely high. We have reason to believe that this ascription of high social standing to physicians is a sign of respect. It is interesting to note, in passing, that studies made by sociologists in eight foreign countries suggest that physicians are held in extremely high esteem in almost every advanced nation.

Turning to a different series of questions, our respondents were asked to what or to whom they gave credit for the improvement during the past thirty years in the chances of having good health. Many people answered this question in terms of generally improved living and working conditions for the population as a whole, the marked reduction in the prevalence of extreme poverty, the increased medical sophistication of the lay public, and the discovery of new drugs. Still, a majority of our sample credited the medical profession along with these other factors. Doctors were seen as being better trained, more knowledgeable, and better equipped than they were thirty years ago, and the public seems to accord the profession due credit for medicine's advances.

We also submitted a list of eighteen traits to the members of our sample. They were then asked to select those traits which characterize the kind of doctor they themselves liked best and those which characterize the kind of doctor liked least. A majority of the respondents selected as among the most desirable traits: "Very up-to-date." This trait of modernity was selected as desirable second most frequently among the eighteen traits, being outranked only by "Takes his time." Meanwhile, half the respondents selected "Old-fashioned" as an undesirable trait for a physician. "Old-fashioned" ranked second only to "Expensiveness" as an

undesirable trait. Thus we are forced to question the notion that people look back with nostalgia at the "good old days" and yearn for the family doctor of their childhood.

Unquestionably, the public's image of today's doctors as being far better equipped than their predecessors to cope with illness may in large part account for the aforementioned preferences. Evidence from a 1958 NORC nationwide survey of urban residents[2] confirms, in this connection, the findings of the 1955 study. Ninety per cent of the respondents thought that "doctors today know a lot more about treating sicknesses than they did thirty years ago," while an even greater majority thought that "the medicines we have today are much better than they were thirty years ago." These views, coupled with the salience to the patient of his physician's technical competence, serve as at least circumstantial evidence in support of the hypothesis that the esteem in which doctors are now held is based to a large degree on pragmatic considerations. In other words, people are satisfied with doctors because, relative to the past, doctors are now successful in their treatment of many conditions.

The question still remains whether the technical efficiency of contemporary physicians need be viewed as serving to compensate for otherwise deteriorated physician-patient relationships or whether this relationship has not been particularly suffering deterioration in the first place. It has been held that doctors today are more impersonal and cold with their patients than used to be the case.[3] Yet in the aforementioned 1958 NORC study, only one-third of the respondents thought that doctors today tend to take any less interest in their patients than did doctors of thirty years ago, and, in fact, almost half of the respondents thought that the situation had actually improved over the past.[4] It must be admitted that the doctor-patient relationship may have been at its best far longer ago than the 1920's. Perhaps the reference period in our survey questions involved a time at which the relationship had already become less satisfactory than it had been. But it seems highly unlikely that the public makes such fine distinctions in responding to survey questions. Thirty years ago and the last two or three generations are probably essentially equivalent in the interview situation. If this assumption is granted, then we must conclude that the bulk of the population does not romanticize the doctor-patient relationship of the past at the expense of the contemporary relationship.

Whether people generally still expect or require the same degree of personal intimacy and paternalism in their relations with doctors as people did several generations ago is at least open to question.[5] But whatever the degree of intimacy and warmth people would generally like to see shown by their doctors, only a minority seem to be disappointed by the present situation and yearn for the "good old days."

In passing, it might be noted that the medical profession's public relations were never as glorious as they are sometimes made out to have been. There is evidence that many people were aware in the past that doctors' healing abilities were at best severely limited and that their ethics were not beyond reproach. Shryock has shown that American doctors were in considerable disrepute up until the early twentieth century.[6] Apparently the situation did not improve very much even at that time. An extremely unsystematic survey conducted in and around Chicago during 1922 claimed to find a fantastic amount of hostility against the medical profession.[7] Doctors were described as rapacious, pompous, arrogant, inconsiderate, and so on. This was at a time when, according to some contemporary authorities, the medical profession was supposedly accorded nothing but deification by a grateful populace.

Actually, the medical profession has probably been rather widely esteemed at least for the last several generations. But throughout that time the opinion has been frequently expressed that the public's acceptance of the profession is currently on the downgrade. This situation is reminiscent of a well-known exchange between the editor of Punch and a disenchanted reader. In response to the reader's complaint that Punch was no longer as good as it used to be, the editor retorted, "It never was."

Returning to 1955, we followed up our query concerning which traits the respondent desired in his doctor with the following questions: "Would you say there are many doctors who are all of those things you mentioned, only a few doctors like that, or hardly any?" A majority of the respondents answered "Many," and only five per cent said "Hardly any." Here again we see that a substantial part of the public finds its desires satisfied by the doctors of today.

We also asked questions concerning the past medical experiences of the respondent and his acquaintances, the respondent's satisfaction and dissatisfaction with his own and his family's recent experiences with doctors, his feelings about the fees charged by doctors, his feelings about the quality of medical service available in his locality, his feelings about his own regular doctor, and several other related subjects. While there was certainly some criticism voiced against them in response to each of the questions, by and large, doctors received an overwhelming vote of confidence. I might add that the general tenor of our results has been confirmed by several other surveys conducted both prior and subsequent to ours.[8]

In general, doctors themselves appear to be aware of their favorable position in the eyes of the public. For instance, we asked our sample of physicians: "Do you think most patients give the physician too much credit, about the right amount, or too little credit for his part in their recovery?" Fully eighty-six per cent of the physicians responded, "Too much," or "About right," while only eleven per cent said, "Too little." This would seem to confirm our assessment of the public's feelings on this matter.

I hope that in the process of demolishing the straw man of public antagonism toward the medical

profession, I have not committed the opposite excess of creating the image of a totally satisfied public. There is, as I have pointed out earlier, substantial criticism concerning a number of particular facets of medical care. But a crucial qualification in this regard is that this criticism is tempered by two considerations: First, there is a tendency to absolve doctors from full responsibility for many aspects of medical care which laymen find irritating. Situational factors like a shortage of doctors or excessive demands by other patients are frequently held to account for shortcomings in the doctor-patient relationship. Second, many of the criticisms involve aspects of medical care which the patients consider to be of only secondary importance. Patients generally seem to be most concerned that the doctor be competent and personally pleasant toward them. Patients seem to overlook certain peripheral flaws and foibles if they have confidence in their doctor's ability and if they feel he treats them in a friendly fashion.

Nevertheless, I'd like to examine briefly some of the correlates of critical attitudes toward doctors, for whatever light this may shed on the situation. First of all, we find that criticisms are generally considerably more prevalent among respondents of lower socio-economic status than among respondents of higher status. This is hardly an earthshaking surprise, since it is in accord with all past research and speculation on the subject. Still, we should like to understand better the processes which underlie this relationship. We have not progressed far enough in the analysis of our data to say anything definitive about this, but please allow me to offer a few alternative speculations. First of all, there is evidence that people of lower socio-economic status actually experience poorer medical care than do people of higher status. Lower-class patients more frequently receive their medical care from less well-qualified physicians, for economic and ecological, as well as social and cultural, reasons.[9] In addition, they may sometimes not be treated quite as well as wealthier patients. These experiences are likely to color their conception of the medical profession as a whole and thereby induce the aforementioned antipathy.

A second possible explanation centers on the fact that lower-status individuals are more negative in their reactions, also, to many non-medical institutions than are individuals of higher status.[10] This may be due to historical influences operating through the sub-culture of the lower classes, to a negativism in general outlook induced by a hard lot in life, to limited access to the more gratifying constituents of institutions, or to a number of similar factors. In any event, the lower-class disaffection from the medical establishment is in line with its reaction to other contemporary institutions.

A third explanation is of a more methodological character: It is known that less-educated people are more prone to accept stereotyped notions than are the more educated. Thus, we may well be exaggerating the extent of dissatisfaction that our lower-class respondents actually feel. It should be made clear that the correlations between class position and attitudes toward physicians are in many instances of only a moderate magnitude to begin with, so whatever exaggeration has taken place could markedly affect the meaning of the results.

Older people were generally more critical of physicians than were younger people. Of course, older people have, on the average, much lower incomes and have had much less formal education than have younger people, and so the previous considerations are applicable here. In addition, there may be generational differences in the appreciation of the medical profession in reflection of historical changes in cultural norms. Another possible explanation lies in the fact that older persons suffer primarily from chronic illnesses. In truth, the medical profession is powerless to influence the course of many of the infirmities of the aged in comparison to what it can do for younger patients. Thus, the dissatisfaction expressed by the older segments of the population may be grounded on the relative inability of doctors to be of much help to them.

We also find that those who consider themselves to be in poor health are more critical of doctors than those who consider themselves in better health. Since lower-status and older persons consider their health to be much worse than do higher-status and younger people, the previous raft of hypotheses is relevant here. In addition, those who designate themselves as unhealthy have tremendously more first-hand experience with doctors than do those who feel healthy. Familiarity actually seems to breed contempt. The frequent users of medical care have had a far greater exposure to the risk of the occurrence of exasperating experiences with the medical establishment.[11] They have thus had much more opportunity to become disaffected with certain features of medical care than have individuals who turn to it more rarely. Several other factors also seem relevant. The chronically ill obviously cannot, by definition, be completely cured by their doctors. Also, an unhealthy person tends to be highly involved emotionally in his relations with his doctor; flaws in that relationship are, therefore, likely to be more salient to him than to a healthier individual, and, anyhow, the ill tend to be chronic complainers. We can thus see that there is no shortage of explanations for the correlation between health status and attitudes. In fact, it's rather surprising that this correlation is not of greater magnitude than it actually is.

I might add that even though socio-economic status, age, and subjective health status are heavily intercorrelated, we cannot view the association between any one of them and attitudes toward the medical profession as totally spurious. The partial correlations stand up well enough to consider each as an independent variable.

In closing this brief discussion of the correlates of dissatisfaction with the medical establishment, we might ask what difference it makes how people feel about doctors. At the beginning of this paper, it was indicated that many health educators consider negative attitudes toward doctors as constituting a serious barrier to the medical

attendance of illness. There is unquestionably some truth to this supposition, but we have some rather curious evidence to the contrary. As was suggested earlier, familiarity breeds contempt, so it is not too surprising that high utilization and disaffection with the medical profession should be concomitant with each other. This means that people with quite negative attitudes toward doctors utilize their services a great deal, so we immediately see that the negative attitudes are not an insurmountable barrier. Still, we might suppose that the individual's feelings toward doctors act as an intervening variable between his perception of his state of health and his utilization of physicians' services. All other things being equal, a person who views doctors as having few faults should utilize their services more frequently than a more antagonistic person. Well, we have correlated attitudes toward doctors with the number of times a doctor was seen during the past year, controlling by a rather refined index of perceived medical needs. This partial correlation still turns out to be remarkably close to zero. In general, it seems that if a person recognizes that he is ill, he will generally consult a doctor no matter what he thinks of the profession as a whole. Lord Byron, of all people, apparently recognized this fact about a century and a half ago. He wrote, in his Don Juan:

> This is the way that physicians mend or end us,
> Secundum artem: but although we sneer
> In health -- when ill, we call them to attend us,
> Without the least propensity to jeer.
>
> Canto X, St. 42.

I shall not here attempt to go beyond the explanation implicit in Lord Byron's formulation. All I can do is refer you to the forthcoming volume by Paul Sheatsley and myself, in which, it is to be hoped, there will appear an even more credible explanation of the low correlation between attitudes and behavior.

FOOTNOTES

1. Ben Gaffin and Associates, Inc., What Americans Think of the Medical Profession (Chicago: Ben Gaffin and Associates, Inc., 1955).

2. Paul N. Borsky, Motivations toward Health Examinations, National Opinion Research Center Report No. 70 (Chicago: 1959). (This survey was conducted for the National Health Survey of the U.S.P.H.S. and was restricted to urban residents 18-65 years of age. The exclusion of the older segment of the population probably did not affect the results reported here by more than a few percentage points.)

3. a) Earl L. Koos, " 'Metropolis' -- What City People Think of their Medical Services," American Journal of Public Health, 45 (1955), 1551-1557.

b) E. W. Hassinger and R. L. McNamara, Relationships of the Public to Physicians in a Rural Setting, University of Missouri Agricultural Experiment Station Research Bulletin 653 (1958).

4. Paul N. Borsky, op. cit., p.81.

5. While most people attach a great deal of importance to their doctors' having a "pleasing personality" and on their "taking an interest" in their patients, this does not necessarily imply a longing for the same degree of personalization and intimacy as was supposedly widespread during earlier periods. It takes more than the mere absence of grouchiness, curtness, and disinterestedness to constitute a primary relationship. Actually, so little is known about precisely how close a relationship with their doctors people presently have, presently desire, used to have, or used to desire, that this matter is completely open to conjecture.

6. Richard H. Shryock, "Public Relations of the Medical Profession in Great Britain and the United States: 1600-1870," Annals of Medical History, New Series, 2 (1930), 308-339.

7. B. C. Keller, "The Laity's Idea of the Physician," Illinois Medical Journal, 44 (1923), 13-17; and "Discussion of Miss Keller's Paper," 17-20.

8. a) C. R. Hoffer, D. L. Gibson, et al, Health Needs and Health Care in Michigan, Michigan State College Agricultural Experiment Station Special Bulletin 365 (1950).

b) Survey in 1951 sponsored by the Macon County (Illinois) Medical Society.

c) Ben Gaffin and Associates, Inc., op.cit.

d) E. Chen and S. Cobb, "Further Study of the Nonparticipation Problems in a Morbidity Survey Involving Clinical Examination," Journal of Chronic Diseases, 7 (1958), 321-331. (See Items 3 and 4 in Table IV, p. 325. The authors of the paper consider the results as indicating a surprisingly wide prevalence of negative attitudes toward physicians, but the standards underlying this judgment may be unrealistically severe.)

e) Survey in 1958 of Hopewell, New Jersey, sponsored by the American Academy of General Practice.

9. O. W. Anderson and J. J. Feldman, "Distribution of Patients Hospitalized for Surgery in the United States from July 1952 to July 1953," Bulletin of the American College of Surgeons, 43 (1958), 236-241.

10. Since there are also a number of institutions toward which lower-status individuals have more positive reactions than do individuals of higher status, this is hardly a significant consideration. The issue is raised here simply to emphasize the importance of viewing attitudes toward the medical establishment within the context of attitudes toward other institutions.

11. Frequent users of medical care are, of course, also more likely to experience favorable incidents than are infrequent users. While the favorable experiences appear to maintain the individual's general confidence in the medical profession, they do not allay the specific criticisms engendered by the bad experiences.

DISCUSSION

By: Sam Shapiro, H.I.P.*

As the discussant on this program, I am in a somewhat equivocal position, one might say a compromised position. For the past 4 - 5 years, I have been associated in one way or another with a number of the research projects carried out by HIF-NORC and as Walt Simmons has mentioned, for the past year or two, we at H.I.P., have been under contract with the NHSP to conduct one of their important methodological studies - the cross-checking of medical record information regarding medically attended illness against household survey reports of these illnesses.

However, I am not particularly concerned about these associations unduly coloring my views since I am also in the position of a consumer of the data being produced by the two organizations. Actually, their records of accomplishment are already so clear that I would be remiss if I were to ignore them. There is no serious discussion today of medical care financing or health insurance that does not utilize the results of H.I.F's research and the prospect is that H.I.F. will continue to play a critical role in finding solutions to the problems that plague the health insurance and medical care fields.

In the case of the NHSP, we have already seen the unusual happen - the release of a multiplicity of survey results within a few months of the completion of the field work. This undoubtedly is one of the values of the continuous sample design described by Walt Simmons. Also, major strides have been made in generating methodological studies in the brief period of NHSP's existence.

Having H.I.F. and NHSP on the same program inevitably leads to a joint consideration of their research activities. Both have a continuing concern with health and medical care - they are to be in business, hopefully, for a long time. Both have been dependent on the household survey as the primary source of information. And both have been producing base line data with an eye principally to national consumption.

There are some obvious differences of a substantive and methodological nature between the 2 organizations. H.I.F. has concerned itself with economic issues and with attitudinal and perceptional influences on health and medical care behavior. The NHSP, on the other hand, is currently concerned with measurement of morbidity, impairment, disability, volumes of medical care, types of care and the like. To be sure there is some overlap, but I think there is little danger of the 2 organizations getting in each other's way.

There are research issues that present special problems for governmental agencies to explore, which a non-governmental group can investigate <u>almost</u> without inhibition. Jack Feldman's paper is a case in point. It is focused on the probing of attitudes and the searching of socio-psychological correlates of behavior in the medical care field. These are research prongs that will continue to fall, I believe, in the province of "private enterprise."

The preview of the type of data Sheatsley and Feldman are incorporating in their book, strongly suggests that they will have a wide audience. Despite some overtones of pessimism about what the health educators are likely to get out of the material - especially since it seems to fly in the face of dogma - I think the findings are of great interest and use. The fact that a sizable minority in the population have negative attitudes on the components of doctor-patient relationship covered in the survey is important to people involved in dispensing medical care on an organized basis. It gives them a perspective on the problems they face and a framework for investigating their own situations.

I do not want to engage in a critical review of the specific points made in Feldman's paper but I do wonder whether the tentative conclusion of no association between attitudes towards doctors and receipt of medical care can be accepted without further analysis. There may be some question about the measure that was used to assess the influence of attitudes on behavior. Frequency of doctor visit is a very crude measure which may well conceal responsiveness to a set of symptoms. I think Feldman put his finger on the issue when he said "it seems that if a person recognizes that he is ill, he will generally consult a doctor no matter what he thinks of the profession as a whole." But, the point is when does a person recognize that he is ill? Do his attitudes towards the medical profession influence this recognition?

To return to my main theme, the comparison of the H.I.F. and NHSP programs, interspersed with a few critical comments, there are some interesting differences between the 2 in methodology and coverage. H.I.F. has used national probability samples for most of its major inquiries but has not hesitated to use local settings when these could help illuminate particular problems that would be difficult to investigate nationally. All of the studies are of an ad hoc nature - one-shot enterprises, with an emphasis on issues that are immediate. There is a tendency, therefore, to speed up the whole process of methodological development with calculated risks taken. There is also a tendency not to make its experiences and knowledge gained generally available. For example, in the national cost study, information reported on interview regarding hospitalization and costs was checked against hospital records. The findings of this reliability check would be of interest to other groups using the household interview as the source of data on hospitalization. But the details of this study have not been published.

One other study currently underway has the potential for providing unique data on the reliability of medical care information obtained via the household interview. In this investigation, 2 sources of information were used, with considerable overlap. The sources were the household interview and the records of physicians' services and hospital care in the health insurance plans where the families were enrolled. H.I.F. can perform an extremely useful service for the field as a whole by providing the opportunities to study

these materials in detail. Any future efforts elsewhere to exploit health insurance plan records as a source of data would profit from this experience.

NHSP is apparently even more heavily committed at this point than H.I.F. to the production of national data, with the household survey the primary mechanism. Regional data, although relegated to a secondary position, are not to be ignored, and it is to NHSP rather than H.I.F. or any other non-governmental group that we will be looking for regional information. The resources required to produce such data quickly outstrip the capacity of other groups - as witness the difference in sample size between H.I.F. studies and NHSP - about 3,000 households in H.I.F. and 35,000 households in NHSP.

Even with the production of regional data there will always be a gap in what NHSP can provide for the consumer at the local level. I do not know who should be concerned with the question of how to go from the national product to the local situation, but it does need attention and in time NHSP should be interested in advancing ideas on ways to resolve the problem.

Simmons has also mentioned that the development of trend data is an integral part of the program. This gives the organization an aura of permanence but 'permanence' cannot be bought with trend data. The longevity of NHSP will be directly related to its ability to contribute a quantitative basis for dealing with specific national problems of an immediate and long term nature. This, of course, will require a flexible program utilizing a variety of approaches, possibly including longitudinal type studies and follow-back studies aimed at amplifying a host of issues including social and economic consequences of illness, who takes advantage of rehabilitation programs, knowledge regarding these programs, etc.

Many of these issues do not lend themselves to a simple expansion of the regular interview but require special questionnaires quite different from the usual census type. Walt Simmons and Forrest Linder have put the matter succinctly in this way- "The Survey is a program of surveys, which use different approaches and have different end objectives as both the techniques and the needs for data evolve." What is left unsaid is that the NHSP has to take an aggressive role in clarifying these needs and that the development of special studies has to proceed expeditiously.

A distinguishing feature of the NHSP is the special position given methodological research. A primary objective is to find ways of improving the quality of the data. But this raises the question of how reliable the data have to be and how much of the resources should be devoted to the improvement of data. For example, if it is found that a measure of prevalence of diabetes is off by 20% are we to conclude that the figure is grossly deficient, moderately deficient, or satisfactory. Obviously there is no answer unless we know how the figure is to be used and the consequences of a specified error. This comes back to the point made previously - definition of the uses of the statistics is of critical importance.

There is one final comparison that I want to make between the 2 organizations. This has to do with background and modus operandi. The development of H.I.F.'s research program is really the product of 2 men - George Bugbee and Odin Anderson with a frequent strong assist from NORC. Whatever philosophy of action or framework or grand design for research that has emerged, is their handiwork. There are no advisory committees, no policy restrictions of serious consequences, no axes to grind. In short, the millennium! But, since part of my function is to needle, I wonder whether an occasional meeting with a group of experts in the field might not be advantageous as a way of supplementing the points of view Anderson now obtains informally. I am not suggesting a permanent advisory committee but a sounding board that has no "official" status.

NHSP came into existence with a blueprint which contributed greatly to its birth. I am referring to the document prepared in 1952 by the Subcommittee on National Morbidity Survey and called, "Recommendations for the Collection of Data on the Distribution and Effects of Illness, Injuries, and Impairments in the U.S." On rereading this report the other day, I was struck by the fidelity of the product to the model. There are to be sure, many areas which NHSP has not yet become involved in and many techniques mentioned that have not been developed or applied. The document encompasses activities that would take many years to carry out. The major problem is to determine priorities and the changes that experience dictates. It would be interesting to see the document rewritten 5 years after the NHSP started operating.

In conclusion, I believe that the consumers of morbidity and medical care data in this country have much to look forward to in the products of NHSP and H.I.F. We can only hope that a satisfactory division of labor continues and that the field will profit not only from the substantive material that is produced but from the methodological advances that will be made.

* Health Insurance Plan of Greater New York

VI

RESEARCH TOOLS OF THE SOCIAL SCIENTIST

Chairman, William Hodgkinson, Jr., American Telephone & Telegraph Company

Purposes of Scaling Techniques and the Choice Among Them—Robert P. Abelson, Yale University

The Pseudoinverse of a Rectangular Matrix and its Statistical Applications—Thomas N. E. Greville, Office of the Quartermaster General

Methods of Measuring Differences in Social Classes—Theodore R. Anderson, Yale University

PURPOSES OF SCALING TECHNIQUES AND THE CHOICE AMONG THEM

Robert P. Abelson, Yale University

Scaling can basically be defined as the establishment of rules by which a set of numbers can be assigned to a set of magnitudes of a property. Or, more simply, one can talk of assigning sets of numbers to set of objects which embody different magnitudes. In psychology, the objects may be people, physical stimuli, attitude statements, objects of preference, such as aesthetic objects or commodities, and so on.

A psychological scale using people as objects might be a scale of intelligence, a scale of some personality trait, or a scale of position on a controversial social issue; using stimuli, one might generate scales of subjective brightness loudness, heaviness, etc; a scale of attitude statements might specify the degree of pro-ness or con-ness of each; a scale of object preference would express the perceived value or utility of each object. There are other varieties of examples which could be given, but one thing is quite clear without proliferating examples; psychological scaling is at the outset a risky and speculative venture. The properties of intelligence, personality attributes, attitude, value, and even sensation are not well enough understood for quantification to be an easy and straightforward matter. One hopes, of course, that the attempt to scale a given psychological property will lead to some understanding of the property. But understanding in turn is often pre-requisite to making a reasonable try at scaling. It is a circular affair, like an inexperienced actor trying to break into show business.

Urgent need, however, has inspired the development of a variety of procedures for psychological scaling. I will discuss some of them briefly, without attempting to give a systematized account of the entire sweep of methods. This has been done very excellently in the recent book by Warren Torgerson (1958).

First of all, let me exclude from present consideration those methods which attempt to assign numbers to attributes of persons. There is in principle a duality between scales of persons derived by having persons respond to stimuli (e.g., test items) and scales of stimuli derived by having them judged or responded to by persons. (Mosier, 1940; Coombs, 1956) In practice, however, there are qualitative differences between scales of people and scales of stimuli. People are more complex than stimuli; it is harder to conceptualize properties of people than properties of stimuli. People are more changeable than stimuli; next year's scale of people is apt to be different from today's. People have less conveniently available time than stimuli; it is easy to use the same stimuli again and again to learn more of their properties, but not so easy to use the same people again and again. In addition, there is a curious sampling conundrum upsetting the duality -- if you want to make a stimulus scale more reliable, you can simply use more people to respond to or judge the stimuli, sampling these people from a specified population; but if you want to make a person scale more reliable, you often cannot sample more stimuli for people to respond to -- you must create or select new stimuli from an unspecified population.

The upshot of these considerations (and others) is that scales of stimuli should be more stable than scales of people, as apparently they are. I discuss all this because it will become important later.

In discussing specific scaling procedures, I will want to distinguish between those techniques which make assumptions about the psychological nature of the responses to the stimuli and those which merely seek useful quantitative scales without establishing a formal model.

Certain specific scaling procedures

Paired Comparisons

Experimentally, paired comparisons involves presenting successive pairs of stimuli to subjects who judge which member of each pair is darker, tastier, more preferred, or whatnot.

The method known to psychologists for analysing such data comes from Thurstone's Law of Comparative Judgment. (Thurstone, 1927) Briefly, this law (or better, hypothesis) states that any stimulus of judgment or choice activates response processes which are to some extent variable, between individuals or within individuals; further, that the distribution of potential responses in Gaussian. The probability that stimulus A will be chosen over stimulus B depends upon the means and standard deviations of the A-distribution and the B-distribution and upon the assumed correlation between responses when A and B are presented concurrently. The scale value of each stimulus is defined as the mean of its response distribution. The matrix of proportions of times each stimulus is chosen over every other is sufficient to determine within an arbitrary linear transformation the scale values of the stimuli. The normal distribution assumption is necessary to get the analysis off the ground, although other distribution forms are equally plausible. In fact, the paired comparison method developed by Bradley and Terry (1952) though couched in somewhat different terms, can in part be viewed as the equivalent of the Thurstone model, except that the logistic distribution is used instead of the normal distribution (Gridgeman, 1955).

These two methods, then, make some claim to an understanding of a reality underlying the choices. These claims lead to internal consistency checks on the methods, so that

applications yield more than scales; they provide tests of the models (Bradley, 1954, 1955; Mosteller, 1951). Unfortunately these tests happen not to be very sensitive to minor departures.

One other paired comparison method, due to Scheffé (1952) requires that the subject not only choose a member of each pair, but also specify the degree to which the member is the more preferred. Scheffé's method also makes assumptions, but these are in the spirit of statistical assumptions rather than psychological assumptions. Other methods, employing ranks or rankits, or the ingenious manipulations of paired comparisons matrices suggested by Kendall (1955), usually make no pretense to being formal models of the judgment process.

Categorical or single stimulus methods

The single stimulus methods present subjects with one stimulus at a time and require a numerical rating of the stimulus. The placements can be taken as they stand, or else some model can be imposed on the data.

A popular model in psychology is the so-called successive intervals model (Edwards and Thurstone, 1952; Diederich, Messick, and Tucker, 1957). This derives again from Thurstone's judgment model. The boundaries between adjacent ordered categories are assigned numerical values on the same underlying continuum as the Gaussian response distributions. There are internal consistency checks on the method, as with the Thurstone paired comparisons model.

Guttman scales

Guttman scaling (Guttman, 1941, 1947) is primarily an assumption about the nature of reality rather than a scaling technique per se. A set of stimuli or items are presented to a set of individuals, who respond either positively or negatively to each item (say, the item might be a statement and the individual is to agree or disagree with it). It is assumed that there exists an ordering of the items such that if an individual agrees with a given item he also agrees with all items beneath the given one in the ordering. Thus the entire set of items constitutes a kind of staircase, with individuals grouped according to which level along the staircase they occupy. Empirically it always turns out that the staircase pattern is not perfect; an index called the Reproducibility of the scale specifies the proportion of responses which conform to the perfect pattern, and investigators are usually happy when this index reaches .90. It has become a social science parlor game to see what new types of responses achieve scales with acceptable reproducibilities. The chief criticism of this game is that as a model of the nature of reality, the Guttman scale is rarely convincing -- the Reproducibilities generally lie about half-way between chance level and the perfection required by the model -- and when the requirement of perfection is far from being met, then the method loses its special appeal as a means of obtaining clean scales. It becomes no less arbitrary a procedure than the next man's.

Magnitude estimation

It has been argued by S. S. Stevens and others (Stevens, 1957; Stevens and Galenter, 1957) that scaling methods based upon the use of variability of response, like paired comparisons and successive intervals, are misguided. The use of intra-subject confusion on which to erect the edifice of measurement is decried. If you want estimates of subjective magnitudes, says Stevens, ask the subject directly. Thus to scale loudnesses, for example, one might give the subject a preliminary tone and tell him to call its loudness 10. Then the subject attaches numerical values to subsequent stimuli in accordance with his subjective estimate of their loudness. A related device in the field of consumer preference would be to ask the subject how many dollars he would be willing to pay for each of a number of items.

The magnitude estimation methods have been making considerable recent headway in psychophysics. It appears that on a large number of physical continua such as loudness, brightness, heaviness, etc., there is a common function relating subjective magnitude to objective magnitude. The continued use of the method depends heavily upon the discovery of such broad empirical relationships, for there is no internal consistency check on the method to verify its validity as a model. Indeed it is not intended to be model at all, but rather a straight-forward means for generating numerical scales.

Other methods

There are a variety of scaling methods devised by Coombs (1950, 1952, 1954, 1958) which are similar to Guttman scaling in that perfection of response pattern is required, but which by and large are more flexible. In particular, they allow for multi-dimensional scales, and offer the possibility of scales of objects as viewed by single judges, rather than sets of judges. The level of measurement achieved is not numerical, but something short of this -- so-called ordered metric measurement.

Another technique resulting in ordered metric scales uses a lottery to establish relative scale positions of objects (Siegel, 1956). Subjects have been asked such piquant questions as, "Would you rather: a) admit Negroes to your school for sure, or b) elect a 50-50 bet, with heads signifying that one of your friends marries a Negro and tails that Negroes be allowed to ride in local buses"? A subject selecting alternative a) is inferred to hold a greater value or utility difference between intermarriage and school desegregation than between school desegregation and bus desegregation.

Further new techniques may soon flow from a model of choice behavior conceived by Duncan Luce (1959). The basic axiom of this model is

essentially that when one object is to be chosen from a set of objects, the ratio of the probability of choice of A vs. B is independent of the number and nature of other objects presented along with A and B for choice.

Studies comparing scaling methods

Though not all methods have been compared with all other methods, a number of comparison studies have been made. Mosteller (1958) compared a variety of assumptions for the underlying distribution in the paired comparisons model -- rectangular, arc sine, normal, exponential, and a t-distribution with high tails -- and found that it makes almost no difference which one is used. The correlations between the resulting scales run from .9965 to .99998. The scales may be characterized as very weak quadratic transformations of one another.

Jackson and Fleckstein (1957) compared a variety of experimental and analytic methods involving paired comparisons. The results were extraordinarily insensitive to choice of method. A representative correlation is .975. Bliss, Greenwood and White (1956) made even more extensive comparisons of methods, including ranking techniques as well as paired comparisons. Again the picture is one of fantastically high correlations among outcomes, especially when differing analytical methods are used on the same experimental data. When different types of experimental data are compared, as with Scheffé's technique vs. a version of Kendall's technique the correlation between scales drops to its lowest, a rock bottom value of .990. Jones (1958), Gulliksen (1953), and others have compared different analytical solutions for successive interval data. The correlations are in the .99 range. Kelley, Hovland et. al.(1955) found a correlation of .910 between paired comparison and successive intervals values. This even in a situation where the category assignments required for the successive intervals method were very deviantly performed by the judging group.

It is not surprising that paired comparisons and successive intervals scales are very closely related, for after all, they are both based on Thurstone's judgment model. Nor is it surprising that paired comparisons scales correlate highly with scales derived from rank ordering of the stimuli. (Incidentally, these correlations can be astronomical. In one study by Ross (1955), these two methods correlated .998). Transitivity of choice (in a statistical, rather than an absolute sense) is an unexceptioned finding in investigations to date (Davidson & Marschak, 1957; Davis, 1958). No matter whether rank ordering of objects is done by separate pairs or all at once, the results are essentially the same. There is good reason to expect, too, that paired comparisons scales extracted by Thurstone's method or the Bradley-Terry method will prove almost identical. Thus many of the high correlations between scales derived by different methods can be rationalized as mathematical necessities or as natural consequences of mundane empirical regularities.

However, there are some correlations between scales which are not so readily explained. Three disparate methods have been applied to scale a set of nine color chips multi-dimensionally -- that is, where scale values on more than one dimension are assigned to each stimulus. The three methods were: a successive intervals method extended to the multi-dimensional case (Messick, 1956) the so-called complete method of triads due to Torgerson (1952) and Shepard's method (1957) which uses probabilities of confusion in learning responses to the stimuli. Shepard (1958) has given the correlations between the three methods on two dimensions as around .98. Other striking examples of high correlation using disparate methods occur in psychophysics. A most interesting example is a recent study by Galanter and Messick (1958). They applied both successive intervals scaling and the direct magnitude estimation method to a set of stimuli varying in physical loudness. The psychological loudness determined by the two methods correlate .896. I will say more about this study later. A study by Benson and Platten (1956) comparing these same two methods in the domain of preferences yields a correlation of .968.

There are other studies which could be cited, but the long and short of the situation is that it is almost impossible to find two scaling methods which when applied to the same stimuli will yield anything less than a very high correlation. It appears that if the objects of a given psychological domain possess sufficient underlying order so that some one scaling method is reliable, then any other reasonable scaling method will reveal essentially the same order. Any knife that cuts at all, even the bluntest, will expose the same corpus. Mind you, I am talking about scales of stimuli, not about scales of people. Method comparisons on the latter yield correlations that run eighty-ish down through fifty-ish or lower. But as I have already indicated, there are several reasons for expecting a certain instability in people-scales.

What is the nature of the underlying stimulus constancies upon which different methods converge? Is it that a stimulus scale basically amounts to a rank order and that all methods reveal the true rank order? No, for even with a rank order correlation of unity among n objects, the product-moment correlation can be as low as $1/\sqrt{n-1}$, or .33 for ten objects, .20 for twenty-six objects. Even if one deals not with minimum possible correlations but with expected or average correlations in some sense, it is still evident that the obtained correlations in comparisons of scaling methods are much higher than could be accounted for by postulating shared rank order alone. Something even stronger than mere ordinality must typically underly the set of stimuli. There are in principle many gradations between ordinal measurement and cardinal measurement. A very interesting pursuit would be to try to find the level of measurement coordinate with the magnitude of

correlations commonly found between scaling methods. Recent work by Tukey and myself may conceivably shed some light on this question (Abelson & Tukey, 1958).

The objection may be raised that product-moment correlation is not the most appropriate indication of degree of equivalence between two scales. A very high correlation occurs when one scale is but a mild transformation of the other, yet the distortion might be the thing which most interested us. A very high correlation would also occur if two scales were linearly related save for a small number of deviant stimuli displaced from the line of realtionship. This displacement itself might be the phenomenon of interest. Examples exist both of mild transformation and of idiosyncratic discrepancies or "bumps".

In order to know whether scaling method makes a difference, we must inquire into the investigator's purposes. For some purposes, the choice of a particular scaling technique may not matter at all, while for others it may.

In order to draw some kind of a sample of purposes from the population of current purposes, I have inspected most of the scaling articles in the psychological literature, 1954-1957, listed in a previous comprehensive review (Messick & Abelson, 1957). For each published study, I noted the investigator's purpose, and asked whether the conclusions would have been different had he chosen some other scaling method. I assumed that the worst that could have happened via change in method would have been the introduction of a few slight bumps in the scale or else a mild transformation of the entire scale. I classified the answers to the question "Would it have made a difference?" into three categories: yes, no, and maybe. In the abstract for this paper, the figure 10% is given for the percent of "yes" answers. At this juncture I feel that this percentage, though not incorrect, is superficial. It takes no account of the relative importance of various studies. I think my best course at this point is to give examples of studies varying in the degree to which the scaling method made a difference.

At one extreme are studies in which the purpose is to verify the existence of a gross effect upon judgments of experimental or natural conditions. For instance, it has been demonstrated that when two foods are tasted at one time point and preference judgments elicited at a later time point, the difference in degree of preference is far less than if the preference judgments are given at the time of tasting (Schwartz & Pratt, 1956). Any method of preference scaling would doubtless show the same effect. A study comparing the basic values of college students in the U. S., India, and other countries showed huge differences in value-orientation (Morris & Jones, 1955). Any scaling method would have found the essence of these differences. And so on. Of course, where experimental conditions differ not grossly, but subtly, there is more chance for scaling method to make a difference. However, suppose that two scaling methods differ only in that one is a mildly non-linear transformation of the other. Common experience with the analysis of variance would indicate that sharp differences in outcome are not at all likely to occur under these circumstances (though, to be sure, the analysis may feel more satisfying in one version than in another). The choice between methods is then based upon experimental convenience, richness of by-product information, and the aesthetics of the analysis.

At the other extreme are studies where the scaling method may make a crucial difference. I will cite four of these, typifying the kind of purposes involved.

The previously cited study by Kelley et. al. had both Negro judges and white judges scale the favorableness of statements about Negroes via paired comparisons. It turned out that the scales produced by the two sets of judges were almost perfectly linearly related except for three statements which in the white scale were displaced slightly but discernably above where they were in the Negro scale. Successive interval scaling did not reveal this effect. The displacement happens to be convincing in this case because the three displaced statements concerned "separate but equal" treatment of Negroes. The Negro judges did not consider these so favorable to Negroes, relatively, as did whites.

Another kind of crucial dependence upon scaling method occurs when the investigator wants to know the functional dependence of a psychological scale upon a physical or natural scale. Davidson, Suppes, and Siegel (1957) scaled the utility of very small amounts of money by a lottery method similar to the one mentioned before. With most subjects, utility turned out to be very nearly linear in money. Had another variety of scaling method been used, utility might have turned out, say, a square root function of monetary amount.

In the study by Galanter and Messick mentioned earlier, a successive intervals scale of loudness yielded a rather different function for the relation between subjective and objective loudness than did the magnitude estimation method (even though the two methods correlated .896).

In these two examples, scaling method makes a difference to the extent that the resulting function makes a difference. But here we encounter a new consideration. The successive intervals method yields a logarithmic psychophysical law; the magnitude method yields a power law. One may go from the latter to the former by a logarithmic transformation and come back with an exponential transformation. The two methods are equivalent in the larger sense that it is possible to go from one to the other Who is to say which method is "better"? The issue seems again to come down to a question of convenience.

One further example may clarify the picture. In a study by Cliff (1956), evaluative adjective, and adjective-adverb combinations were scaled for intensity by the successive intervals method. Cliff then sought the relationship between the judged favorableness of the combinations, such as "rather evil", "de-

cidedly charming", etc. and the individual components. He concludes that adverbs exert a multiplicative effect upon adjectives. The quantitative support for this conclusion is really remarkably good. Let us imagine the magnitude methods applied to the same stimuli yielding scale values exponentially related to Cliff's values. His law of adverbs would then be lost. One would need the extra insight, "Take the logarithm of the scale values and then there will be a multiplicative law". The choice between scaling methods yielding different functions only begins to matter when these functions are imbedded in still further operations.

One may also stop and ask, "How is it that one method comes to yield a scale that is, say, logarithmically related to another? What is going on in the judgment process anyhow? This question can be asked of the successive intervals and magnitude methods and is as yet unanswered to everybody's satisfaction. One needs a superordinate model of the judgment process which will explain both types of scale outcomes. Thus, as has happened in other areas of psychology, one starts out with an interest in psychological content matter, constructs measurement devices to deal with that content, and ends up studying the interaction between measurement and measuree instead of the original content matter. Explicit models of judgment or choice processes, like Luce's, thereby assume great importance.

Summary

In most applications of psychological scaling techniques to objects other than people, the results are monumentally indifferent to the choice of scaling method. For most purposes, the choice of method should be made in terms of convenience. Three exceptions to this generalization are: 1) When the conclusions depend upon slight bends or bumps along the scale of stimuli, 2) When a functional relationship is sought between the stimulus scale and some external scale, provided that the function enters into some further consequences, 3) When the interest is not really in scale values at all but rather in a model of what is going on in the response, judgment, or preference process.

When looking for bumps or bends, one wants to choose the method that reveals meaningful bumps clearly and at the same time suppresses false or phrenological bumps. When looking for functional relationships, one wants to choose the scale methods yielding the neatest relationships. When testing models, one of course uses methods appropriate to test the particular model at hand.

It is extremely doubtful that any single method will prove superior on all counts. We will probably have to continue to live with too many methods, differences that make no difference, and too few models.

References

Abelson, R. P. and Tukey, J. W. (1958) Efficient utilization of non-numerical information in quantitative analysis. (In preparation).

Benson, P. H. and Platten, J. H. (1956) Preference measurement by the methods of successive intervals and monetary estimates. J. appl. Psychol., 40, 412-414.

Bliss, C. I., Greenwood, Mary L., and White, Edna S. (1956) A rankit analysis of paired comparisons for measuring the effect of sprays on flavor. Biometrics, 12, 381-403.

Bradley, R. A., and Terry, M. E. (1952) The rank analysis of incomplete block designs. I. The method of paired comparisons. Biometrika, 39, 324-345.

Bradley, R. A. (1954) Incomplete block rank analysis: On the appropriateness of the model for a method of paired comparisons. Biometrics, 10, 375-390.

Bradley, R. A. (1955) Rank analysis of incomplete block designs III. Some large-sample results on estimation and power for a method of paired comparisons. Biometrika, 42, 450-470

Cliff, N. (1956) The relation of adverb-adjective combinations to their components. ONR Technical Report. Princeton, N. J.: Princeton University.

Coombs, C. H. (1950) Psychological scaling without a unit of measurement. Psychol. Rev., 57, 145-158.

Coombs, C. H. (1952) A theory of psychological scaling. Ann Arbor: University of Michigan Press. Eng. Res. Bull. no. 34.

Coombs, C. H. (1954) A method for the study of interstimulus similarity. Psychometrika, 19, 183-194.

Coombs, C. H. (1956) The scale grid: some interrelations of data models. Psychometrika, 21, 313-330.

Coombs, C. H. (1958) An application of a nonmetric model for multidimensional analysis of similarities. Psychol. Reports. (In press).

Davidson, D., and Marschak, J. (1957) Experimental tests of stochastic decision theory. Cowles Foundation Discussion Paper #22. New Haven, Conn.: Yale Univ. (Mimeo).

Davidson, D., Suppes, P., and Siegel, S. (1957) Decision making: An experimental approach. Stanford, Calif.: Stanford Univ. Press.

Davis, J. M. (1958) The transitivity of preferences. Behavioral Sci., 3, 26-33.

Diederich, Gertrude, Messick, S. and Tucker, L. (1957) A general least squares solution for successive intervals. Psychometrika, 22, 159-173.

Edwards, A. L. and Thurstone, L. L. (1952) An internal consistency check for scale values by the method of successive intervals. Psychometrika, 17, 169-180.

Galanter, E. and Messick, S. (1958) The relation between category and magnitude scales of loudness. (In preparation).

Gridgeman, N. T. (1955) The Bradley-Terry probability model and preference tasting. Biometrics, 11, 335-343.

Gulliksen, H. (1954) A least squares solution for successive intervals assuming unequal standard deviations. Psychometrika, 21, 125-134.

Guttman, L. (1941) The quantification of a class of attributes: A theory and method of scale

construction. In Horst, Paul, et. al., The Prediction of Personal Adjustment. New York: Social Science Research Council.

Guttman, L. (1947) The Cornell technique for scale and intensity analysis. Educ. Psychol. Msmt., 7, 247-280.

Jackson, J. E. and Fleckstein, Mary (1957) An evaluation of some statistical techniques used in the analysis of paired comparisons data. Biometrics, 13, 51-64.

Jones, L. V. (1957) Methodology of preference measurement. Quartermaster Food and Container Institute Report P1101 #4. Chicago: Univ. of Chicago.

Kelley, H. H., Hovland, C. I., Schwartz, M. and Abelson, R. P. (1955) The influence of judges' attitudes in three methods of attitude scaling. J. Soc. Psychol., 42, 147-158

Kendall, M. G. (1955) Further contributions to the theory of paired comparisons. Biometrics, 11, 43-62.

Luce, R. D. (1959) Individual choice behavior: A theoretical analysis. New York: Wiley (In press).

Messick, S. (1956) An empirical evaluation of multidimensional successive intervals. Psychometrika, 21, 367-375.

Messick, S. and Abelson, R. P. (1957) Scaling and measurement theory. Rev. Educ. Res., 27, 487-497.

Morris, C. and Jones, L. (1955) Value scales and dimensions. J. abnorm. soc. Psychol., 51, 523-535.

Mosier, C. I. (1940) Psychophysics and mental test theory. Fundamental postulates and elementary theorems. Psychol. Rev., 47, 355-366.

Mosteller, F. (1951) Remarks on the method of paired comparisons: III. A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed. Psychometrika, 16, 207-218.

Mosteller, F. (1958) The mystery of the missing corpus. Psychometrika, 23, 279-289.

Rimoldi, H.J.A., and Hormaeche, Marceva. (1955) The law of comparative judgment in the successive intervals and graphic rating scale methods. Psychometrika, 20, 307-318.

Ross, R. T. A linear relationship between paired comparisons and rank order. (1955) J. exp. Psychol, 50, 352-354.

Scheffé, H. (1952) An analysis of variance for paired comparisons. J. Amer. Stat. Assn., 47, 381-400.

Schwartz, N. and Pratt, Cornelia H. (1956) Simultaneous vs. successive presentation in paired comparison situation. Food Res., 21, 103-108.

Shepard, R. N. (1957) Stimulus and response generalization: a stochastic model relating generalization to distance in psychological space. Psychometrika, 22, 325-345.

Shepard, R. N. (1958) Stimulus and response generalization: Tests of a model relating generalization to distance in psychological space. J. exp. Psychol, 55, 509-523.

Siegel, S. (1956) A method of obtaining an ordered metric scale, Psychometrika, 21, 207-216.

Stevens, S. S. (1957) On the psychophysical law. Psychol. Rev., 64, 153-181.

Stevens, S. S. and Galanter, E. H. (1957) Ratio scales and category scales for a dozen perceptual continua. J. exp. Psychol, 54, 377-411.

Thurstone, L. L. (1927) A law of comparative judgment. Psychol. Rev., 34, 273-286.

Torgerson, W. S. (1952) Multidimensional scaling I. theory and method. Psychometrika, 17, 401-419.

Torgerson, W. S. (1958) Theory and methods of scaling. New York: Wiley.

THE PSEUDOINVERSE OF A RECTANGULAR MATRIX AND ITS STATISTICAL APPLICATIONS

By: T. N. E. Greville, Office of the Quartermaster General

SUMMARY

The connection between inversion of matrices and solution of nonsingular systems of linear equations is well known. The statistician, however, is often concerned with systems of equations in which the number of equations exceeds the number of unknowns and there is no exact solution. In such cases the least squares solution of the system is usually sought, and the classical matrix theory is of little avail.

In 1920 E. H. Moore announced a generalization of the notion of inverse of a matrix, which provides a generalized inverse or "pseudoinverse" for rectangular matrices, as well as for singular square matrices. Bjerhammar in 1951 and Penrose in 1956 have shown that this pseudoinverse is related to the least squares solution of an inconsistent system of linear equations in a way analogous to the relationship of the classical inverse to the solution of a nonsingular system.

In the present article two possible applications of this concept to statistical procedures are suggested. These relate to (1) the computation of multilinear regression coefficients and (2) least squares curve fitting, with particular reference to the fitting of polynomials. In the latter application the procedure suggested here has an advantage over the use of orthogonal polynomials in that unequal spacing of the arguments does not increase the amount of calculation required. Among other possible uses of the pseudoinverse not discussed here is its application to bivariate interpolation.*

A recursive algorithm is described by which one can derive from the pseudoinverse of a given matrix that of a second matrix obtained by the addition of a single column. Thus one computes first the pseudoinverse of the first column of the coefficient matrix, then that of the first two columns, and so on until the pseudoinverse of the entire coefficient matrix is obtained. In the regression application, this makes it possible to arrange the variables in decreasing order of their probable importance in the regression equation, and to stop the process when it appears that the introduction of further variables will not have a significant effect. Similarly, in fitting a polynomial, the process is arranged so as to fit polynomials of successively higher degree, and one can stop when it appears that the most suitable degree has been reached. In either case the residual variance is easily obtained as a by-product.

1. INTRODUCTION

It is of course well known that any square matrix A with nonzero determinant has a unique inverse A^{-1} such that

$$AA^{-1} = A^{-1}A = I , \tag{1}$$

where I denotes the unit matrix or identity matrix having 1's along its principal diagonal and 0's elsewhere. It seems to be not so well known that in 1920 [1,2] the eminent American mathematician E. H. Moore announced a generalization of the inverse concept to include rectangular matrices and those with vanishing determinant. Little notice was taken of Moore's discovery for about 30 years, but during the past decade the properties of this generalized inverse or pseudoinverse have been vigorously explored by Bjerhammar [3,4,5], Penrose [6,7] and Hestenes [8].1/ It is found to have some useful applications in numerical analysis and statistics. Of the latter, probably the most obvious are in connection with multiple regression and least squares curve fitting.

It is the purpose of the present expository article to indicate these statistical applications and to describe a simple numerical algorithm for computing the pseudoinverse of a given matrix. No claim to originality is made; everything in this article is explicit or implicit in the work of Bjerhammar, Penrose and Hestenes.

A knowledge of the elementary properties of matrices is assumed.2/ In particular, the reader should keep in mind that a vector can be thought of as a matrix of one column (or row), and that the row-by-column rule for multiplying two matrices together implies that, in a matrix product

$$AB = C ,$$

each column of C is a linear combination of columns of A and each row of C is a linear combination of rows of B. Extensive use will be made of the notions of vector spaces and orthogonality. A fuller version of this article, including a brief exposition of these concepts and proofs of certain important properties of the pseudoinverse, can be obtained from the author in mimeographed form.

2. SOLUTION OF SYSTEMS OF LINEAR EQUATIONS

It is well known that a system of linear equations

$$\sum_{j=1}^{n} a_{ij} x_j = b_i \qquad (i = 1, 2, \ldots, m) \tag{2}$$

can be written compactly as a matrix equation

$$Ax = b , \tag{2}'$$

where A denotes the matrix (a_{ij}), x is the vector (i.e., single-column matrix) whose elements are the values of the variables x_j which consti-

tute a solution of the system, and b is the vector whose ith element is b_i. If A is nonsingular (i.e., if m = n and its columns are linearly independent), it has a unique inverse A^{-1} satisfying equation (1), and moreover the system (2) or (2)' has a unique solution given by

$$x = A^{-1} b . \tag{3}$$

The statistician is often concerned with systems of equations in which m > n and there is no exact solution. In such a case, Bjerhammar [3,4] and Penrose [7] have shown that the "best" solution in the sense of least squares is given by

$$x = A^{\dagger} b ,$$

where $A^{\dagger}$ is the pseudoinverse of the rectangular matrix A. The definition and computation of the pseudoinverse will now be taken up.

3. DEFINITION AND PROPERTIES OF THE PSEUDOINVERSE

Any nonzero real matrix A of rank r can be expressed as a product

$$A = BC , \tag{4}$$

where the r-column matrix B and the r-rowed matrix C are both of rank r. To show this, let B be any matrix whose columns form a basis for the column-space of A. Then a matrix C exists such that A = BC, and C is of rank r, since the rank of a product cannot exceed the rank of any factor. Since the columns of B and the rows of C are linearly independent, the matrices $B^T B$ and CC^T (where the superscript T denotes the transpose) are positive definite, and therefore nonsingular.

We now define the pseudoinverse for the matrices B and C (and generally for rectangular matrices of maximal rank) as follows:

$$B^{\dagger} = (B^T B)^{-1} B^T \qquad C^{\dagger} = C^T (CC^T)^{-1} . \tag{5}$$

It will be noted that, for a nonsingular square matrix, these expressions reduce to the classical inverse. We have also

$$B^{\dagger} B = CC^{\dagger} = I , \tag{6}$$

and moreover $B^{\dagger} A = C$, which, together with (4), shows that the row-spaces of A and C are identical. It follows that the rows of C form a basis for the row-space of A.

For the nonzero real matrix A, we now define the pseudoinverse[3] as

$$A^{\dagger} = C^{\dagger} B^{\dagger} . \tag{7}$$

Finally, for completeness, we define the pseudoinverse of an m x n zero matrix as an n x m zero matrix.

In consequence of this definition, the row-space and column-space of $A^{\dagger}$ are the transposes of the column-space and row-space, respectively, of A. If A is m x n, evidently $A^{\dagger}$ is n x m; thus both the products $AA^{\dagger}$ and $A^{\dagger} A$ can be formed. Moreover, these products have interesting properties. First, we note that each is symmetric and idempotent (i.e., equal to its own square). It is easily verified that $I - AA^{\dagger}$ and $I - A^{\dagger} A$ are also symmetric and idempotent. We can show that the product $AA^{\dagger}$ is the same for any two matrices A_1 and A_2 having the same column-space. For, equations (4), (6) and (7) give $AA^{\dagger} = BB^{\dagger}$. Since the same matrix B can serve for both A_1 and A_2, the result follows.

Let u denote any vector in Euclidean m-space and consider the vectors:

$$x = AA^{\dagger} u , \qquad y = u - x = (I - AA^{\dagger})u .$$

It is evident that x is a vector in the space S_c, the column-space of A. Moreover, y is orthogonal to this space. For, if z is any vector of S_c, there exists a vector w, such that z = Aw; and, in view of the symmetry and idempotency of $AA^{\dagger}$,

$$y^T z = u^T (I - AA^{\dagger}) Aw = 0 ,$$

since (4), (6) and (7) give $AA^{\dagger} A = A$. This decomposition of u into a vector of S_c and a vector orthogonal to S_c is unique. To show this, let u = $x_1 + y_1$, where x_1 is in S_c and y_1 orthogonal to S_c. Then there exists a vector w_1, such that $x_1 = Aw_1$, and $AA^{\dagger} u = AA^{\dagger} Aw_1 = Aw_1 = x_1$, showing that $x_1 = x$.

We shall call the vector x the projection of u on S_c, and the matrix $AA^{\dagger}$, which has been shown to be characteristic of the space (since it is the same for all matrices A having this column-space), will be called the projector on S_c, and will be denoted by P_c.

Similar remarks apply to the matrix $A^{\dagger} A$ with regard to left multiplication by row-vectors, and it will be called the projector on the row-space S_r of A, and will be denoted by P_r.

4. APPROXIMATE SOLUTION OF INCONSISTENT SYSTEMS OF LINEAR EQUATIONS[4]

Consider the problem of approximating an arbitrary vector u by a vector v which is restricted to the extent that it must belong to a given vector space[5] S. We shall take the differences between corresponding components of u and v, and shall say that the approximation is "best" when the sum of the squares of these differences is a minimum. This sum of squares, which we shall

denote by q, is given by

$$q = (u - v)^T (u - v) . \tag{8}$$

Now, we have

$$u = Pu + y ,$$

where P is the projector on S and y is orthogonal to S. Therefore,

$$u - v = (Pu - v) + y .$$

Since $Pu - v$ belongs to S, it is orthogonal to y, and therefore equation (8) reduces to

$$q = (Pu - v)^T (Pu - v) + y^T y .$$

This is the sum of two positive terms, the second of which is independent of v, and is clearly a minimum when the first term is made to vanish by choosing $v = Pu$. In other words, the projection of u on S is the best approximation in S to u.

Now, let us return to the consideration of the system of equations represented by (2) or (2)'. An exact solution is possible if and only if b belongs to the column-space of A. When this is the case, the solution is unique when A is nonsingular, and is given by (3). If A is singular, the general solution is

$$x = A^\dagger b + y , \tag{9}$$

where y is any vector orthogonal to the row-space of A. It is clear that this is a solution, since

$$Ax = P_c b + Ay = b ,$$

since, under the hypotheses, $P_c b = b$ and $Ay = 0$. To show that this includes all solutions, we first note that any solution x can certainly be expressed in the form (9) if y is unrestricted. Substitution of this expression in (2)' then gives $b + Ay = b$, or $Ay = 0$, showing that y is indeed orthogonal to the row-space of A.

Since the two vectors in the right member of (9) are orthogonal, the length of the vector x is a minimum when we take $y = 0$. Thus, the solution of minimum vector length is given by

$$x_m = A^\dagger b . \tag{10}$$

Of greater interest to us, however, is the case in which b is not in the column-space of A, so that no exact solution is possible. In this case, we consider as the "best" solution the vector x for which the length of the vector $Ax - b$ is a minimum. This is tantamount to the usual least squares criterion, and implies, as we have already seen, that

$$Ax = P_c b ,$$

where P_c is the projector on S_c, the column-space of A. Since $P_c b$ is in S, the solution is given by (9) with b replaced by $P_c b$, and is therefore

$$x = A^\dagger P_c b + y = A^\dagger b + y ,$$

since $A^\dagger A A^\dagger = A^\dagger$ by (4), (6) and (7). Thus we have shown that when there is no solution, (9) and (10) give the "best" solution in the sense indicated.

5. STATISTICAL APPLICATIONS

Perhaps the most obvious statistical application is to multiple regression. Let a variate y depend on n variates $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$, and let it be required to determine the coefficients a_j in the regression equation

$$y = \sum_{j=1}^{n} a_j x^{(j)} .$$

It is assumed that corresponding numerical values $y_i, x_i^{(j)}$ are given for $i = 1, 2, \ldots, m$. If y denotes the column-vector whose ith component is y_i, a the column-vector whose jth component is a_j, and X the matrix $(x_i^{(j)})$, the regression coefficients are given by

$$a = X^\dagger y . \tag{11}$$

If the columns of X are linearly independent, as will usually be the case, the least squares regression equation is unique. Otherwise, there will be many solutions which yield the minimum value for the sum of the squared residuals. Of these possible solutions, (11) then gives the one for which the sum of the squares of the coefficients a_j is smallest.

Let (x_i, y_i), $i = 1, 2, \ldots, m$, be a set of points to which a curve $y = f(x)$ is to be fitted. It is stipulated that $f(x)$ is to be a linear combination of n given functions $g_1(x), g_2(x), \ldots, g_n(x)$: thus

$$f(x) = \sum_{j=1}^{n} a_j g_j(x) .$$

The coefficients a_j are to be determined so as to minimize the quantity

$$S = \sum_{i=1}^{m} \left[y_i - f(x_i) \right]^2 .$$

This covers many, but not all least squares curve fitting situations. The simplest and most usual case is that of fitting a polynomial of degree $n - 1$, for which $g_j(x) = x^{j-1}$.

If y and a are defined as before and Q denotes a matrix such that the element in the ith row and the jth column is $g_j(x_i)$, then

$$a = Q^\dagger y .$$

If u_i denotes the fitted ordinate corresponding to the given ordinate y_i and u is the vector whose ith component is u_i, we have

$$u = Qa = QQ^\dagger y = P_c y ,$$

where P_c is the projector on the column-space of Q.

6. RECURSIVE ALGORITHM FOR OBTAINING THE PSEUDOINVERSE OF A MATRIX

Equations (5) and (7) are not very practical for computational purposes. The writer has given elsewhere [13] the following algorithm[6] for obtaining the pseudoinverse of a matrix. Let a_k denote the kth column of a given matrix A, and let A_k denote the submatrix consisting of the first k columns. Then the pseudoinverse of A_k is of the form

$$A_k^\dagger = \begin{bmatrix} A_{k-1}^\dagger - \gamma_k c_k \\ c_k \end{bmatrix}, \tag{12}$$

where

$$\gamma_k = A_{k-1}^\dagger a_k, \tag{13}$$

and the last row c_k remains to be determined. In its determination two distinct cases arise, according to whether or not a_k belongs to the column-space S_{k-1} of A_{k-1}. In other words, it must be ascertained whether or not the space spanned by the first $k-1$ columns of A is enlarged by the addition of the kth column. Now a_k belongs to S_{k-1} if and only if the projection of a_k on S_{k-1} is equal to a_k itself: in other words, if

$$A_{k-1} \gamma_k = a_k . \tag{14}$$

If (14) is not satisfied

$$c_k = (a_k - A_{k-1} \gamma_k)^\dagger , \tag{15}$$

while if (14) is satisfied

$$c_k = (1 + \gamma_k^T \gamma_k)^{-1} \gamma_k^T A_{k-1}^\dagger . \tag{16}$$

Formulas (12) to (16) constitute a recursive procedure for obtaining successively $A_2^\dagger$, $A_3^\dagger$, ..., starting with $A_1^\dagger$. Both for the initial determination of $A_1^\dagger$ and for the evaluation of the right member of (15) a formula is required for the pseudoinverse of a single-column matrix a. It follows from (5) that this is

$$a^\dagger = \begin{cases} a^T & (a = 0) \\ (a^T a)^{-1} a^T & (a \neq 0). \end{cases} \tag{17}$$

For the purpose of statistical applications, some "streamlining" of the algorithm can be effected by noting that in these situations it is unnecessary to obtain the pseudoinverse explicitly. Rather, what is wanted is the "best" solution $x = A^\dagger b$ of an inconsistent system $Ax = b$. The algorithm can be modified to give $A_k^\dagger b$ for $k = 1, 2, \ldots$ successively. To this end it is convenient to define a matrix A' obtained by enlarging A through the addition of two columns on the right: (i) the vector b and (ii) a total column, which is the sum of all the preceding column vectors. Then (12) gives

$$A_k^\dagger A' = \begin{bmatrix} A_{k-1}^\dagger A' - \gamma_k (c_k A') \\ c_k A' \end{bmatrix} . \tag{18}$$

The penultimate column of this matrix is $A_k^\dagger b$, while the final column should be the sum of the preceding column vectors if the arithmetic has been correctly performed. Moreover, (13) shows that γ_k is the kth column of $A_{k-1}^\dagger A'$.

In order to obtain $c_k A'$ for use in (18) we must first compute

$$a_k - A_{k-1} \gamma_k . \tag{19}$$

If this vector vanishes, (16) shows that

$$c_k A' = (1 + \gamma_k^T \gamma_k)^{-1} \gamma_k^T A_{k-1}^\dagger A' .$$

If (19) does not vanish, it equals $c_k^\dagger$ by (15). By (17) we then have

$$c_k A' = (c_k^{\dagger T} c_k^\dagger)^{-1} c_k^{\dagger T} A' . \tag{20}$$

If we first compute the vector $c_k^{\dagger T} A'$, we note that its kth element is $c_k^{\dagger T} a_k$. By (13), $A_{k-1} \gamma_k = A_{k-1} A_{k-1}^\dagger a_k$ is the projection of a_k on the column-space of A_{k-1}. Therefore, $c_k^\dagger$ as given by (19) is orthogonal to this space. Thus $c_k^{\dagger T} A_{k-1} \gamma_k = 0$, and consequently, $c_k^{\dagger T} c_k^\dagger = c_k^{\dagger T} a_k$. It follows from (20) that c_k is obtained from the computed vector $c_k^{\dagger T} A'$ upon "normalizing" it by dividing by its kth element. With these explanations, (18), (19) and (20) constitute the recursive procedure desired.[7]

For example, if m sets of corresponding values of n statistical variables $x_1, x_2, \ldots, x_n$ are given, and it is required to compute regression coefficients of x_1 against $x_2, x_3, \ldots, x_n$, the matrix A' is formed so that its first column consists of all 1's, and the 2nd to nth columns exhibit the successive values of the variables $x_2, \ldots, x_n$, respectively. The $(n+1)$th column is the vector x whose components are the corresponding values of x_1, and this is followed by the total column. Then $A_1^\dagger x$ is the mean value of x_1, while at the kth stage of the process $A_k^\dagger x$ is a vector whose components are the coefficients in the regression equation

$$\hat{x}_1 = b_k + b_{12\cdot 3\ldots k} x_2 + b_{13\cdot 24\ldots k} x_3 + \ldots + b_{1k\cdot 2\ldots k-1} x_k .$$

The components of $A_k A_k^\dagger x$ are the values of x_1 predicted by the regression equation, while $x^T(x - A_k A_k^\dagger x)$ is the sum of the squares of the errors of estimate. Thus, if there is doubt as to how many of the variables should be included in the regression equation (and if one is fortunate in choosing the order in which the variables are introduced), this method shows at a glance how much the coefficients change as the less significant variables are brought into the equation, and, if desired, the reduction at each step in the standard error of estimate. It will be noted also that the vector γ_k given by (13) exhibits the coefficients in the regression equation of x_k against $x_2, x_3, \ldots, x_{k-1}$.8/

Consider now the problem of least squares fitting of a polynomial. Let (x_i, y_i), $i = 1, 2, \ldots, m$, be the set of points to which a polynomial is to be fitted, let A be the n-column matrix (x_i^{j-1}), and let y denote the vector whose ith component is y_i. Then, if the algorithm is applied, the vector $A_k^\dagger y$ exhibits the coefficients of successive powers of x in the least squares polynomial of degree $k - 1$, while $A_k A_k^\dagger y$ exhibits the values of the fitted polynomial corresponding to the given abscissas x_k, and $y^T(y - A_k A_k^\dagger y)$ is the sum of the squared residuals. The latter quantity, of course, may be used in testing to see what degree of polynomial is most suitable.9/

Extensive tables are available [16] to facilitate the use of orthogonal polynomials in fitting a least squares polynomial to data with equally spaced arguments, but they are of no avail when the abscissas are irregularly spaced. It is to be noted that the procedure described here makes no assumption about the spacing of the arguments. Further, the recursive nature of the process obviates the need to make use of orthogonal polynomials directly. If desired, however, they can easily be obtained as a by-product.10/ If $\gamma_k^{(i)}$ denotes the ith component of γ_k, the set of polynomials

$$(21)\qquad \begin{aligned} p_0(x) &= 1 \\ p_k(x) &= x^k - \sum_{i=1}^{k} \gamma_k^{(i)} x^{i-1} \\ &\qquad (k = 1, 2, \ldots, m - 1) \end{aligned}$$

constitutes an orthogonal set over the discrete domain $(x_1, x_2, \ldots, x_m)$. In other words,

$$(22)\qquad \sum_{i=1}^{m} p_j(x_i)\, p_k(x_i) = 0 \qquad (j \neq k).$$

To show this, we first note that j and k are interchangeable in (22), so that we can assume without loss of generality that $j < k$. Thus, (22) will be established if we can show that

$$\sum_{i=1}^{m} x_i^j\, p_k(x_i) = 0 \qquad (j = 0, 1, \ldots, k - 1).$$

But, in view of (21) and of the definition of A, this follows from the fact previously noted that the vector $a_k - A_{k-1}\gamma_k$ is orthogonal to S_{k-1}.

FOOTNOTES

*This possibility was suggested to the writer by William Hodgkinson, Jr., of the American Telephone and Telegraph Co.

1/The first two writers mentioned were unaware of Moore's work until it was brought to their attention by Rado [9].

2/An excellent brief treatment of the subject, more than adequate for the understanding of this article, is given in Chapter 1 of [10].

3/The pseudoinverse can be defined in several other ways [2,6,8,12, and the fuller, mimeographed version of this paper]. The present approach, which is probably the simplest, was suggested to the writer by A. S. Householder, whose assistance is gratefully acknowledged.

4/See also [3,4,7,11].

5/See also [12].

6/In essence, this algorithm is an abbreviated form of a particular case of the method of matrix inversion by biorthogonalization proposed by Hestenes [8]. See also [13,14].

7/A further "streamlining" is possible by working with the symmetric matrix $A'^T A'$, which, in essence, merely exhibits the usual "normal" equations. This kind of procedure is easily explained without reference to the pseudoinverse, and is probably the simplest approach for small-sized calculations. In large-scale calculations, it has the disadvantage that, if the "recursive" feature is retained, certain key quantities in the computations (e.g., those which we have called "normalizing" factors) are obtained as differences between large, and almost equal, numbers, and accuracy is rapidly lost.

8/A numerical example of the application to multiple regression is given in the fuller, mimeographed version of this article. It is not reproduced here because the actual arithmetic is essentially the same as that involved in other methods of calculating multiple regression coefficients.

9/See [15], pp. 461-465.

10/See also [17].

REFERENCES

[1] Moore, E. H., Bulletin of the American Mathematical Society, 26 (1920), 394-395.

[2] Moore, E. H., "General Analysis, Part I," Memoirs of the American Philosophical Society, 1 (1935), especially pp. 197-209.

[3] Bjerhammar, A., "Application of calculus of matrices to method of least squares; with special reference to geodetic calculations," Transactions of the Royal Institute of Technology (Stockholm), No. 49 (1951), 1-86.

[4] Bjerhammar, A., "Rectangular reciprocal matrices with special reference to geodetic calculations," Bulletin Géodésique, 1951, 188-220.

[5] Bjerhammar, A., A Generalized Matrix Algebra, National Research Council of Canada, Division of Applied Physics, Ottawa, 1957. Pp. 24.

[6] Penrose, R., "A generalized inverse for matrices," Proceedings of the Cambridge Philosophical Society, 51 (1955), 406-413.

[7] Penrose, R., "On best approximate solutions of linear matrix equations," Ibid., 52 (1956), 17-19.

[8] Hestenes, M. R., "Inversion of matrices by biorthogonalization and related results," Journal of the Society for Industrial and Applied Mathematics, 6 (1958), 51-90.

[9] Rado, R., "Note on generalized inverses of matrices," Proceedings of the Cambridge Philosophical Society, 52 (1956), 600-601.

[10] Rao, C. R., Advanced Statistical Methods in Biometric Research, New York, John Wiley and Sons, 1952.

[11] Greville, T. N. E., "The pseudoinverse of a rectangular or singular matrix and its application to the solution of systems of linear equations," SIAM Review, 1 (1959), 38-43.

[12] Greville, T. N. E., "On smoothing a finite table: a matrix approach," Journal of the Society for Industrial and Applied Mathematics, 5 (1957), 137-154.

[13] Greville, T. N. E., "Some applications of the pseudoinverse of a matrix," to be published in SIAM Review, 1 (1959).

[14] Dent, B. A., and Newhouse, A., "Polynomials orthogonal over discrete domains," SIAM Review, 1 (1959), 55-59.

[15] Snedecor, G. W., Statistical Methods, 5th Edition, Ames, Iowa, Iowa State College Press, 1956.

[16] Anderson, R. L., and Houseman, E. E., Tables of Orthogonal Polynomial Values Extended to N = 104, Research Bulletin 297, Agricultural Experiment Station, Iowa State College of Agriculture and Mechanic Arts, Ames, 1942.

[17] Forsythe, G. E., "Generation and use of orthogonal polynomials for data fitting with a digital computer," Journal of the Society for Industrial and Applied Mathematics, 5 (1957), 74-88.

METHODS OF MEASURING DIFFERENCES IN SOCIAL CLASSES

By: Theodore R. Anderson, Yale University

1. Statement of the Problem

A social class, as the term is used in modern sociology, is a complex phenomenon. It consists of many social categories (such as occupations) which bear a particular relation to a fixed scale (such as a prestige, an income, or a composite scale). This paper will not discuss social classes directly, but rather will focus on the problem of measuring differences between social categories out of which class groupings may arise. Socially, how different are two occupations, or two ethnic groups, or two religions? This paper will present one means of answering these questions.

Suppose for the moment that a matrix describing the social difference between each pair of, say, occupations, were available. It is highly unlikely that such a matrix would have unit rank; that the differences could be resolved into a space of one dimension or into a single directional scale. Two or more dimensions would almost certainly exist. It is convenient to call the space characterized by these hypothetical dimensions a social space. In keeping with this terminology, the difference between any two categories may be described as the social distance between the categories. Using these terms the problem discussed here is that of measuring the social distance between any two categories which exist within a social space.

A means of measuring the social distance between categories should prove useful to a variety of research problems. A matrix of such social distances represents one means of describing a social structure, such as a community, in quantitative terms. Such a description, in turn, would permit studies of temporal change in structural characteristics. At present it is difficult to do more than to speculate about answers to many structural change questions of considerable interest. For instance, is anything like a polarization of occupational roles into opposing camps occurring within the United States, or is the direction of change toward greater homogeneity? An effective measure of social distance could be used to provide an answer.

Such a measure should also prove useful in developing what might be called a social ecology. For example, social mobility might be subjected to the same analysis, via mathematical models, that is currently occurring in connection with physical migration. At the present time, breaking a social mobility matrix into its theoretical components is virtually impossible, largely due to the absense of a sound multi-dimensional measure of social distance.

It is clear that a measure of social distance will not, in itself, solve the above problems. Rather, it will merely make their solution possible. It is also clear that a complex social structure will contain many parameters other than those characterizable as social distances. Never the less, a sound measure of social distance should have considerable utility. This paper will present such a measure after briefly considering other possible ways of solving the problem. The proposed solution will be illustrated using relatively simple empirical data. Finally, an example will be presented showing how the derived distances may be used to solve a research problem.

2. General Considerations

The term social distance, as it is used here, is equivalent to the phrase, behavioral difference. That is, if the members of two groups perform precisely the same behaviors in all contexts, or behavioral domains, they occupy the same position within social space. The more similar are the behaviors, the closer in space are the groups. The problem of measuring social distance, or behavioral difference, reduces initially to the problem of selecting a behavioral domain within which to observe similarity and difference. By a behavioral domain is meant some population of behaviors such that all the behaviors have something in common. That which is common, for instance, might be a relation to a specific institutional order. Thus, all behaviors bearing upon property might form a domain, or all behaviors bearing upon the mass media of communication.

In general, there are two basic types of behavioral domains. First, there are those sets of behaviors through which individuals in one category or group are related to individuals in other categories or groups. Such behaviors include interactions between groups, interchanges of goods, services, and people, and evaluations of group members by members of other groups. The second type includes all other behavioral domains. To be most useful, measures of social distance should be generated from the second type of domain and not from the first type.

The reason for this statement is simple. Virtually all of the important hypotheses which might be tested in part through the use of social distance measurements involve the relational domains (interactions, interchanges, and evaluations). Do members of groups which are closer in social space interact more frequently? Do

members of groups which are far apart feel more hostile toward each other? Is social mobility related to social distance? To be able to answer questions such as these it is crucial that the measure of social distance be generated out of behaviors which are conceptually independent of the ones in question.

Of course, each of these questions implies an hypothesis which could form the basis of a measure of social distance. Thus, a matrix which states the frequency of interaction between social categories could be transformed into one stating the social distance between each pair of categories using the assumption that the frequency of interaction is inversely related to social distance. Virtually all existing measures of distance in social space are based upon some such transformation. Such measures are presumably reasonably accurate. However, to the extent that they use intergroup relations to generate the social distances, they cannot be used to test relational hypotheses. It is the opinion of this author that social distance should be measured within other domains because most of the hypotheses worth testing lie in the relational domains.

No attempt will be made here to survey the literature of existing measures. Such measures are either drawn from relational domains (e. g., the Bogardus measure of social distance (1)) or are measures along only one direction in social space (e. g., the vast number of measures of socio-economic status). The measure proposed here avoids these limitations, in that it is based on behaviors not in the intergroup domains and is a general, non-directional, measure of distance.

3. The proposed Solution

Social distance is essentially equivalent to behavioral difference. Groups are socially distant to the extent that their members behave differently. There is, however, no general measure of the amount of difference between unlike activities. Therefore, the key indicator of social distance between two groups is the difference in the proportion of persons performing the same activity in each group. In particular, consider a set of mutually exclusive and exhaustive categories within a population of persons. Occupations, religions, and ethnic groups are examples. Consider two such categories, say the ith and kth. Consider also a set of behaviors, which need not be mutually exclusive (the same person may perform one or more of the behaviors). Let the jth behavior be one of this set. Finally, let the proportion of persons in the ith and kth categories who perform the jth act be P(ij) and P(ik). The social distance between i and k is assumed to be closely related to the difference between these two proportions.

The difference between these proportions, however, is an index of social distance only to the extent that it is shared by or common to other behaviors (in the factorial sense). Social space may be characterized according to this assumption by creating the matrix specified by the typical element, P(ij), where i ranges over all categories and j over all behaviors. This matrix may be called the data matrix.

Consider a column of the data matrix; that is, a matrix consisting of the elements, P(1j), P(2j), P(3j), ..., P(nj). This matrix specifies the relative frequency with which the behavior j occurred within each of the various categories under study. In other words, it specifies the relative distribution of this particular behavior through the categories. This matrix may, thus, be called the pattern of the behavior, j. Each behavior, of course, has such a pattern. The analytical problem is to discover a set of independent patterns which reproduce the common elements in the original set of observed patterns.

The procedures, usually called factor analysis, for producing such a set of independent behavior patterns are well known and need not be discussed at length here. The correlations between all possible pairs of behavior patterns are first determined. This correlation matrix is then factor analyzed. The factors which emerge are the independent, common behavior patterns from which the correlation matrix may be reproduced, and from which the original behavior patterns (the data matrix) may be reproduced, save for specific elements.

Such a set of independent behavior patterns constitutes what is here called a social space. The position of each category within this space may readily be determined from the factor loadings. The result is a characterization of social space, and a description of the position of the categories within it. The matrix of distances between each pair of categories may be determined either by simple algebra using the data matrix and the factor loadings or by measuring distances directly on a representation or map of the social space. Social structures may be identified either through the distance matrix or in terms of the overall shape of the categories in the social space.

4. An Illustration

To facilitate an understanding of this solution, the method was applied to a set of readily available data. This empirical analysis is primarily an illustration. In particular, it is based upon too few observations to be highly reliable in detail as a description of social

space. Further, the factor analysis is only approximate. For illustrative purposes, however, the analysis is entirely adequate.

The population, in the illustration, consists of employed persons living in 10 randomly selected census tracts of Akron, Ohio in 1950. Living in a tract is a behavior. Thus, there are 10 possible behaviors corresponding to the 10 census tracts. These behaviors happen to be mutually exclusive, though they need not be. Social categories are defined as the various occupations (professional, managerial, etc.) in the one column census classification, cross-classified by sex. The problem is to measure the social distance between these occupations. (2)

Table 1. Per Cent of Persons (Classified by Occupation and Sex) Living in each of Ten Randomly Selected Census Tracts in Akron, Ohio, 1950.

Sex and Occupation	Census Tract										
	A-5	A-8	B-8	C-6	D-3	F-1	F-3	F-5	F-7	G-2	Total
Male											
Professional	2	3	4	3	1	19	26	15	3	24	100
Managerial	3	3	3	4	1	23	23	13	4	23	100
Clerical	7	10	11	8	3	7	25	14	8	8	101
Sales	4	5	4	4	2	13	27	16	6	19	100
Crafts	8	13	10	10	6	4	24	11	8	7	101
Operatives	12	18	10	13	12	1	16	6	10	2	100
Service[a]	15	20	3	7	15	3	15	6	13	3	100
Laborer	15	23	3	8	21	3	11	4	10	3	101
Female[b]											
Professional	12	3	5	3	1	18	19	16	4	18	99
Managerial	7	7	3	5	4	15	19	13	6	21	100
Clerical	5	8	9	8	3	8	26	17	7	10	101
Sales	7	11	7	8	3	6	26	16	9	8	101
Crafts	11	7	12	9	9	4	16	12	10	11	101
Operatives	16	19	9	14	14	1	13	5	9	2	102
Pvt hsld wkrs	12	23	2	2	21	6	8	3	9	15	101
Service	16	20	6	9	14	4	12	5	11	3	100

a/Includes private household workers.
b/Female laborers omitted (too few in number for stability).

Source: 1950 United States Census of Population, Akron, Ohio Census Tracts, 1950 Population Census Report, Vol. III, Chapter 1. United States Government Printing Office, 1952. Figures adapted from Table 2.

The first step is to create the data matrix (see Table 1) which indicates the proportion of persons in each occupation living in each tract. The next step is to correlate each pair of columns in this matrix, thus producing the correlation matrix (see Table 2). From the correlation matrix it is clear, for instance, that tracts A-5 and A-8 are very similar in occupational distribution (r = .79). This fact may be confirmed by examining the first two columns of the data matrix. The third step is to factor analyze the correlation matrix. The factor analysis (via the centroid method) of this matrix yielded two factors (see Table 2) which represent orthogonal directions in social space. The final step is to use these factors to create a map of social space (shown as Figure 1).

Some comments about the procedure and the findings are in order. First, the factors themselves are, initially at least, only arbitrary directions in social space. The distance between each pair of occupations remains constant through any orthogonal rotation. No particular meaning need be attached to each factor; that is, the factors need not be named as long as attention is focused on the relations among the occupations. If the problem is shifted from measuring social distance to explaining how existing distance relations come into being, then an identification of the factors would be appropriate. From this point of view, it is fairly obvious from the map that Factor I measures something that is very similar to the socio-economic status of the occupation. Factor II, based on only two behavior patterns (B-8 and C-6), is less reliably measured and more difficult to identify. It is possible that it is related to the institutional structure of the community, with service personnel low (from private household workers to professionals) and material goods personnel high (from operatives to craftsmen) on this factor. Such an interpretation is highly tentative at best, however.

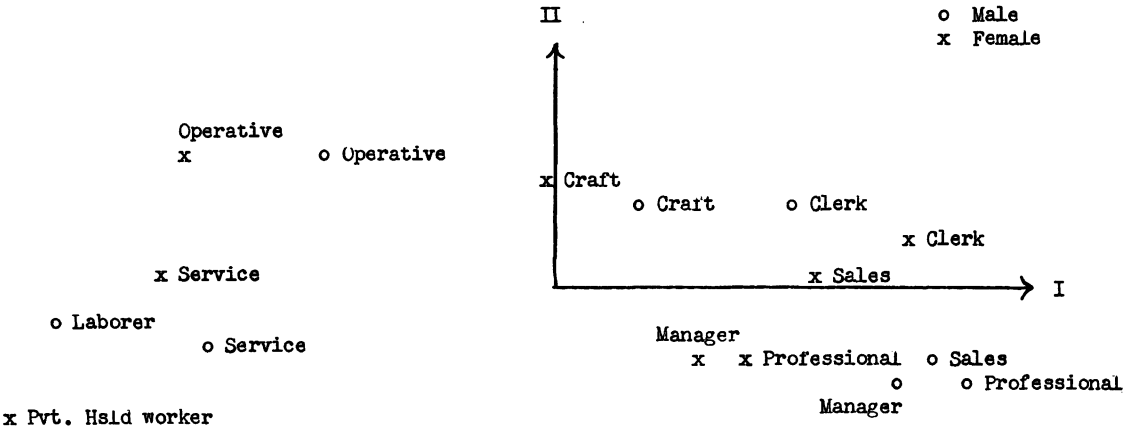
Second, it is likely that there is at least a third factor underlying residential behavior. The small number of tracts included in this illustration, however, precluded even its tentative identification. Third, it will be observed that the occupations as a whole form a rough semi-circle in the space. This shape cannot be interpreted at the present time, but might possibly prove important in cross-cultural comparisons or in trend analyses.

Table 2. Correlation Matrix based on the Akron, Ohio Census Tract Data Matrix and a Rotated Factor Matrix which Approximately Reproduces the Correlation Matrix.

	Correlation matrix[a]										Factor matrix[a]	
Census Tract	A-5	A-8	B-8	C-6	D-3	F-1	F-3	F-5	F-7	G-2	I	II
A-5		79	05	23	81	-69	-86	-80	75	-74	-87	12
A-8			-06	44	94	-80	-77	-91	82	-77	-93	18
B-8				68	-16	-46	20	15	22	-44	12	76
C-6					29	-76	-19	-36	58	-82	-32	84
D-3						-69	-90	-95	74	-62	-98	-02
F-1							49	62	-90	92	70	-68
F-3								90	-59	44	93	19
F-5									-68	59	97	04
F-7										-86	-76	46
G-2											67	-70

a/ Decimal signs omitted.

Figure 1. Map of Social Space B ased upon Factorial Structure.



Finally, it will be observed from the map that the high status occupations are more tightly clustered (especially among males) than are the lower status occupations. Thus, the distance from laborer to craftsman is as great as the distance from craftsman to professional. Male professionals, managers, and salesmen are very close together as compared to male laborers, servicemen, and operatives. This finding is not entirely consistent with existing studies of the social hierarchy, which tend to assume greater differentiation near the top of the scale. In interpreting the finding, however, it must be remembered that each occupation is a figure within this space such that the point on the map is the central tendency of the figure. It is possible that the upper occupations are less homogeneous (i.e., would have larger figures) than are the lower occupations. Comparing males and females in the same occupation gives some idea of the spread within each occupation. Discovering the shape of the figures which characterize each occupational category is a problem for future research.

5. An Application

A map of social space, and the distances which can be derived from it, are useful only in so far as they permit the resolution of problems which are otherwise difficult or impossible to solve. To illustrate how this map, or one like it based on more extensive data, might be used, it was decided to apply the P/D (or population divided by distance) model of interaction and interchange to a matrix of social mobility, using social distances derived from the map of social space. Rogoff's (3) 1940 intergenerational mobility matrix, which lists the occupations of about 10,000 sons in Indianapolis classified by the occupations of their fathers, provided the observations. (See Table 3 for details.) The expected values in Table 3 were generated from the model and the map of social space. The average accuracy of reproduction is about 92%,

Table 3. Social Mobility of Sons Observed by Rogoff, Indianapolis, 1940, and Expected by a Model Using Population Divided by Social Distance, where Distance was Raised to the 2/3rds Power.

Frequency

Son's occupation	Father's occupation: Professional		Managerial		Clerical and sales		Skilled		Semi-Skilled		Service		Unskilled	
	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.
Professional	—	—	133	196	141	112	167	182	70	74	23	21	28	32
Managerial	40	68	—	—	83	100	117	157	63	60	25	18	20	26
Clerical and Sales	152	145	368	377	—	—	520	636	262	218	83	61	94	88
Skilled	100	77	172	194	165	209	—	—	279	308	78	75	111	105
Semi-skilled	59	66	238	154	179	149	734	643	—	—	125	141	216	198
Service	17	13	44	31	35	29	141	107	99	97	—	—	43	64
Unskilled	15	14	30	33	26	30	154	108	81	97	27	46	—	—
Per cent accuracy	91		90		90		91		94		91		94	

Source: (Of observations) Natalie Rogoff, Recent Trends in Occupational Mobility, Glencoe, Ill., The Free Press, 1953: Table in Jacket.

Each expectation is of the form, $aP/D^{2/3}$, where a is a constant of proportionality, P is the total number of sons in each occupation, and D is the social distance between occupations as measured from the map (Figure 1) of social space.

which means that only 8% of the observations would have to be shifted to achieve perfect reproduction. The fact that results as accurate as these occurred despite the unreliability of the measured distances, the shift from Akron to Indianapolis, and the shift from 1950 to 1940 (not to mention the unreliability in the observations) suggests that this method of measuring social distance has considerable promise.

It is important, however, not to exaggerate the importance of the findings to date. Before this procedure can be considered confirmed two types of research need to be performed. First, studies of residential behavior should be conducted in several cities to see if the spatial pattern of the occupations remains essentially constant. Second, studies within other behavioral domains should be conducted, again to see if the results are stable. Investigations along both of these lines are currently in the planning stage.

Footnotes

(1) See, for example, E. S. Bogardus, "Measuring Social Distance," Journal of Applied Sociology, 1925, 9, 299-308 and E. S. Bogardus, The New Social Research, Los Angeles, R. J. Miller, 1926.

(2) For an alternative method of analyzing the same data with roughly the same objectives see Otis Dudley Duncan and Beverly Duncan, "Residential Distribution and Occupational Stratification," American Journal of Sociology, LX, 5, March 1955, 493-503.

(3) Natalie Rogoff, Recent Trends in Occupational Mobility, Glencoe, Ill., The Free Press, 1953: table in jacket.

VII

RECENT SURVEYS OF FACTORS AFFECTING FAMILY SIZE

Chairman, Dudley Kirk, Population Council

The Growth of American Families: Results of a National Survey—Arthur A. Campbell and Pascal K. Whelpton, Scripps Foundation for Research in Population Problems and Ronald Freedman, University of Michigan

Interim Report on the Study of Future Fertility of Two-Child Families in Metropolitan America—Philip G. Sagi, Robert G. Potter, Jr., and Charles F. Westoff, Office of Population Research, Princeton University

Family Planning in Medical Practice—Sydney S. Spivack and Florence A. Ruderman, Columbia University

Discussion—Philip M. Hauser, University of Chicago

Discussion—William J. Gibbons, Fordham University

THE GROWTH OF AMERICAN FAMILIES: RESULTS OF A NATIONAL SURVEY

By: Arthur A. Campbell and Pascal K. Whelpton, Scripps Foundation for Research in Population Problems, Miami University; and Ronald Freedman, Professor of Sociology and Research Associate of the Survey Research Center, The University of Michigan

The study Growth of American Families yields a number of important generalizations about family building patterns and expected family size among the white population of the United States. This study was done jointly by the Scripps Foundation for Research in Population Problems of Miami University and the Survey Research Center of the University of Michigan. The main findings are presented in the book Family Planning, Sterility, and Population Growth by Ronald Freedman, Pascal K. Whelpton, and Arthur A. Campbell, to be published by McGraw-Hill in the spring of 1959.[1]

The 2,713 women included in our sample are a representative cross section of white wives, 18 to 39 years old, living with their husbands or with husband temporarily absent in the armed forces. The Survey Research Center interviewed these wives in the spring of 1955. They were questioned intensively about their pregnancy history, physiological limitations on their fertility, their use of contraception, their expectations regarding family size, and various socioeconomic topics such as religious preference, educational attainment, and income. Although the interviews were long--80 minutes on the average--they were, in general well received by the respondents. The women showed interest in the subjects covered--particularly those concerning their own family growth--and often volunteered information not asked for in the questionnaire. Very few of the wives who were interviewed refused to answer questions on the presumably "sensitive" topic of contraception. Only ten declined to say whether or not they had ever tried to limit family size or space their pregnancies. In contrast, over 100 refused to give any information about their husband's income. The willingness of the wives to discuss methods of contraception and impairments of the reproductive system leads us to believe that even more information about such personal topics than was gathered in this study can be collected in future surveys.

Fecundity impairments.--One of the basic areas covered in our study deals with impairments of the reproductive system. How widespread are such impairments, and how severe are they?

It should be noted first of all that our information about fecundity is limited to that given by the wives interviewed. At best, it represents the wife's report of her doctor's opinion. A systematic medical investigation of such a large sample was, of course, out of the question. Our reliance on the women's answers undoubtedly introduces some biases in our estimates of the extent of subfecundity. These biases are probably downward--in part because some women may have been reluctant to report any known physiological defects, but largely because the use of contraception prevents some couples from discovering that they have reproductive impairments. For example, a couple that is unable to have any children will not discover this fact as long as they use contraception, unless they have obvious physical indications of subfecundity.

Even though our estimates of reproductive impairments are minimal, it is apparent that subfecundity is widespread. About one in three couples with wife aged 18 to 39 can be classified as Subfecund[2]--that is, their capacity to have children in the future is either entirely lacking or substantially below normal. However, this proportion is partially a function of the way we have defined subfecundity. Four categories were established:

1. The Definitely Sterile couples are those who cannot have another pregnancy. For most such couples, the basis for classifying them as Definitely Sterile is an operation on the husband or wife making conception impossible.

2. The Probably Sterile couples are those for whom a birth in the future is considered improbable, rather than impossible, on the basis of the wife's report of her doctor's opinion.

3. The Semifecund couples are those who knew of no physiological condition limiting reproduction, but who did not conceive at a "normal" rate while contraception was not being used.[3]

4. The Indeterminate couples cannot be classified as Fecund or Subfecund on the basis of our information and criteria. Like the Semifecund, they failed to conceive at a "normal" rate when contraception was not used. However, even though they did not report using contraception, they did report using a douche after intercourse for cleanliness only. Since we did not ask about the type of douche, regularity of use, or how soon after intercourse it was used, we have no basis for evaluating its contraceptive effect. It is impossible to say, then, whether the abnormally low rate of conception was due to douching or to impaired fecundity.

The combination of the four groups just defined is designated as the Subfecund. The remaining couples, for whom we have no reason to suspect impaired fecundity, constitute the Fecund group. Although we are sure that some of the Indeterminate couples are Fecund, we include all of them among the Subfecund partly in order to counterbalance the suspected inclusion among the Fecund of some couples with undiscovered fecundity impairments.

The proportions in the various fecundity groups are shown in Table 1 both for the total sample and for wives married 15 years or more. The proportion who are Subfecund increases steadily with duration of marriage and with age.

How does subfecundity affect the national fertility picture? We cannot give a precise answer to this question, but the general answer is clear: subfecundity has a relatively minor affect on average family size in the United States. We estimate that if all fecundity impairments were to be eliminated, the number of births probably would rise by about 10 to 15 per cent, other things being equal.

The effect of subfecundity on our birth rate is so small because most Subfecund couples have at least one birth. Among Subfecund couples married

15 years or more, only 17 per cent had had no children; the average number of children ever borne was 2.5, or 1 child fewer than the 3.5 borne by Fecund couples who had been married 15 years or more.

Although most couples classified as Subfecund had been able to have at least one birth, it is clear that subfecundity is by far the major cause of childlessness. Again, confining our attention to couples who had a chance to test their fecundity during 15 years of marriage, we find that 10 per cent are childless and that nearly all of these childless couples (96 per cent) are Subfecund. Couples who voluntarily remain childless for as long as 15 years are extremely rare.

TABLE 1. PER CENT DISTRIBUTION BY FECUNDITY STATUS FOR ALL COUPLES AND FOR COUPLES MARRIED 15 YEARS OR LONGER[4]

Fecundity status	All couples	Couples married 15 years or longer
Total: Number	2,713	509
Per cent	100	100
Fecund	66	42
Definitely Sterile	10	24
Probably Sterile	7	8
Semifecund	12	19
Indeterminate	5	7

Is fecundity related to socioeconomic status? Apparently not to any great extent. When couples are classified by income or rural-urban background we find no systematic differences in fecundity. When we use education as the basis of classification, however, we find that the less educated are more likely to be Subfecund than are the better educated. This may seem surprising, because the less educated usually have more births. Why is it that they are also more likely to be Subfecund? We think that this apparent inconsistency arises largely because the less educated make less use of contraception and therefore have more opportunity to discover fecundity impairments than do the better educated. As we noted before, many couples discover fecundity impairments only when they are not using contraception. We doubt that there are any basic biological differences between educational groups that would lead to substantial differences in their fecundity.

In general, then, we cannot explain differences in the fertility of the major socioeconomic groups by variations in fecundity. There is, however, one important fertility differential that is related to fecundity--specifically, wives who are gainfully employed have smaller families than those who are not employed. This is partly due to the fact that working wives are more likely to be Subfecund than are nonworking wives. A higher incidence of subfecundity is not the only reason for the lower fertility of working wives, however, for we also find that among couples with no fecundity impairments, wives who work have fewer children than those who do not work. This is true regardless of age or duration of marriage.

The use of contraception.--A large majority of the wives interviewed said that they approved of contraception and that they had already used it or intended to do so. Only 21 per cent of the wives said that they and their husbands never had used and never would use contraception; the majority of such couples (15 per cent of the total sample) were Subfecund. Only 6 per cent of all couples were Fecund and intended never to use contraception. It is possible, of course, that some of the couples in this small minority may not put their expressed intentions into practice.

Within the meaning of the term contraception we include all methods except sterilization that couples use to avoid conception. Thus, the term encompasses periodic continence (or the rhythm method) advocated by the Catholic Church, abstinence from sexual intercourse, withdrawal (or coitus interruptus), and the various appliance and chemical methods. We included douche if it was used with contraceptive intent, but not if it was used merely for cleanliness.

The generalization that most Fecund couples use contraception can be extended to all major socioeconomic groups (Table 2). In general, there are differences in the proportions using contraception, but these differences are confined to a relatively narrow range. As would be expected, a smaller proportion of Catholics than of non-Catholics use contraception. Nevertheless, a large majority of Fecund Catholics (80 percent) have used or intend to use contraception. As we shall see later, this does not necessarily mean that they are violating the teachings of their Church with respect to contraceptive practices. Education and income are also related to the use of contraception--couples with lower status having the lower proportions of Users. Even among the lower status groups, however, a large majority have used contraception or intend to do so.

Most couples begin to use contraception at an early stage of family growth. Many start at the

TABLE 2. PERCENTAGE OF FECUND COUPLES WHO HAVE USED OR WHO INTEND TO USE CONTRACEPTION, BY SELECTED SOCIOECONOMIC CHARACTERISTICS

Socioeconomic characteristics	Have used contraception	Have used or intend to use contraception
Total	83	90
Wife's religion:		
Protestant	88	94
Catholic	70	80
Jewish	95	96
Wife's education:		
College	91	94
High school, 4 yrs	85	92
High school, 1-3 yrs	79	88
Grade school	68	78
Husband's income		
$6,000 or more	93	95
$5,000-$5,999	90	95
$4,000-$4,999	85	92
$3,000-$3,999	83	90
Under $3,000	71	84

time of marriage. Many of the others begin after the first pregnancy. For example, among couples who had had three pregnancies, one-third had begun using contraception before the first, and a quarter had begun after the first but before the second. Of the remaining couples, 20 per cent began use after the second or third pregnancy, and 23 per cent had not yet begun.

Because so many couples begin to use contraception at marriage or soon thereafter, the proportion of Users rises rapidly in the early years of marriage and then remains relatively constant at a high level. Among Fecund couples, for example, the proportion who are Users rises from 69 per cent for those married fewer than 5 years to 88 per cent for those married 5 to 9 years. For longer durations, there is very little increase in this proportion.

How successful are couples in using contraception? About one-fourth of the Users reported one or more accidental conceptions--that is, conceptions occurring in spite of the use of contraception. The proportion rises rapidly as families grow; over half of the Users with four or more pregnancies have had accidental conceptions.

It must not be supposed, however, that all couples who have accidental pregnancies have more children than they want. For some couples, an accidental pregnancy may simply be one that occurred earlier than planned. For example, among those couples whose most recent pregnancy was accidental, two-thirds wanted the pregnancy at a later date.

Very few of the couples (only 13 per cent) had had pregnancies that were unwanted either when they occurred or later. From this point of view it appears that Americans are quite successful in avoiding too many pregnancies, even though a substantial minority have had accidental conceptions.

It should be noted that our estimates of accidental and unwanted pregnancies are probably minimal. This opinion is based on the assumption that some women were reluctant to admit that they had not used contraception successfully, and that others did not like to say that any of their children were unwanted. Nevertheless, the low proportions who did report accidental or unwanted conceptions suggest that most couples can and do avoid having more children than they want.

With respect to the methods of contraception, condom and diaphragm are the two most widely used. Forty-three percent of the couples using any method had tried condom and 36 per cent diaphragm. These are also the most effective methods, judging from the relatively small proportions of accidental conceptions reported for couples using only one or the other of these methods. The third most popular method is periodic continence (or rhythm), which had been used by one-third of the couples who tried any method. Douche ranked fourth, with 28 per cent reporting its use, and withdrawal ranked fifth with 15 per cent. These percentages add to more than 100 because many wives reported that more than one method had been used.

There has been much interest in how closely Catholics conform to the teachings of their Church regarding methods of contraception. The Church regards periodic continence and abstinence as acceptable (if used appropriately), but condemns other methods. Among all couples with Catholic wives, 70 per cent had conformed to the teachings of the Church either by not using any method of contraception or by limiting the method used to periodic continence or abstinence. Among users of contraception, 47 per cent of the couples with Catholic wives reported only methods approved by the Church. The proportion of Catholics who conform to Church doctrine is higher among the better educated. This may reflect their greater familiarity with the teachings of the Church.

How families are planned.--The couples covered in our study have followed a variety of family planning patterns. The specific patterns differ greatly in detail, but our main findings can be summarized by referring to three broad groups:

1. Completely Planned: these are couples who either used contraception continuously since marriage and had no pregnancies, or who deliberately planned all pregnancies by interrupting their use of contraception.

2. Partially Planned: these are couples who did not plan all of their pregnancies by interrupting contraception, but who did not have more pregnancies than they wanted.

3. Excess Fertility: these are couples who did not want their most recent pregnancy either when it occurred or later.

The couples who plan every pregnancy by discontinuing the use of contraception are a distinct minority (Table 3). Only one-fifth of the couples planned their families so carefully. Apparently, a substantial proportion of couples begin marriage with the intention of timing the birth of every child, but for one reason or another fail to do so as married life progresses. We find, for example, that almost one-third of those married less than five years can be classified as Completely Planned. The proportion for couples married 15 or more years, in contrast, is only one-tenth. This difference may also be due in part to the fact that more younger couples know about and use the more effective methods of contraception.

There are very few couples who are so unsuccessful in their efforts to plan family growth that they have too many pregnancies. Only 13 per cent are classified as Excess Fertility. The proportion in this group is very low early in married life, but rises to over one-fifth for couples married 15 or more years.

Thus, the majority of couples are neither very careful nor very careless planners. Approximately two-thirds are included in the intermediate group--the Partially Planned. Some of these couples had accidental conceptions. Some had never used contraception but had not yet had too many pregnancies. Some had one or two pregnancies before beginning to use contraception and then began in order to avoid any more pregnancies or to time the occurrence of those that were wanted.

As would be expected, the distribution of couples between the three planning groups varies with family size (Table 3). Couples with few children are more likely to belong to the Completely Planned group than are other couples. The proportion who plan family growth very carefully declines rapidly as the number of children increases. Among couples with 6 or more children, there are none with Completely Planned fertility. At the other end of the planning scale, the

TABLE 3. PER CENT DISTRIBUTION OF ALL COUPLES BY FERTILITY PLANNING STATUS, BY NUMBER OF LIVE BIRTHS

Number of births	Number of couples	Fertility planning status				
		Total	Completely Planned	Partially Planned	Excess Fertility	Not ascertained
Total.....	2,713	100	19	66	13	2
0.....	419	100	37	59	2	2
1.....	603	100	25	68	6	1
2.....	843	100	18	72	8	2
3.....	468	100	9	70	19	2
4.....	190	100	3	68	27	2
5.....	104	100	1	49	48	2
6 or more.....	86	100	0	41	56	3

proportion with Excess Fertility rises with family size. Among couples with 6 or more children, over half are in this group of unsuccessful planners.

Expected family size.--It is clear from the widespread use of measures to control fertility that the major immediate determinants of average family size in the United States are the individual decisions of millions of couples about how many children to have. Although it is doubtful that many couples enter marriage with the intention of having a specific number of children and then have exactly this number, most couples have at least a rough idea about how many children they will have. This is particularly true of couples who, like most of those in our sample, have already had some experience with bearing and raising children. One of the main purposes of this study was to find out how many births such couples expect to have. It is hoped that this kind of information will provide a better basis for making population projections than we have had previously.

In replying to our questions regarding expected completed family size, many of the wives gave a range rather than a single number. The averages of their minimum and maximum replies are 2.7 and 3.3 births, respectively. Their most likely number of births averages 3.0. These numbers include an average of 2.1 births that had already occurred before the time of the interview in 1955.

We cannot yet judge how accurate these expectations are in terms of actual performance. We hope that it will be possible soon to begin to collect the kind of information we need for validation.

One striking fact that emerges from this study is the wide agreement on the desirability of relatively small families. Three-quarters of the wives expected to bear 2, 3, or 4 children. Only one-eighth expected fewer, and only one-eighth expected more. This convergence on the 2- to 4-child family contrasts sharply with the wide distribution of families by size found among women who completed their childbearing a few decades ago. For example, the wives who were born in the early 1890's and who had most of their children between 1910 and 1940 had an average of 3.0 births, which is the same as the number expected by the wives in our sample. However, fewer than half of the older wives had only 2 to 4 children. The proportions who had no births or only one and who had 5 or more births, on the other hand, are higher than comparable proportions of our sample who expect these numbers of births. Obviously, the distribution of families by size is becoming narrower.

We find this convergence on the 2- to 4-child family in every major socioeconomic group. The average numbers of births expected by wives having different socioeconomic characteristics do vary, but not widely. The greatest difference we have found between the expectations of wives in major socioeconomic groups is the familiar urban-rural differential. Wives living on farms expect an average of 3.7 births, as compared with 2.8 for wives living in the 12 largest cities. The difference between these averages does not seem great when we consider the wide variations that are possible.

In general our data on expectations suggest that socioeconomic differences in family size will continue to narrow. We infer this tendency from the family size expectations of wives belonging to different age groups. For example, among the older wives (35 to 39 years old), the urban-rural differences in expectations is fairly large: 4.2 births are expected by farm wives and 2.5 births by wives living in the 12 largest cities. But among the younger wives (18 to 24 years old) there is no difference between the expectations of the wives in these two residence categories. We find similar tendencies toward a narrowing of socioeconomic differences for wives classified by educational attainment, husband's income, and husband's occupation. It is quite apparent that Americans are continuing to become more alike with respect to family size.

The one characteristic for which we did not find a tendency toward convergence is wife's religion. The expectations of the Protestant wives do not vary significantly with age; on the average those in each age group expect about 2.9 births. Among Catholic wives, however, the younger wives expect more births than do the older wives (3.8 births for the 18-24-year olds and 3.1 for the 35-39-year olds). This may mean that the difference between Protestant and Catholic birth rates will increase. We do not think that this will happen, however. Instead, we think that some of the younger Catholic wives have somewhat exaggerated ideas about how many children they will bear. Our basis for this opinion is that many of the older Catholic wives, unlike older Protestant wives, reported that the number of children they wanted was relatively high at the time of marriage, but became lower as they had more experience with bearing and raising children. It is quite possible that the younger Catholic wives will revise their

childbearing expectations downward in the same manner.

Our data on fertility expectations indicate that average family size will continue to increase. Until recently there was a definite secular decline in family size in the United States. Wives born in the early 1870's had an average of 4.4 births. Average family size gradually declined to a low of 2.4 births for wives born between 1906 and 1910. Wives born more recently have already reversed this downward trend. Our data suggest that the average number of births per wife will increase to about 3.0 for women born in 1931-35. This is a medium projection based not only on the data from our study, but also on estimates of the fertility of two other groups: (1) white women who had not yet married by 1955 and were not represented in our sample, and (2) nonwhite women.

An increase in average family size from 2.4 to 3.0 births is not large: neither is an expected slight increase in the proportion married. But together these trends have important implications for the future growth of the United States. Our medium series of population projections show a decline in the rate of natural increase to about 10 or 11 per 1,000 in the next decade, and then an increase to about 13 per thousand. This rate of growth would give us over 300 million people by the end of this century and nearly 600 million by the middle of the next century. Clearly, the moderate-sized families Americans are now having are large enough to maintain fairly rapid population growth.

These long-range projections are cited simply to show what could happen if average family size increased to a specific level. They cannot be regarded as predictions, because if there is one fact that our study has underlined it is that fertility has become a highly volatile phenomenon. With family size largely under voluntary control, birth rates can fluctuate widely and rapidly. As far as forecasting short-range population growth is concerned, however, we think that the information we have obtained concerning expected family size will prove quite useful. Since the climate of opinion regarding family size can change, and probably will change, we think it is desirable to collect information on expectations periodically and use it to revise population projections from time to time. As we learn more about the basic variables affecting fertility and about the relationship between expected and actual family size, we shall be able to make better forecasts of population growth. It is hoped that the present study will prove to be an important step in this direction.

[1]The publisher has granted permission to present this summary of material that will appear in the book.

[2]The words Fecund and Subfecund and certain other terms relating to fecundity and family planning are capitalized to indicate that they are used in a special sense in this study.

[3]Couples were classifed as Semifecund if (1) they failed to conceive during one or more long periods when contraception was not used (a long period is defined as three years for wives who have been pregnant and two years for wives who have not been pregnant) and (2) the average interval between births, if any had occurred, was three or more years. The time intervals used in these criteria are obviously arbitrary, but they are longer than the average length of time required to conceive for couples who do not have serious fecundity impairments.

[4]If wife married more than once marriage duration is measured from her first marriage.

INTERIM REPORT ON THE STUDY OF FUTURE FERTILITY OF TWO-CHILD FAMILIES IN METROPOLITAN AMERICA

By: Philip C. Sagi, Robert G. Potter, Jr., and Charles F. Westoff
Office of Population Research, Princeton University

Introduction

This paper is a brief report on three aspects of a study now in progress known as "The Study of the Future Fertility of Two-Child Families." The study is under the administrative auspices of the Milbank Memorial Fund and the technical direction of the Office of Population Research, Princeton University, with financial assistance from The Carnegie Corporation and The Population Council.[1] Before presenting summaries of (1) the methodology, (2) some substantive findings, and (3) future plans, it might be well to spend a few moments reviewing the background of this current study.

The study is viewed by many as a successor and even a continuation of the Indianapolis Study of Social and Psychological Factors Affecting Fertility. Certainly the two studies share many substantive emphases. There also exists a considerable continuity to research and advisory staffs. Charles F. Westoff and Robert G. Potter, Jr. are associated in a research capacity with both studies. Moreover, in planning the Princeton Study a conscious effort has been made to avoid certain features now accepted as faults in the former study. In the Princeton Study it was hoped that:

1) the process of deriving specific testable hypotheses from fewer, more general hypotheses would give a greater unity to the study

2) by reducing the number of fertility determinants investigated, greater effort could be expended on the problem of measuring each fertility determinant

3) by enlarging the sample size above the 860 "relatively fecund" couples interviewed in the Indianapolis Study, more detailed analyses would be possible

4) by permitting as broad a definition of the population as feasible the general applicability of the findings would be enhanced, and

5) through the choice of a longitudinal design problems associated with post factum data collection would be minimized.

Despite the conscientious efforts to achieve each of these five goals, the first three problems have remained to haunt the present study. In the course of initial theoretical work, in main the efforts of C. F. Westoff and E. G. Mishler, many specific hypotheses were derived from four general propositions.[2] However, the problems of measurement ultimately forced enough changes in definitions of the variables involved that it is now presumptuous to consider the hypotheses as integrated to the extent initially conceived. Certain emphases have changed, generally in the directions that improved the opportunities for measurement. Three pretests preceded the main survey but the scarcity of appropriate and valid instruments capable of being fitted into a crowded hour and twenty minute interview and supplementary questionnaires, the costs in time and money of repeated pretests, the status of knowledge with regard to brief paper-pencil type measurements - these and other difficulties have meant qualities of measurement short of desired standards. Finally, the field costs per couple proved so high, that the sample size, originally set at 1500 couples, had to be reduced to 1165, resulting in a less than desired improvement over the sample size of the Indianapolis Study.

In other fundamental respects the present study has broken with the precedents set by the earlier one. A longitudinal design was introduced as the only genuine solution to the problem inherent in post factum data collection. In the first interview, information about presumed determinants of fertility was collected together with information about fertility intentions. After an appropriate lapse of time, couples are to be reinterviewed to determine how far these intentions have remained constant, how successfully they have been carried out, and which of the potential determinants best predicts fertility performance. Another basic deviation was prompted by the theoretical work of Westoff and Mishler. Their work suggests that some of the determinants of fertility intentions change with parity and that a sound way to delimit the research effort would be to concentrate on a single parity. In view of the recent U. S. demographic situation, the most strategic point in family building was perceived to be the transition between the second and third child. No restrictions have been placed on religion, and indeed, religious affiliation is turning out to be an important attribute. Finally, ecological changes since 1940 have proceeded so far that it seemed wise to base the sample upon the population residing in standard metropolitan areas rather than the central cities.

Methodology

Once the decision to restrict the population to recent two parity couples had been made, the principal problem regarding sampling revolved about the feasibility of reaching a rather rare type of couple spread thinly throughout seven metropolitan areas. This problem was further complicated by such additional eligibility criteria as native birth, single marriage, white race, the absence of plural births and more than one spontaneous abortion. Without the use of birth registration records, it is likely that no economical solution would have appeared.

A systematic 20 per cent sample of all women with second births occurring in September 1956 was drawn from each of seven S.M.A.'s. The resultant sample size consisting of 2891 cases was intentionally greater in number than the planned number of interviews. Since the ultimate number of interviews had been fixed at approximately 1150, the sampling design had to be flexible enough to permit the required number of interviews without affecting the proportionate representation and the random character of the sample in each S.M.A. It was estimated, on the basis of pretest experience, that 50 per cent more names than interviews desired should be allocated to each S.M.A. since some of the potential respondents would refuse to cooperate, or would be ineligible on characteristics not shown on birth certificates.

Because of possible variations among S.M.A.'s on various attrition rates, the sample of names in each S.M.A. were further randomly subdivided into two samples. The first of these samples contained only about 75 per cent of the required number in each S.M.A. The second sample in each S.M.A. became a source of names to fill out the allocated number of interviews in each S.M.A. based on experiences with the first sample. In all, 1709 names were chosen from the first and second samples. From these 1709, came the 1165 successfully interviewed couples. There were in all only 83 refusals. The remaining attrition which reduced the 1709 to 1165 successfully interviewed mothers is summarized in Table 1.

Though these personal interviews averaged an hour and 20 minutes and intensively covered details of contraceptive behavior, marital relationships, income, and religious behavior, the refusal rate was encouragingly low (6%), perhaps attesting to the efficacy of the interview technique as well as the changing mores.

Two additional, return-by-mail questionnaires were given to each women interviewed as a means of reducing the prohibitive costs of lengthy interviewing sessions. A total of 938 or 80.5% of the 1165 wives returned both of these shorter self-administered 20-minute forms. Those not returning one or both forms tended to be Catholics, persons of low educational attainment, low income, short marriage duration, and short preferred third birth interval. The converse was true for those returning these forms with a greater than chance Jewish representation among the cooperative respondents.

Processing these data involved the typical preliminary steps of scoring, coding, card punching, verification and tabulation. Extensive quality controls were maintained throughout. These steps, the drawing of the sample, as well as the field interviews were ably carried out by National Analysts of Philadelphia. In addition, such statistics as means, variances, and correlation coefficients as well as all intermediate statistics were computed by means of IBM 650.[3]

Perhaps the most distinctive features of the data processing employed in this study is the commitment to correlational analyses wherever feasible and the heavy reliance on electronic computer work. In all, some 80,000 coefficients of correlation were computed for combinations of religious and occupational class breaks as well as all marginal totals. It should be understood, however, that our analyses are not limited to Pearsonian r's nor are we attempting to digest any more than a small fraction of the 80,000 or so coefficients spewed forth by our overly zealous machine.

Substantive Findings

The results summarized in this paper are necessarily only a selected few of all findings to date. However, the findings we have chosen to report include "tests" of hypotheses from each of the four areas included in the original theoretical scheme.

The first area is concerned with the relationship between religion and fertility. Since religions are known to be differentially concerned with fertility and fertility controls, such variables as religious persuasion, religiosity, church attendance, and informal religious orientation may be viewed as indices of conformity to various religious teachings.

The specific hypothesis of differences in fertility desires by religion is one that finds support in our data. Catholics desire the largest families and Jews the smallest with Protestants intermediate. The same rank order prevails for fertility-planning success and time taken to have children.

These observed differences of course fail to account for variation in fertility desires within religion. Some of this internal variation is explained in terms of varying degrees of religious belief and behavior within each religion. While religions do vary in the contents of religious teachings, people differ in the degree to which they espouse or practice the religious teachings. An indication of the efficacy of this explanation can be inferred from Table 2 in which we present correlations between formal and informal religious behaviors and fertility desires. While these correlations are not particularly high, they do buttress the arguments relating fertility and fertility desires to religious teachings. A further manifestation of the relationship is found in the association between the extent of denominational education and fertility desires (Table 3). Since denominational schooling tends to be a Catholic phenomenon, our sample only permits comparisons within the Catholic group. The fertility desires of Catholic wives increase with years of denominational education

to a mean of five children desired in the college educated group. However, the cause-effect relationship cannot be unraveled in our survey data. Denominational training may be selective, with the most religious of the parents sending their children to special schools. Further investigations of these relationships are being planned for the second interview with special attention to be devoted to measurements of religiosity of parental home and ethnic background.

The second area of investigation is concerned with hypothesized relationships between fertility and the degree of familistic orientation. As examples, wife's marital adjustment, adjustment to mother role, and liking for children are viewed as indexing the wife's acceptance of her traditional role as mother and housekeeper. The underlying assumption is that wives with many interests outside the home would experience the greatest difficulties in accepting the traditional role and would therefore desire the smallest families, feel least adjusted to marriage and mother role, and express the least interest in children.

The evidence in support of these hypotheses is fairly weak. Correlations between wives' fertility desires and the adjustment variables are presented in Table 4. In all but one case, the correlations are in the proper direction, acceptance of traditional role being associated with high fertility. The magnitudes of correlations, none being greater than .18, leave much to be desired but are still intriguing given the crude measurements. While we hesitate to interpret the variations between religion and class shown by these correlations as other than sampling quirks, it is interesting to speculate over the associations found among Catholics. Perhaps the traditional wife role is most acceptable to the highly religious Catholic wife and mother. (This seems reasonable enough in the light of the high fertility desires of denominationally educated Catholic wives.) Further speculations along these lines, linking the acceptance of the traditional wife role to other variables, for example to husband-wife dominance pattern, fails to be supported by our data.

The third area of interest concerns personality traits. The most general but untestable form of the hypothesis asserts that fertility control and fertility desires are related to certain types of personality characteristics. These personality characteristics are conceived as sets of dispositions supporting or inimical to childbearing and fertility control. Specific hypotheses, for example, involve variables such as generalized manifest anxiety, ambiguity tolerance, impulse gratification, and need achievement. Typically, the expected relationships between these variables and fertility desires or the ability to control fertility rest on rather direct one step reasoning. Persons exhibiting high anxiety are overly concerned with their own problems and, it is argued, such concern reduces fertility desires because having children demands an expenditure of care and affection. In the area of control of fertility, the ability to control or defer sexual gratification appears as a plausible requisite for effective contraceptive behavior. Each of the personality traits measured is linked to fertility or control of fertility by arguments no more elaborate than the two just presented.

Though analysis of these data is not as yet complete, a general observation may be made. The correlations between measures of wives' personality characteristics and various fertility variables (Table 5) indicate that these personality traits, as measured, do not or only barely affect fertility. Despite the low correlations (which we would prefer to attribute to measurement errors) the directions of the correlations tend to follow the expected pattern though again in not too convincing a manner.

These rather disappointing results are somewhat offset by findings in the fourth and last general area of investigation covered in this paper. This area is primarily concerned with the control of fertility in both number and spacing of births. There is therefore a heavy emphasis on use and effectiveness of use of various types of birth control measures.

One of the major hypotheses is the assertion that religions and classes differ in the effective use of contraception. More precisely, recognizing the existence of Catholic prohibitions on means of birth control, Catholic control of spacing is hypothesized as less effective than non-Catholic control. Blue-collar classes were thought to be poorer contraceptors than white-collar classes. As expected (see Tables 6 and 7) religion turned out to be an important factor in the control of spacing. Class differences, on the other hand, were less important than expected. In fact, among Catholics class differences in contraceptive failure rates and proportions of successful contraceptors are in directions opposite from expectations. White-collar classes have higher failure rates and smaller percentages of successful contraceptors among Catholics. This class reversal for Catholics is in part due to the fact that the more religious Catholics tend to be concentrated in the white-collar class and use the less effective contraceptive measures.

The data in Tables 6 and 7 are supplemented with additional information that may be of some general interest. It is evident from a comparison of the failure rates of first and second birth intervals, that contraceptive effectiveness increases with parity. The stability of the proportion of successful contraceptors between birth intervals, 65 and 64 per cent, is somewhat

misleading for the increase in average intended interval from 18.3 to 22.6 months actually implies that couples successfully practiced contraception for a longer period of time.

These few results, to reiterate, have been presented simply in order to give an idea of the scope of this research. It is perhaps worth noting, as a way of summarizing this section on substantive findings, that religion has emerged as the most important determinant of fertility found in the study. Religion affects fertility desires, birth spacing, and contraceptive practices, to a greater extent than any other social category we have isolated to date, including class.

Future Plans

As previously indicated, plans call for a second interview. Field work on this second phase of the study is currently scheduled for the spring of 1960. Prior to this date, interview and questionnaire procedures will be pretested on the same samples used for the pretests preceding the first interview.

Detailed comment about the second phase is not yet possible since the work of drafting specific plans has not progressed beyond the outline stage. Obviously, among the many things that will have to be considered will be the solutions of problems associated with the loss of informants due to various sources of attrition. Residential mobility, death, divorce and separations will create problems of relocating the original 1165 couples and eliciting the required cooperation. Depending in part on the extent of these various attritions will be the intensity of effort spent on locating couples that have moved out of the seven metropolitan areas, the length and form of the interview schedule, and the intensity of the follow-up program for the less cooperative and not-at-home categories of respondents. We have not decided yet if the resurvey will be restricted to couples still residing in the S.M.A. of first residence, or all couples regardless of residence. The latter, if that is our final choice, will obviously involve greater expense and effort. The exact determination of resource allocation depends on this decision as well as the final definition of the objectives of the second phase.

Some of these objectives can be cited tentatively: at present, the single most important concern is whether couples have had their third pregnancy or not. This will permit a test of the prediction validity of stated fertility desires. In any event, third pregnancy information will permit substituting a behavioral criterion for fertility desires in prior analyses involving fertility desires as the dependent variable. Emphasis will probably be placed on changes in fertility intentions as conditioned or not conditioned by third pregnancy experiences. This is a rather important question, for a great deal of our fertility research is conducted on the assumption that fertility desires remain relatively stable and are associated with ultimate family size.

At the same time, there will be some attempt at measuring changes in the economic circumstances of couples in the interim between interviews. There is particular interest in view of the recent recession and whether it had any detectable effect on fertility performance or fertility desires. In other areas, we will elaborate on themes either neglected or found promising in the analyses of first phase data. These include more data on religion, family background, closeness of kinship and friendship ties, and social participation. Other new areas are almost certain to include questions designed to get at respondents' knowledge of the time of ovulation and the extent that such knowledge is used to hasten conception. To some extent the discrepancy between actual birth interval and desired birth interval may be explained by differing lengths of time required to conceive.

These do not exhaust all objectives discussed to date. However, the question of what will or will not be included in the final schedule or schedules is uncertain at this point. The final decisions hinge on an appraisal of the results of phase one, the experiences with the pretest samples and attritions, all of which are also part of the future.

[1]Acknowledgments are due many people who have been connected with this study in one capacity or another since its beginning. The contributions of Elliot G. Mishler, a member of the research staff for the first three years of the study, are immeasurable. Ansley J. Coale, Clyde V. Kiser, and Frank W. Notestein have been involved very closely with the course of the study at the Office of Population Research, forming part of an advisory committee consisting also of: Ronald Freedman, Philip M. Hauser, Dudley Kirk, Frank Lorimer, Donald Marquis, Frederick Osborn, Lowell J. Reed and P. K. Whelpton.

[2]1) Fertility desires and future fertility are positively associated with a familistic as opposed to a work orientation.
2) Fertility desires and future fertility are inversely related to the degree of personality needs that are satisfied by relationships inimical to child rearing.
3) Fertility desires and future fertility are expressions of but one value in the field of socially acceptable and often competing values. (For example, fertility and fertility desires are therefore hypothesized as inversely related to mobility aspirations.

2(con'd)
4) The discrepance between ultimate fertility and current fertility desires, apart from the fact that the latter is susceptible to change while the former is irreversible, is in large measure a function of variations in fecundity, psychological availability of contraception, and fertility planning success.

[3] Jonathan Robbins, at that time with The Population Council, was graciously loaned to us for this computing stint.

Table 3. Family-size preferences of Catholics by education and parochial school attendance.

Years of Parochial School Education	Wives' Family Size Preferences				Husbands' Family Size Preferences			
	Less than H.S.4	High School Graduates	Attended College	Total	Less than H.S.4	High School Graduates	Attended College	Total
None	3.2	3.4	3.4	3.4	3.5	3.4	3.4	3.4
Some	3.5	3.6	4.4	3.7	3.3	3.5	4.3	3.8
All	3.4	4.2	5.1	4.2	3.5	4.0	4.1	3.8
Total	3.4	3.7	4.4	3.7	3.4	3.6	3.9	3.6

Table 4. Correlations between Fertility Desires of Wives and Adjustment Variables by Religion and Class.

	Adjustment to Marriage	Adjustment to Mother Role	Liking for Children	N
Protestant	.06	.10	.12	388
Mixed Catholic	.08	.01	.13	70
Catholic	.10	.18	.16	370
Jewish	.15	-.01	.12	110
White-collar	.11	.08	.13	471
Blue-collar	.04	.16	.14	467
Total **	.08	.12	.14	938

* Test of homogeneity of correlation coefficients fail to indicate significant heterogeneities ($\alpha = .05$) in any combination of religion-class subgroupings.

** All correlations based on total sample appear to be significantly different from zero and in the expected direction.

Table 5. Correlations Between Measures of Wives' Personalities and Fertility Variables for Total Sample N = 938 [1]

Personality Measures	Wives' Fertility Desires	Fertility Planning Success
Manifest anxiety	-.07	.05
Ambiguity tolerance	.11	-.02
Impulse gratification	-.02	.02
Need achievement	.04	.02
Nurturance needs	.06	-.01
Social manners interest	-.01	-.02
Work alone preference	-.02	-.03
Self awareness	-.00	-.05
Compulsiveness	-.11	.03

[1] Tests of homogeneity of correlations between religions and class groupings failed, with $\alpha = .05$, to justify subsample analyses involving religion or class breaks.

Table 6. Contraception Before First Pregnancy by Class and Religion.

Class - Religious Group	Months of Contraceptive Exposure	Intended Interval[1]	Percentage of Successful Contraceptors[2]	Failure Rate[3]
Jewish	1,960	20.6	81	9.8
Mixed Catholic	606	21.2	50	29.7
Catholic	1,968	13.4	61	35.4
White-collar	752	14.3	51	46.3
Blue-collar	1,216	13.0	68	28.6
Protestant	4,224	20.1	58	27.3
White-collar	2,574	21.5	59	25.6
Blue-collar	1,650	18.2	58	28.4
Total	8,758	18.3	65	25.5

[1] Average number of exposure months per successful contraceptor.

[2] Percentage of contraceptors deliberately stopping contraception in order to conceive.

[3] Total failures divided by total contraceptive exposure, all multiplied by 1,200.

Table 7. Contraception After First Pregnancy by Class and Religion.[1]

Class - Religious Group	Months of Contraceptive Exposure	Intended Interval	Percentage of Successful Contraceptors	Failure Rate
Jewish	2,941	22.2	85	7.8
Mixed Catholic	1,434	22.9	58	25.9
Catholic	7,640	24.7	62	21.7
White-collar	2,981	24.0	59	23.8
Blue-collar	4,659	25.2	63	20.3
Protestant	7,901	20.6	60	24.6
White-collar	4,070	20.9	65	21.5
Blue-collar	3,831	20.2	55	27.9
Total	19,916	22.6	64	21.1

[1] Includes both contraception following first birth and miscarriage.

Table 1. Completion rates for total sample by S.M.A.

	Total Sample	N.Y.	Chi.	L.A.	Phila.	Det.	Pitts.	S.F.
Total No. of names used	1709	572	313	289	168	178	107	82
% distribution	100	33.5	18.3	16.9	9.8	10.4	6.3	4.8
No. completed interviews	1165	406	207	169	120	118	88	57
% distribution	100	34.8	17.8	14.5	10.3	10.1	7.6	4.9
Number of refusals[1]	83	35	15	19	7	5	0	2
% distribution	100	42.2	18.1	22.9	8.4	6.0	0.0	2.4
Refusal rate[2]	5.7	7.2	5.8	8.4	4.7	3.2	0.0	2.6
Miscellaneous attrition[3]	66	24	15	17	3	5	1	1
% distribution	100	36.4	22.7	25.8	4.5	7.6	1.5	1.5
No. moved out of area	103	27	26	27	8	10	3	2
% distribution	100	26.2	25.2	26.2	7.8	9.7	2.9	1.9
Attrition rate (moves)[4]	6.0	4.7	8.3	9.3	4.8	5.6	2.8	2.4
Number of ineligible respondents	292	80	50	57	30	40	15	20
% distribution	100	27.4	17.1	19.5	10.3	13.7	5.1	6.8
Ineligibility rate[5]	20.0	16.5	19.5	25.2	20.0	25.3	14.6	24.7

[1] Includes 3 respondents "too busy" to be interviewed.

[2] Number of refusals divided by number of completed interviews plus number of ineligibles.

[3] The total 66 cases includes 19 where no one was home (despite up to 4 calls); 18 where no such address could be located; 16 where there was no respondent contact; 11 cases of illness; 1 vacant dwelling; and 1 language difficulty.

[4] Number of moves divided by number of names used.

[5] Number of ineligible respondents divided by number of completed interviews plus ineligibles. The ineligible respondents appearing in this table are those found ineligible at the time of interview only. They do not include the cases eliminated at the Vital Statistics Offices.

Table 2. Correlations[1] of wives' and husbands' desired family size with measures of religiousness by religion and class.[2]

Religion and Class	Between wives' family-size preferences and:			Between husbands' family-size preferences and:	
	Frequency of church attendance	Informal religious orientation	Religiosity	Frequency of church attendance	Informal religious orientation
All couples	.27	.18	.20	.25	.15
Protestant total	.05	.04	.02	.06	-.03
Mixed Catholic total	.17	.19	.27	.15	.14
Catholic total	.28	.22	.26	.18	.18
Jewish total	-.11	.14	.08	.03	.16
White-collar total	.30	.22	.21	.30	.21
Blue-collar total	.25	.14	.20	.20	.08
White-collar class:					
Protestant	.05	.04	-.04	.07	-.05
Mixed Catholic	.03	.27	.53	.17	.47
Catholic	.23	.26	.29	.32	.30
Jewish	-.09	.13	.08	-.01	.18
Blue-collar class:					
Protestant	.03	.02	.06	.01	-.04
Mixed Catholic	.27	.17	.10	.13	-.16
Catholic	.27	.15	.28	.07	.08
Jewish	-.21	.15	.08	.23	.07

[1] Correlational values required for different levels of statistical significance are shown in the appendix.

[2] The correlations vary among the eight independent subsamples at the .01 level for the wife's religiosity index and the husband's informal religious orientation; at the .05 level for the wife's church attendance; but do not vary significantly for the remaining two variables.

FAMILY PLANNING IN MEDICAL PRACTICE

By: S. S. Spivack and Florence A. Ruderman, Bureau of Applied Social Research, Columbia University

There is a large and growing body of knowledge about the popular use of contraception, and about the demographic, social, and psychological variables associated with fertility. Two of the most important studies in this field, The Growth of American Families, by Whelpton and Freedman, and The Future Fertility of Two-Child Families, by Westoff, Potter, and Sagi, are to be discussed here today. However, in one important area our knowledge has lagged. There has been almost no research into the relevant opinions and actions of the medical profession, a group in a strategic position to disseminate information and to influence popular views on contraception and family limitation.

A study of a national sample of obstetricians and general practitioners, conducted by Dr. Alan F. Guttmacher in 1945, indicated that the great majority of doctors approved of contraception, and recognized social and psychological as well as purely medical indications for its use. Unfortunately, however, this study had a return rate of only 22%, and, being based on a mail questionnaire, was necessarily limited in scope. Besides the questions of extent of approval and of accepted indications, there are other aspects of contraception-counseling in medicine that we should like to know about. For example, what are the circumstances under which a doctor will introduce the subject of contraception rather than waiting to be asked? How do doctors respond to patients' religious or moral problems concerning the use of contraception: and what personal, social, or professional characteristics are associated with different patterns of medical behavior in this area?

Our study was designed to explore some of these broader questions about family planning in medicine. It is based on interviews with 551 doctors in the three apparently most relevant fields of practice (those in which questions or problems concerning pregnancy, child spacing, and family size seem most likely to occur): Obstetrics-Gynecology, Internal Medicine, and General Practice. The interviews, which took on the average one-and-a-half hours, consisted mostly of check-list items, but included also a considerable number of open-end questions. The interviewing took place in summer, 1957, and was conducted by the field staff of the National Opinion Research Center. Nearly all respondents were classified by the interviewers as cooperative and interested, and only one doctor broke off the interview prior to completion.

In designing the study we were particularly interested in the possibility of contextual or environmental effects. We wanted to know, for example, whether rural doctors act differently, or have different opinions, from urban doctors, on whether doctors in areas with large Catholic populations differ from those where there are few Catholics. For this reason we selected six-sample areas, each representing a different demographic constellation. Two of the areas are cities with population over 300,000; two are small-town areas; two are rural. On each of these levels of urbanization the two areas are roughly comparable in median income, level of education, and birth rate; but one has a low proportion of Catholics in the population (from 5% to 23%), the other, a relatively high proportion (from 38% to 41%). (For the sake of brevity, we will hereafter refer to these as "non-Catholic" and "Catholic" sample-areas.) In the two large cities, the sample included all Obstetricians-Gynecologists and about one-third of the Internists and GPs. In the other four areas, all of the doctors in the three relevant fields of practice were included in the sample.

The over-all refusal rate was 12%: 9% among non-Catholics, 19% among Catholics.

In this report we consider some of the factors which are related to participation in counseling patients on family limitation and contraceptive procedures. As a measure of such participation or "involvement" we use an index composed of five items. These are: how often the doctor provides marital counseling for patients; how often he counsels brides on emotional and sexual adjustment in marriage; how often he is asked for information on contraception by patients generally and how often by brides; and, lastly, how often he himself introduces the subject of contraception in pre-marital examinations.

All of the items in the index are related. For example, doctors who say they frequently suggest contraception also tend to say they are frequently asked for it; doctors who say they seldom introduce the subject are likely to say they are seldom asked; and so on. Thus the Index reflects a high degree of correspondence between the doctor's behavior and that of his patients (or between the doctor's perceptions of both). This correspondence may be due to patient self-selection: patients may choose to go to doctors with values similar to their own and to avoid those with very different views. Or it may be that a receptive attitude on the doctor's part encourages patients to ask for information on contraception, while a less receptive attitude inhibits such requests. And, of course, the stimulation or inhibition may work in the other direction: for example, a doctor who wants to advise on contraception may hesitate if he feels the patient is uninterested or disapproves. Whichever pattern it follows, it seems likely that there is considerable mutual, i.e., doctor-patient, effect in this very sensitive area of practice.

The following are a few of our findings on variables related to doctors' involvement in family planning and contraception-counseling. Unless otherwise indicated, these findings hold in each of the six sample-areas.

1. Doctor's Religion. In answer to a question on whether or not they approved of chemical and mechanical contraception, 98% of the non-Catholics, and 27% of the Catholics, said they did. (The few non-Catholic doctors who did not approve--an older group than the sample as whole--expressed reservations on medical grounds, such as

fear that prolonged use might cause sterility, and in some cases a preference for some other, more "natural", method of birth control. The Catholics, by contrast, indicated disapproval primarily on moral or religious grounds.) In view of this great difference in approval of contraception, it does not seem surprising that religion itself produces the largest differences in Involvement in family-planning counseling. In each specialty-group, and in each sample-area, Catholic doctors are considerably less involved in counseling on family planning and limitation than non-Catholics. Thus, for the sample as a whole, 58% of the Catholics are low in involvement, and only 15% high, while for non-Catholics these figures are 29% and 36%, respectively (Table 1).

It is interesting that the Catholic doctors who do approve of chemical and mechanical contraception are only somewhat more likely than those who do not to be involved in contraception-counseling. (Fifty-three percent of those who approve are Moderate or High in involvement, compared with 40% of those who disapprove.) Thus, Catholic doctors' personal attitude toward "artificial" contraception (approval or disapproval) is not the sole, or even the major, variable in their low rate of involvement in counseling on family limitation.

2. Area or Locality. While we find a number of differences among the six sample-areas, there is only one area-related variable that has a consistent effect on involvement, and this appears only among Catholics. Involvement rates among Catholic doctors are always higher in non-Catholic than in Catholic areas. That is, in the large cities, small-towns, and rural counties, Catholic doctors' involvement in giving advice on contraception and related questions is always higher in the non-Catholic than in the paired Catholic area. For example, in the large city with few Catholics in the population, 37% of the Catholic doctors are low in involvement; while in the other city, of the same size, comparable median income, etc., but with a high proportion of Catholics, 68% are low in involvement. In the three non-Catholic areas, an average of 33% of Catholic doctors are low in involvement; in the three Catholic areas, an average of 70% (Table 2.)

Moreover, this is not simply a function of the larger proportion of Catholic patients in the doctor's practice in Catholic areas: when proportion-of-practice Catholic is controlled, Catholic doctors in the non-Catholic sample areas still appear to be more involved in family planning than those in Catholic areas. Thus it seems to be the general religious climate of the community, and not sheer frequency of interaction with Catholic patients, that influences doctors' behavior. And, as the Involvement Index also reflects patient-behavior (and Catholic doctors in our sample tend to have predominantly Catholic practices), this finding suggests that Catholic patients themselves may feel or act differently about contraception and family limitation, depending on the "Catholicity" of the area in which they live.

3. Age. We find that older doctors -- particularly those over 55 -- are less likely to be involved in contraception-counseling than younger doctors (Table 3). An age difference appears in each specialty-group. However, among non-Catholics, the only important difference is between doctors over 55 and those under this age, while young and middle-age doctors are about the same in involvement. Among Catholics, however, there is a steady decline in involvement with age. Thus the youngest Catholic doctors (those under 40) are most involved, the middle group (doctors between 40 and 54) are less so, and the oldest doctors (those 55 and over), are least involved.

Several factors may be at work here. Perhaps the younger doctor is personally closer to problems of family size and child-spacing, and therefore most interested or sympathetic. Perhaps patients find it easier to discuss such problems with younger doctors. Or it may be that the sharp drop in involvment among doctors over 55 is simply part of the more general narrowing of interests and activities that comes with age. Finally, it also seems possible that this relationship may reflect a trend, in the society as a whole, toward fuller acceptance of birth control, and a parallel trend within the medical profession toward greater recognition of its importance. (A general societal trend toward greater acceptance of birth control may have begun earlier, or have reached a peak earlier, among non-Catholics than among Catholics. This would account for the different age patterns observed in the two groups.)

4. Status. It seemed particularly important to know whether doctors who are prominent in the profession differ from others in their views on family planning and contraception-counseling. High status doctors are likely to be the most influential, and therefore this, too, could provide a clue to trends and tendencies within the profession. The index used to measure status was composed of such items as the nature of the doctor's hospital affiliation (senior or chief, active or associate, etc.), having formal accreditation in his specialty, and holding offices in professional societies. As Table 4 shows, status is related to involvement: the higher the doctor's professional status, the more involved he is likely to be in giving advice on contraception.

In part this finding reflects specialty-group differences, because, in general, specialists have both higher involvement rates and higher status than GPs. However, the relationship between status and involvement remains when we control for specialty; i.e., it is found within each specialty-group, and in fact is strongest among (non-Catholic) Internists.

Thus the association of status with involvement suggests that participation in contraception-counseling, or the medical orientation it implies, is a positive value in the medical groups sampled, and that the higher the doctor's professional standing, the more likely he is to have incorporated into his professional behavior the role of adviser on family planning and contraception.

Summary. This paper reports some initial findings in a study of physicians' attitudes and behavior in advising patients' on birth control and contraception. It is based on interviews with 551 doctors, in Obstetrics-Gynecology, Internal Medicine, and General Practice, in six different areas of the country.

An Index of Involvement in Family Planning and Contraception-Counseling, composed of five interview-items on the frequency with which the doctor is approached for advice on contraception and family planning, and with which he gives or offers such advice, showed the following relationships:

1. Involvement in contraception and related counseling is considerably higher among non-Catholic doctors than among Catholics.

2. Among Catholics, involvement is only slightly affected by the doctor's own approval or disapproval of chemical and mechanical contraception; i.e., even among approving Catholics, involvement is lower than among non-Catholics.

3. Involvement among Catholic doctors is significantly related to the "Catholicity" of the area; it is always higher in areas where Catholics are a small minority, and lower where they are a larger proportion of the population. This relationship remains even when the proportion of Catholics in the doctor's practice is controlled. ("Catholicity" of area has no significant or consistent effect on the involvement rates of non-Catholic doctors.)

4. Both among non-Catholics and Catholics, younger doctors are more likely to be involved in counseling on family-planning than older doctors. Among non-Catholics, however, the main difference is between doctors 55 years or older, and those less; among Catholics, there is a more marked trend; Catholics under 40 are the most involved.

5. Doctors of higher professional status, as measured by an Index using nature of hospital appointment, formal accreditations, office-holding, etc. are more likely to be involved in counseling on family planning than doctors of lower professional status. This relationship is found in each specialty-group and among Catholics as well as non-Catholics.

These findings suggest the following tentative interpretations:

1. The fact that the more prominent doctors are most involved in family planning, and that younger doctors are more involved than older ones suggests that the profession generally may be moving toward greater acceptance of counseling on birth control and family planning, as significant medical functions.

2. Despite their obvious and important differences in regard to birth control and contraception in medicine, Catholic and non-Catholic doctors alike may be subject to similar pressures, and similar societal and professional trends. (This is suggested by the relation of involvement to age and status; among both non-Catholic and Catholic doctors - and among Catholics, its relation to religious environment: cf. Points 3, 4, and 5 above.)

Table 1. INVOLVEMENT IN CONTRACEPTION - COUNSELING, BY DOCTORS' RELIGION

Percent in Each Involvement Group

	Low	Moderate	High
Non-Catholics (417)	29	35	36
Catholics (134)	58	27	15

Table 2. CATHOLIC DOCTORS' INVOLVEMENT IN CONTRACEPTION - COUNSELING, IN NON-CATHOLIC AND CATHOLIC AREAS

Percent in Each Involvement Group

	Low	Moderate	High
Cities:			
Non-Cath. (19)	37	21	42
Catholic (44)	68	21	11
Towns:			
Non-Cath. (17)	41	41	18
Catholic (16)	81	6	12
Rural:			
Non-Cath. (14)	21	65	14
Catholic (24)	67	29	4
Mean:			
Non-Cath.	33	40	27
Catholic	70	20	10

Table 3. INVOLVEMENT BY AGE

Percent in Each Involvement Group

	Non-Catholics			Catholics		
Age	Low	Mod.	High	Low	Mod.	High
39 (114)	26	35	38 (42)	50	31	19
40-54 (184)	22	35	43 (67)	55	27	18
55+ (118)	43	34	23 (25)	72	24	4

All Doctors

	Low	Moderate	High
39 (156)	32	34	33
40-54 (251)	31	32	37
55+ (143)	50	32	18

Table 4. INVOLVEMENT AND STATUS

Percent in Each Involvement Group

	Non-Catholics			Catholics		
Status	Low	Mod.	High	Low	Mod.	High
Low (105)	44	39	17	(35) 71	23	6
Mod. (183)	32	34	34	(72) 53	32	15
High (75)	15	38	47	(22) 46	27	27
Very high (53)	11	23	67	(5)		

	All Doctors		
	Low	Moderate	High
Low (140)	51	35	14
Moderate (255)	38	33	29
High (97)	21	36	43
Very high(58)	14	21	65

DISCUSSION

Philip M. Hauser, University of Chicago

The three papers presented at this session, although well done in terms of the present state of knowledge and technique, individually and collectively testify to the deplorable state of knowledge about factors affecting family size. Each of the papers is based on studies which may be described as significant, ambitious, and carefully and effectively pursued. Yet each, by reason of the state of the art and science involved rather than the competence and qualifications of the study directors or the authors, testifies to the ignorance more than the knowledge of the demographer, sociologist and statistician about the variables which account for family size, differentials in size, and family planning.

My discussion focuses largely on the Princeton study presented by Mr. Sagi,partly because it was the only paper made available to me in advance of this meeting and partly because my observations apply with about equal force to the others.

First, it must be noted that the Princeton study embodies a number of substantial gains over previous research in the field. It gives evidence of the accumulation of knowledge, both substantive and methodological, in the design of the study and in the nature of the findings. It represents an important step forward in utilizing a longitudinal approach, and in providing for the reinterview of respondents to check verbal reports by means of reports of subsequent actual behavior. It introduces more effective controls than previous fertility studies in concentrating on a single parity, and is ingenious in focusing on a point in family development which is critical in determining whether there will be relatively large or small families. Other innovations include the study of entire metropolitan area units, which are increasingly the crucial population agglomerations on the American scene; the utilization of an electronic computer which has made possible some 80,000 coefficients of correlation to keep the study personnel occupied; and the achievement of a relatively low refusal rate, 6 percent--low for the nature of this undertaking. Such a low refusal rate necessarily points to an effective coordination of research design and field operations.

Findings of the study some of which are presented in this paper are organized around the impact of religiousity, familistic orientation, personality characteristics, and the control of fertility on family size. A relatively important finding is indicated in the greater effect of religion than of class differences on fertility behavior. The findings speak for themselves and call for little further additional comment here other than the basic one, which documents my reference above to the deplorable state of knowledge in this general area. That is, it is doubtful whether this study will materially increase the explanation of variance in fertility behavior above the 20 percent level achieved in the Indianapolis fertility study. It is the 80 percent area of ignorance which represents the target for students of fertility phenomena in the decades ahead; and from the looks of things it may well be decades before we materially improve our explanation of the variance in family size.

Let me turn next to the consideration of a few of the problems which are involved in this and similar studies and to their basic general limitation.

In my judgement, perhaps the basic problem which afflicts studies of this type, the problem which characterizes much of social science, is the inadequacy of the metrics which are available for many of the significant social and social-psychological variables. Undoubtedly, much of the failure to account for a greater proportion of the variance in family planning behavior and family size lies not so much in failure of theory and concepts as in the inability of investigators, as yet, to obtain good measurements of the social and social-psychological variables which are involved. This is a problem, of course, which afflicts much of social science and one which the demographer, in spite the relative hardness of his data and methods in other respects, shares with other social scientists when he attempts to study sociological and social-psychological variables in relation to demographic variables. This study, like many others, highlights the need to concentrate on ways and means of obtaining better metrics of social variables.

Second, another basic problem, harassing in studies of this type, is the relatively great cost of obtaining information by means of the interview method, even though the most recent advances in sampling are employed. High costs in the context of relatively scarce funds necessarily means frustration to investigators of fertility behavior. In the years which lie immediately ahead, it is doubtful that methodological developments will do much to reduce cost factors. If progress is to be made, therefore, in continued researches in this field, ways must be found to increase the resources allocated and available for research of this type. It should be possible to obtain the needed funds for researches into fertility behavior. With the resurgence of national population growth and its implications for the future population of the United States, it would seem that few problems have greater national import both from the standpoint of domestic policy and the place of the U.S. in the world order.

Third, the study will probably document the limitations of correlation analysis which one of my old teachers, Professor Thurstone, once referred to as a "confession of defeat." Adequate funds which would permit larger samples might well permit a more direct analysis of the data in multi-dimensional cross tabulations and, thus,

more directly unscramble confounded variables which hinder researches of this type.

Finally, I should like to call attention to a more basic limitation of research in this area, the limitation of "historicism." Studies of this type are necessarily conducted over a span of a few years at the most, in a specific context--social, economic, and political. The findings which are obtained must be regarded as representing a point on a secular trend line, a point on a possible business cycle, a point on a possible large deviation from trend represented by a major event such as war. Moreover, findings of this type must be interpreted in light of the nature of the American population in mid-twentieth century--an admixture of ethnic and racial groupings, most of them in a relatively early stage of acculturation, accommodation, and assimilation, to a common national life; and all of them caught in a swirl of rapid social change characterized by increasing and accelerating rates of urbanization and metropolitanization. Considerations of this type obviously point to major limitations to generalizations drawn beyond the specific studies.

The study of the Growth of the American Family presented by Messrs. Whelpton, Campbell, and Freedman is, in essence, subject to the same observations as those made above. The study contains much which represents addition to the fund of knowledge. But it is subject to the same type of specific and general limitations as those to which I have referred above in respect of the Princeton study. Perhaps the most important observation that can be made about the GAF study is that perhaps its most significant conclusion will not be achieved unless it is successful in the additional financing necessary to follow up and check verbal responses with actual behavior.[1]

The Spivak and Ruterman study has touched on a relatively neglected as well as important field. It contains findings of considerable significance not set forth nor documented as well before, namely: (1) the religion of a physician is an important element in the determination of the nature of his medical practice in the family planning area; (2) the medical practioner may follow, rather than lead, in social change in respect to family planning.

In this study, as in the others, the historical context undoubtedly greatly affects the ability of the authors to generalize beyond their specific population.

In conclusion, it should be observed and emphasized that each of these three studies represents an important step forward in achieving a better understanding of factors affecting family size. The teams which have been responsible for the design and operation of the researches have displayed competence, ingenuity, and creativity, both theoretical and methodological in the pursuit of these investigations. The studies, however, must ~~be~~ reviewed not only against the background of what we know but against the background of what need to know. The studies represent progress in the light of the past, but, nevertheless, one must acknowledge the ignorance which all of us in population studies share in respect to the factors which account for family size.

---

[1]Subsequent to the presentation of this discussion the enterprise did receive additional grant of funds from the Rockefeller Foundation to permit the follow-up study.

DISCUSSION

By: William J. Gibbons, Fordham University

All three papers presented reflect the recent advances in methodology and research design of studies on psychological factors affecting family size. In fact, the projects they in part report have contributed significantly to progress in the development of hypotheses for such studies and in improvement of the testing instruments. Moreover, their success in handling with reasonable objectivity, the inherent values of different religious groups is itself worthy of note. 1/

In the "Growth of American Families" paper by Campbell, Whelpton and Freedman, some problems of analysis and interpretation inevitably follow upon the categories used. The definition chosen for subfecundity, for example, leaves relevant questions unanswered. It does not reveal the extent to which known or suspected sterility is the result of voluntary choice or of surgical operation for non-contraceptive purposes.

The "semifecund" category, probably needed under the circumstances, may not take adequately into account the naturally "normal" range of the intervals between pregnancies, apart from efforts to control conception. Put another way, with equal logic those with very short intervals could be placed in an "excessively-fecund" category. Though a minor point, this illustrates how very limited is our knowledge of fertility potential and the various factors, other than contraception, which influence frequency of births.

Likewise, the "indeterminate" category poses problems. It does not indicate extent to which douching "for cleanliness only" may conceal unexpressed contraceptive expectations and desires. We have here an area of possibly confused motivation and latent values. As with other practices which affect fecundity indirectly, this merits study.

Among Catholic couples individually committed to periodic continence, there occur at times discrepancies between expectations and performance because of value conflicts. Failures in personal self-control, essential to the method, can become an important factor in the outcome. How one can measure post factum this element, and then isolate its influence from method-failure is not clear.

The report on the Princeton Study by Sagi, Potter and Westoff, offers instructive findings on cultural-religious factors. The shorter interval among Catholics before first births, may reflect a desire to demonstrate willing acceptance of parenthood as such, as well as a fairly general non-use of contraception during the first years of marriage. The longer average interval between first and second births is specially noteworthy.

Use of the term "desires" in connection with anticipated number of children was not a fortunate choice. In the case of Catholics especially, the differences between actual desires and the expectations may be quite significant, and both may be different from ultimate performance. Given truly improved methods of detecting fertile periods, these gaps between desires, expectations and performance could eventually narrow.

Categories relating to amount of education under religious auspices are somewhat less clear than desirable. Designation of schools simply as parochial or denominational raises difficulties. Breakdown by type might reveal variations according to quality and intensity of religious instruction. A direct relationship between years in the schools and depth of theological-moral knowledge should not be assumed. More refinement in this connection presupposes adequate size of subsample, as well as additional hypotheses regarding degrees of understanding of authentic Church positions.

Of special interest is the indicated differential in desired family size as between the male and female graduates of Catholic colleges. Such a finding merits further exploration, and also gives warning against too hasty generalizations on the motivational effects upon students of religiously oriented education as such.

The study of attitudes of medical practitioners, by Spivak and Ruderman, throws new light on the religious and cultural factors at work. It bears out what was long suspected, that changes over time in medical training have affected the views of doctors on their role as family advisors.

But the finding which stands out in their report is the relationship between degree of involvement and the characteristics of the region wherein the doctor practices. Additional studies may be in order, however, to see what, if any, selectivity factors are at work in leading the Catholic doctors to choose particular areas. Also, the medical school in which they studied possibly is of some relevance.

Use of the term contraception in all three studies, without clearcut breakdown into methods which Catholic norms accept and those they reject, may to some extent bias the results. Because the connotation to Catholics in general is narrower than to sociologists, careful explanation of terms seems necessary. Similarly, the more generic term birth control awakens different emotions and concepts in varying groups. This confusion in terminology is now increased with advent of oral drugs to forestall ovulation, a procedure which has not met with Catholic approval. 2/

To get at additional factors operative in the thinking and decisions of Catholic couples, close attention to their degree of understanding the approved norms will be necessary. And there is also needed greater differentiation between what they would like, what they accept intellectually, and what, on occasion, they may do. Some earlier attitudinal studies were not sufficiently precise in this regard. It is encouraging to note that today's papers manifest considerable care as to definitions. The above comments are made merely in the interests of further advances in clarity.

1/ For an early discussion on the role of norms, see: W.J. Gibbons, "The Catholic Value System in Relation to Human Fertility" in Studies in Population (ed. by G.F. Mair), Princeton Univ. Pr., 1949, p.108-134.

2/ For discussion of the drugs, see "Physiologic Control of Fertility: Process and Morality" by W.J. Gibbons and T.K. Burch, American Ecclesias-Review, 138:246-77 (April 1958). Bibliography.

VIII

ECONOMICS AND POPULATION CHANGE

Chairman, Frederick F. Stephan, Princeton University

Population Growth and Economic Development in the U.S.S.R.—Warren W. Eason, Princeton University

Business and Babies: The Relations Between Economic Fluctuations and the Birth Rate—Dudley Kirk and Dorothy L. Nortman, Population Council

Some Observations on Migration and Economic Opportunity—Dorothy S. Thomas, University of Pennsylvania

Discussion—Demitri B. Shimkin, George Washington University and Bureau of the Census

POPULATION GROWTH AND ECONOMIC DEVELOPMENT IN THE U.S.S.R.

By: Warren W. Eason, Princeton University

It has now been thirty years since the start of the industrialization drive in the Soviet Union and the launching of the First Five-Year Plan; and it will be just fifteen years more, according to the official view, before the Soviets overtake the leading capitalist countries in per capita production and establish the "material-technical" basis for communism.[1] Without challenging the realism or meaning of this prospect, but acknowledging the rapidity of change which has taken place so far, it would seem that enough time has passed for there to begin to be evident basic relationships between population growth and economic development under Soviet conditions. I would like to outline for you briefly what these relationships seem to be to the present, and also to consider some future possibilities.

First, on a technical note, I shall use as population and labor force data, Soviet figures themselves and also estimates which I have made at some length and laboriously on the basis of fragmentary information in Soviet sources. The length of the paper precludes description of methods or discussion of possible magnitudes of error, except in a few cases; and for details I refer you to the usual unpublished manuscript.

The need to understand relationships between population growth and economic and social change is apparent to Soviet social scientists. "Population growth and changes in its structure," wrote Riabushkin at the World Population Conference in 1954, "are determined by socio-economic conditions; while, on the other hand, the planning of socio-economic measures must certainly take into account population indices and trends."[2] At the same time, available materials have been notably weak in offering any discussion or analysis of such relationships. This shortcoming is compounded for us by the fact that since 1928 there has been published what is clearly only a portion of the data on the population and labor force available to the Central Statistical Administration of the U.S.S.R.

Nevertheless, the materials which are available give us something to go on, and I propose to examine the evidence under the following four headings: first, changing population size and structure under changing economic and social conditions; second, implications of changes in population growth for the pattern of economic development; third, relationships between population growth and labor force trends; and fourth, implications of changes in population growth and labor force trends for changes in labor policy. Please forgive me if, in the time at my disposal, I am something less than comprehensive under each of these headings.

I. Changing Population Size and Structure Under Changing Economic and Social Conditions

In the period as a whole since 1928, the demographic aspects of Soviet population growth have assumed the following general pattern: first, a lowered rate of increase of the total population, including lowered birth and death rates; second, a relative decline in the population age 0-15, a relative increase in the population age 16-59, and constance to moderate relative increase in the population age 60 and over; and third, a reduction in the number of males relative to females.

If preplan peacetime characteristics of population growth had been maintained, one would have expected many, if not all, of these indices to have remained essentially unchanged or even to have moved in the opposite direction. The socio-economic forces which operated to change established growth patterns seem to have been of both a temporary and more long-lasting nature.

Starting from a peacetime growth rate of 2 per cent or more per year, with a birth rate in excess of 40 per thousand and a death rate near 20 per thousand, the first of the temporary forces we can locate as having transpired immediately after the start of the plans, during the early 1930's. The data are not complete, but they indicate a sharp fall in the rate of population increase during 1931-1934, and perhaps even an absolute decrease in the population during one or more of these years; they also suggest that this was more a question of higher death rates than lower birth rates. The temporary if rather acute nature of these forces is suggested by the fact that birth and death rates by the later 1930's, when general conditions were more stable, were near preplan levels.

We can only speculate on the nature of the forces affecting population growth in this way during the early 1930's, but the following characteristics of the period would seem to be relevant: the rapid collectivization of agriculture and the reported violent reaction of the peasants; the rapid rate of rural-urban migration; the reformation of regulations concerning marriages and divorces and birth control; and the widespread food shortages of 1932-1933.

The second force of a temporary nature affecting population growth since 1928 is, of course, World War II. If Soviet estimates of their total population in the postwar period are anywhere near correct, the implication for population losses during World War II verges on the incredible. This itself makes some analysts both here and in the Soviet Union feel that when the total population from the forthcoming all-Union census of population of January 15, 1959, is known, it will reveal a higher total than that implied by the official estimate of 200 million as

of April, 1956. In any event, these and other data taken at face value suggest population losses in excess mortality and reduced fertility during World War II and the years immediately thereafter of an order of 40 million.

If the Soviets were to publish their own estimates of the population by age and sex we would undoubtedly see the effects of World War II, and we could chart some of the subsequent demographic and economic effects with respect to the population of working and reproductive ages. In the absence of such data, I have tried to estimate the population by age and sex assuming the mutual interconsistency of the major pieces of statistical information, such as birth and death rates, average life expectancy, etc., and using also a modified "projection" from a 1940 base population. The result as of 1955 compared to 1939 is a sizeable relative decrease in the population age 0-15, from 38 per cent to 29 per cent of the total; a corresponding increase in the population age 16-59, from 55 to 63 per cent of the total; and a slight increase in the population age 60 and over, from 7 to 8 per cent of the total.

The smaller share of children reflects the wartime conditions and the fact that the peacetime birth rate is now less than in prewar years. This appears in spite of the effect of high military losses during the war toward lowering relatively the group age 16-59. The wartime and peacetime birth-rate effect thus outweighs, so to speak, the effect of military losses on the age structure of the population. This is reinforced by the fact that there entered the working ages between 1950 and 1955 the persons born during the period 1935-1939, when the birth rate was high -- persons who, moreover, were of an age during the war to escape military service. The number of males per 100 females, according to my estimates, declined from 92 to 88, between 1939 and 1955. (The corresponding ratio as of 1945 is estimated to be 85.)

Before tracing some of the implications of these wartime changes in population structure for the future growth of the population, let me give some attention to what is apparently a more fundamental change in the pattern of population growth under Soviet conditions. I refer to the lower peacetime birth and death rates and increased life expectancy revealed by annual data for the 1950's. These data show a crude birth rate of about 25 per thousand compared to more than 40 per thousand on the eve of the plans; a death rate of about 8 compared to 18; and average life-expectancy of 67 years compared to 44 years. Standardization for changes in age structure since preplan years would lower the recent birth rate and raise the death rate, which is to say that birth and death rates can be said to be now about one-half of the corresponding preplan rates.

Soviet sources have not provided in any sense a real analysis of these changes. With respect to the lower death rate, a number of sources refer simply to "the growth of public health measures and the improvement of the well-being of the workers."[3] We really can say more than that. It may be shown by a hypothetical calculation that approximately one-half of the decline in the crude death rate from 18 to 8 per thousand is due to the decline in infant mortality, which is reported for 1957 at 45 deaths per one thousand births. Incidentally, this extremely rapid decline in infant mortality, from a level of about 184 in 1940, is rapid enough to cause some to raise a question about the comprehensiveness of the recent figure, a question also raised with respect to other vital statistics. Unfortunately the evidence is very obscure and its examination beyond the scope of this paper, although it may be noted that in conversations with Soviet statisticians, these data were said to be quite comprehensive. In any event, whatever doubts we may have, the evidence does not seem enough on which to base substantial adjustments.

In addition to the effect of infant mortality on the decline in overall mortality, several additional points per thousand in the decline may be ascribed to the aforementioned changes in the age structure of the population, insofar as this reflects the lower birth rate itself. Fewer children born means that fewer children are subjected to the hazards of mortality in the early years.

Finally, the remaining points of the decline in the death rate would represent increased chances of survival for members of the population past the very young ages.

Only one recent source has come to my attention which includes any attempt to explain the decline in the crude birth rate.[4] Even in this case a prewar date is selected for comparison (1940) which minimizes the extent of the decline, which is then ascribed to (1) the lower rate of infant mortality (leading, we can presume, to the need or desire to have fewer children), (2) the increase in the proportion of urban residents, and (3) changes in the sex ratio resulting from war-time population losses. It may be shown, again, by hypothetical calculations, that in all probability these factors fall considerably short of explaining the full extent of the decline. The calculations imply that married women in urban and probably also in rural areas now are having significantly fewer children than formerly, a condition which, as far as it is apparent in reduced average size of families, is confirmed by recent visitors to the Soviet Union.

In the absence of interpretations of the lower birth rate in available Soviet sources, we are forced to speculate on the underlying reasons by relying heavily on analogy with experience in other countries. This subject is also beyond the scope of this paper, although the following factors in the Soviet case which suggest themselves may be listed: the rapid rate at which the educational level of the population has been raised; the larger share of work falling on women, in the cities as well as in the countryside; and the continuation of the shortage of urban housing.

II. Implications of Changes in Population Growth for the Pattern of Economic Development

I should like to be very brief in this section of the paper, having in mind what remains to be said in the third and fourth sections.

The fact that the rate of population growth, starting almost with the beginning of the plans, has been considerably lower than preplan, is of considerable interest with respect to the "population dilemma" facing many of the other industrializing countries of the world. This is the dilemma of achieving a given rate of growth of production, the effectiveness of which on a per capita basis is reduced or even eliminated by an increased rate of population growth due to lower death rates with high or only moderately declining birth rates. The Soviets have not experienced such a dilemma, although the transferability of their experience to other countries is questionable. The initial, sharp decline in population growth rates was of a temporary nature, and with several undesireable implications. The effects of World War II are unique, and the net result in per capita growth rates from this source must also take into account the destruction of plant and equipment.

With respect to longer-run considerations, perhaps forces of a universal nature are implied by the presently lower birth and death rates, in which case the fact that this has taken place in a relatively short period of time is very important. However, until we know the underlying reasons for this trend, and in particular can isolate the effects of World War II itself on the timing of the decline in the birth rate, we may not be justified in seeking general applicability of the Soviet case.

Finally, it would seem that economic growth in the Soviet Union has been enhanced by the trends since 1928 in the age distribution of the population. The larger share of the population of working ages has tended to reduce relatively the demands for consumption on production. The effect has been to permit a higher share of productive effort relative to the given population of working ages to go for investment (nonconsumption) than would have been the case if the preplan pattern of population growth had been maintained.

III. Relationships Between Population Growth and Labor Force Trends

Changes in the size and composition of the labor force result from changes in the size and composition of the population and from changes in the labor force per cent of population by age and sex. A relatively broad definition of "labor force" is convenient for interpretting Soviet data, so as to include persons "having an occupation," but without specific time reference such as the census week in use in the United States. The effect of this definition, essentially one according to "usual occupation," is to include "in the labor force" those working full-time in industry as well as those with a seasonal pattern of work in agriculture. Under this definition, Soviet conditions on the eve of the plans supported a relatively high percentage of the population "in the labor force."

Specifically, more than half of the total population was classified as "having an occupation" in the 1926 census -- including 91 per cent of urban males and 97 per cent of rural males age 16-59, and 45 per cent of urban females and 85 per cent of rural females age 16-59. Unfortunately, the Soviets have published virtually nothing since the beginning of the plans on the percentage of the population in the labor force; but it would seem from indirect evidence that the percentage is only moderately if at all lower. What would otherwise be a downward effect due to rural-urban migration, together with a declining percentage of children and young people in the labor force, has apparently been partly compensated for by a higher percentage of urban males and females in the labor force. Even the downward effect with respect to young people is modified by the fact that, under the broad definition of "labor force," virtually all rural youths and many urban youths 12 years of age and older, despite increased school attendance, would continue to work on farms in the summer, and would therefore be considered to "have an occupation."

At the same time, the yearly average number of persons actually working or employed has increased per cent of the total labor force. The reason is the large-scale migration of labor from rural to urban areas, i.e., from areas where many months of the year were typically spent idle, to more or less full-time employment.

The direct effects of population changes on the labor force include the following: (1) The lowering of the sex ratio as of the early 1930's and due to World War II has tended to increase the number of females per cent of the labor force. (2) The changing age composition has tended to increase the total labor force per cent of the total population. This latter effect has been the stronger in the 1950's, when there entered the labor force persons born during the later 1930's, when the birth rate was relatively high and infant mortality lower than in preplan years.

By the same token, the rate of increase of the labor force can be expected to be slowed down drastically in the coming five or ten years as persons born during World War II -- when the birth rate was low and infant mortality high -- enter the working ages. In the past, the population of working ages 16-59 has increased by about 1 or ? million persons per year, including increases of more than 2 million per year in the most recent period, 1950-1955. Between 1955 and 1960 the rate of increase will continue high, at about 1.5 million persons per year. However, between 1960 and 1965, according to my projections, the increase will be of an order of only a few hundred thousand persons per year. Only thereafter, when groups born after World War II come of age, will effective growth be resumed, with an increase of about 1.5 million per year to 1970 and 2 million to 1975.

Even this is a slower rate of increase than in earlier years, due to successively higher numbers in the labor force. The deficit of males over the next twenty years, of course, will be gradually reduced.

The rate of flow of new labor into production in the Soviet Union in the near future will thus have rather different characteristics from the past. The Soviet planners must certainly not be unmindful of these impending changes, but official data on the subject have not been released, and there is almost no discussion of the subject in available materials. One response to the sharply declining rate of increase of the population of working ages would be to attempt to increase the percentage of the population in the labor force, especially with respect to females. There is some evidence of thinking along these lines, at least by academic people. However, if I am correct in assuming that the percentage of the population in the labor force is already very high, the possibility for compensation in this direction is limited.

This is to say that the Soviets will be under mounting pressure to treat labor as an increasingly "scarce" commodity, a development, now accentuated by demographic forces, related to the continuously rising stock of capital.

This is also to say that the Soviets must face a serious alteration of the age-group "balance" within the labor force. As the next decade passes, the following pattern of distribution by age will develop:

(1) The senior group in the labor force, from which, in addition to others, are drawn top managerial personnel, as well as technical and skilled personnel with accumulated years of experience, is the group born anywhere from about 1895 to 1920. Some of the males in this group were old enough at the time to have been subjected to the military hazards of World War I; some were born during the Civil War, when the birth rate was low; and almost all were of an age to have been subject to military service in World War II. For all of these reasons, the population in these age cohorts is relatively small.

(2) The middle group in the labor force over the next decade will include persons born between the early 1920's and World War II. For most of these years, birth rates were relatively high, and the majority of males in this group were of an age to have escaped military service during World War II. These are the members of the labor force in the next decade who are "in transition" to positions of responsibility, and who are otherwise acquiring experience in all types of jobs. For the reasons listed, this will be a relatively large group in the labor force.

(3) The junior group in the labor force over the next decade, those entering the labor force, will comprise persons born during years of relatively low birth rate and high infant mortality rate (1940-1950), and will therefore be of relatively small number.

The impact of these developments in the matter of planning for labor utilization will be considered below.

IV. Implications of Changes in Population Growth and Labor Force Trends for Changes in Labor Policy

Several major directives have been issued recently by the Soviet government which affect the recruitment, training and utilization of labor. The directives cover, first, the transformation of the educational system and the establishment of priority in full-time advanced training for those with production-line experience; and, second, the reduction of the work-day from seven to six (and in some cases five) hours, without reducing per-man productivity. Although there is no space in this paper to examine the official reasons given for the policy changes, I would like to examine what appear to be rather substantial demographic considerations. These considerations seem important enough to merit our attention, even though there is virtually no discussion of this aspect of the question in Soviet materials.

The essence of the reorganization of the educational system is to create eight-year schools, in a sense (but not literally) to replace the present seven-year schools, and to make eight-year education obligatory for all; and to create eleven-year schools, in a sense (but not literally) to replace the present ten-year schools. Although eight years of education are now obligatory, the number who may avail themselves of the 9th, 10th and 11th years on a full-time basis is strictly limited and on the condition of merit. Furthermore, priority for entrance into higher training (universities and institutes) is now to be given to those who not only have completed eleven years but who also have had production-line experience. This is to say that many who complete eleven years without interruption will be forced to go to work before continuing their education full-time; and that those who go to work after eight years will have to achieve the 9th through 11th years while they work. Widened facilities, particularly within factories, are being developed.

The formula for priorities in entrance to higher education is yet to be worked out. In discussions which I had recently with Soviet educators, following the publication of Khrushchev's thesis in November, a figure of some 80 per cent was mentioned in several instances as the proportion who would be entering given institutions of higher education with not only eleven years of secondary training but also at least two years of experience in production. This compares with a present proportion of an order of 20 per cent, and the transition to the new system will take place within a period of a few years. The entrance procedure in each case, I am told, will be altered to achieve the desired proportion in favor of persons with work experience.

The effect of this program is not to change

the overall rate of increase of the labor force, except to delay entrance into full-time work by one year or so due to the fact that eight-year schooling is now obligatory. Younger persons forced to work before entering advanced training will simply replace the older ones who leave for training.

For this reason the program will serve to increase the average age of the body of students receiving advanced training in universities and institutes. In the first place, it will attract initially the older people who have been on the production line a number of years and who until now have not been able to compete academically with those coming directly from secondary schools. In the second place, it will delay the entrance into higher education of individuals now in the secondary schools.

The effect of the new program in drawing people immediately from the relatively large "middle group" in the labor force will tend to redress any "imbalance" in the proportion of trained and untrained people which occurred when this group passed through the ages heretofore devoted to advanced training. It will also tend to maintain a "balance" in terms of formal skills and age structure between the middle group and the relatively small junior group in the labor force. Thus, by giving the people in the middle age cohorts an opportunity to raise their qualifications, the program will tend to bring the rate at which the advanced schools turn out people of given ages more into line with the rate at which these people are being supplied to the national economy.

The directives with respect to the reduction of hours of work appear as part of a long-run objective the attainment of which was delayed when hours were increased substantially during World War II. The condition that this reduction in hours must take place in a given enterprise without reducing per man productivity seems to be an administrative technique for stimulating individual directors to raise the technological and organizational levels of the enterprise's operation.

The demographic question that comes to mind is this: why, in view of the imminent decline in the rate of increase of the population of working ages, have the Soviets picked this time to reduce hours of work? Will it not aggravate the labor supply problem? In the most obvious sense it will. On the other hand, if we take into account the age structure of the labor force to which I have referred in terms of the three groups, and the fact that the rate of increase of the labor force in the future will not for many years again be as high as in the past, this reduction of hours may not necessarily be ill-timed.

First, to reduce hours at anytime in the future will be to do so in the face of a slower rate of increase of the labor force than in the past. Furthermore, the rate of increase of the labor force over the past seven years or so has been unusually high. Taking this seven-year period together with the next seven years produces an average increase of about one million persons per year, or not significantly below the increase for most years since 1928.

Second, as the program is carried out over the next few years, the hours of labor reduced will be primarily those supplied by the relatively large "middle group" in the labor force. To execute the program at the present time, therefore, serves to accentuate the need to adjust to a long-run reduction in the rate of increase of labor supply, by cutting immediately into the relatively large labor supply that has come on the scene over the past five or ten years.

Since this is a meeting of statisticians, I would like to end my presentation on a statistical note, by calling your attention to the fact that within the next month, on January 15, 1959, the Soviets will conduct their first census of population in twenty years. There is some reason for hoping that more of the data from this census will be published than from the 1939 census. A certain optimism in this connection is enhanced by the gradual appearance in recent years of more statistics in all fields and by the discussion of some of these statistics in the academic journals. With the publication of the data from the census, therefore, we might also hope that discussions will widen to the area, heretofore virtually ignored, comprising relationships between population and economic and social variables.

Footnotes

1. Pravda, November 14, 1958.

2. T. V. Riabushkin, "Social Aspects of Population Structure and Movement," Proceedings of the World Population Conference, 1954 (New York: United Nations, 1955), Volume V, p. 1031.

3. MinZdrav SSSR, Zdravookhranenie v SSSR: statisticheskii spravochnik (Moscow: 1957), p. 9.

4. M. Ia. Sonin, "Ob aktual'nykh voprosakh vosproizvodstva trudovykh resursov SSSR," Akademiia Nauk SSSR, Voprosy sotsialisticheskogo vosproizvodstva (Moscow: 1958), pp. 258-259.

BUSINESS AND BABIES: THE INFLUENCE OF BUSINESS CYCLES ON BIRTH RATES

By: Dudley Kirk and Dorothy L. Nortman, Population Council

The recent economic recession has revived interest in the effects of economic fluctuations on birth rates. Both in popular thinking and in the scholarly literature it is assumed that the low birth and marriage rates of the 1930s were attributable to the great depression. Correspondingly, the continuing marriage and baby boom of the past decade is commonly thought to reflect the prosperity of the postwar years. There have been predictions that the 1957-58 recession would be followed by a significant recession in births.(1)

While there has been general awareness that vital events are influenced by economic conditions, there have been relatively few published studies that apply statistical measurement to this relationship. There have been excellent reasons for this, both in the defects of the basic data and in the methodological problems of measurement. The basic data are now improved, but the problems of measuring and interpreting the degree of covariance between economic data and vital statistics have not been entirely resolved. Among the more serious of these problems is serial or auto-correlation in time series, an influence that tends to exaggerate the apparent covariance.

Studies by Thomas, and particularly the well-known paper by Galbraith and Thomas(2), attempt to avoid this error by using trend deviations rather than absolutes in analysis of the relation between business cycles and births. Using trend deviations these authors found a correlation of .80 between the Bureau of Labor's suggested index of factory employment and total births for the years 1919-1937, with births lagged one year. The authors state their conclusion that

> "Marriages are 'controlled' during depressions, and, within marriages, births of all orders are likewise controlled. Since birth rates of higher orders are overweighted with births to the lower income, occupational, and educational classes, it is clear that the birth control movement has penetrated deep into the social structure during the past two decades.(3)

The present paper will: (1) summarize the application of similar methods to annual data in the entire period since World War I,(4) and (2) extend the analysis to monthly and quarterly data using both correlation and reference cycle techniques.

Correlation Analysis

The present study correlates trend deviations of economic measures (as independent variables) to measures of nuptiality and natality (as dependent variables). This method greatly reduces the serial correlation that often inflates the apparent covariance in time series, for example, when this covariance is measured by absolutes, by annual percent change, or by deviations from moving averages.(5)

For the test of the relationship between vital rates and business activity by the conventional correlation technique, three economic indicators are used: (a) real per capita personal income, (b) the Federal Reserve Board's index of industrial production and (c) non-agricultural employment. Each of these has been related for the period 1920-58 (omitting years most directly affected by World War II) with measures of nuptiality and natality, i.e., marriages per 1000 unmarried women 15 to 44 years of age and births per 1000 women 15 to 44. The use of these data reduces the variability resulting from changes in the age structure of the population. The relation between economic fluctuations, nuptiality and natality are shown in Chart 1. In this chart per capita income in constant dollars is used as representative of the economic indicators.

It will be noted that since World War I the three variables have had quite different trends, indicated by the three dissimilar lines plotted on Chart 1.

1. Economic Activity. There have been wide fluctuations around a strong upward historical trend of growth, which for personal income has averaged over 2% per year. Personal income, and indeed most indices of economic activity, were far below this historical trend throughout the decade 1930-1940, far above it during the war years, and close to the historical

REAL PER CAPITA PERSONAL INCOME, NUPTIALITY & FERTILITY AND RESPECTIVE TRENDS ... 1920-1958

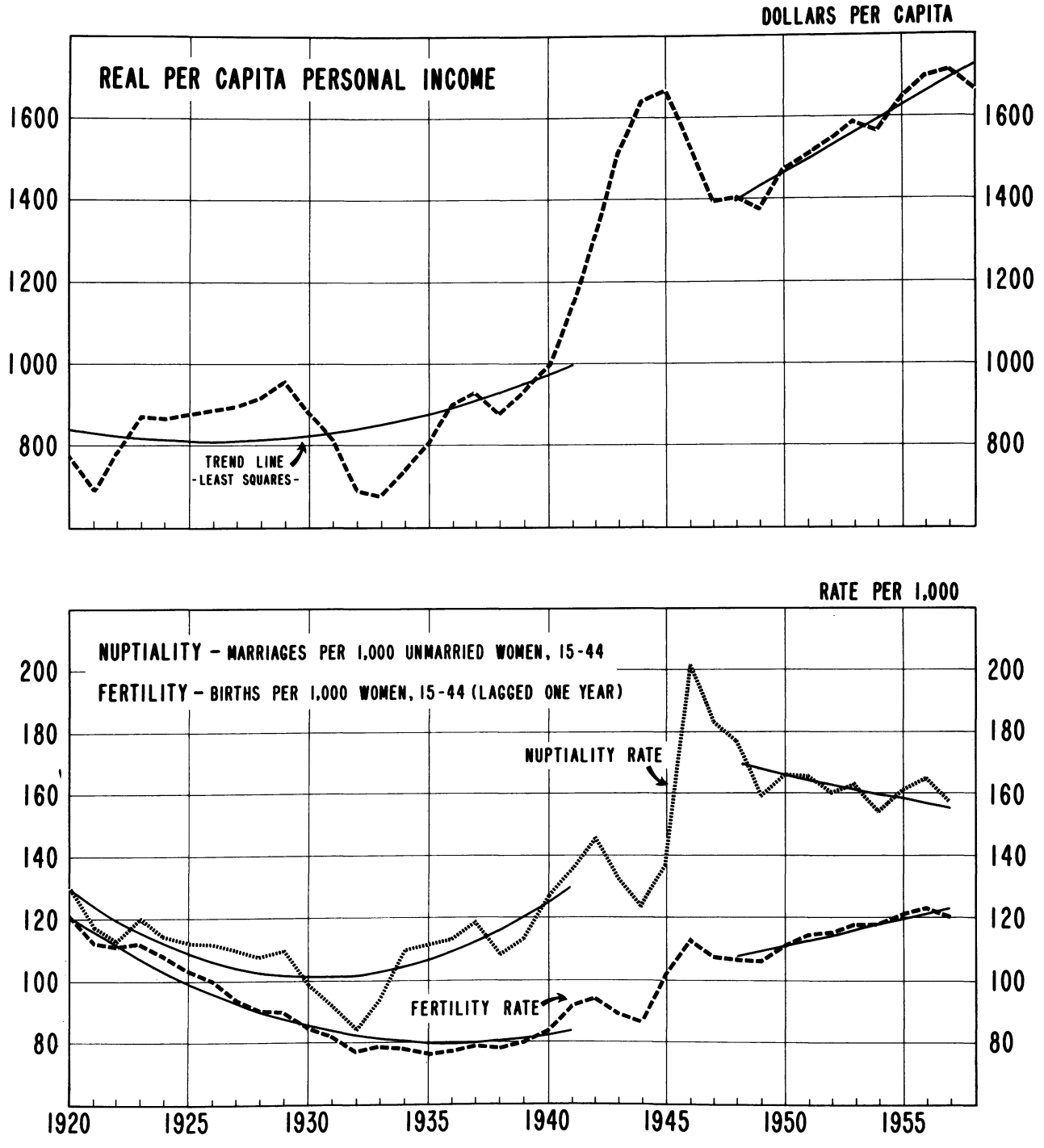


CHART I

trend during the 1920s and since World War II.

2. Nuptiality. Marriage rates were drifting downward in the 1920s, and experienced a very sharp dip in the worst depression years, followed by a recovery so strong that by 1940 "postponed" marriages had effectively been made up. This is evidenced by the fact that the proportions of women married at each age in 1940, at the end of the depression decade, were very similar to the proportions in 1930 at the end of a decade of prosperity. In the years 1940-1946, marriage rates were of course sharply influenced by the threat and fact of the draft, the outbreak of war, mass conscription, and finally postwar demobilization. Since World War II nuptiality has drifted downward but has remained at much higher levels than before the war. The higher birth rates of the postwar period are attributable in part to the higher proportions married.(6)

3. Natality. The plotted fertility rates shown on Chart 1 basically describe a shallow U-shaped curve, somewhat distorted by events connected with World War II. There is less year-to-year variability than in the economic indices or in nuptiality.

It will be evident that there is only general correspondence between the major trends of the three variables. The downward trend of nuptiality and fertility is usually interpreted as a reflection of the spread of voluntary control of family size among progressively larger segments of the population. The downward drift of fertility was accentuated, but not drastically, by the headlong economic decline of the early '30s. At its low point in 1933 the fertility rate was only 14% below its 1930 level, a year which reflected the peak prosperity and employment conditions of 1929. During the preceding three years, 1927-1930, reflecting a period of economic prosperity, the decline in fertility had been 10%, suggesting on this rough basis of comparison that the depression brought about a decline of only 4% more than otherwise would have occurred. After 1933 the historical decline of fertility was checked and stabilized so that fertility was higher than would have been expected with the continuation of pre-depression trends.

It would be unreasonable to argue from these major trends that the prosperity of the 1920s "caused" the fertility declines of that period and that the depression brought about the end of this decline in the 1930s. But conversely the pattern of major trends does not support the belief that the depression "caused" the low birth rates of the 1930s. While it is true that marriage and birth rates were relatively low in the depression years and relatively high in the years of prosperity, there is no clear indication of causal relationship, so far as the major trends are concerned.

The statistical measurements of covariation employed in this study require the removal of secular trends and the measurement of deviations from these trends. The precise definition of trends inevitably involves arbitrary elements. For present purposes least squares lines were fitted for all series, in each case parabolas for the data for the period 1920-1941 (1921-1942 for the birth rates) and linear least squares lines for the postwar materials. The trend lines are plotted in Chart 1 and the results of correlation analysis are shown in the first column of Table I.

The possible effects of using different trend lines were tested by computing a set of correlation coefficients based on deviations from assumed historical geometric rates of growth for income (2.3 per cent per year) and industrial production (3.9 percent), the measure of employment being represented by unemployment as a percent of the civilian labor force. The results are shown for comparative purposes in Table I. The two sets of correlations are generally consistent, and in some instances identical, giving confidence that the particular method of determining trend is not a decisive factor in the results.

The correlation data suggest the following generalizations:

1. There is a high degree of correlation between year-to-year fluctuations in economic indicators, in nuptiality and in fertility.

2. The relationship between the economic indicators and fertility was very stable for the interwar period, with no important difference where different independent variables or trend lines were employed. Not unexpectedly the coefficients in this series approximate the figure of .80 obtained by Galbraith and

TABLE I

CORRELATIONS OF PERCENTAGE DEVIATIONS FROM TRENDS OF FERTILITY, NUPTIALITY AND ECONOMIC INDICES

Period	Independent Variable	Dependent Variable	Correlation Coefficient[a/] Method I	Method II
1920-41	Per capita income	Fertility[b/]	.77	.74
	Industrial production	"	.76	.73
	Employment[c/]	"	.76	-.73
	Per capita income	Nuptiality	.68	.76
	Industrial production	"	.76	.74
	Employment[c/]	"	.72	-.72
1948-57[d/]	Per capita income	Fertility[b/]	.66*	.86
	Industrial production	"	.78	.79
	Employment[c/]	"	.57**	-.65*
	Per capita income	Nuptiality	.79	.79
	Industrial production	"	.66*	.66*
	Employment[c/]	"	.63*	-.45***
1920-41	Nuptiality	Fertility[b/]	.55	.49*
1948-57	"	"	.41***	.30***

*, **, Significant at .05, .10 level respectively.*** Not statistically significant. Unless otherwise indicated all values significant at .01 level.

a/ Methods as described in text. b/ Lagged one year. c/ For Method II, unemployment as percent of civilian labor force. d/ 1947-1956 for Method II.

Thomas for their series (using somewhat different variables) in the period 1917-1937.

3. The relationship between economic indicators and nuptiality for the interwar period is also quite stable, with a range of coefficients from .68 to .76. The results give confidence that the method of determining the trend lines is not a decisive factor in the measurements of the interrelationship in the interwar years.

4. There is much greater variability in the coefficients for the past decade, perhaps owing to the shorter series and lower statistical reliability of the results. The selection of years in this shorter series makes a significant difference in the results. There is some suggestion that the relation of employment or unemployment to nuptiality and fertility is lower in the postwar than in the interwar period. This is a result to be expected in view of the fact that even during postwar recessions a relatively small part of the total population was directly affected by unemployment per se. Otherwise, the relationships seem to be of the same general order of magnitude as in the interwar period. The coefficients of correlation between nuptiality and fertility are the lowest in the series, both for the interwar and the postwar series.

Partial and multiple correlation analysis for the interwar period is presented in Table II. The results suggest that when the economic variables are held constant there is effectively no correlation between nuptiality and fertility. In other words, the data imply that the influence of nuptiality on trend deviations in fertility is a secondary effect of economic fluctuations.

The correlations between economic indicators and fertility (i.e. .76-.77) in the interwar period suggest that the economic indicators explain some 58-59% of the variance in fertility series. These data combined with the simple correlation between nuptiality and fertility (.55, Table I) suggest that about 30 per cent or roughly one half is exercised through nuptiality and the remainder

TABLE II

PARTIAL AND MULTIPLE CORRELATIONS FOR TREND DEVIATIONS, 1920-1941[a/]

Variable	Correlation Coefficient
	Partial
Nuptiality Constant	
Per Capita Income & Fertility	.65
Industrial Prod. & "	.64
Employment & "	.62
Per Capita Income Constant	
Fertility & Nuptiality	-.06*
Industrial Production Constant	
Fertility & Nuptiality	-.08*
Employment Constant	
Fertility & Nuptiality	.004*
	Multiple
Fertility & Nuptiality with	
Per Capita Income	.77
Industrial Production	.76
Employment	.76

* Not statistically significant.
a/ Fertility lagged one year throughout.

TABLE III

CORRELATIONS FOR TREND DEVIATIONS 1947 QUARTER 3 THROUGH 1958 QUARTER 3[a/]

Variables		Coefficient
Independent	Dependent	
Per Capita Total Personal Income	Fertility	.48
	Nuptiality	.58
Per Capita Disposable Income	Fertility	.33*
	Nuptiality	.54
Index of Industrial Production	Fertility	.59
	Nuptiality	.66
Unemployment as Percent of Civilian Labor Force	Fertility	-.43
	Nuptiality	-.45
Nuptiality	Fertility	.42

* Significant at .05 level. Unless otherwise indicated all values significant at .01 level.

a/Data seasonally adjusted, by quarter, with fertility lagged nine months.

through the direct influence of economic conditions on fertility.

The coefficients for the postwar period (not shown) are of questionable significance, but indicate a somewhat lower level of influence of economic conditions, especially employment, than for the interwar period.

A final step in the correlation analysis was to apply it to monthly and quarterly data. Because of the difficulty of obtaining a monthly seasonally adjusted series of nuptiality rates, correlation coefficients were computed based on seasonally adjusted data by quarters. As in the case of the annual data, linear least squares lines were assumed to represent the secular trend, in this case for the period from the third quarter of 1947 through the third quarter of 1958. Correlations of percent deviations from these trends, in which fertility rates were matched with the economic indicators and nuptiality nine months earlier, are presented in Table III.

The coefficients are generally lower than for the annual data. The highest relationship was found with the index of industrial production, which tends to support the hypothesis suggested earlier in this paper that in a prosperous era, the general economic climate may influence fertility more than specific factors such as employment. It is interesting that the coefficients for nuptiality are higher than those for fertility, indicating that marriages may well be more immediately responsive to changes in economic conditions than births. The computed line of regression indicates positive conformity to changes in the economic variable, in the ratio of one to five for fertility rates and about three to five for nuptiality rates with the index of industrial production.

The above figures relate to aggregate and period data. The aggregate fertility index used, for example, masks divergent secular trends in the rates for the several parities, as shown in Chart 2.(7)

It will be noted that the parity rate for first births was first to turn upward from the long secular decline that culminated in the low fertility of the thirties. The upturns of parity rates

REAL PER CAPITA PERSONAL INCOME, NUPTIALITY RATES & GENERAL FERTILITY RATES BY 1ST, 2ND, 3RD AND HIGHER ORDER PARITY

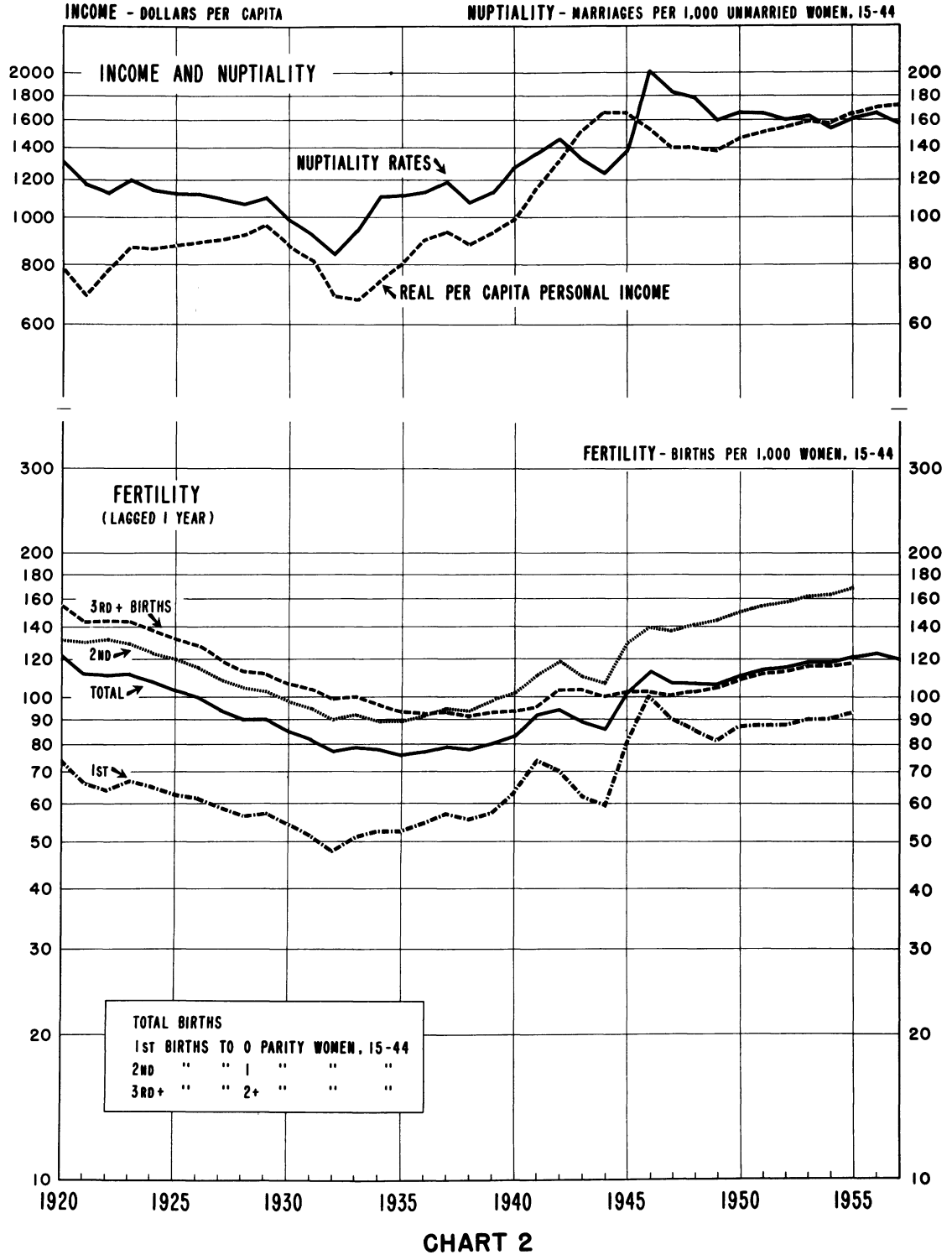


CHART 2

for first and second births were harbingers of the general rise of fertility in the postwar period. The parity rate for third and higher orders continued to decline through the decade, reaching its lowest point in 1939. This sequence would be expected with the continuing spread of family limitation to larger segments of the population and the reduction in the number of larger families.

It will be observed in Chart 2 that variability in parity rates for first births, and to a less extent second births, are chiefly responsible for the year-to-year fluctuations in general fertility, and it is clear that these are closely associated with the incidence of marriage. The aggregate rates for the third and higher orders of births to women having two or more children show little year-to-year variability. When the more immediate influence of marriage is removed, as in parity rates for higher orders of births, there is a clear secular trend, modified only slightly in its annual variations by economic fluctuations, and even by World War II.

Reference Cycle Analysis

The correlation analysis used in the preceding section of this report has important methodological weaknesses as a basis for statistical inference. As noted above the fitting of trend lines inevitably is arbitrary and may well produce vagaries in the results. The postwar period is too short to justify drawing firm conclusions from correlation data. Finally correlation techniques do not adequately measure the influence of critical turning points in cyclical behavior.

An alternative method is the so-called "reference cycle" approach developed by the National Bureau of Economic Research.(8) It has the advantage that it does not depend upon any specific economic series but involves an examination of the degree and direction of movement of the dependent variable at crucial reference dates pre-determined by the Bureau to represent the turning points of cycles of economic activity. Besides affording various measures of the cyclical behavior of the series under consideration, the technique yields an index of conformity to reference cycles analogous to the coefficient of correlation.

The use of short-term data is an essential aspect of the reference cycle approach. In the absence of a monthly series of fertility rates, the crude birth rate (i.e., births per 1000 population) by months since 1919, seasonally adjusted and corrected for under-registration, was employed.

A striking feature of the birth rate series is the absence of cyclical pattern. The birth rate fluctuates within a relatively narrow range around its secular trend and without the marked peaks and troughs that characterize the measures of economic activity.(See Chart 3).

The National Bureau of Economic Research has identified eight business or reference cycles in the period from the trough of March 1919 to the trough of April 1958. For comparative purposes the birth rate series was also divided into eight groups, each one starting and end-ind nine months after the initial and terminal troughs of the reference cycles. Table IV presents the dates of each cycle and the average birth rate during the period concerned.

TABLE IV

AVERAGE BIRTH RATES
DURING SPECIFIED BUSINESS CYCLES

Cycle	Month and Year		Average
	Initial Trough	Peak	Birth Rate
1	12/19	10/20	27.8
2	4/22	2/24	26.0
3	4/25	7/27	23.8
4	8/28	5/30	20.3
5	12/33	2/38	18.8
	War Years Omitted		
6	7/46	8/49	25.4
7	7/50	4/54	25.0
8	5/55	4/58	24.9

The reference cycle technique further involves the subdivision of each cycle into nine stages; stages I, V and IX are the initial trough, peak and terminal trough respectively; stages II, III and IV are the expansion phase divided into three equal segments; stages VI, VII and VIII are three contraction periods of equal duration. The monthly data during each cycle are reduced to nine values,

TABLE V

REFERENCE CYCLE DATA FOR BIRTH RATES, 1919-1958

Cycle[1]	Reference Cycle Relative at: Initial Trough	Peak	Terminal Trough	Aver. Monthly Change in Ref. Cycle Relatives During: Phase: Expansion	Contraction	Preceding Expansion	Succeeding Expansion (Sign)
1	104.7	97.1	92.8	-.76	-.24	.52	Negative
2	99.2	100.8	98.5	.07	-.16	-.23	Positive
3	107.7	98.6	92.0	-.34	-.51	-.17	Negative
4	108.0	105.6	92.4	-.11	-.31	-.20	"
5	99.6	102.3	100.0	.05	-.18	-.23	"
6	92.8	96.2	95.5	.09	-.06	-.15	"
7	97.1	100.1	99.0	.09	-.08	-.17	"
8	99.5	99.1	97.1	-.01	-.22	-.21	....
Av.	101.1	100.0	95.9	-.12	-.22	-.10	....
Av.Dev.	4.3	2.2	2.7				

Index of Conformity to Reference Cycle				
Expansions	0			
Contractions	...	100		
Cycles, Trough to Trough	...	...	75	
Cycles, Peak to Peak	...	...	...	71
Cycles, Both Ways	...	...	73	

1/ For dates, see Table I

one for each stage, to represent the average value per stage as a percent of the average value of the birth rate during the entire cycle.

On the basis of these nine percentage values per cycle (called reference-cycle relatives) it was found that for the eight cycles combined the crude birth rate rose by .2 percent per month from stage I to II, declined by .4 percent per month from II to III, showed no change to stage IV and again declined by .2 percent per month between stages IV and V. In the contraction phase, the monthly rate of change was .1, -.3, -.3 and -.2 percent respectively between successive stages V to IX. These figures indicate first the relatively minor intra-cycle variability of the crude birth rate, second, the tendency to decline from stage to stage, which is the result of the marked secular decline from the early 1920s to World War II.

As can be seen in Table V above, in four of the eight cycles, namely 2, 5, 6 and 7, the value of the birth rate at the business cycle peak exceeded its standing at the initial trough. During the contraction phase, the value at the terminal trough was always less than its level at the peak. In 7 of the 8 full cycles the average monthly rate of decline was greater during the contraction than the corresponding expansion phase. The algebraic signs yield indices of conformity to business cycles of 0, plus 100 and plus 75 respectively for the expansion, contraction and full cycle phases.

The index of conformity represents the percentage improvement over chance that the variable will move in accordance with fluctuations in business cycles.

According to the standards applied by the National Bureau of Economic Research, an index of conformity of more than 50 is indicative of close conformity.(9) The average index of conformity for birth rates, at 73, indicates a close relationship. But the index of conformity is computed from direction rather than

INDEX OF INDUSTRIAL PRODUCTION AND CRUDE BIRTH RATES FOR POST-WAR PERIOD

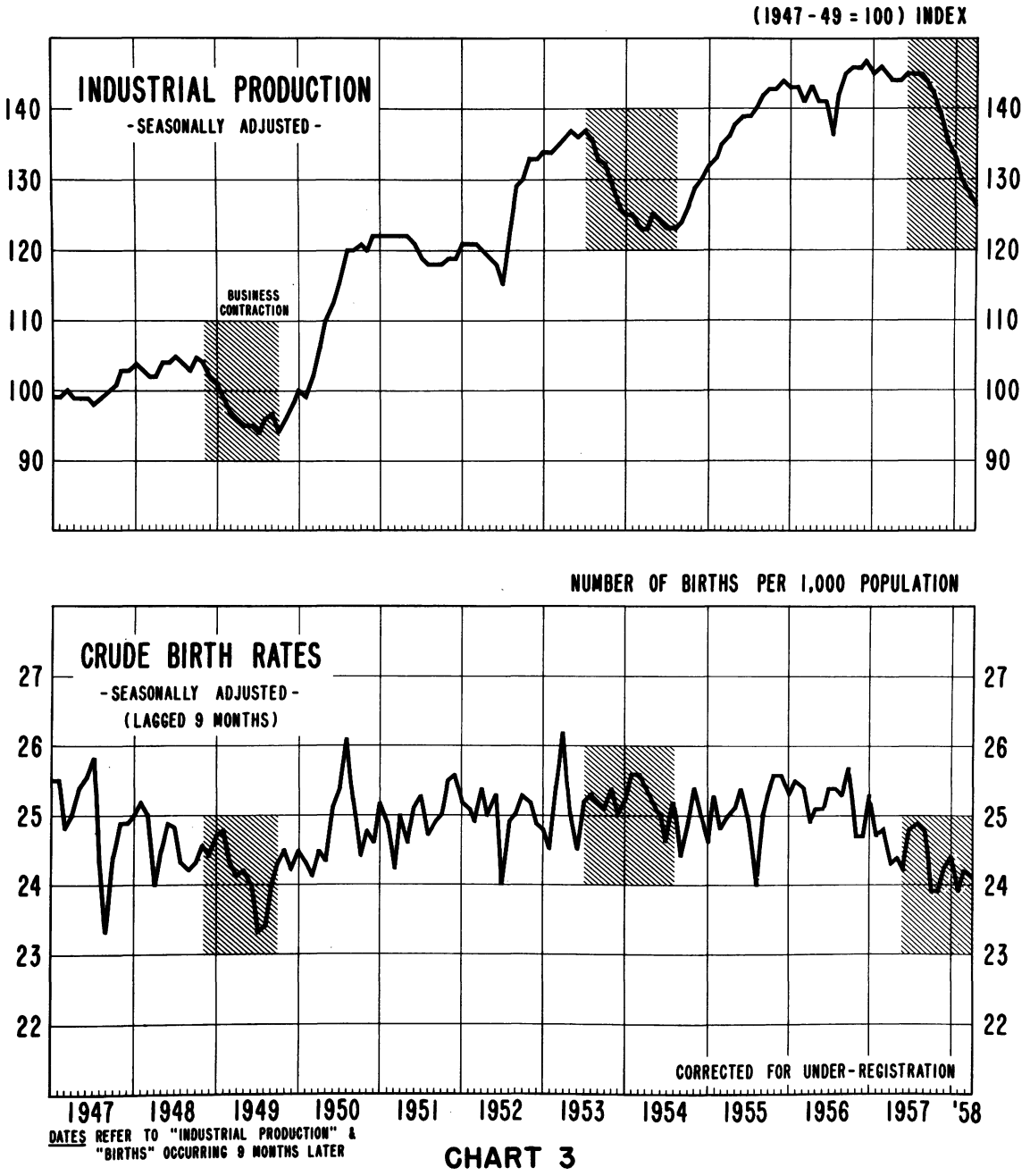


CHART 3

amount of change. As noted above the average amount of change in birth rates in concordance with phases of the business cycle is small.

The rather tenuous nature of the relationship is suggested by the postwar data shown in Chart 3. The birth rate was responsive to business contraction in 1949 and to the outbreak of the Korean War in 1950. It seems to have responded briefly but strongly to the dip in economic activity in mid-1952 but scarcely at all to the recession of 1953-54. Even the sharp business contraction of 1957-58 had had only mildly depressing effect on

the birth rate up to the end of 1958. In fact much of the drop in the birth rate reflected in lower 1958 figures apparently occurred as a result of a fall in the marriage rate antedating the onset of the recession. Much of the variability in the birth rate appears to be unrelated to economic cycles, and the birth rate has in any case fluctuated within narrow limits during the past decade.

Conclusions

In this study the influence of business cycles on natality has been examined with standard correlation and reference cycle techniques. At best these methods supply statistical inference and not direct measures of causality. Within this and other limitations the analysis suggests the following conclusions.

1. The level of natality (whether measured by crude birth rates or fertility rates) is chiefly determined by stable secular trends, and not by cyclical influences. These secular trends may well be quite independent of trends in economic activity. Thus the data of this study do not confirm the view that natality was low in the 1930s because of the depression.

2. Year-to-year fluctuations in natality are to a large extent a function of economic cycles. The results of this study confirm the generally held view that nuptiality and fertility respond sensitively to economic conditions. The correlation between fertility and the economic indices used in this study are high. They suggest that economic conditions control rather more than half of the annual variance of fertility from its secular trend, the degree of control varying relatively little with the economic index, the choice of trend and the period covered. A possible exception is the low correlation of employment and unemployment with fertility in the postwar period. The reference cycle analysis corroborates correlation analysis in that both show a substantial degree of covariance between business cycles and natality.

3. Monthly and quarterly data for natality are not as highly correlated with economic indices as are annual data and cyclical data used in reference cycle analysis.

4. While fertility responds sensitively to business cycles in its trend deviations, changes in economic conditions are not accompanied by changes in fertility of comparable magnitude. Thus over the whole period studied a trend deviation of 4 percent in real per capita income produced a trend deviation of only 1 percent in fertility. Thus sharp business contractions, such as the 1957-58 recession, may be expected to produce much smaller fluctuations in fertility. Quick judgments that there will be a major recession in births following on the recent business recession may well prove to be premature and lacking in perspective.

FOOTNOTES

(1) Population Reference Bureau, "Recession in Births?", Population Bulletin, vol.14, no.6, October 1958.

(2) Virginia L. Galbraith and Dorothy S. Thomas, "Birth Rates and the Interwar Business Cycles," Journal of the American Statistical Ass'n. Dec.1941, pp.465-476.

(3) Ibid., p.473.

(4) This section of the paper draws heavily from materials presented in a paper by Dudley Kirk, "The Influence of Business Cycles on Marriage and Birth Rates in the United States" in Proceedings of the Conference on the Interrelations of Demographic and Economic Change, December 5-7, 1958, being published by National Bureau of Economic Research.

(5) This problem is discussed at length in the above paper, footnote 4.

(6) Cf. Wilson H. Grabill, Clyde V. Kiser, and Pascal K. Whelpton, The Fertility of American Women, New York, John Wiley & Sons, 1958, pp.369-70.

(7) Fertility rates by parity for Chart 1 were kindly supplied by P.K.Whelpton and Arthur A. Campbell.

(8) Arthur F. Burns and Wesley C.Mitchell, Measuring Business Cycles, National Bureau of Economic Research, Studies in Business Cycles No.2, 1947.

(9) Burns & Mitchell, op.cit. p.123

SOME OBSERVATIONS ON MIGRATION AND ECONOMIC OPPORTUNITY

By: Dorothy Swaine Thomas, University of Pennsylvania

In 1951, Simon Kuznets and I planned a series of estimates of population redistribution and of indicators of economic growth in the United States for the period 1870-1950, with the state as the spatial unit and the decade or multiples of decades as the temporal units. And, in 1957, our colloborators, Lee, Miller, Brainerd, and Easterlin published series of reference tables, based mainly on estimates and refinements of census data, accompanied with detailed discussions of the procedures by which these estimates were brought into reasonable conformity with our conceptual framework.[1]

We conceived of economic growth and population redistribution as linked by a continuous chain of interdependent variables. On the one hand, the growth of population that accompanies economic growth might in itself stimulate migration from more densely to more thinly settled areas and, to the extent that the movement resulted in discovery and opening up of natural resources valuable to the settled area would provide an attraction to further migration. More important, in recent times, would be the effect of differential technological progress upon the distribution of economic opportunities through structural changes involved in industrialization and urbanization, which have proceeded so rapidly and so specifically that the vital processes of birth and death could play but a minor role in adjusting the distribution of population to economic opportunities in different parts of the country. In consequence, migration was conceived as the main mechanism by which adjustment to differential economic opportunities could be maximized. On the other hand, if migration were insufficient or proceeded at too slow a pace from areas of lesser toward those of greater economic opportunities, differential population growth might prove to be an impediment to economic growth. From this standpoint, also, migration becomes the main mechanism of adjustment in offsetting economic disparities.[2]

The present paper collates two of our series, as a first approximation to an approach that we hope to develop in detail in later publications. Easterlin's estimates[3] of service income per worker, along with Leven's and Department of Commerce estimates are taken as indicators of the level of economic opportunity in each state as of three dates, namely 1900, 1920, and 1950, and from these estimates its position relative to every other state was determined crudely in six categories: positive differentials of 20 percent or more, 10-19 percent, and less than 10 percent; negative differentials of less than 10 percent, 10-19 percent, 20 percent or more.

Massachusetts, for example, had an average service income per worker (in 1929 prices) of $1162 in 1900; Kentucky, an average of $608; Minnesota, of $983; Nebraska, of $1076; Washington, of $1236; California, of $1361; and Montana, of $1630. Massachusetts was, therefore, classified as having a positive income differential of 20 percent or more in relation to Kentucky, of 10-19 percent in relation to Minnesota; of less than 10 percent in relation to Nebraska; and negative differentials of less than 10 percent, 10-19 percent, and 20 percent or more in relation to Washington, California, and Montana, respectively. A similar procedure was used for classifying the income differential of every state in relation to every other state in 1900; and again, for the two succeeding dates, 1920 and 1950.

The only migration series that could be utilized historically to determine migration streams, that is, the movement from specific states of origin to specific states of destination is that derived from data on state of birth of the residents of each state. Series for native whites and for native nonwhites, assembled by Lee from census sources,[4] were therefore utilized as follows:

The change in the number of persons born in each state was computed for residents of each state from 1880 to 1900, from 1900 to 1920, and from 1920 to 1950. For the states cited above, in respect to Massachusetts, for example, the following computation was made for native whites living in Massachusetts in 1880 and 1900:

Living in Massachusetts

Born in:	1880	1900	Change 1800 to 1900
Kentucky	465	884	419
Minnesota	309	1474	1165
Nebraska	57	494	437
Washington	11	161	150
California	742	1810	1068
Montana	5	139	134

When, as in the above example, positive migration balances emerged,[5] they were allocated to the appropriate income-differential class, that is 419 to the category in which the destination income was 20 or more percent higher than that of the origin 1165 to the category in which the destination income was between 10 and 19 percent higher than that of the origin, 437 to the category in which the destination income was less than 10 percent greater than that of the origin; and 150, 1068, and 134 to the appropriate negative income-differential categories.

Summation of positive migration balances for every state by color of migrants and contiguity of states yielded the following distribution in terms of thousands of migrants:

NATIVE WHITES	1880-1900	1900-1920	1920-1950
All states	4,264	6,927	14,046
Contiguous states	2,113	3,207	4,788
Noncontiguous states	2,151	3,720	9,257
NATIVE NONWHITES			
All states		859	2,533
Contiguous states		340	456
Noncontiguous states		518	2,077

In Table I are shown the percentage distributions in income-differential classes of these migrations, by color of migrants and contiguity of states, for three time intervals for native whites and two[6] for native nonwhites; and these categories are graphed on the accompanying chart.

The chart shows clearly that the cumulative migration of native whites has, in general, proceeded from states of lower to states of higher average service income per worker, with such "economically-oriented" migration accounting for roughly 60-70 percent of the total. An upward trend from the late decades of the nineteenth century to the more recent period is apparent, the values for the three periods being, respectively, 58 percent, 64 percent, and 69 percent. For each period there is an observable tendency for longer-distance migrations (between noncontiguous states) to be more economically-oriented than for those from shorter distances (between contiguous states), with however a downward trend in the margin between the two. Thus the economically-oriented migrations between noncontiguous states exceeded those between contiguous states by 16.7 percentage points in the first period, but only by 5.7 percentage points in the second period, and 5.6 percentage points in the last period. The chart shows equally clearly that the proportions of economically-oriented migrations among native nonwhites greatly and significantly exceeded those among native whites for both time periods for which computations are available, amounting to no less than 90 percent in 1900-1920, contrasted with 64 percent for whites; and 94 percent in 1920-1950, again contrasted with 69 percent for whites. For nonwhites, as for whites, there is an increase over time and a distance differential, favoring the noncontiguous states.

In view of the fact that we are dealing with average state incomes and with aggregate measures of migration, the strength of the observed relationship is impressive. To test its persistence, a crude sort of variance analysis is carried through by successive subtractions of migration streams between specific origins and destinations where forces other than service-income differentials are known, or believed to have been important determinants of internal migration.

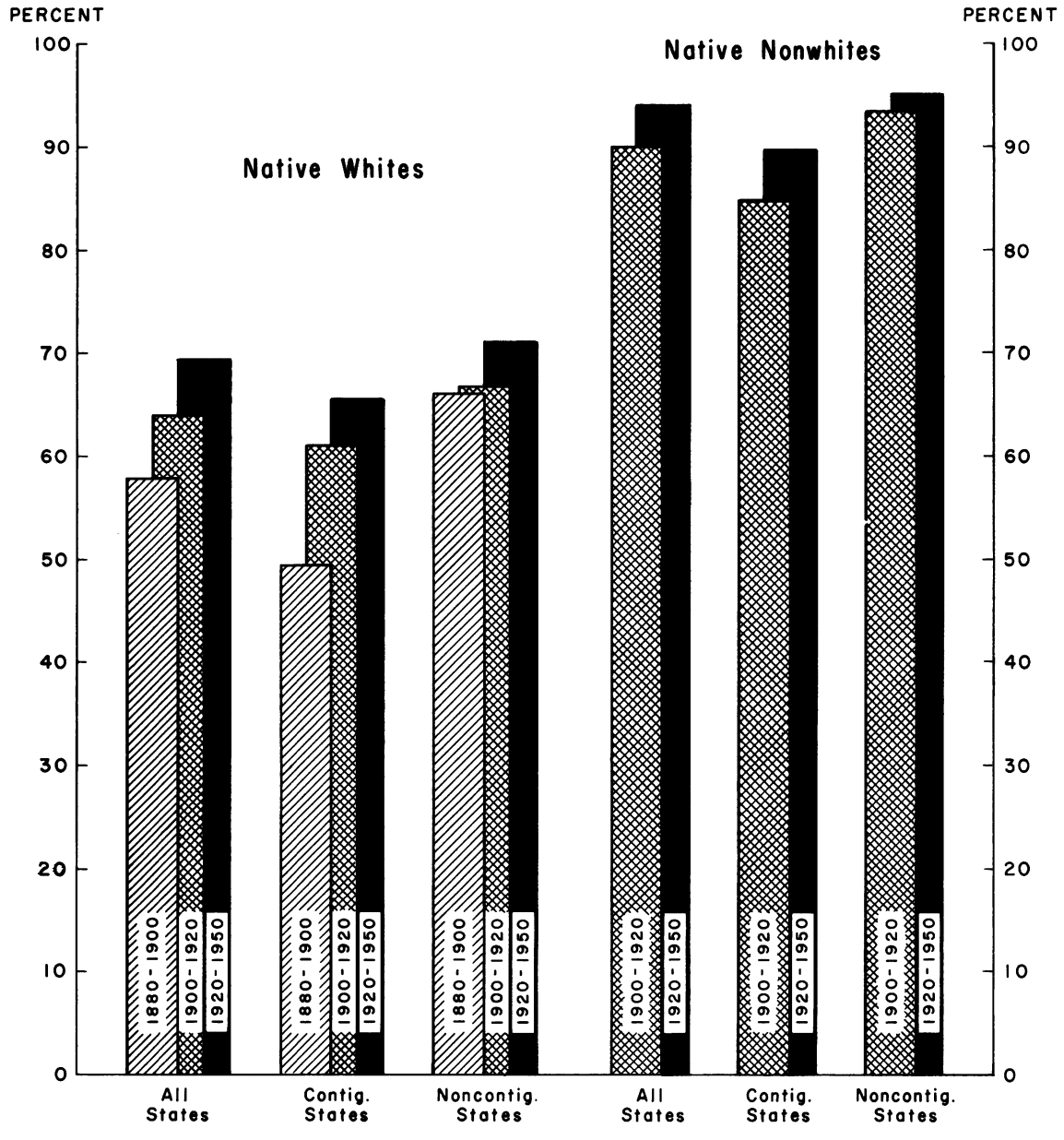
Examining first the specific income-differential classes for native whites in the first period, it will be noted that although for "all states" each positive income-differential class exceeds the corresponding negative class, there is a surprising piling up of migration to states with a markedly unfavorable differential, that is -20 or more, and that, for contiguous states, this negative class is appreciably larger than the corresponding positive class. That this distortion was caused largely by migration to newly opened states, is suggested by the second line in the 1880-1900 panel where migration gains of Oklahoma, the Dakotas, and Arkansas are excluded. The "opportunities" sought by migrants to these states were at this time probably represented less by the level of current state service income per worker than by anticipations of higher income in the course of the state's economic development that followed the exploitation of new resources.

By the turn of the century, apparently "non-economic" migration to new lands had played out, and other distorting factors were operating. Among these has been movement to Florida of persons attracted by other than purely economic conditions, at least to the extent that these are measured by service income-per-worker differentials. Among these also is the expansion of metropolitan areas across boundaries into states having lower levels of service income per worker, with the result that migrants may not be allocated properly to the state where they actually receive their income. This misallocation has been especially important in recent years in connection with Maryland and Virginia migrants, many of whom work in the District of Columbia;[7] and it is probably also a factor in the migration of the New York born to New Jersey and Connecticut. To remove the distortion caused by primarily climatic and suburban migration, migration to Florida, Maryland and Virginia from all states and the net migration to New Jersey and Connecticut of the New York born was subtracted from corresponding totals and percentage distributions recomputed, as shown in the second lines on the panels for 1900-1920 and 1920-1950. These exclusions tend to raise the proportions of "economically-oriented" migrations, in general, and among noncontiguous states, but, contrary to expectation, have little or no effect on the birthplace-residence movement among contiguous units. The net effect of the exclusion of migration to the new-resource states during the first period was to increase the proportion of economically-oriented migration to all other states by 5 percentage points on the average, 4 for contiguous and 6 for noncontiguous states. During the second period, the exclusion of essentially suburban migration, along with that to Florida, had an average effect only about half as great, that is, by about 2 percentage points for contiguous and 3 for noncontiguous states; and in the last period these exclusions had no observable effect on economically-oriented migration to contiguous states after the adjustment was made, but a very marked effect on migration to noncontiguous states where a seven-point increase in economically-oriented migration raised the overall proportion to 78 percent.

The question now arises as to possible biases in the opposite direction with especial reference to the dominant position of California which, unlike Florida, ranked very high in the service-income-per-worker series at all three periods,[8] but like Florida undoubtedly attracted migrants for climatic and other hedonistic

MIGRATION GAINS OF STATES WITH HIGHER SERVICE INCOME PER WORKER THAN STATES OF BIRTH OF RESIDENTS

as Percentage of Total Interstate Birth-Residence Gains by Contiguity and Color
United States: 1880-1900, 1900-1920, 1920-1950



anb

reasons. In fact, California's share in all interstate migration gains[9] of native whites rose spectacularly during the period under consideration from less than 5% in 1880-1900, to 13% in 1900-1920, to 26% in 1920-1950; and in gains among noncontiguous states from 8% at the earliest period to 23% in the middle years to 38% during the last three decades. Without implying that the extensive movement to this state is predominently noneconomically determined, it seems appropriate to evaluate migration among other states when this dominant influence is removed. The resulting calculation, shown in the third line of each panel has no appreciable effect in the first period but reduces the proportions of economically oriented migration by 4 percentage points in the second and by about 9 in the last period. Correspondingly, economically-oriented migration among noncontiguous states was reduced by 8 percentage points in 1900-1920 and by 14 in 1920-1950. All vestiges of a distance differential disappeared, and the slope of the upward trend during the twentieth century was greatly diminished.

Adjustments of the nonwhite series for migration gains of California as well as of Florida, Maryland, and Virginia are shown, on the second page of Table I, to be comparatively small and to have no marked effect upon either the level or the trend or the distance differential in the very high overall proportions of economically-oriented migration.

Implicit in the preceding discussion is the assumption that state-of-birth data adequately represent streams of migration from specific origins to specific destinations during specified periods of time. To the extent that interstate migration has proceeded in stages rather than directly from state of birth at the initial to state of residence at the terminal year of each time interval, this assumption is patently untenable. It is hoped, however, that further light will be thrown on this process by analyses that are now under way of net intercensal interstate gains and losses in comparison with the present series.

Footnotes

1. Everett S. Lee, Ann Ratner Miller, Carol P. Brainerd and Richard A. Easterlin. Population Redistribution and Economic Growth, 1870-1950. Volume I. Methodological Considerations and Reference Tables, (directed by Simon Kuznets and Dorothy S. Thomas) American Philosophical Society, Philadelphia, 1957. pp. XIX +759

2. Simon Kuznets and Dorothy S. Thomas. "Internal migration and Economic Growth" in Proceedings of the 1957 Annual Conference of the Milbank Memorial Fund, New York, 1958. Pt.III. pp 196-211.

   Dorothy Swaine Thomas. "Age and Economic Differentials in Interstate Migration" Population Index, Princeton, October, 1958. pp. 313-325

3. Easterlin's series on "Service income per Worker" are shown, in current prices, on p. 754 of the work cited in Footnote 1. They have since been expressed in terms of constant (1929) prices for use in the forthcoming Volume II of Population Redistribution and Economic Growth, and the figures quoted in the present paper are from the "constant prices" series. His definition of service income is "the sum of wages and salaries (excluding employee contributions to social insurance and 'other labor income' such as cash sickness compensation, etc.) and proprietors' income with imputed rents of farm dwellings included in the agricultural component of service income." Ibid p. 703

4. Everett S. Lee, State of Birth of the Native Population, 1870-1950, three dittographed volumes. University of Pennsylvania, Philadelphia, 1953, pp 687

5. Negative balances are not susceptible to analysis of migration streams, inasmuch as the destination of the migrants cannot be determined.

6. Because of small numbers and questionable reliability of the data, computations were not carried through for nonwhites for 1880-1900.

7. Easterlin, op. cit., omitted the District of Columbia from his state-income estimates. The present paper, therefore, excludes D.C. from all calculations.

8. Easterlin's concept of service income conforms to an "income originating" rather than an "income received" basis.

9. California ranked 4th from the top among the states on service income per worker in 1900; 3rd in 1920; and 4th in 1950, whereas Florida ranked 8th from the bottom in the first subperiod; 6th from the bottom in the second; 10th from the bottom in the last subperiod.

TABLE I

Percentage Distribution of Birth-Residence Migration Gains by Color of Migrants, Contiguity of States and Differentials in State Service Income per Worker at End of each Interval: 1880-1900; 1900-1920; and 1920-1950

Migration Interval and Contiguity of States	Percent excess (+) or deficiency (-) of service income per worker in state of residence over that in state of birth								
	NATIVE WHITES								
1880-1900	+20 or more	+10 to +19	+0 to +9	-0 to -9	-10 to -19	-20 or more	All +s	All -s	TOTAL
All States	28.3	14.5	15.1	12.6	7.8	21.7	57.9	42.1	100.0
excl. Dak., Ark., Okla.	31.6	15.3	16.0	14.5	8.7	13.9	62.9	37.1	100.0
excl. also Calif.	29.4	15.2	16.4	15.1	9.1	14.8	61.0	39.0	100.0
Contiguous States	18.0	11.6	19.9	16.0	9.6	24.9	49.5	50.5	100.0
excl. Dak., Ark., Okla.	18.7	14.1	20.7	19.3	11.8	15.4	53.5	46.5	100.0
excl. also Calif.	18.9	13.8	20.7	19.2	11.9	15.5	53.4	46.6	100.0
Noncontiguous States	38.5	17.3	10.4	9.3	6.1	18.4	66.2	33.8	100.0
excl. Dak., Ark., Okla.	44.0	16.6	11.4	9.8	5.6	12.6	72.0	28.0	100.0
excl. also Calif.	40.6	16.7	11.8	10.7	6.2	14.0	69.1	30.9	100.0
1900-1920									
All States	22.2	22.2	19.7	15.1	9.3	11.5	64.1	35.9	100.0
excl. Md. Va., Fla. N.Y.met.	23.3	22.2	20.9	15.9	8.0	9.7	66.4	33.6	100.0
excl. also Calif.	18.5	21.3	22.2	17.5	9.3	11.2	62.0	38.0	100.0
Contiguous States	14.7	23.2	23.2	16.3	14.1	8.5	61.1	38.9	100.0
excl. Md., Va., Fla., N.Y. met.	15.6	22.1	24.9	17.9	11.5	8.0	62.6	37.4	100.0
excl. also Calif.	15.9	21.3	25.0	18.1	11.6	8.1	62.2	37.8	100.0
Noncontiguous states	28.7	21.3	16.8	14.0	5.2	14.0	66.8	33.2	100.0
excl. Md., Va., Fla.,	29.7	22.4	17.6	14.2	5.1	11.0	69.7	30.3	100.0
excl. also Calif.	21.3	21.2	19.1	17.0	6.7	14.7	61.6	38.4	100.0
1920-1950									
All States	16.9	30.6	21.9	13.5	6.2	10.9	69.4	30.6	100.0
excl. Md., Va., Fla., N.Y. met.	18.5	31.5	23.7	14.1	5.7	6.5	73.7	26.3	100.0
excl. also Calif.	17.6	24.8	22.0	18.2	8.1	9.3	64.4	35.6	100.0
Contiguous States	9.7	34.5	21.4	21.3	9.2	3.9	65.6	34.4	100.0
excl. Md., Va., Fla., N.Y. met.	11.2	31.2	22.9	22.3	8.4	4.0	65.3	34.7	100.0
excl. also Calif.	11.6	30.4	22.1	22.9	8.8	4.2	64.1	35.9	100.0
Noncontiguous States	20.5	28.5	22.2	9.5	4.7	14.6	71.2	28.8	100.0
excl. Md., Va., Fla.,	22.4	31.7	24.1	9.8	4.2	7.8	78.2	21.8	100.0
excl. also Calif.	22.9	19.8	21.9	14.0	7.5	13.9	64.6	35.4	100.0
	NATIVE NONWHITES								
1900-1920	+20 or more	+10 to +19	+0 to +9	-0 to -9	-10 to -19	-20 or more	All +s	All -s	TOTAL
All States	67.8	14.0	8.2	2.2	3.7	4.1	90.0	10.0	100.0
excl. Md., Va., Fla.	72.1	8.9	8.9	2.3	4.0	3.8	89.9	10.1	100.0
excl. also Calif.	71.6	8.9	9.1	2.4	4.1	3.9	89.6	10.4	100.0
Contiguous States	41.4	28.3	15.2	2.5	7.9	4.7	84.9	15.1	100.0
excl. Md., Va., Fla.	46.7	17.9	17.9	3.1	9.1	5.3	82.5	17.5	100.0
excl. also Calif	46.8	17.6	18.0	3.2	9.1	5.3	82.4	17.6	100.0
Noncontiguous States	85.2	4.7	3.5	1.9	1.0	3.7	93.4	6.6	100.0
excl. Md., Va., Fla.	86.9	3.7	3.7	1.8	1.0	2.9	94.3	5.7	100.0
excl. also Calif	86.7	3.7	3.6	1.9	1.1	3.0	94.0	6.0	100.0
1920-1950									
All States	70.0	18.0	6.0	2.6	1.2	2.2	94.0	6.0	100.0
excl. Md., Va., Fla.	74.9	13.0	6.6	2.6	1.1	1.8	94.5	5.5	100.0
excl. also Calif	77.8	9.2	6.7	2.9	1.3	2.1	93.7	6.3	100.0
Contiguous States	16.6	55.4	17.7	5.9	3.2	1.2	89.7	10.3	100.0
excl. Md., Va., Fla.	24.6	35.4	26.2	8.2	4.3	1.3	86.2	13.8	100.0
excl. also Calif	25.1	34.8	26.4	8.0	4.4	1.3	86.3	13.7	100.0
Noncontiguous States	81.8	9.8	3.5	1.8	0.7	2.4	95.1	4.9	100.0
excl. Md., Va., Fla.	82.8	9.6	3.5	1.7	0.6	1.8	95.9	4.1	100.0
excl. also Calif.	87.5	4.6	3.1	1.9	0.8	2.1	95.2	4.8	100.0

Discussion

By: Demitri B. Shimkin, George Washington University and Bureau of the Census

Professor Eason has ably summarized the salient problems facing demographic research on the Soviet Union today. In my opinion, his formulations are sound. Consequently, my discussion will seek to complement rather than to criticize his paper. It will attempt to identify and to sharpen through more detailed review a number of his key generalizations.

The first of these is the characterization of Soviet population history as a discontinuous phenomenon, immensely affected by catastrophes. These included World War I, postwar epidemics, and the Civil War; the collectivization crisis; and finally, World War II and the backbreaking recovery drive of the later 1940's.[1] The effects of the last two were most devastating; as Professor Eason has noted, the population deficit which they brought about may be as high as 40 million. By far the largest component of this deficit is among those born from 1941 through 1948. For example, a recent Soviet educational planning document presents a model as of 1953[2] in which the survivors of 1944 births are only a third as numerous as those of 1940 births, while even the survivors of 1948 births are only two-thirds as numerous. This reduction reflects both sharply lowered fertility and enormous infant mortality; thus, according to official data, 75 percent of the infants born in Leningrad during the siege and starvation year of 1942 died.[3] Military killed and died of wounds and disease during or soon after the war probably approximated 10 million. Among civilians, the Jews, infants, the aged, and those subjected to the double hazard of German and then Soviet slave-labor camps suffered most severely. In general, such phenomena as the current labor shortage and the lowered fertility of the Soviet Union's western areas, particularly the Ukraine, can best be visualized as continuing demographic effects of World War II.

Another of Professor Eason's major conclusions is that, since 1950 at least, basic changes have taken place in both the mortality and the fertility rates of the Soviet population. In all likelihood, the crude death rate is, today, less than 10 per 1,000; however, the official figure of less than 8 and, above all, the official life expectancy datum of 67 years appear dubious. My reasons are as follows: first, the completeness of registration of vital statistics in the Soviet Union is uncertain and the vital rates may reflect quite incomplete coverage or even an unrepresentative sampling. It is especially doubtful whether much accurate information on age of death is extant. Also, as collateral evidence of the coverage of Soviet vital statistics, it may be noted that, prior to 1958, the largest detailed Soviet morbidity studies had covered only Moscow (in 1926, 1937, and 1947) and nine large industrial centers (in 1947). Furthermore, the planned coverage of the 1958 studies comprised 25 million persons in 120 cities (over a quarter of the urban population), but only about 3 million persons in 84 rural *rayons* (or less than 3 percent of the rural population).[4]

Second, the sharp difference between urban and rural medical facilities (in 1956, 7.0 medical treatments per man-year, and one X-ray per 10,000 persons in cities; 1.6 medical and 1.4 "fel'dsher" treatments per man-year, and one X-ray per 37,000 persons in rural areas)[5] suggests considerably greater urban-rural mortality differences than those actually reported.[6] Recent data on morbidity are pertinent in this regard. Thus, the tuberculosis infection rate in a West Russian rural rayon in 1955 was 30 per 1,000, compared to 9.5 per 1,000 in Minsk. For circulatory diseases the corresponding rates were 80 and 47.[7]

Third, even in urban areas, the conditions for substantial rates of enteric infection and of tuberculosis are still evident. For example, the Soviet urban water supply in 1955 was less than 25 gallons per person per day.[8] Again, in 1956, only 12.7 percent of all retail stores handling perishable foods were equipped with electric refrigeration.[9] The extreme shortage of Soviet housing (a per capita space less than 30 percent of the inadequate French level)[10] and especially the customary barracks housing of juvenile workers must contribute heavily to endemic tubercular infection. Although considerable progress in controlling tuberculosis has been claimed for selected cities,[11] continued Soviet expansion in the number of tuberculosis specialists argues for a gloomier, non-publicized projection of tubercular morbidity.[12]

In the Soviet drop in birth rates, the major operative factor has been the deficiency of males of reproductive age. For example, among Ukrainian collective farmers 65 percent of the "able-bodied" workers in 1950, and 62 percent in 1955, were women.[13] In addition, I believe that family formation has been minimized and abortion maximized by the wide use of sexually segregated barracks housing.[14] A variety of local reports suggest, in the aggregate, that a fifth to a sixth of the Soviet urban population lives in such housing, which is commonest in the Urals and Siberia. Many students, young workers drafted from the countryside, and, in very congested areas, in-migrant families, are put up in such quarters.[15] Finally, the available data indicate little change in the size of the *conjugal* family either in the city or, since 1938, in the country.[16] In short, I would regard the Soviet decline in fertility, to date, as an exteriorly caused, quite reversible, phenomenon, rather than the result of a fundamental change in values.

Another aspect of the decline in Soviet fertility should be stressed. As James Brackett has reported in an important paper,[17] regional and ethnic differences in birth rates are very great. In northwest and western Russia, the rate of population increase is about one percent annually, in contrast to about 1.5 percent in Central Russia and Siberia, and to 2.5 to 3.0 percent in Kazakhstan, Central Asia, and the Caucasus. The future political implications of this trend are indeed great. They are acerbated by the limited economic development, the strongly persistent cultural patterns, the Russian-native economic stratifi-

cation, and the absence of significant out-migration which characterize Central Asia and the Caucasus today.

I concur heartily with another conclusion, namely, that population pressure should not be a problem for the Soviet Union in the foreseeable future. This point needs elucidation in view of the slow progress of Soviet agriculture, even considering the marked advances realized since the great crisis of 1953. Briefly put, the limitations on Soviet agriculture reflect bad management, underinvestment and inadequate incentives far more than deficiencies in natural endowment. Reasonable prudence and competence among Soviet leaders are the essential requisites for an adequate if not varied diet for the Soviet population, even if agriculture remains a source rather than a beneficiary of transfers within the economy.

This point also needs qualification. Over-all adequacy does not exclude hardship for considerable segments of the population at lower income levels. In this regard, the belated advent of minimum-wage laws and the continuing absence of social-security coverage for the collective farmers, who constitute two-fifths of the Soviet labor force, are noteworthy.

A fourth generalization made by Professor Eason infers that, historically, Soviet economic growth has been fundamentally dependent upon vigorous expansion of the labor input, via maintenance of high labor force participation rates and the reduction of overt or disguised unemployment. To this I would add the large-scale transfer of labor resources to priority sectors, both through migration and through intensified demands by the State. Illustrative of such intensification has been a 70 percent rise, since 1937, in the labor days worked by each able-bodied collective farmer. In the former year, the State demanded 194 labor days,[18] or somewhat fewer man-days a year, a level which permitted, concurrently, active maintenance of private agricultural and handicraft activity. Today, the private sector is genuinely marginal.

I must also note that, if the studies made by colleagues and me at the Bureau of the Census are right, Soviet productivity gains even in the all-important area of industry have been rather modest. In 1950, man-year productivity among industrial wage workers was some 60 percent higher than in 1928. Between 1950 and 1956, it increased over 40 percent, a rapid rate; in 1957 and the first half of 1958, however, the annual growth of productivity was under 3 percent annually, a fact partly ascribable to a shortened work week. In general, therefore, labor input appears to be the critical variable in Soviet economic development.[19]

However, present policies, while influenced by demographic factors, are not exclusively determined by them. In my opinion, the changes in education and such paradoxes as work-week reductions in a time of labor shortage reflect Khrushchev's desire to gain the support of the urban worker. Khrushchev may be, in a sense, a political champion of the Soviet masses against the managerial classes. This image is especially needed in view of the continuing pressures to accelerate urbanization by reducing wage-work possibilities in the countryside.

To summarize, the analysis of population movements, in the Soviet Union at least, depends upon close scrutiny of a variety of related socio-economic and political phenomena. Here especially, demography must be considered as a socio-biological interaction. I hope that Professor Eason's paper and my discussion have aided in the formulation of this basic point.

Footnotes

[1] A recent quantitative survey of Russian population history is J. N. Biraben, "Essay on the Demographic Evolution of the U.S.S.R.," Population, vol. 13, no. 2b, pp. 29-62, 1958.

[2] V.M. Dmitriyev, et al, Organizatsiya i planirovaniya vseobshchego desyatiletnego obrazovaniya (Organization and Planning of Universal Ten-Year Education), Izd. Akad. Pedagog. Nauk RSFSR, Moscow, 1955, esp. pp. 62-67.

[3] S. A. Novosel'skiy, "Life Expectancy and Mortality of Infants 0 - 1 in Leningrad in the Period After the Blockade," Gigiyenia i sanitariya, vol. II, no. 7/8, pp. 1-5, 1946.

[4] L. A. Brushlinskaya and M. M. Mazur, "Let us Use the All-Union Population Census for Studying Morbidity," Sovetskoye zdravookhraneniye, 1958, no. 4, pp. 29-34.

[5] Data calculated from Ts.S.U. pri Sov. Min. SSSR, Narodnoye khozyaystvo SSSR v 1956 godu. (National Economy of the U.S.S.R. in 1956)(Moscow: Gosstatizdat, 1957), p. 25; and Min. Zdrav. SSSR, Zdravookhraneniye v SSSR (Public Health in the U.S.S.R.) (Moscow: Gos. Izd. Med. Lit., 1957), pp. 109-114. As Dr. M. B. Shimkin has noted in a private communication, the general use of B.C.G. in the U.S.S.R. renders patch tests for tuberculosis ineffective. Hence, X-rays are vital for this purpose.

[6] Officially reported crude death rates for 1956 are Kuybyshev city 8.9, remainder of oblast' 9.2; Novosibirsk city 11.3, remainder of oblast' 12.0; Leningrad city 6.7, remainder of oblast' 7.8 (compiled by Michael Roof).

[7] V. I. Berlin, "An Attempt at Studying the General Morbidity of the Population of a Rural Rayon," Sovetskoye zdravookhraneniye, 1958, no. 4, pp. 49-53; N.N. Govor, "The Morbidity of the Population of Minsk in 1955 and 1956," ibid., 1958, no. 7, pp. 23-29.

[8] Calculated from A. Kogan, "Water Supply and Sewer Systems in the R.S.F.S.R.," Zhilishchno-kommunal'noye khozyaystvo, 1956, no. 2, pp. 9-10.

[9] Sovetskaya torgovlya, 1957, no. 11, p. 19.

[10] G. Grigoroff, "Population and Housing in the U.S.S.R.," Population, vol. 13, no. 2b, pp. 77-88, 1958, esp. p. 78.

[11] N. N. Dolge, "Dynamics of Tuberculosis Morbidity and Mortality in R.S.F.S.R. Cities," Sovetskoye zdravookhraneniye, 1958, no. 4, pp. 45-49.

[12] Min. zdrav. SSSR, 1957, op.cit., p. 58.

[13] Calculated from average male and female labor-day input figures in Ts.U. pri Radi Min. S.R.S.R., Narodne gospodarstvo ukraynskoy RSR (National Economy of the Ukrainian S.S.R.) (Kiyv: Der. Stat. Vid., 1957), p. 293.

[14] Another factor promoting abortions is the inadequate protection of unmarried mothers and deserted wives in Soviet law, cf. V. Gsovski, Soviet Civil Law, 2 vols. (Ann Arbor: University

of Michigan, 1948), esp. I:111-136; also New York Times, October 7, 1956, and M. G. Field, "The Relegalization of Abortion in Soviet Russia," New England Journal of Medicine, 255:421-7, 1956.

[15] The basic regulations on barracks are given in I. Kozlov, et al, Sbornik postanovleniy i instruktsiyi po zhilishchno-bytovym voprosam (Collection of Decrees and Instructions Concerning Questions of Housing Economics) 2nd ed., (Moscow: Profizdat, 1952), pp. 108-124; on alcoholism and allied problems of Soviet dormitory life, see M.G. Field, "Drink and Delinquency in the U.S.S.R.," Problems of Communism, vol. IV, no. 3, pp. 29-38, 1955.

[16] In 1926, the average size of the urban family was 3.9 persons (4.1 persons for those headed by males; 3.2 persons for those headed by females), cf. Gosplan SSSR, Vsesoyuznaya perepis' naseleniya 1926 g. (All-Union Census of Population, 1926), 55: 2-3, 26-47, 52-53, Moscow, 1931. For 1956, Ioffe mentions "about 20 million" urban families, in a total urban population of 87 million (Planovoye khozyaystvo, 1956, no. 6, p. 55). If 15 to 20 percent are deducted as an allowance for single persons in barracks, the average size per family may be estimated at 3.5-3.6. Such a decline from 1926 is ascribable largely to the increase in female-headed families. This estimate has been confirmed by the Soviet sample census of 1957 which covered 1,021,000 persons, 87 percent urban, cf. P. Pod"yachikh, "On Developments in a Sample Census of Population," Vestnik statistiki, 1958, no. 1, pp. 25-42, esp. p. 34 which indicates an average of 4 per family.

For rural areas, there has been a considerable decline in household (dvor) size. In 1927, a sample census of 615,370 households established an average size of 5.1 persons, including 2.45 "able-bodied", i.e., aged 16 - 60 for males, 16 - 55 for females. See Ts.S.U. SSSR, Statisticheskiy spravochnik SSSR za 1928 (Statistical Handbook for 1928), pp. 92-145, Moscow, 1929. In 1938, Soviet collective farm households averaged 4.4 persons, including 1.9 "able-bodied," according to Laptev, as quoted by W. Eason, The Agricultural Labor Force and Population of the U.S.S.R., 1926-1941 (Rand Research Memorandum 1248, 4 May 1954), p. 78. For 1956 an average collective-farm household size of 4.0 persons, including 1.9 "able-bodied," can be calculated from statistics published in Ts.S.U., 1957, op.cit., pp. 17, 19, 141, 203, and 218.

Interpretation of this trend must take into account not only the high postwar proportion of female-headed, smaller households, but also an increasing number of units (even separate individuals) created by the partition of extended families to maximize land allotments. See, G. S. Maslova, "Culture and Economic Life of a Collective Farm in the Moscow Area," Sovetskaya etnografiya, 1951, no. 1, pp. 39-62, esp. p. 54; L.N. Terentyeva, "On the Way to a Prosperous and Cultural Life (in Latvia)," ibid., 1951, no. 2, pp. 85-104, esp. pp. 91, 104.

[17] J.W. Brackett, "Differential Growth of Nationality Groups in the Soviet Union," (given at the American Anthropological Association Meeting, Washington, D. C., November 20, 1958; available from the Foreign Manpower Research Office, U. S. Bureau of the Census).

[18] N. Jasny, The Socialized Agriculture of the U.S.S.R. (Stanford Univ. Press, 1949), page 411; Ts.S.U. pri. Sov. Min. SSSR, Dostizheniya sovetskoy vlasti za 40 let v tsifrakh (Accomplishments of the Soviet Regime Over 40 Years in Figures) (Moscow: Gosstatizdat, 1957), p. 165.

[19] D. B. Shimkin and F. A. Leedy, "Soviet Industrial Growth," Automotive Industries, vol. 118, no. 1, pp. 48-59, 122.

IX

LABOR, STATISTICS, AND THE COST OF LIVING

Chairman, Theodore W. Schultz, University of Chicago

Factors in the Rising Cost of Living—Aryness Joy Wickens, U. S. Department of Labor

Is Labor the Culprit?—Peter Henle, AFL-CIO

Is This a New Type Inflation?—George P. Hitchings, Ford Motor Company

FACTORS IN THE RISING COST OF LIVING

By: Mrs. Aryness Joy Wickens, U. S. Department of Labor *

So much has already been said this year on so many platforms on the subject of rising prices that I hardly know where to begin. Certainly, the data now available offer little opportunity to present ideas which are both novel and profound. However, as one who has had some responsibility for the establishment of some of these statistics and who knows something of their limitations and shortcomings, I may be able to call your attention to a few fundamental points which are often overlooked in general discussions and to point out a few pitfalls in analyzing trends for the use of the public and of those policy makers whose decisions can contribute to the amelioration of the problem of rising prices.

The most commonly used measure of domestic inflation is prices paid by consumers, for the standard of living of the entire population is affected by the rise or decline of prices paid at retail. Consumer prices are at the end of the whole economic chain. They embody the effects of many preceding costs, and many prices: the prices of raw materials, costs of fabricating and packaging, the using up of capital goods, charges for transportation, the cost of wholesale and retail distribution, excise taxes, and the like. Most economic forces sooner or later impinge upon consumer prices in some way, and analysis of the various waves of price increases in recent years will in itself provide part of the answer to the question assigned to me to answer--the factors affecting rising costs.

Consumer prices today are a little more than double what they were before the war broke out in Europe in 1939. The American consumer's dollar is worth about 48 cents in prewar terms. Interestingly--but not surprisingly--other prices and values have risen in much the same range. Wholesale prices are up by about 130 percent, with farm prices up about 150 percent and industrial goods up 115 percent from 1939. Real estate prices, to select only one more of a large variety of prices of equities, seem to have gone up even more--urban residential by about 200 percent, and farm land values by about 220 percent.

Let me remind you at this juncture that conclusions are too often affected by the time period selected; you get different results if you choose 1929, 1939, 1947, or 1956, as your starting point. Another common problem, which affects the conclusions arrived at in the vast amount of gratis literature now circulating on the subject of prices, is the use of rates of change which do not take account of the relative importance of the particular commodity or area of the economy rather than the use of aggregates, of percentage changes rather than of points in the increase in a total index. No matter how you look at it, however, and whatever time period you use, every kind of desirable goods or service has a much larger price tag today than most of us thought likely not too long ago. Thus it can be said that much of the adjustment to higher prices, induced by war and postwar inflation, is now thoroughly imbedded in the cost, value and price structure of the economy.

All of the countries of the free world have had to wrestle with the problem of inflation during and since the war. In Western European countries, retail prices have advanced more rapidly, both in the immediate postwar period and since 1953, than in the United States, notwithstanding their exercise of notable public and private restraints upon the factors affecting prices. The principal reason for this difference is that the United States, undamaged by the ravages of war, did not have to divert from consumption to reconstruction as large a proportion of its postwar industrial and agricultural output as did the European countries.

That the inflation both here and abroad is the offspring of war is now so commonly recognized that I need not labor the point. Prices have risen with virtually every major war in modern history. The shortages of supply and excess of demand created by the diversion and destruction of resources--natural, industrial and human--and by the pyramiding of government debt and the expansion of the money supply have inevitably affected prices and costs. The last two wars--World War II and the Korean episode--have been no exceptions. Increased prices have been attended by higher charges, higher wages, higher profits and higher money values for virtually all forms of equities.

In fact, it is surprising that we controlled prices in the United States as well as we did during World War II. About a third of the rise of 109 percent in consumer prices from 1939 to the present occurred before the end of the war. Of this a substantial part occurred before the United States actually declared war, and in the year before price controls were made effective, particularly controls of food prices. For a considerable period during the war, prices were held down with the aid of price controls, rationing, rent controls, and other forms of controls--at both the producers' goods and the consumers' goods levels.

In the immediate postwar years, however, we must admit that we made some mistakes as rationing controls were removed in 1945 and as price controls were removed in 1946, when supplies of goods were not yet adequate to meet the pent-up demands backed by the vast amounts of savings and the vast credit resources accumulated during the war. This proved to be the case. Prices spiraled upward and rose even more than they had done during the war years, to reach a peak in 1948. This second wave accounted for another third of the rise, so that by 1948 prices had risen two-thirds of the way to the height which they have now attained. Following this rapid spurt in prices, there was a mild

* With Hyman L. Lewis, U. S. Department of Labor

decline in 1948-49, coincident with a recession in the level of business activity. Recovery, however, was already underway before Korea, and the new demands for goods and services for defense, coupled with sustained private demand, drove prices still higher. By 1951 this third wave pushed consumer prices up to 87 percent above 1939.

These first three waves of price increases since the outbreak of World War II are, thus, clearly attributable to classic fiscal and monetary factors--a combination of war inflation and postwar demands, with the absence of strong controls in the postwar period. But the worrisome point is not that prices went up in 1946-48 nor that they went up further in 1950-51, but that they have not returned at least part of the way to prewar levels, as was true in most earlier postwar periods. We count the number of years after the cessation of hostilities, and we ask, where is the typical postwar deflation?

An obvious explanation, so obvious that we sometimes take it for granted and leave it out of our considerations, is that today we do not have a typical postwar letdown in armament outlays. True, there is no active shooting; but there is a vast amount of international tension, and a large proportion of government expenditures in all countries is going toward armaments and various defense activities. These are inflationary, in a different push-pull sense than is commonly used for those two badly overworked verbs. They pull labor out of activities which produce goods and services for immediate civilian consumption; they push up the prices of many durable goods, from raw materials through to finished products; and they create government debt and civilian purchasing power without in themselves adding to the supply of goods and services available for distribution. Yet, while they may possibly account for the failure of prices to dip, neither defense outlays nor new warlike developments can explain our fourth wave of price increases, the one which began in 1956, the one I describe as a creeping inflation. In this most recent period there has been a rise of 8 percent, culminating in a peak in the summer of 1958. Since that time the over-all consumer price index has remained comparatively stable.

Such a rise is not entirely unusual. There have been other non-war periods in the past when prices have risen. But coming after a four-year period when over-all values held virtually steady instead of suffering at least a moderate postwar decline, this rise has contributed to an economic climate in which many people--both in business and in labor circles--have been led to assume that prices will, if anything, go up and up rather than level off or go down. The new viewpoint on prices has brought into sharp focus a number of differences in the way in which our economy behaves today as compared to earlier years.

Let us examine some of these differences. One is that there has been a great burgeoning of demand, partly because of a rapid increase in the population, but partly also because of sharply advancing standards of living and a redefinition of what is a necessity and what a luxury, especially in the U.S. Adding to the demand has been a great social leveling process, designed to improve the position of the lower income groups, through tax policies, social security, and other Government devices. There has come to be a general recognition, in countries where men are considered as individuals, of the right to improve their levels of living. There has been the growth and acceptance of strong organizations in both business and in labor, and of new procedures by which both government and organized economic groups have combined to slow down what were formerly considered natural economic forces. There has been an acceptance of the idea of softening economic ups and downs. These procedures in the United States range from Government floors under wages and prices in order to assure incomes for certain groups in the population, to mass collective bargaining. Similar developments have gone on throughout the free world.

The emphasis on human values as the major criterion of economic policy has had an important manifestation in the "full-employment" orientation of this and other western countries, under which government buttresses a declining economy and thereby so modifies the relationships of supply and demand as to temper downtrends in prices.

Another institutional development, which was in evidence before World War I but which has been increasingly important, is the gradually increased control by a growing number of producers over the prices of their products. Farm products and other raw materials are steadily declining in importance in relation to fabricated goods. Their prices are normally more volatile because their supply often cannot be immediately adjusted to demand. But in our economy today other types of goods and services--produced by a much more complicated process, involving a higher proportion of labor cost and yet more easily controlled with reference to supply and therefore prices--have steadily increased.

Thus, for a variety of reasons, some old and many new, we now have an economic system in which price advances are relatively easily facilitated, but price declines are braked. I think we do not fully understand these forces, nor can we yet appraise their effects upon the economic "laws" on which we were all brought up.

I am not at all sure, however, that these basic factors have had as much influence upon the economic thinking of the man in the street--the businessman or the employee--as a more noticeable development, the apparent contradiction of prices continuing to rise even while business was shrinking. This is the kind of development every housewife is aware of. She doesn't need a statistician to tell her. Many good people have made the assumption that prices go down when business declines. Some do, of course, but even when they do decline, there is always a lag in prices of finished goods. Consumer prices are generally the slowest of all to reflect the turns of the business cycle, largely because there are so many built-in and virtually fixed costs. Increasingly, certain list prices do not change at all, although

quality may improve, or sale prices may be "shaded".

It is worthwhile to compare the movement of prices in the four most recent recessions. From September 1937 to September 1938, the total consumer price index declined by 3.5 percent, entirely because of food prices; there were substantial increases in rents and new car prices and small increases in nondurable goods, public transportation and miscellaneous services, but these were offset by a good-sized drop in apparel and textile housefurnishings and small decline in fuels and miscellaneous durable goods.

Similarly, from September 1948 to September 1949, the consumer price index fell by 2.6 percent, with food alone accounting for a drop of 2.2 percent in the total index. As in the 1937-38 recession, there was a sizable drop in apparel and textile housefurnishings, together with small ups and downs in a variety of the other segments in the index, but a sizable rise in rent.

Coming to the 1953-54 downturn, we find a very small net drop in the consumer price index, only 0.4 percent, mainly because food prices also fell only very slightly; changes elsewhere all were small and offsetting. The last recession was the only one in which over-all consumer prices rose--by a substantial 2.1 percent from September 1957 to September 1958. It is also the only recent recession in which food prices rose. The food situation, you will remember, was affected by severe weather damage in southern growing areas and small marketings of cattle and hogs.

Food prices, thus, took up the slack in most earlier recessions. They did not this time, partly because of the accident of weather. But we must remember that agricultural prices--by action of the government as well as by long-run economic prices--are becoming not only less flexible but also less significant in the grand total. We cannot therefore look forward to agricultural prices providing the only flexible price element in every future business cycle. If this is our only source of price decline--it is not enough.

Another important note which emerges from the comparison of price trends during recessions is that in the most recent downturn there were no commodities which affected the total index on the downside by as much as one-tenth of one percent; in fact, only the apparel and textile housefurnishings dipped at all, and all other groups showed measurable increases. Thus, while the price rise of the 1957-58 recession is not as much cause for anxiety as has been commonly believed, it nevertheless does point to the conclusion that many prices are less flexible on the downside than they had been in earlier periods of declining demand. We must grant, of course, that consumer buying power was well sustained in this period.

As I said in the beginning, consumer prices are the most commonly used and most convenient statistical series for use in discussing inflation. Many economists have long been accustomed to say that they are a symptom rather than a cause. This is no longer entirely true, however, partly because of formal wage escalation through contracts, and informal escalation through common acceptance of the thesis that everyone--the barber, the nurse, and even the social security recipient--is entitled to the protection of his purchasing power against price changes, and to the maintenance of his relative standing in the income scale. All told, there are now subject to cost-of-living adjustments at least 4 million workers under union agreements and more than 300,000 unorganized workers. Actually, however, wage increases in various major industries have tended to keep pace with each other, whether or not there was a contractual escalation.

Services constitute another instance in which wage and price trends are closely linked. Prices of services have either held steady or risen since 1935, with the exception of only two brief periods. During the past recession, they accounted for an increase of nearly one percent in the consumer price index. The rising cost of services is often considered to be a direct reflection of labor costs or professional fees, on the assumption that there is little room for productivity improvement. This is not entirely true, however, as there do appear to have been substantial gains in some service areas, such as dry cleaning. Actually, the cost of services (less rent) has risen less in the past two decades than has the rest of the CPI; charges for certain services are still "catching up".

Many other types of costs have also increased--transportation for example, and mail, and all the regulated utilities; and many of these costs show up as services in the consumer index and do not appear at all in their own name in the wholesale indexes. Depreciation and other forms of capital consumption is one of these latter. These costs have gone up sharply because of higher original costs and because of more rapid write-offs.

Wages, which obviously are also a price, have also increased persistently and pervasively over the past two decades, and faster than commodity prices. Factory wage rates--adjusted as best we now can for overtime and changes in industrial composition--have gone up more than 200 percent since 1939, and this does not include the large variety of other labor costs which we characterize as fringe benefits. Some of the increase is associated with rising productivity, some of it with adjustments to rising costs of living. Much of it, unfortunately, has been only a paper raise, in view of the declining value of the dollar.

Actually, the upward course of wages has been so intertwined with changes in other prices and in demand that I do not think it is possible with our present data to get a general, economy-wide conclusion about the effect of wages on final prices. In some years, such as those immediately after the war when demand and productivity were rising strongly, prices rose more rapidly than wages, including fringe benefits. In more recent years, with productivity gains

small and demand tapering off while employee compensation was rising, the reverse has been true. Demand, institutional forces, long-term contracts, assumptions about trends in productivity--all of these have played their role. However, since wages and salaries are relatively a larger share of the GNP than are other single factors in costs, employee compensation can rise by relatively small amounts and still add more in dollars to prices than a proportionately higher rise in other factors of less importance.

Whether the most recent development--for labor costs to rise more rapidly than productivity gains--constitutes a new trend or not, I have no way of knowing. But, in looking back over the span of years, I find a very interesting fact: Despite all the new institutions and the new rigidities, the share of national income going to labor has not changed significantly, either since the 'twenties or in the past decade. When business is relatively poor, the share going to employee compensation rises, mostly because profits drop sharply; when business is improving rapidly, the share to profits increases, and the share to labor correspondingly drops. But these are merely short-run developments. There is no evidence that either labor or capital, as a group, has lost or gained at the expense of the other for very long. Instead, the data on shares of the national pie suggest strongly that there is some sort of stability in the economy which we do not yet fully comprehend. It may well be that attempts of either labor or capital to get ahead of the other--i.e., to reduce the ratio of profits or to add to capital through increasing the selling price--may work (if at all) for only a short period of time. In any event, much of the past gains have turned out to be illusory; to the extent that dollar incomes have gone up faster than physical output--to that extent have dollars lost their value. Money gains which exceed the real gain in output are soon wiped out; that is a truism which too many people have forgotten for too long a time in what history may prove is a fool's paradise.

There are, however, groups of people within these broad categories who have by no means kept up with the procession and who are generally at a disadvantage. This includes some wage-earners and especially salaried workers, whose earnings always lag; those whose savings from an earlier day are in fixed dollar assets like bonds, life insurance policies, etc.; those living on relief or social security payments, which never keep pace with fast-moving prices; and, in general, the "little man" without capital. Thus the distribution of the gains in national product within big groups in the population needs to be taken into account.

Prices, wages, profits, and productivity are not all the facets of the inflation problem, however. There are also taxes--both income and excise--which have an important influence. There is monetary policy affecting the supply of money versus the supply of goods and property. We must consider foreign aid, the debt structure, and also the role of the huge unregulated non-banking institutions. And there is, finally, public psychology which now seems to be stampeding in one direction. There are thus many Hamlets in this play. There is no one devil in the piece.

IS LABOR THE CULPRIT?

By: Peter Henle, AFL-CIO

In recent years, a new phrase, "wage inflation",has been welcomed into the vocabulary by newspaper editorial writers, private research organizations, and even some eminent economists, as well as traditionally anti-union employer groups. Not since the "open shop" days of the 1920's have the wage and collective bargaining objectives of organized labor been questioned so rigorously as they are today.

As one example, and I cite this merely as illustrative, let me quote the vice-president of a prominent auto firm who testified earlier this year before a subcommittee of the U.S. Senate Judiciary Committee. He concluded his 68-page statement on inflation with the following sentence: "I agree with you, Mr. Chairman, that inflation is domestic Public Enemy No. One and I would call it what it is, Wage Inflation."

This, I need hardly state, is not my point of view. I do not feel that inflation is "domestic Public Enemy No.1" which must be solved before we can beat the Russians, reach the moon, or even put across a new breakfast cereal.

Nor would I describe whatever inflation there may be by the term "Wage Inflation" or for that matter, by the term "Profit Inflation." In fact, my preference would be to scrap the word "Inflation" or even "inflation" to characterize the price movements that have been taking place in the American economy.

Admittedly, this is a dissenting opinion. The voices of public opinion makers appear to speak with ever more authority in arguing the opposite case.

Yet I suggest that the trend of prices in the postwar American economy does not justify such an alarming point of view.

Let us consider, for example, the Consumer Price Index. From June 1946 (Index 79.8, 1947-49 = 100) when price controls were abandoned to November 1958 (Index 123.7), the Index has risen 55 percent, an average of 3.6 percent compounded annually.

While this is hardly a record of price stability, it should be noted that an annual increase of 3.6 percent is a far better record than that compiled during the same period by the economies of practically all other countries. In fact, an international comparison of price changes between 1947 and 1957 shows that the annual rate at which the value of money has depreciated in the United States is lower than 21 of the 24 nations being compared. 1/

When this overall price increase is examined in more detail, it is clear that by far its largest proportion has been the result of special circumstances arising either from the aftermath of World War II or the Korean hostilities.

The postwar price increases have been concentrated in three relatively short periods of time.

a. The two years from June 1946 to June 1948

b. The one year from June 1950 to June 1951

c. The two years from March 1956 to March 1958

In these three periods covering less than half the 12-year period, the Index rose 93 percent of the entire postwar rise. For more than half the postwar period the price level has been relatively stable.

Moreover, 74 percent of the postwar price rise occurred during the first two of these periods (1946-48 and 1950-51) when specific inflationary demands arising from World War II or the Korean conflict can be held responsible for the pressure on prices. In fact, since 1951,consumer prices have risen by only 1.5 percent a year.

Nevertheless, the concern over inflation persists, cultivated assiduously by newspaper headlines, financial columnists, and Congressional committees. Much of the more recent discussions of this type have singled out union wage policies as the villain in the piece.

The most recent period of price increases ended in the spring of 1958. Perhaps because they could find no simple key to this recent upward price movement, some economists have decided that the major factor responsible has been union wage pressure. They have pointed out the absence of any major inflation-breeding influences, such as war or threat of war, which would produce higher prices. They have argued that union pressure has forced continually higher wages, that these wage increases have gone beyond gains in productivity and thus caused businessmen higher costs and forced higher prices. The natural conclusion is that "something needs to be done" to curb organized labor.

What evidence is available to judge the validity of this contention? It seems to me that such evidence could be sought in two ways: by examining price movements of individual items to try to ascertain to what extent they are union-caused and by a broader, economy-wide analysis of wages, other costs, prices, and productivity.

EXAMINING PRICE MOVEMENTS OF INDIVIDUAL ITEMS

With regard to the first method I have tried to analyze the price movement of individual items in the Consumer Price Index for the two-year period of rising prices, March 1956-March 1958. (Before and after this period, the Index was relatively stable.) If this charge against the unions is correct, it should be possible to document it by discovering that pressure for increasing prices has been most serious in those industries in which unions and union-won wage increases have played a prominent role.

The results of this analysis are included as an Appendix to this paper. Here practically every item in the Index is grouped according to its place in the economy. The normal breakdowns provided by the Bureau of Labor Statistics are not adequate for this purpose since they group under one heading a number of different products or services from diverse industries, and with diverse patterns of union organization.

For each item listed information is included for price changes 1956-57 and 1957-58 and the total increase for the two years is distributed among the various groups of items.

Such an analysis makes possible a closer look at the influence of unionism on price changes during these two years.

In a rough way, the items in the Consumer Price Index can be divided in two categories, those products and services in which unions play a prominent role in wage determinations, and those in which unions do not. Such a listing would be as follows: 2/

Relatively Unionized Section of the Economy

Newspapers

Labor Services
Amusement

Alcoholic Beverages
Tobacco Products

Metal Products
Oil, Chemical, Rubber and Pottery Products

Public Utilities

Housing

Textile Mill Products

Wood and Paper Products

Apparel

Relatively Non-unionized Section of the Economy

Professional Services
Finance and Insurance
Hospital Care
Miscellaneous Services (legal, banking, burial)

Perishable Foods
Non-perishable Foods
Food away from Home
Beverages

Government Services and Taxes

When price changes for these two groupings are compared, the results give little support to the contention that the highly unionized industries have been largely responsible for higher living costs. In fact, price increases in those industries in which a large proportion of workers are organized into unions and in which collectively bargained wage settlements receive prominent attention, are significantly lower than in those areas of the economy where unions are either weak or non-existent.

Average prices for the unionized sector increased 5.2 percent during the two-year period, while for the non-unionized sector the increase was 11.1 percent.

While the unionized sector comprises 62 percent of the total Index, it accounts for only 44 percent of the total price increase.

Surely, this comparison is a clear indication that the influence of union-won wage increases has played but a small role in the price movements during this two-year period. Obviously, many other factors have been at work on the pricing process.

A few details, perhaps, will make clear not only why unionism has not been a major factor in this price rise but also how complex the causal factors are which operate on the variety of prices which consumers normally pay.

The highest increase for any group of commodities and services is listed for perishable foods. In fact, increases in this sector account for almost one-third of the two-year rise in the Index. While union organization may be prominent in handling some of these foods after they leave the farm, certainly the major influence on prices has been specific crop conditions that have affected the market price for these commodities. The freeze affecting the citrus fruits in Florida and the drought conditions on the Great Plains affecting meat prices, have both had a far greater influence on food prices than union organization or wage pressure.

The second largest increase is recorded for the price of newspapers. While labor costs are certainly one factor in the business of running a newspaper (and this has been included in the unionized sector of the economy), it would appear that the sharp rise in price has been the result of many forces accumulating over a number of years. Newspapers obviously cannot be subject to frequent repricing. There has been a reluctance to move away from the newsstand price of 5¢ a copy. The price rise over this two-year period came only after a three-year period when prices remained almost stable. Thus the sharp price rise has to be viewed as a reaction from demand and cost factors that have been accumulating for several years, rather than an indication of special demand or cost pressures in this particular two-year period.

The next highest price increase listed is for finance and insurance services. This includes interest on the homeowner's mortgage, and automobile and property insurance. Here the effect of higher interest rates becomes quite marked. The higher premiums charged for automobile insurance reflect a generally increased level of claims as well as higher repair costs caused by the structural characteristics of the newer model cars.

Another group showing a sharp increase during this two-year period is hospital care. In the case of hospitals, a number of factors seem to be at work. The increase in the figure for hospital care may reflect, not only increases in price, but also the more highly skilled technical services and improved equipment required in modern hospitals. The group hospitalization component may reflect the increased extent to which families have to resort to hospital care. While wage and salary costs form a large part of hospital expenses and have increased substantially, part of this increase reflects a higher proportion of skilled workers rather than an increase in wage rates. Moreover, as is generally realized, wage rates for hospital employees have been among the lowest in our society.

Another sharp increase is recorded for government services and taxes. The Consumer Price Index does not measure federal or state income taxes, but it does include various state and local taxes including real estate property taxes and auto registration fees. These prices increased quite markedly in the two-year period. Some of this increase may be related to higher wages (not that unions are not a major factor among state and local government employees) but basically the increase has been caused by expanding demands for state and local government services.

Moreover, there is some real doubt whether any index can accurately measure the price increases for government activities. The Index does measure the increased payments which the average city worker has to make for local government services. However, if state and local governments have been improving the quantity and the quality of services rendered to the taxpayer, the higher taxes paid reflect not only an increase in price, but also a better product for the taxpayer.

Some of the products or services in which unions play an important role show a relatively high increase in price for the two-year period. This is true, for example, of pottery, home repairs, and transportation equipment. However, there are other products in which union organization is equally strong such as electric machinery, public utilities, textiles and apparel in which the price increase was well below the average for the total index.

This brief run-down of price changes during the past few years indicates the complex nature of the pricing process. Today's American economy offers an almost bewildering variety of products (and therefore prices) to the American consumer. Specialized factors affect the demand, supply, or both, of different products and services to different degrees.

No attempt has been made here to argue that wages do not represent an element of cost to the firm or that the level of prices is independent of changes in wage rates. What is argued is that the weight of the evidence obtained by examining specific price changes over the past two years does not support the charge that union-won wage increases have been a major factor causing higher prices during the past two years.

EXAMINING ECONOMY-WIDE MOVEMENTS

The other type of evidence that must be considered relates to broad movements of wages, other costs, prices, and productivity. Specifically, it has become the economic fashion to argue that recent wage increases have been far out of line with the economy's increase in productivity.

The issue leads into an area where the technicians have to take over. Not only are the basic data often deficient but the task of developing proper ratios to determine productivity or unit costs becomes involved in numerous technical problems of measurement.

The Bureau of Labor Statistics has put together for the Joint Economic Committee yearly indexes for the postwar period indicating labor and non-labor costs, prices, employee compensation, and productivity.

What do these figures show?

To this observer, the single most impressive feature about them is that unit labor and non-labor payments have increased about at the same rate during the postwar period. Since the two shares are roughly equal, this means that the non-labor part of the price tag including such items as capital consumption allowances, interest payments, rental income, and profits has been creating just about as much pressure on costs and

prices as labor costs.

Not only have unit non-labor costs moved roughly the same as unit labor costs, but for certain non-labor items, particularly capital consumption allowances and charges for interest, the increase in the past few years is far higher than the increase for total employee compensation.

Comparison between wages (real compensation in constant purchasing power) and productivity (real product per man-hour -- all persons total private sector) can be made for any postwar time period. Percentage changes will vary with the particular period chosen, but for the postwar period as a whole, (1947-1957) these figures show wages still lagging behind productivity.

In all these calculations, employee compensation is considered as one lump sum. This obscures some very real changes that have been taking place in the composition of employment.

The fact is that the composition of the "employee" group has shifted markedly through the rapid growth in the number of workers on salary status. The increase has been particularly noteworthy among professional and technical employees but also includes large numbers in the office, clerical, and sales groups. Along with the increase in salaried workers has come an absolute as well as a relative decline in production workers as newer labor-saving equipment is introduced. The 1958 recession has accentuated this shift. In manufacturing, for example, the all-time peak of 14 million production workers was reached in August 1953. Today there are 2.1 million fewer. During the same time span, the number of non-production workers has risen from 3.5 to 3.8 million.

This change has a real impact on the productivity statistics. Although the new force of technicians is expected to contribute its full share to future increases in productivity, the initial result of adding non-production workers to the payroll has had the effect of limiting the currently recorded improvement in productivity.

It's difficult to see how these economy-wide movements of wages, costs, and prices can be interpreted to support the charge of Wage Inflation. Although it may be true that real wages advanced more than productivity during the two-year period 1955-57, other two-year periods could easily be chosen to show the opposite story. When the figures are viewed in the context of the entire postwar period, it is clear that employees have not gained a greater share of the benefits of productivity than other groups in society.

Moreover, the fact that wages increased more than productivity in 1956 and 1957 was not caused by excessive wage increases that were above average. Real average hourly earnings in manufacturing, for example, increased by 5.4 percent from 1955 to 1957, well within the average rate of productivity improvement for the postwar economy. Rather it was the rate of productivity improvement for these two years that was abnormal, since it was below the average for the economy during the postwar period.

Even though these were years of extraordinary high capital investment, the more efficient new plant and equipment was not immediately translated into equivalent advances in productivity. Among possible reasons for this are the following:

(1) The high cost of getting a new plant into full production together with the natural time lag involved before the newer equipment has reached its most efficient level of production.

(2) The sharp increase in the number of salaried workers as previously indicated, temporarily affecting adversely the productivity equation.

(3) The failure of consumption to grow as rapidly as the nation's capacity to producr. The result was that the economy was not operating at close to capacity, its most efficient level of production.

In any event, there is little reason to believe the low productivity rates of 1956 and 1957 will set a pattern for the future.

It now appears likely that this relatively low rate of productivity increase is already giving way to more rapid increases. Thus any gap that may have developed between the rate of productivity advances and employees' compensation will be eliminated as productivity returns to its normal postwar level.

IMPLICATIONS FOR PUBLIC POLICY

I have tried to indicate why I believe that inflation is not America's Public Enemy No.1, and why it certainly cannot be characterized as Wage Inflation. This does not mean that rising prices are not causing some serious problems in some important segments of the American economy, or that we have to stand back helplessly without taking any action to meet these problems. While I doubt that inflation deserves the label, "Public Enemy No.1," perhaps it could be called, "No.1 Juvenile Delinquent." Rather than requiring the entire resources of the FBI and other law enforcement agencies, the job of keeping the price rise within manageable bounds can be accomplished more effectively by measures to curb specific types of anti-social behavior.

Instead of complaining about the general rise in prices, it would be more fruitful to give more individual attention to the behavior of prices in specific segments of the economy.

For example, we need to know a great deal more about the price of medical care. This component of the Consumer Price Index has risen the most sharply of any during the past few years. Are these rising prices simply a response to the increased demand for medical services? In any event, are there governmental policies that can be undertaken that could lead to easing this price pressure?

Another problem area is the relatively few critical sections of industry where a small number of firms can establish prices without the same regard for market conditions that limit the price decisions in the more competitive industries. The basic steel industry is perhaps the most prominent of these. While steel prices are not directly represented in the Consumer Price Index, the implications of steel pricing policies go far beyond its particular section of industry, since steel products are used by practically every other industry in the economy.

If these are some of the problem areas, what types of actions can be taken to meet the problems they disclose?

Naturally, it is easy to diagnose rather than cure. One encouraging sign is the fact that some of these problems may have a built-in correction factor. After all, the price level has stabilized in recent months. Although the casual newspaper reader would hardly realize it, consumer prices have been practically stationary since March, and wholesale prices for almost a year.

Nevertheless, there are actions that can be taken to keep to a minimum any inflationary inclination that the American economy may have. As far as government is concerned, this analysis suggests that any search for a magic formula to apply to all prices would be a fruitless one. Instead, a more rewarding task might be to spend more time examining a whole host of government and private policies which affect the prices for particular products. Examining specific laws and policies affecting specific prices may appear a more roundabout method of attacking the problem, but it could prove to be the more effective if it leads to removing obstacles to more competitive pricing and thus assuring a more effective functioning of the pricing process. Among the government activities which could be so scrutinized are the following:

Anti-Monopoly Legislation--Although some economists have minimized the impact of antitrust action, the recent decision in the Bethlehem-Youngstown Steel case shows that this type of legislation still has its teeth. While antitrust laws can hardly be used to prevent specific price increases, the importance of maintaining as competitive an economy as possible in today's era of bigness cannot be over-estimated.

Fair Trade--Recently, there has been increased support for federal legislation to authorize resale price maintenance. Here is one specific area where the Federal Government can help keep prices competitive by failing to pass legislation.

Agriculture--Farm prices have been relatively weak during the postwar period. Yet there is something wrong with a program which yields ever-mounting surpluses of farm products even under an Administration devoted to lower price supports. Isn't there some way whereby the government's support program for agriculture can be translated into lower prices to consumers and still provide income-support to growers?

Taxation--Many specific taxes have an impact on consumer prices. The recent Congress took the healthy step of eliminating the three percent excise tax on freight transportion which provided an uneconomic addition to business costs. There may well be similar excise taxes which should be reduced or eliminated. The lost revenue could then be recaptured by closing existing loopholes in the Federal income tax structure.

Monetary and Credit Policy--Despite its deficiencies, this remains an important weapon against demand inflation, although those in charge of utilizing this weapon have, I believe, over-estimated its usefulness and have been too zealous in applying it. Some way must be found to modify the use of this weapon so that its impact is not concentrated in some areas of the economy (e.g.housing) while other areas, including many larger corporations, are practically immune to the operations of the central bank.

In the last analysis, we must remind ourselves that price stability still must remain just one of several objectives for government economic policy. It would obviously be unwise to adopt government policies which might yield a stable price level, but at the same time would limit economic growth and produce a higher level of unemployment. As others have suggested, perhaps the most effective guide to government policy for price stability is not to get in the way of economic growth, and to make certain that mistakes in monetary policy such as were made in the summer of 1957 are not repeated.

In the area of labor-management relations, many of those arguing that labor is the "culprit" have specific remedies to suggest. In a number of respects their views seem rather puzzling. Let me cite one example. A great hue and cry has been raised about the inequitable effects of union rivalry for higher wages. The theory is that high wage settlements generated by intense union competition has created inequities among lower-paid workers and produced additional pressure for wage increases that cannot be satisfied without increasing prices.

What is the remedy for this supposed malady? The answer, among a large group of experts, seems to be -- break up the unions and reduce the scope of collective bargaining. I do not know to what extent this may weaken unions, but the one result that is almost inevitable would be a heightening of the wage competition among union groups. To

someone worried about wage inequities, the result would be a nightmare.

Moreover, there should be some hesitation about tampering with the American system of collective bargaining which, though obviously not perfect, has proved an efficient and democratic mechanism for wage determination. While partisans from both the labor and management sides have been arguing for many years regarding the extent of government intervention in the collective bargaining process, both groups demonstrably prefer the process of mutual accommodation in wage-setting to a system with greater government intervention.

The collective bargaining process has proved flexible to changing economic circumstances. The American system with its emphasis on local or company bargaining rather than national collective bargaining, yields a great diversity of wage settlements. In effect, most of the wage bargains have been fashioned with an eye to the specific conditions prevailing in the industry, locality, or firm concerned. Experience in the textile industry, for example, demonstrates how collective bargaining results are affected by economic conditions. The extent to which particular wage settlements have become the "pattern" for other industries is probably less today than it was 10 years ago.

Of course, it will always be possible for economists to find particular collective bargaining settlements which they feel have increased wages at a higher rate than productivity. But the positive values of collective bargaining should not be lightly sacrificed. The crucial question is whether the system of collective bargaining -- not an isolated case or even groups of cases -- persistently produces a condition in which wage gains are achieved at the expense of the general welfare.

On the basis of the American experience since World War II, I do not see any compelling reason for altering the basically voluntary character of wage settlements negotiated through collective bargaining.

1/ The First National City Bank Letter, June 1958, p. 71

2/ There may be some disagreement regarding the degree of union influence in some of these industries. For example, textile mill products have been listed as "relatively unionized" because collective bargaining settlements are given prominent attention even t hough the majority of textile plants are not unionized. On the other hand, the processing and distribution of nonperishable foods and beverages is often done by union members,but unions are not normally involved in the preparation of the basic product. Very little change in the price increase recorded for the groups as a whole would result if any items such as these were transferred to the opposite column.

APPENDIX - 1

CHANGES IN CONSUMER PRICES *
March 1956-March 1958

(Groups listed in order of price increase)

Industry Classification	Weight December 1955	Percent Increase[a/] 1956-58	1957-58	1956-57	Percent of total price increase
TOTAL INDEX	100.00	7.4	3.82	3.43	100.0
NEWSPAPERS	.99	16.6	14.8	1.5	2.2
SERVICES	14.59	10.0	5.0	4.7	19.9
Labor services	5.03	7.2	3.1	4.0	5.0
Professional services	3.04	6.7	2.6	3.9	2.8
Finance and insurance	2.74	16.0	10.9	6.7	6.0
Amusement	1.61	11.5	6.8	4.4	2.5
Hospital care	1.34	14.1	7.8	5.8	2.6
Miscellaneous (legal, banking, burial)	.83	b/	b/	b/	1.1
FOOD, LIQUOR, TOBACCO	32.91	10.0	6.4	3.5	44.9
Perishable foods	13.68	17.2	13.1	3.8	31.9
Nonperishable foods	8.45	5.3	2.1	3.2	6.1
Food away from home	4.73	7.1	3.4	3.5	4.6
Alcoholic beverages	2.38	3.7	0	3.7	1.2
Tobacco products	1.98	6.4	5.1	1.2	1.7
Beverages	1.69	- 2.5	- 8.1	6.1	- 0.6
GOVERNMENT SERVICES AND TAXES	1.64	9.9	6.6	3.1	2.2
Taxes	1.38	11.6	7.7	3.7	2.2
Government services	.26	0.8	0.8	0	---

* The figures in this table differ slightly from the table included in the author's paper to the Joint Economic Committee published in The Relation of Prices to Economic Stability and Growth, Commentaries Submitted by Economists from Labor and Industry, Joint Economic Committee, October 31, 1958. The above table reflects the use of more refined statistical techniques in computing the various group indexes. Differences between the two tables are slight.

a/ As computed from the individual items in the quarterly index published by the Bureau of Labor Statistics for a subsample of 19 cities. Total index increases, thus computed, differ slightly from those shown by regular monthly series, which were as follows: 1956-58, 7.5%; 1957-58, 3.7%; 1956-57, 3.7%.

b/ Not priced. Price increase for major group imputed to this item.

APPENDIX - 2

Industry Classification	Weight December 1955	Percent Increase[a] 1956-58	1957-58	1956-57	Percent of total price increase
OIL, CHEMICALS, RUBBER AND POTTERY PRODUCTS	7.40	6.2	1.1	5.2	6.3
Petroleum and coal products	3.98	4.7	- 2.0	6.8	2.5
Chemical products	2.88	8.4	4.8	3.4	3.3
Rubber products	.35	2.9	4.5	- 1.5	0.1
Pottery	.19	12.0	3.7	8.1	0.3
METAL PRODUCTS	11.28	5.3	0.7	4.4	8.1
Transport equipment	4.31	11.0	1.3	9.4	6.5
Electrical machinery	3.04	- 3.2	- 2.5	- 0.8	- 1.3
Fabricated metal products	2.49	4.7	1.8	2.8	1.6
Miscellaneous manufactured goods (Appliances)	1.44	6.8 - 2.9	4.4 - 2.2	2.3 - 0.8	1.3
PUBLIC UTILITIES	4.95	5.3	3.3	1.9	3.6
Water, gas, electricity	2.40	3.4	2.1	1.3	1.1
Transit and railroad fares	1.46	8.8	5.8	2.9	1.8
Communications	1.09	4.9	2.8	2.1	0.7
HOUSING	13.31	4.9	3.0	1.8	8.8
Home purchase	6.08	4.5	3.6	0.9	3.7
Rent	5.89	4.2	2.0	2.1	3.4
Home repairs	.94	11.6	5.0	6.3	1.5
Housing away from home	.40	b/	b/	b/	0.3
TEXTILE MILL PRODUCTS	1.36	2.9	0.3	2.5	0.5
Housefurnishings	.75	0.4	- 0.9	1.4	--
Floor coverings	.55	4.9	1.6	3.2	0.4
Other	.06	15.0	3.6	11.0	0.1
WOOD AND PAPER PRODUCTS	2.40	3.5	0.1	3.4	1.1
APPAREL	9.17	1.9	0	1.9	2.4
Textile	7.82	1.2	- 0.3	1.4	1.2
Leather	1.35	6.1	1.4	4.6	1.1

APPENDIX - 3

Detailed Classification of Items in Consumer Price Index

Item	Weight December 1955
NEWSPAPERS	.99
SERVICES	14.59
Labor services	5.03
dry cleaning & pressing	1.26
laundry	.81
domestic	.56
shoe repairs	.15
auto repairs	1.19
men's haircuts	.70
beauty shop services	.32
TV repairs	.04
Professional services	3.04
physicians' fees	1.92
dentists' fees	.85
optometrist	.27
Finance & Insurance	2.74
mortgage interest	1.64
auto insurance	.90
property insurancw	.20
Amusement	1.61
motion picture admissions	1.61
Hospital care	1.34
group hospitalization insurance	1.12
hospital room rates	.22
Miscellaneous (legal, banking, burial)	.83
FOOD, LIQUOR, TOBACCO	32.91
Perishable foods	13.68
beef and veal	1.89
pork	2.01
lamb	.16
frankfurters	.63
poultry	.94
fresh fruits & vegetables	2.67
dairy products	3.97
eggs	1.41

APPENDIX - 4

Item	Weight December 1955
Non-perishable foods	8.45
cereals & bakery products	3.23
canned luncheon meat	.22
canned & frozen fish	.55
vegetables & fruits (canned, dried, frozen)	1.65
tomato soup	.37
beans with pork	.16
sweet pickles	.23
tomato catsup	.10
fats & oils	.92
misc. (gealtine)	.10
sugar & sweets	.92
Food away from home	4.73
Alcoholic beverages	1.69
beer	1.45
whiskey	.93
Tobacco Products	1.98
cigarettes	1.84
cigars	.14
Beverages	1.69
coffee	1.17
tea	.15
cola drinks	.37
GOVERNMENT SERVICES AND TAXES	1.64
Taxes	1.38
real estate	1.08
auto registration	.30
Government service	.26
postage	.26
METAL PRODUCTS	11.28
Transportation equipment	4.31
new cars	2.85
used cars	1.46
Electrical machinery*	3.04
refrigerators	.76
washing machines	.49
vacuum cleaners	.21
toasters	.19
sewing machines	.17
television sets	.80
radios	.35
electric light bulbs	.07

*"Electrical machinery" consists mainly of appliances. Classified under the usual CPI grouping, however, "appliances" exclude electric light bulbs & include stoves.

APPENDIX - 5

Item	Weight December 1955
Fabricated Metal Products	2.49
stoves	.48
water heaters	.78
cabinet kitchen sinks	.13
faucets, sink	.33
saucepans, aluminum	.30
razor blades	.14
tools	.33
Miscellaneous manufactured goods	1.44
toys	.28
sporting goods	1.16
CHEMICAL, RUBBER, AND POTTERY PRODUCTS	7.40
Petroleum & coal products	3.98
gasoline	2.40
motor oil	.22
solid fuels & fuel oil	1.36
Chemical products	2.88
laundry soap & detergents	.64
toilet soap	.22
prescriptions & drugs	.81
toothpaste	.20
face powder	.11
shaving cream	.06
face cream	.12
shampoo	.11
home permanents refill	.05
exterior housepaint	.56
Rubber products	.35
tires	.35
Pottery products	.19
dinnerware	.19
PUBLIC UTILITIES	4.95
Water, gas, electricity	2.40
water	.37
gas & electricity	2.03
Transit & railroad fares	1.46
transit fares	1.18
railroad fares, coach	.28
Communication	1.09
telephone	1.09

APPENDIX - 6

Item	Weight December 1955
HOUSING	13.31
Home purchase	6.08
Rent	5.89
Home Repairs	.94
repainting rooms	.29
repainting garage	.16
refinishing floors	.18
reshingling roof	.31
Housing away from home	.40
TEXTILE MILL PRODUCTS	1.36
Housefurnishings	.75
towels, bath	.07
sheets, muslin	.19
curtains	.15
blankets, wool	.09
bedspreads, cotton	.08
drapery fabrics, cotton	.17
Floor coverings	.55
rugs, wool axminster	.12
carpets, wool broadloom	.24
rugs, felt base	.13
rugs, rayon or cotton	.06
Other	.06
sanitary napkins	.06
WOOD AND PAPER PRODUCTS	2.40
Wood Products	2.03
furniture & bedding	1.72
porch flooring	.31
Paper Products	.37
toilet tissue	.21
cleansing tissue	.12
paper napkins	.04
APPAREL	9.17
Textile	7.82
Men's and boy's apparel	2.94
women's and girl's apparel	4.08
other apparel	.80
Leather	1.35
shoes	1.35

IS THIS A NEW TYPE INFLATION?

By: Geo. P. Hitchings, Ford Motor Company

The type of inflation experienced in recent years is different from past inflations. Traditional inflations involved a sharp rise in prices, reflecting excess money demand for goods and services relative to supply. This excess demand was made possible by expansion in the supply and/or use of money. In recent years, however, the rise of prices has been more gradual and for the most part has not been based upon excess money demand. Instead, price increases have been based largely upon a steady rise in costs forced from the supply side.

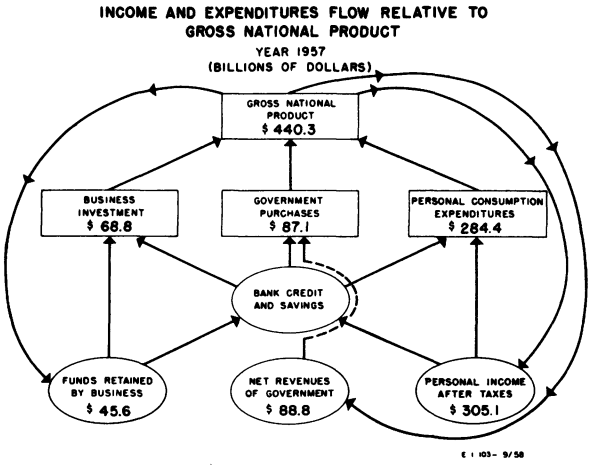
Careful statistical analysis is required to disentangle cause and effect relationships in rising prices. This is because, with a rise in prices, there is also a rise per unit of production in:

1. Expenditures.
2. Incomes (which are also costs of production).
3. Money usage.

The problem is to determine which factor initiated the rise and which factors followed.

These identity relationships are illustrated in the chart below for income and expenditure flow relative to gross national product.

Chart 1



Total production is bought by some customer in the form of business investment, government purchases, or personal consumption. In 1957, for example, the $440.3 billion of gross national product was absorbed by:

$68.8 billion - Business investment (Residential construction, Other private construction, Producers' durable equipment, Change in business inventories, and Net exports).

$87.1 billion - Government purchases (Government employee payrolls, Public construction, and Other purchases of equipment, supplies, and services from business.

$284.4 billion - Personal consumption expenditures (Durable goods, Nondurable goods, and Services).

In turn, the revenue received from sale of this production flowed to these groups as follows:

$45.6 billion - Funds retained by business (Depreciation and other capital consumption allowances, Undistributed profits, and Government transfer payments to foreign countries).

$88.8 billion - Net revenues of government (Tax revenues less the total of Government interest, Transfer payments, and Net subsidies).

$305.1 billion - Personal income after taxes (Personal income less tax payments).

The slight difference between total income and production (or expenditure) is due to discrepancies in source data. Conceptually, they are equal.

Although income and expenditure balance in total, individual spending units or groups spend either more than, or less than, their income for a given period. The central category of bank credit and savings shown in the flow chart makes possible such variation for individual components, within the framework of equality in total income and expenditure.

The identity relationships in this flow chart make it clear that when prices rise, there is an equivalent percentage rise per unit of production in total expenditures, total income, and total money usage. It is necessary to examine internal relationships within the aggregate flow, therefore, to determine cause and effect.

One of the primary factors to examine, in view of its importance in traditional inflations, is the relationship of money supply to physical quantity of production.

During World Wars I and II, and the immediate postwar periods, there was a substantial rise in the quantity of money relative to production. In 1915-20, demand deposits and currency rose 108 per cent, compared with an increase of about 14 per cent in quantity of production. Time deposits (which are not used directly as a medium of exchange but which are readily converted to money) rose nearly 72 per cent. Similar monetary inflation occurred in 1940-45, when demand deposits and

currency increased 142 per cent, compared with about 53 per cent for production. Time deposits were up 75 per cent. After the war, in 1945-48, the money supply increased further in the face of slightly lower production levels.

Time Periods	Money Supply vs. GNP in Constant Dollars (% increase)			Factors Affecting Deposits & Currency - Total Banking and Monetary System (increase in billions of $)				
	GNP in Constant Dollars 1/	Demand Deposits & Currency	Time Deposits	Total Deposits & Currency	Loans & Investments excl. U.S. Govt. Securities	U.S. Govt. Securities less U.S. Treas. Balances	Gold & U.S. Treas. Currency less Treas. Cash	Less: Foreign Bank Deposits, Capital, & Misc. Accounts
Price Increase								
June 1915-June 1920	13.9	108.1	71.5	18.9	16.3	3.9	.7	2.0
June 1933-June 1937	44.9	60.1	19.6	15.8	.8	10.4	5.4	.7
Dec. 1940-Dec. 1945	52.6	142.1	74.7	80.8	5.8	79.6	-.8	3.8
Dec. 1945-Dec. 1948	-6.7	9.0	18.7	18.3	20.8	-5.7	5.4	2.2
Dec. 1949-Dec. 1951	16.8	12.0	4.8	16.2	21.3	-2.4	-1.6	1.1
Dec. 1954-Dec. 1957	12.1	3.1	18.4	18.0	31.9	-13.0	1.3	2.2
Price Stability								
June 1921-June 1929	57.6	25.9	72.5	18.4	17.8	1.1	1.3	1.8
June 1937-Dec. 1940	12.2	37.7	7.1	13.4	.5	3.5	11.4	2.1
Dec. 1948-Dec. 1949	-0.1	-0.4	1.9	.7	2.5	-.7	.2	1.3
Dec. 1951-Dec. 1954	6.2	7.9	22.5	23.7	22.7	5.8	-.2	4.5
Price Decline								
June 1920-June 1921	-8.5	-12.4	4.7	-2.2	-1.4	-.4	.5	.9
June 1929-June 1933	-30.4	-26.8	-24.3	-14.0	-21.1	4.1	.2	-2.8

1/ Year-to-year comparisons.

The expansion of deposits and currency during 1915-20 and 1945-48 arose largely from increased loans and investments to business and individuals. During 1940-45, on the other hand, the increased money supply stemmed from government borrowing to finance the war.

These periods of substantial price inflation were clearly the result of excess demand fed by expansion of the money supply relative to production. During World War II, the excess money supply was held in check by the substitution of rationing and price controls for free markets. Once free markets returned, this excess money, coupled with the large volume of other liquid assets readily convertible into money, made possible money demands well in excess of the available supply of goods and services.

An increase in the quantity of money relative to production was not a primary factor in the other periods of rising prices since 1915. During 1933-37, rising prices represented more of a recovery from depression levels rather than excess demand pressing against productivity capacity. They also reflected government policies deliberately aimed at such recovery of prices.

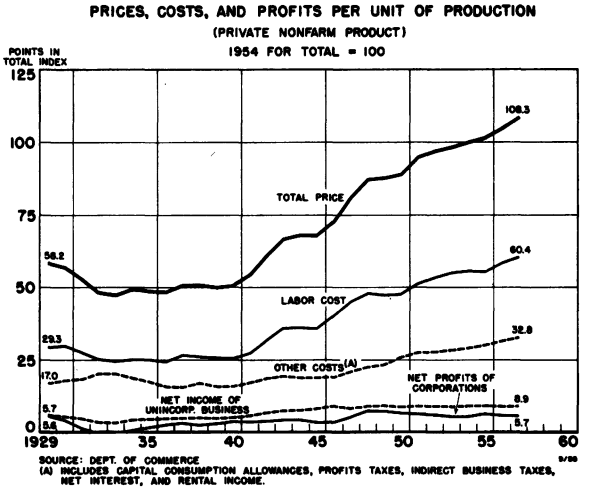
In the 1949-51 price rise, excess demand was again the driving force. After the outbreak of war in Korea, business and consumers bid for commodities and services that they feared would become in short supply and higher priced. This was accomplished through faster turnover of money rather than an increase in its quantity relative to production.

The recent general price rise of 1954-57 was marked neither by expansion of the money supply relative to production, nor by broad demand pressures against available supplies. Demand pressures were a factor only in limited segments of the economy, and then only for part of the period. For the broad range of consumer goods and services over the period as a whole, however, excess demand was not a factor.

The weight of evidence points to conditions on the supply side rather than on the demand side as the initiating force behind general price increases over recent years. This is not to deny the role of demand in determining relative price changes among various commodities and services. Nor is it to deny that sufficient demand and money usage was required to pay generally higher prices. These flowed from the price rise, however, rather than the reverse.

The initiating force in steadily rising prices for nonfarm production since 1951 has been the increase in costs of production. This is illustrated in the chart below where the aggregate average price for nonfarm production is broken down into component costs and profits.

Chart 2



The chart is based on U.S. Department of Commerce data for gross national product exclusive of government and agriculture. Dollar revenues from production are divided by production in constant dollars to obtain the implicit price, costs, and profits per unit of production. The average price is shown in terms of an index (1954=100), with costs and profits shown as component points in the total index. Costs and profits shown on the chart do not add exactly to the total price because of excluding a small residual for business transfer payments (such as consumer bad debts and gifts to nonprofit institutions), net subsidies of government enterprises, and the statistical discrepancy between income and expenditure measures of production value.

Unit labor costs (which include total employee payrolls and fringe benefits) accounted for

two-thirds of the aggregate price rise from 1951 to 1957. Unit labor costs increased because pay rates and fringe benefits expanded at a more rapid rate than production per employee man-hour.

Other unit costs also moved up steadily over this period. This rise was concentrated in depreciation, interest, and indirect taxes. Higher unit labor costs in construction and capital equipment industries, as well as in government, were a factor also in these cost increases.

Income remaining for business owners did not rise over this period relative to production. Net income to owners of unincorporated business per unit of production was virtually unchanged. Corporation unit profits declined, both on a before-tax and after-tax basis. The chart shows unit profits on an after-tax basis because, to the stockholder, this represents income before taxes. The investor pays income taxes on dividends and capital gains taxes on gains realized from investment of undistributed profits. Corporation profits taxes are a cost to the investor of doing business just as any other expenditure. In diagnosing the 1951-57 period, however, it is not essential to distinguish between before- and after-tax figures. There was a decline on either basis relative to production.

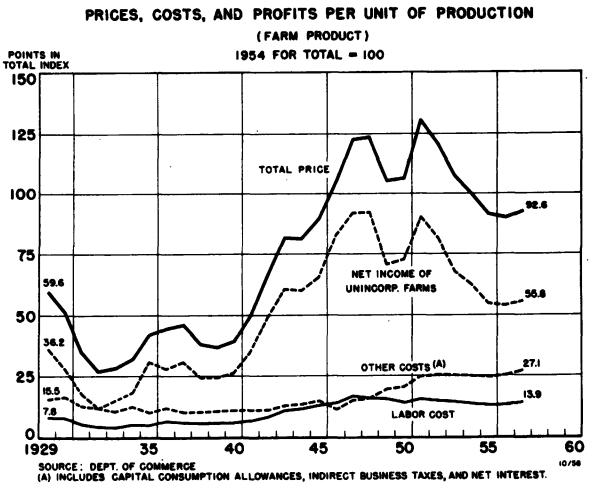
This type of income distribution from production does not prevail in a traditional inflation. Where demand is the driving force, unit profits will rise along with unit costs. Excess money demand normally bids up prices more rapidly than costs. The prospect for higher profits also leads to competitive bidding for employees, construction, capital equipment, and investor funds.

The decline in unit profits for recent years clearly indicates that neither general excess demand nor business pricing for higher profit margins was responsible for rising prices on nonfarm production after 1951. The generating force was a push-up in unit costs, largely the result of attempts to move ahead more rapidly on general pay rates and fringe benefits than could be validated by greater efficiency of production for the nonfarm economy as a whole. Business profits were not large enough to absorb these cost increases even if profits had been wiped out.

A strikingly different picture has prevailed since 1951 in farm prices, costs, and net incomes per unit of production.

In contrast with the nonfarm segment, farm unit costs (hired labor, depreciation, etc.) account for a relatively small portion of the price and have been fairly stable since 1951. Farm prices declined sharply from 1951 to 1955, as temporarily heavy demands receded and supplies increased. The decline in price was reflected in lower net income to farmers per unit produced. It should be noted, however, that farm prices and unit net income had previously risen more sharply during 1940-48.

Chart 3



It is not surprising that farm prices and incomes show greater fluctuation than the total nonfarm sector of the economy. Farm supplies cannot be geared as readily to demand as most nonfarm products and services because:

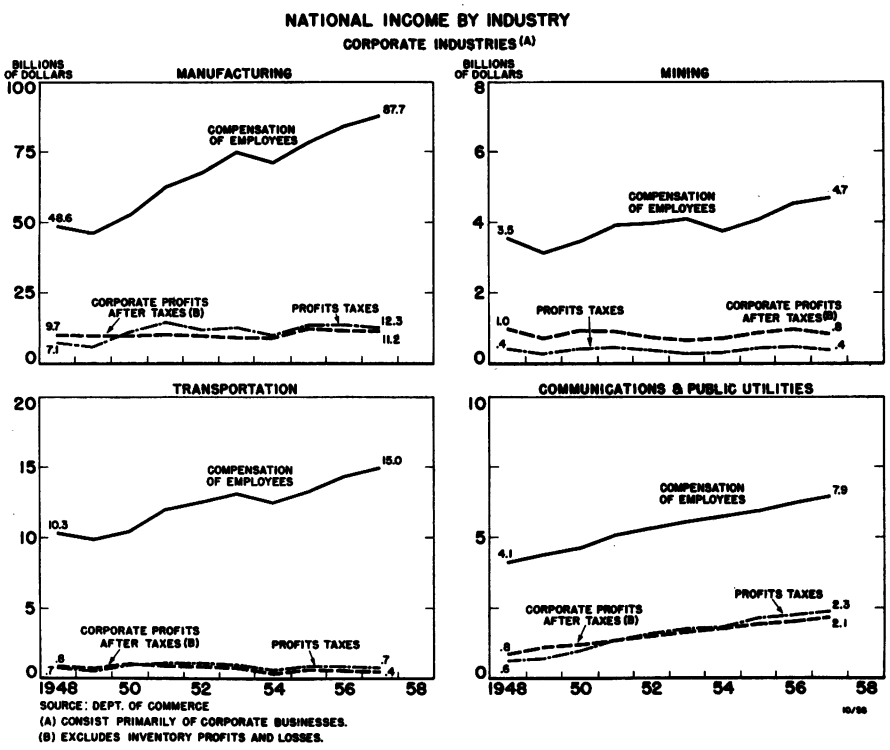
1. Farm production is influenced to an important extent by weather conditions.
2. The production cycle is longer than most nonfarm items.
3. Low unit costs for marginal production relative to price encourage extra production, particularly with government price-support operations.
4. Perishable farm products cannot be held off the market. They must be marketed for whatever price they will bring.
5. The large number of producers makes for less stability in production and prices. It is more difficult to effect necessary changes in production.

Further light on the inflation problem can be obtained by examination of the major component industry groups within the nonfarm economy. Unfortunately, data are not yet published on an industry breakout for the components of gross national product in current dollars and in constant dollars. Computations thus cannot be made for unit prices, costs and profits comparable with the two preceding charts for the total private nonfarm segment and for the farm segment of the economy. It is hoped that such data can be made available in the future.

Data are available, however, for national income originating in the major industry groups. This makes possible a comparison of employee compensation with net profits and profits taxes. Chart 4 illustrates the data for industries dominated largely by the corporate form of business.

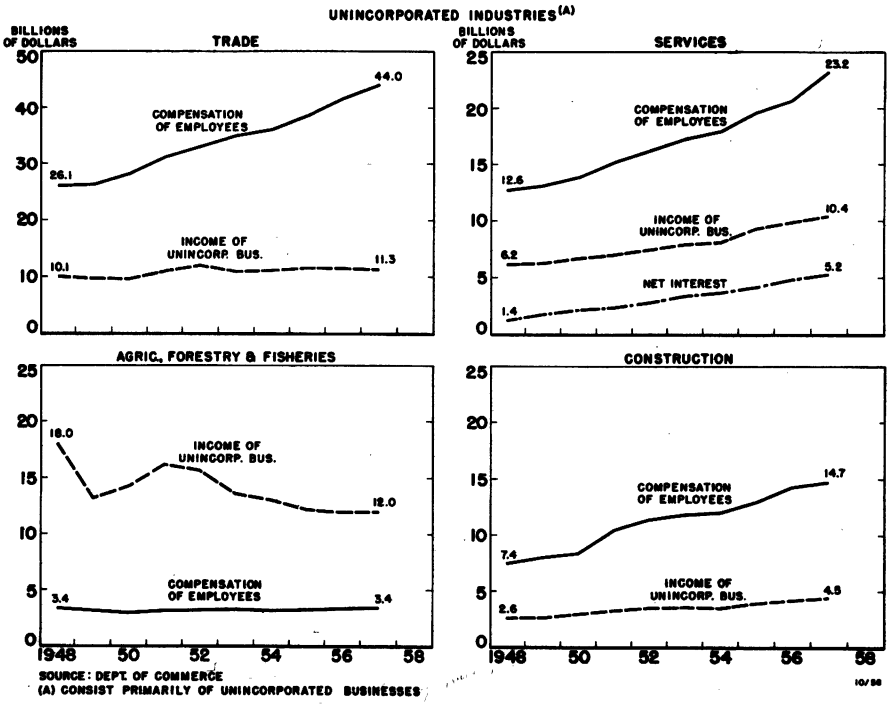
In manufacturing, mining, and transportation, corporation profits have obviously not been a

CHART 4



SOURCE: DEPT. OF COMMERCE
(A) CONSIST PRIMARILY OF CORPORATE BUSINESSES.
(B) EXCLUDES INVENTORY PROFITS AND LOSSES.

CHART 5



SOURCE: DEPT. OF COMMERCE
(A) CONSIST PRIMARILY OF UNINCORPORATED BUSINESSES

factor in higher prices over recent years. Manufacturing corporation profits after taxes were $1½ billion higher in 1957 than in 1948, compared with a $39 billion increase in employee compensation. 1/ The absolute level of employee compensation in 1957 was nearly 8 times as large as net profits. It is clear then that higher manufacturing prices in the aggregate did not reflect higher profits, and that profits are not large enough to cover much of a pay increase. The 15 per cent rise in profits was far short of the 41 per cent rise in physical quantity of manufacturing production shown by the Federal Reserve Index.

In mining and transportation, profits actually declined over this period, while employee payrolls went up substantially. Obviously, profits had nothing to do with higher prices in these segments.

Only in communications and public utilities did aggregate profits participate appreciably in rising revenues. Net profits rose by $1.3 billion from 1948 to 1957, while employee compensation increased $3.8 billion. Profits are also larger in absolute amount relative to employee compensation in this segment of the economy, because so much of the production process represents capital facilities input. Prices charged by these corporations are, however, subject to government regulation. It can be assumed that pricing to obtain higher profits was not a factor even for communications and public utilities.

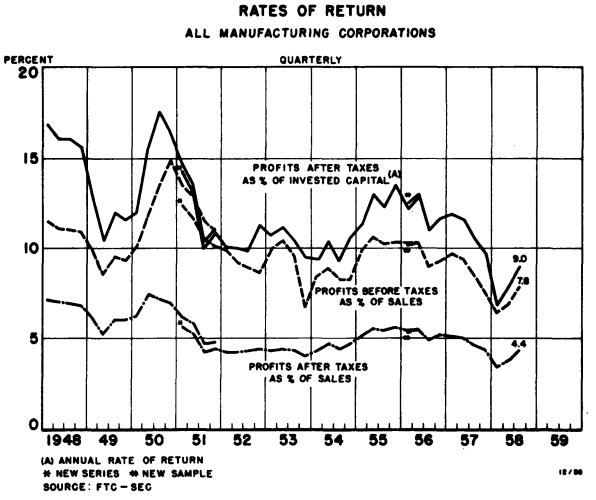
For the industries dominated by unincorporated business firms, trade (wholesale and retail) shows a picture comparable with manufacturing. (See Chart 5.) Employee compensation 2/ increased by $18 billion from 1948 to 1957, while income to the unincorporated owners rose only slightly more than $1 billion. The 12 per cent rise in aggregate owner income was undoubtedly less than the increase in value of services performed in constant prices. Here again, higher profits had nothing to do with higher prices.

The picture for agriculture in terms of unit labor costs and owner income has been covered in the earlier chart for farm product. Income to owners declined, while total hired worker payrolls remained unchanged. Owner income in agriculture is much larger than employee compensation because hired workers contribute only a small portion of the value added on the farm. The owner and his unpaid family help account for the great bulk of man-hours worked on the farm, as well as for the capital used in farm production. To a substantial extent, net income to farm owners represents pay for their labor. Only part of this income can be considered as return on capital invested in production.

The results for service industries and construction are more similar to communication and public utilities. Income to the owners of unincorporated firms in these industries has grown in line with employee payrolls, although the absolute level and amount of increase in payrolls has been greater. These industries are not dominated by large business enterprises, which are usually designated as the culprits in pricing policies to obtain high profits.

That pricing to obtain higher profit margins is not the cause of inflation in recent years is even more forcefully indicated by the following chart, which relates profits of all manufacturing corporations to sales and to invested capital. The profit margin on sales is shown both on a before- and after-tax basis. This comparison is useful in periods of changing tax rates.

Chart 6



Selection of proper base dates is most important in analysis of profit margins between two years. Profit margins always shrink in recession years such as 1949, 1954, and 1958, and rise in recovery years. Tax-rate increases after outbreak of war in Korea and reductions after cessation of hostilities affected after-tax comparisons in the last half of 1950 and in 1954.

With due allowances for these impacts, profit margins in recent years have fallen short of comparable postwar years prior to Korea. Even in 1955, profit margins did not return to the levels of earlier prosperous years such as 1948 and 1950. Margins declined from 1955 to 1957 during the period of sharpest price rise since 1951.

Despite these facts on aggregate profits and costs, some observers still argue that pricing policies of large corporations are to blame for inflation. They usually base their arguments on data comparisons which are at best incomplete and at worst a distortion.

Such is the case where price increases are attacked because they bring in more total revenue than the added cost of hourly worker payrolls per unit of production for an individual company or group of companies. Hourly worker payrolls of the individual company are only part of the total cost. There are many other factors in the production process; e.g., salaried workers, capital input, and materials and services purchased from others. The latter in turn represent employee costs and profits of supplying firms.

Wage-rate and fringe benefit increases for a company's hourly-paid workers infiltrate all these other costs as well. Also, there may be a shift in the mix of the various input factors. A comparison of increased cost per unit of production for only one of the input factors relative to price presents a partial and distorted picture.

Another distortion is the comparison of profits from a low-volume year to a high-volume year, or of absolute dollar profits over a period of time without allowance for increased volume or capital investment. Profits fluctuate sharply over the business cycle so that comparisons are valid only if they refer to the same stage in the cycle. The absolute volume of profits in itself is not meaningful in a discussion of prices. Only the profit margin on sales or on invested capital is significant for price purposes.

The relationship of profits to prices is also distorted by use of the most profitable companies as a standard for ability to absorb cost increases without raising prices. A squeeze in profits of these companies also means a squeeze on their less profitable competitors. The ultimate result of such a policy would be to drive out of business all but the most efficient producer. Even if it were possible to equalize the profit position of the most and least efficient producers, the incentive for increased efficiency would be eliminated.

Incomplete analysis also underlies the argument that "administered prices" are to blame for inflation in recent years. The term "administered prices" is far from precise, but generally refers to quoted prices that do not fluctuate frequently and sharply with changes in demand. The seller is not at the mercy of an auction market over which he as an individual has little control.

On that basis, most nonfarm prices are administered. Since nonfarm prices have accounted for most of the price rise since 1951, it obviously follows that this is the area of inflation during this period. As mentioned earlier in this discussion, however, the reason for the differential performance of prices is to be found in costs and in the fact that farm production was not adjusted adequately to demand.

Every seller can "administer" the price of his own product or service. He cannot, however, administer his combined price-volume-cost-profit relationship. He is primarily interested in maximizing his return on investment, rather than receiving a stated price. Price is only one factor in determining return on investment, as shown in the following examples. This is true whether the business is large or small.

In Example A, the seller's cost picture shows a variable cost 3/ per unit produced of $700 and a total dollar fixed cost 4/ of $4,000 at his present productive capacity. His problem is to maximize his return on an investment which we assume to be $10,000. At a price of $1,000 per unit, he can sell 20 units. On this basis, his profit before taxes will be $2,000 or 20 per cent on investment. At a price of $950 a unit, he would have to sell 24 units to make the same total dollar profit (assuming that his productive capacity is ample to handle the greater volume without a rise in fixed cost). In other words, he would have to sell 20 per cent more units at a 5 per cent lower price, an elasticity factor of 4 to 1. Furthermore, his break-even volume would also increase 20 per cent (from 13-1/3 units to 16 units) by the narrowing in price spread over unit variable cost.

PRICE-VOLUME-COST-PROFIT RELATIONSHIPS

EXAMPLE A

	UNIT VOLUME WITH PRICE OF $1,000			UNIT VOLUME REQUIRED AT PRICE OF $950		
	DOLLARS PER UNIT OF SALES	UNITS SOLD	TOTAL DOLLARS	DOLLARS PER UNIT OF SALES	UNITS SOLD	TOTAL DOLLARS
PRICE	$1,000	20	$20,000	$950	24	$22,800
VARIABLE COST	700	"	14,000	700	"	16,800
FIXED COST	200	"	4,000	166-2/3	"	4,000
PROFIT BEFORE TAXES	100	"	2,000	83-1/3	"	2,000

EXAMPLE B

PRICE	$1,000	20	$20,000	$950	22	$20,900
VARIABLE COST	450	"	9,000	450	"	9,900
FIXED COST	450	"	9,000	409	"	9,000
PROFIT BEFORE TAXES	100	"	2,000	91	"	2,000

In Example B, the seller's cost distribution is different. Variable costs are only $450 per unit, but total fixed costs are $9,000. Sale of 20 units at a price of $1,000 will yield the same $2,000 before taxes. In this case, however, only a 10 per cent increase in volume is necessary to offset a 5 per cent reduction in price, or an elasticity factor of 2 to 1. Break-even volume would be raised the same 10 per cent.

It is clear, then, that differential price behavior may be related to differences in cost structure and in elasticity of demand relative to price. Each seller must take into consideration his particular price-volume-cost-profit relationship in setting his price. If he sets a price higher than his competitors, the loss of volume may more than offset his higher unit margin over variable costs. On the other hand, if he takes the lead in setting a lower price, he must gamble on enough increase in volume to offset the lower unit margin. He receives no competitive volume advantage if his competitors meet his price. If all sellers meet the lower price, they will realize lower returns on investment unless the total market is stimulated enough to offset lower margins.

Presumably the concern about "administered prices" as a factor in inflation is that sellers will be able to set higher profit margins for themselves. The previous chart on profit margins in recent years, however, indicates that profit margins have acted as a drag, rather than a push, on prices in recent years.

Conclusion

The underlying factor in the long-run rise of nonfarm prices since 1951 has been increasing costs. Price performance has not been uniform

among various products and services because of differential cost increase and because of differential demand impact on profit margins. The economy was not, however, suffering from general excess demand in this period.

This is a different type of inflation than that experienced in wars and immediate postwar periods. It is more gradual and sustained. It leads to different expectations by business and consumers and has a different impact on the economy.

Traditional tools to combat inflation through restrictive monetary, credit and fiscal policies are effective primarily against excess money demand. They are also partially effective in holding down cost increases, but only at the risk of curtailing demand below levels needed for a prosperous and growing economy.

The recent type of inflation can be combatted only by holding total payrolls in line with production. Rising production is the only source of real purchasing power. Paying out excess dollars creates only inflation, not real economic growth and stability.

Footnotes

1/ Employee compensation includes a small amount for unincorporated firms which cannot be segregated in the national income figures.

2/ Including employee compensation for corporations, which cannot be segregated in the national income data.

3/ Variable costs are those items where total dollar costs vary directly in proportion to volume.

4/ Fixed costs are those items where total dollar costs are the same regardless of volume.

APPENDIX TABLE - CHART 2

PRICES, COSTS, AND PROFITS PER UNIT OF PRODUCTION

(PRIVATE NONFARM PRODUCT)

	Total Price (Index) (1954=100)	Labor Cost	Other Costs 1/	Net Profits 2/ of Corporations	Net Income 2/ of Unincorp. Business
		(Points in Total Index)			
1929	58.2	29.3	17.0	5.6	5.7
1930	56.9	29.7	17.9	4.2	5.3
31	52.6	27.5	18.4	1.0	4.5
32	48.2	25.2	20.4	-2.2	3.3
33	47.3	24.6	20.3	-2.5	3.2
34	49.3	25.1	18.5	0.3	4.1
35	48.8	25.1	17.3	1.6	4.4
36	48.6	24.4	15.7	2.5	4.6
37	50.5	26.7	15.6	3.1	4.7
38	50.6	26.0	16.7	2.3	4.8
39	49.9	25.6	15.7	2.8	4.7
1940	50.4	25.4	15.6	3.7	4.9
41	54.3	27.4	17.2	3.5	5.5
42	60.9	32.0	18.6	3.8	6.5
43	66.6	36.0	19.4	4.2	7.4
44	68.0	36.1	18.8	4.2	7.5
45	67.9	35.9	18.8	3.2	8.0
46	72.8	40.3	18.9	3.5	9.1
47	81.1	45.0	20.9	5.1	8.2
48	86.9	48.0	22.6	7.3	8.9
49	87.6	47.4	23.5	7.1	9.0
1950	88.8	47.7	26.0	6.5	8.6
51	94.9	51.3	27.7	6.3	8.9
52	97.2	53.4	27.9	6.0	8.9
53	98.7	55.0	28.6	5.4	8.7
54	100.0	55.6	29.3	5.3	8.9
55	101.4	55.2	30.2	6.3	9.0
56	104.5	58.3	31.5	5.9	8.9
57	108.3	60.4	32.8	5.7	8.9

1/ Includes capital consumption allowances, profits taxes, indirect business taxes, net interest, and rental income. Does not include business transfer payments, net subsidies of government enterprises to nonfarm business, and the statistical discrepancy between measures of income and production. Total costs, net profits, and net income of unincorporated business will differ from total price by the total of these excluded items.

2/ Excludes profits and losses from inventory revaluation.

Source: Indexes computed by the author from Department of Commerce data on national income and product.

APPENDIX TABLE - CHART 2

COMPONENTS OF "OTHER COSTS" PER UNIT OF PRODUCTION

(PRIVATE NONFARM PRODUCT)

	Total Other Costs	Rental Income	Net Interest	Capital Consumption Allowances	Indirect Business Taxes	Corporate Profits Taxes
	Points in total price index (1954=100)					
1929	17.0	3.5	3.6	4.8	4.2	0.9
1930	17.9	3.4	3.8	5.3	4.8	0.6
31	18.4	3.0	4.1	5.8	5.1	0.4
32	20.4	2.6	4.6	6.6	6.2	0.4
33	20.3	2.0	4.5	6.5	6.8	0.5
34	18.5	1.5	3.9	5.7	6.7	0.7
35	17.3	1.4	3.5	5.2	6.4	0.8
36	15.7	1.2	3.0	4.6	5.8	1.0
37	15.6	1.4	2.9	4.5	5.9	1.0
38	16.7	1.8	3.0	4.8	6.3	0.7
39	15.7	1.8	2.7	4.4	5.8	0.9
1940	15.6	1.7	2.4	4.2	5.7	1.7
41	17.2	1.8	2.1	4.0	5.5	3.8
42	18.6	2.1	1.8	4.1	5.3	5.3
43	19.4	2.2	1.5	4.1	5.4	6.2
44	18.8	2.2	1.3	4.3	5.7	5.4
45	18.8	2.4	1.2	4.5	6.3	4.4
46	18.9	2.6	1.2	4.1	7.2	3.9
47	20.9	2.7	1.4	4.7	7.4	4.6
48	22.6	2.9	1.5	5.4	7.9	4.9
49	23.5	3.3	1.7	6.0	8.3	4.1
1950	26.0	3.3	1.8	6.0	8.4	6.5
51	27.7	3.2	2.0	6.4	8.5	7.6
52	27.9	3.4	2.2	6.9	9.0	6.4
53	28.6	3.3	2.4	7.3	9.2	6.4
54	29.3	3.5	2.7	8.2	9.4	5.5
55	30.2	3.2	2.9	8.4	9.4	6.4
56	31.5	3.1	3.0	8.9	9.9	6.4
57	32.8	3.4	3.4	9.6	10.3	6.1

1/ See pp.133-134 of "The Relationship of Prices to Economic Stability and Growth," Joint Economic Committee, Oct. 31, 1958.

Source: Indexes computed by the author from Department of Commerce data on national income and product.

APPENDIX TABLE - CHART 3

PRICES, COSTS, AND PROFITS PER UNIT OF PRODUCTION

(FARM PRODUCT)

	Total Price (Index (1954=100))	Labor Cost	Other Costs 1/	Net Income of Uninc. Farms 2/
		(Points in Total Index)		
1929	59.6	7.8	15.5	36.2
1930	51.3	7.8	16.3	27.4
31	35.1	5.2	12.5	17.9
32	26.9	4.0	11.6	11.6
33	28.1	3.8	10.3	14.9
34	32.0	5.0	12.1	18.1
35	42.1	4.7	9.8	30.5
36	44.4	6.1	11.8	28.0
37	45.9	5.6	10.0	31.9
38	37.8	5.5	10.1	24.4
39	36.6	5.6	10.4	24.3
1940	39.2	5.9	10.6	26.1
41	49.8	6.7	10.8	34.6
42	65.5	8.1	10.9	49.1
43	81.5	10.9	12.5	60.4
44	81.4	11.6	13.1	60.0
45	89.5	12.8	14.7	65.3
46	104.9	13.9	11.0	82.8
47	122.8	16.7	14.8	92.0
48	123.6	15.8	15.8	92.2
49	105.2	15.7	19.3	70.6
1950	106.2	13.9	20.3	72.5
51	130.4	15.5	24.9	90.2
52	121.0	15.0	25.4	81.6
53	107.4	14.4	25.4	68.1
54	100.0	13.4	25.0	62.5
55	91.6	12.8	24.6	55.0
56	90.1	13.1	25.1	54.1
57	92.6	13.9	27.1	55.8

1/ Includes capital consumption allowances, indirect business taxes, and net interest. Does not include farm corporate profits after tax, profits taxes, and subsidies of government enterprises to farms. Total costs and net farm income will differ from total price by the total of these excluded items.

2/ Excludes profits and losses from inventory revaluation.

Source: Indexes computed by the author from Department of Commerce data on farm income and product.

APPENDIX TABLE - CHART 4

NATIONAL INCOME BY INDUSTRY

CORPORATE INDUSTRIES 1/

(Billions of Dollars)

	MANUFACTURING			MINING		
	Compensation of Employees	Corporate Profits Taxes	Corporate Profits After Taxes 2/	Compensation of Employees	Corporate Profits Taxes	Corporate Profits After Taxes 2/
1948	48.6	7.1	9.7	3.54	.40	.97
49	46.1	5.7	9.6	3.13	.26	.69
1950	52.5	10.9	9.5	3.44	.40	.92
51	62.4	14.4	10.0	3.91	.45	.90
52	67.4	11.6	9.5	3.97	.35	.72
53	74.8	12.4	9.0	4.08	.27	.65
54	71.1	9.6	8.8	3.74	.29	.69
55	78.0	13.1	11.9	4.06	.42	.86
56	83.9	13.2	11.3	4.53	.47	.96
57	87.7	12.3	11.2	4.68	.36	.82

	TRANSPORTATION			COMM. & PUBLIC UTILITIES		
	Compensation of Employees	Corporate Profits Taxes	Corporate Profits After Taxes 2/	Compensation of Employees	Corporate Profits Taxes	Corporate Profits After Taxes 2/
1948	10.29	.68	.79	4.12	.58	.81
49	9.88	.47	.66	4.38	.67	1.06
1950	10.42	.90	.98	4.62	.94	1.15
51	11.97	1.04	.84	5.11	1.34	1.29
52	12.51	1.02	.81	5.61	1.54	1.44
53	13.08	.92	.69	6.13	1.72	1.60
54	12.47	.59	.30	6.46	1.78	1.71
55	13.25	.84	.58	6.85	2.10	1.90
56	14.32	.82	.51	7.44	2.22	1.99
57	14.99	.72	.42	7.92	2.34	2.13

1/ Consist primarily of corporate businesses. 2/ Excl. inventory profits and losses.

Source: Department of Commerce

APPENDIX TABLE - CHART 5

NATIONAL INCOME BY INDUSTRY

UNINCORPORATED INDUSTRIES 1/

(Billions of Dollars)

	TRADE		SERVICES		
	Compensation of Employees	Income of Unincorp. Businesses	Compensation of Employees	Income of Unincorp. Businesses	Net Interest
1948	26.05	10.06	12.57	6.17	1.37
49	26.43	9.81	12.95	6.24	1.66
1950	28.33	9.57	13.82	6.68	2.12
51	31.23	10.94	15.14	7.01	2.36
52	32.99	11.20	16.22	7.45	2.72
53	35.09	10.88	17.30	7.99	3.39
54	36.09	10.99	17.96	8.13	3.63
55	38.62	11.51	19.61	9.38	4.18
56	41.66	11.35	21.58	9.87	4.81
57	44.01	11.30	23.17	10.38	5.21

	AGRIC., FORESTRY & FISHERIES		CONSTRUCTION	
	Compensation of Employees	Income of Unincorp. Businesses	Compensation of Employees	Income of Unincorp. Businesses
1948	3.35	18.03	7.44	2.61
49	3.18	13.19	7.27	2.66
1950	3.01	14.27	8.35	3.00
51	3.18	16.63	10.37	3.28
52	3.22	15.65	11.26	3.51
53	3.24	13.60	11.80	3.58
54	3.18	13.03	12.02	3.54
55	3.23	12.13	12.93	4.00
56	3.35	12.02	14.30	4.26
57	3.45	11.99	14.72	4.50

1/ Consist primarily of unincorporated businesses.

Source: Department of Commerce

APPENDIX TABLE - CHART 6

RATES OF RETURN

(ALL MANUFACTURING CORPORATIONS)

		PROFITS AS % OF SALES				PROFITS AS % OF INVESTED CAPITAL			
		Before Taxes		After Taxes		Before Taxes		After Taxes	
1948	1Q	11.5		7.2		26.8		16.8	
	2	11.1		7.0		25.2		16.0	
	3	11.0		6.9		25.2		16.0	
	4	10.9		6.8		25.2		15.6	
1949	1Q	9.9		6.1		20.4		12.8	
	2	8.5		5.2		16.8		10.4	
	3	9.5		6.0		18.8		12.0	
	4	9.3		6.0		18.0		11.6	
1950	1Q	10.1		6.2		19.6		12.0	
	2	11.8		7.4		24.8		15.6	
	3	13.5		7.6		31.2		17.6	
	4	14.9		6.9		35.6		16.4	
			*		*		*		*
1951	1Q	13.5	12.4	6.1	5.6	32.8	32.0	14.8	14.3
	2	12.8	11.7	5.8	5.2	30.4	29.7	13.6	13.3
	3	11.5	10.5	4.7	4.2	25.5	24.9	10.4	10.0
	4	11.0	10.1	4.8	4.4	25.8	25.3	11.2	10.9
1952	1Q		9.9		4.2		23.6		10.1
	2		9.2		4.2		22.0		10.0
	3		8.9		4.3		20.7		9.9
	4		8.6		4.4		22.2		11.3
1953	1Q		10.0		4.3		24.9		10.7
	2		10.4		4.4		26.4		11.2
	3		9.6		4.3		23.3		10.5
	4		6.7		4.0		15.8		9.5
1954	1Q		8.4		4.3		18.5		9.4
	2		8.9		4.7		19.8		10.4
	3		8.2		4.4		17.5		9.3
	4		8.2		4.7		18.3		10.6
1955	1Q		9.9		5.1		22.3		11.4
	2		10.6		5.5		25.0		13.0
	3		10.2		5.4		23.3		12.3
	4		10.3		5.6		24.6		13.5
				#		#		#	
1956	1Q		10.3	10.2	5.4	5.3	23.4	23.8	12.2	12.5
	2		10.3	10.3	5.5	5.5	23.8	24.2	12.8	13.0
	3			9.0		4.9		20.2		11.0
	4			9.3		5.2		22.3		12.6
1957	1Q			9.7		5.1		22.5		11.9
	2			9.4		5.0		21.6		11.6
	3			8.5		4.7		19.1		10.5
	4			7.6		4.4		16.8		9.8
1958	1Q			6.4		3.4		12.9		6.8
	2			6.8		3.8		13.9		7.8
	3			7.8		4.4		15.9		9.0
	4									

* New Series beginning 1Q 1951; # New Sample beginning 1Q 1956.
(Source: FTC-SEC)

X

IS THERE A SHORTAGE OF SCIENTIFIC MANPOWER?

Chairman, Donald C. Riley, Office of Statistical Standards

AN ECONOMIST LOOKS AT THE ENGINEERING SHORTAGE

By: David M. Blank, Columbia Broadcasting System, Inc.

Surely one of the most popular topics of recent years has been the "engineering shortage." By some accounts this shortage has been among the most serious problems facing the American people. When you try to evaluate the validity of these fears, however, you find yourself in some difficulty because there are several different kinds of shortage that the viewers with alarm have in mind, and the way you go about evaluating the existence or magnitude of the alleged shortage, much less the evaluation itself, will vary depending upon the kind of shortage you are discussing.

Value-Judgment Shortage

One very simple kind of shortage that often underlies public discussion is based upon a judgment that the United States "should", in some sense, have more engineers. The judgment is often based these days on a comparison with Russian output of engineers. Other times it is based on a view that this country should devote more resources to research and development, particularly basic research, and that a lack of engineers (and scientists) is holding this effort back.

This view of the engineering shortage is closely related to similar views held in other fields. For example, a view held by some in the housing field is that the United States "needs" two million new houses a year, almost double the rate at which we have been producing them in the last decade or so, in order to make possible the kind of upgrading of housing standards that the analyst feels ought to take place. A similar argument is often raised about the medical profession, when it is held that the output of physicians ought to be X percent higher than it is in order that some desired ratio of patients to physicians, perhaps that characteristic of New York or California, be reached.

The economist as economist can make only a limited contribution to analysis of such a view, for in essence, this view is based on a value judgment and the economist has little basis on which to argue about value judgments.

However, there are three aspects of the question which the economist can throw some light on. In the first place, he can point out that this value judgment is at variance with that of the market place. And this in itself carries certain implications.

Second, the economist can indicate what the results would be if in fact supply increased as the shortage-viewer wished, and nothing else changed. In the housing field, for example, if the two million houses per year were produced, they could not be sold unless prices were held substantially below cost, and this could not occur for any length of time without a government subsidy. And if new houses were produced and sold at this rate, it would probably result in a collapse of real estate values. To prevent this, the government would again have to provide a subsidy.

The same kind of reasoning holds for engineers. In view of the salary experience engineers have had over the last decade or two, an increase in the number of graduating engineers of the magnitude sometimes talked about in recent years would have created havoc in the profession. Presumably even the increased number of engineers would have found jobs but at what salaries and at what kind of work is another question. Indeed, one important contribution the economist can make is to point out that those who have wanted to increase substantially the number of graduating engineers ought, in all fairness, to insure that there will be means of satisfactorily employing those whom they wish to induce to enter the applied sciences. Presumably, the government is again the only hope in this direction.

A third contribution the economist can make to a discussion of this kind of shortage is an investigation of whether there are aspects of the market mechanism which are not working properly. For, to the extent that such is the case, the inducements produced by the market place may not yield a social optimum.

It is suggested by some[1/] that research and development, particularly basic research, is an activity that falls into this category. The argument is that it is almost impossible for the rewards attainable by research to be fully appropriated by the business units performing the research. Or, to put the matter in other terms, the benefits of research are received by many more than receive their direct financial returns. To the extent this is so, and there seems to be considerable merit to the argument, one can support on economic grounds substantial investment by the government in research.

Price-Control Shortage

A second kind of shortage occurs when there are restraints placed upon the operations of the market, so that price is not permitted to equate demand and supply. This sort of shortage, with which economists are most familiar, characterizes periods in which price (or wage) control is exercised by the government and prices are deliberately kept below equilibrium levels. Under such conditions, demand exceeds supply, at prices current in the market, and direct or indirect rationing of some form performs the rationing function of free market prices.

Closely allied to this kind of shortage is the market situation in which demand is characterized by monopsony, for while there is no excess demand in this situation, market prices remain below free market equilibrium levels.

Neither of these situations is of major significance in the market for engineers. There is no general control of wages in the engineering field, such as there was during World War II. Nor is there any single dominant purchaser of engineering services who could exercise substantial monopsonistic power; the largest single private employer of engineers probably does not employ more than two or three percent of the profession.

The closest we can come to the existence of wage control is the power the Defense Department possesses in the supervision of defense contracts, particularly in cost-plus-fixed-fee contracts characteristic of research and development programs. In the latter case, the Defense Department has the obligation to review all wage and salary schedules.

There is no clear evidence, indeed little evidence at all, that the power of the government over salaries of defense contractors has ever been exercised in a manner which has resulted in salaries in defense industries lagging behind salaries in other industries, or indeed, behind levels that would have been reached without such review. In fact, there is some current complaint of just the reverse, at least from universities, the complaint being that government-financed salaries are unduly high so that resources are being drained away from other, more socially valuable, activities.

But there is a serious question as to whether such control, even if effective, would really involve any generalized form of shortage. For the government, after all, supports directly or indirectly only a fraction, albeit a substantial one, of the applied scientists in the United States. If, in fact, the government had exercised its powers in a manner which seriously limited engineering wages in defense industries, defense contractors would find themselves unable to compete with other employers of engineering skills. But is such a condition a general market shortage or merely an unwillingness on the part of one employer to meet market prices? And surely the solution to this problem would be simply to permit engineering wages to rise in defense industries so that contractors could compete once again in the market place.

Domestic Servant Shortage

A third kind of complaint of shortage, and the one most often encountered in free labor markets, occurs when demand increases relative to supply, at the salaries paid in the recent past. "Then salaries will rise and activities which once were performed by (say) engineers must now be performed by a class of workers who are less well trained and less expensive".[2/]

The complaints of shortage in this situation are usually from consumers of the product or services whose demand is marginal in whole or in part, that is, who find it unprofitable or undesirable at the new higher prices to continue to purchase the commodity or service. To these consumers, a shortage has developed, and the evidence is that they are no longer purchasing the commodity or service in the amounts to which they have been accustomed. The most common example of this kind of complaint is the servant shortage about which there has been so much talk for more than half a century.

If this type of shortage has characterized the market for engineering services, one would expect to find the salaries of engineers to have risen, after adjustment for price changes. What kind of price adjustment to make in this situation is difficult to determine, but one reasonable technique is to use changes in salary levels of occupations in which there is no claim of shortage. When this is done, we find that engineering salaries in recent periods have been substantially lower, relative to salaries of other professions as well as to the labor force as a whole, than they were two or three decades ago. Even in the last decade there is little evidence of any marked improvement of engineering salaries relative to those of other occupations. There is some evidence, however, of a relative rise in engineering salaries around 1951 and 1952, when the effect of the defense program engendered by the Korean War was at its peak, and again around 1956 when the capital investment boom was at its peak.

In summary, the modest relative increases in engineering salaries since the late Forties, when the fear was that there were too many engineers, casts considerable doubt on the existence of a shortage of the magnitude that is implicit in much recent and some current discussion.

On this note, let me conclude with a brief quotation from a letter sent to me by an eminent engineering educator[3/], concerning the market for engineers. This educator stated, in part, that: "while there is a shortage in many categories of the labor force, ...[he] ...doubted that the engineering shortage was more severe than in other areas, such as skilled technicians and nurses... [He] also pointed out that approximately one-third of the young men entering college these days were heading towards engineering and science, and this was not an unreasonably small proportion, because we also need people to do other things than engineering and science... [Finally he] noted that young engineers are widely used for sales and service work, where their engineering is used only as a background. If there was really a critical shortage of engineers to do technical work, [he argued] wage rates would adjust themselves so that the technical jobs would pay more than the quasi-technical, which is not the case... [His] conclusion based on these and other similar points was that it seemed that the seriousness of the shortage of engineers and scientists was probably overrated, and that while engineering was a good career for those who like this type of work and have the necessary qualifications, young people should not rush into engineering merely because of the propaganda about a shortage, or because they think it is a short cut to riches."

1/ A. A. Alchian, K. J. Arrow and W. M. Capron, "An Economic Analysis of the Market for Scientists and Engineers", The RAND Corporation, 1958, pp. 68-71.

2/ David M. Blank and George J. Stigler, The Demand and Supply of Scientific Personnel, National Bureau of Economic Research, 1957, p. 24.

3/ F. E. Terman, Dean of the School of Engineering, Stanford University.

SHORTAGE OF SCIENTIFIC PERSONNEL AS SEEN IN GOVERNMENT

By: Thomas J. Mills, National Science Foundation

When the inevitable day on which I tried to put together some random thoughts for this paper arrived, this speaker was appalled. A session and a subject which sounded so attractive in July from the honeyed lips of the program arranger, Professor Hauser, suddenly presented all kinds of difficulties. The topic, as presented on the program, appeared much too broad to be adequately covered by a representative of a minuscule agency with a rather narrowly specialized interest in Government personnel. I therefore determined to exercise a speaker's prerogative and to select for my discussion those limited segments of the topic with which I am most familiar rather than attempt a balanced presentation. These deal with the Government first as an employer, and next as an instrument through which public policy becomes expressed.

While most of this audience is generally aware of the fact of "Big Government's" influence on employment of scientific manpower, recital of a few measures may serve to frame our subsequent remarks. Unfortunately, our statistics relate almost entirely to the Federal Government. Data on the hundreds of scientists and thousands of engineers employed by State and local governments are almost entirely lacking. (Parenthetically, this situation should not last much longer; a National Science Foundation sponsored survey will be undertaken by the Bureau of Labor Statistics to secure data from a representative sample of State governments at least within the coming year.)

The Federal Government is the direct employer of approximately 140,000 civilians in its engineering, physical sciences, biological sciences, and mathematics and statistics job series. About 100,000 of them are considered to be professionals as distinguished from engineering assistants, laboratory helpers, and other support personnel. They are engaged in all types of scientific and other activities with about one third of the professionals engaged in research and development. Most of them are employed in the three military departments, or in such civil departments as Interior, Agriculture, and Commerce. Their professional duties cover the widest possible range from research and field exploration through development, testing, teaching, technical writing, contract and procurement supervision, weather forecasting, production, etc., to management and administration. They are stationed in all States and in many foreign countries.

As a direct employer of scientists, the Federal Government employs about one in every six. Their fields of specialization reflect the areas of public concern: the extent to which Government is a major employer measures alternative employment opportunity. For example, about two out of every three meteorologists are employed by Government, but only about one in every 20 chemists is so employed. About one out of every 12-15 engineers works directly for the Federal Government. It can readily be seen that conditions of Government service are a major determinant of salary scales and working arrangements of American scientists and engineers.

Another aspect of Government service worthy of comment is its character of a "closed system." Many of its employees will be continuously employed in Government, while changing agencies for which employed. Government service is apparently attractive to many who are perhaps overly impressed by its job security aspects; others are attracted by its less competitive character; still others are influenced by its unique opportunities for research. Many will spend their working lives in Government service without ever seriously considering alternative employment. All of which suggests that the working conditions under which Government scientists and engineers are employed do not have to be strictly competitive with outside employment in terms of financial rewards, since the jobs themselves do not compete in the view of many Government scientists. (The "closed system" characteristics should not be overemphasized, and may well be less influential with the scientists than with others. The increase in scientific personnel in Government, and the growing intercourse between Government and industry and university scientists will tend to reduce this situation over time.)

Salarywise, the averages usually show the Government scientist as commanding more than the university and college faculty scientist, but less than the industrial scientist. The averages obscure experience, training, quality, and differing kinds of responsibility. This is another area sadly in need of better statistical measures. Available fragmentary evidence shows beginning Government salaries lagging behind industrial--although approaching them in selected fields where special incentives have been authorized. For positions at the higher levels of responsibility, Government does not try to compete seriously in salary terms with private industry. A recent survey by the Engineers Joint Council, for example, shows engineers' salaries at the top decile as $16,300 after 20 years of experience. Only a handful of Government engineers--less than one percent--will have attained after 20 years even the GS-15 grade, which commands a $12,800-$14,000 salary range.

There are, of course, no quantitative measures of shortage of scientists and engineers in Government. Rather the supply-demand situation must be deduced from job openings, recruiting difficulties as found in testimony of personnel officers, non-existence of names on employment registers, adoption of new recruiting methods, quality of recruits, establishment of incentives, etc. Since there are not direct measures, it is not deemed particularly profitable to review possible definitions of shortage in order to select an appropriate one in the economic sense. If we

have not satisfactory measures to evaluate an ideal definition, it may be more enlightening to note some of the bits of evidence while devoting most of our energies to development of better basic data.

In terms of the ability of Government to fill existing science and engineering positions, the present shortage situation seems less severe than has been true for the past several years. The business decline of 1957-8 did not result in appreciable lay-offs of industrial scientific and technical personnel, but it did decelerate new recruiting. At the same time larger graduating classes have increased the supply of newly trained workers, for whom competition has been especially severe. Furthermore, salary differentials in the hard-to-fill classifications and a general 10 percent salary increase in 1958 were important factors in recruiting and retention of Government scientists. New recruiting methods and incentives, some of them long standard in industry, are contributing to putting Government in a better competitive position. Such factors include use of paid advertising, payment of transportation expenses, campus recruiting, etc.

Interestingly enough, the present business recovery is already displaying some signs of making the Government's scientific personnel problem more serious. Industrial salaries continue to increase; campus recruiting in 1959 will likely be back at 1956-57 levels. Currently, the new National Aeronautics and Space Agency is expanding and reports some difficulties in recruiting scientists and engineers at even its "supergrade" levels, i.e., at salaries of $19,000 to $21,000.

Before leaving the topic of Government employment, mention should be made of certain reporting developments, which may well improve our information in this area. One of the few hopeful results of a series of committees and activities immersed in problems of Government scientific personnel over the past few years has been recommendations for periodic reports by the so-called "Young Committee"--The Committee on Scientists and Engineers for Federal Government Programs. The Office of the President's Personnel Management Advisor has been instrumental in urging the adoption of some of them. Within the past few months the Civil Service Commission and National Science Foundation have developed a roster of scientists and engineers in Federal service, which should produce at least annual data on numbers, specialties, training, functions, level, etc., for those in the higher grades. Another development has been the review of data requirements by an interagency group of data consumers and producers. While it is too early to evaluate this review, it shows some promise of leading to a reasonably comprehensive body of data on Federal scientific personnel in terms of types, functions, accessions, separations, etc., in the not too distant future.

In addition to its concern for direct employees, Government is expected to assume the broader role of assuring that adequate personnel resources are available to carry out those activities in science and engineering which are deemed important. Whether we like it or not, the American people have accepted as a basis for action the conclusion that we face a serious and continuing threat to our existence as a free Nation. A threat, once considered primarily a military one, is more and more being accepted as a political and economic one as well. Our position of leadership in science and technology is challenged: Government, as the instrument through which public policy is expressed, is looked to for leadership in new and unfamiliar areas.

A heartening result is the prompt development of certain governmental programs which are already showing results. The classic debate on Federal-State relations in the field of education is disregarded. Under the threat of foreign difficulties, the Defense Education Act passed the Congress this year with scarcely a seriously expressed qualm that the Federal Government might detrimentally interfere with education in efforts --as the law states--"to insure trained manpower of sufficient quality and quantity to meet the national defense needs . . ." Even before this legislation, Federal appropriations to the National Science Foundation, the National Institutes of Health, the Atomic Energy Commission, and other Federal agencies were having a significant impact on education. A democracy can act promptly; under strong stimulus, the dogma of a hundred years has been swept aside.

The Federal education support program has been preceded by the phenomenal growth of Government support of research and development. Federal expenditures for this activity will exceed $5 billion this year; as recently as 1950 the total was little more than $1 billion. Most of these funds support industrial R & D, although significant amounts are expended through universities and other non-profit organizations.

Finally, public concern with the progress of science and technology is reflected in the spectacular achievements in missile and satellite developments and in programs designed to organize Government to emphasize more effectively particular branches of science. Millions more than ever called upon the gods of mythology follow the progress of their namesakes in Jupiter, Juno, Atlas, Hercules, Nike, Thor, etc. We eagerly listen for favorable news of our Explorers, Vanguards, and Space Probes. A Government agency is reorganized and becomes the National Aeronautics and Space Agency. These and other actions are indicative of an emerging public policy which holds that science and technology must be strengthened.

All of these developments have necessarily focused attention on the demand, supply and training of scientific personnel. The public official is overwhelmed with opinions, generally unsupported by quantitative facts, as to the availability of engineers. A cut back in an experimental missile program and accompanying plant layoffs are seized upon as evidence of a surplus of technical

personnel. An unprecedented volume of help wanted ads is dismissed as company advertising at the expense of the taxpayer. The problem is stated as one of labor hoarding and poor utilization. At the same time, fragmentary indicators show new engineering graduates are able to command substantial salary differentials over their fellow graduates in non-technical fields. Increasing numbers of job orders for engineers appear in the public employment offices, not exactly the traditional medium for placement of engineers. Surveys of employer hiring expectations continue to show openings which it is expected can not be filled. Salary rates continue to drift upward. In this melee of opinion and part truths, it requires indeed the wisdom of a Solomon and the leadership of a Moses to sort out the facts and formulate the appropriate program.

It is refreshing that a few statistical measures, which throw some light on the demand-supply situation of recent years are now becoming available. Preliminary findings of a sample survey by the Bureau of Labor Statistics show an increase of about 112,000 engineers and 55,000 natural scientists in industrial employment between January 1954 and January 1957. (This total increase of 167,000 scientists and engineers included about 65,000 reported in the function of research and development.) During this same period about 70,000 baccalaureate degrees in engineering were awarded according to U. S. Office of Education data. In other words the employment of professional engineers by industry increased by 42,000 more than the additions to the supply through normal training channels without any allowance for death, retirement from the labor force, etc. Nor can the excess be assumed to have been drained from competitive employments. Fragmentary--again--evidence indicates that both public and educational institution employment of engineers slightly increased during this period. We conclude that industry continues to upgrade technical personnel into engineering jobs, even though the emphasis over the past 20 years has been on recruiting the more formally trained engineer. It is further hazarded that this condition has been forced upon industry through inability to recruit graduate engineers in the numbers sought.

Public policy has accepted the premise that scientific training should be expanded and its quality improved. The answer to the question of "how much" is still debated. At one extreme are those who believe there are vast resources of unexploited talent which should be so trained. No over supply of trained engineers is feared, since such training is considered useful in many pursuits and--as they say--"why be concerned that training in engineering ought to lead to employment in engineering when we do not hold forth a similar standard for majors in history, languages, or literature." On the other hand are those who sincerely feel that an adequate expansion in numbers will be attained through normal population growth, and that maintenance of quality training is most important. In between are those who fear overemphasis on technical training at the expense of other occupations. Say they, "True we need qualified engineers, but we also need high quality teachers, statesmen, social scientists, etc." It is in this welter of opinion that a Government department proposes a program to expand training, the Budget Bureau imposes a further judgment, and the Congress disposes.

We are all generally aware that a complicating factor in the production of scientific personnel is a long "lead time." Secondary school curricula now virtually require a decision in favor of scientific training in the ninth or by the beginning of the tenth grade, if college entrance deficiencies are to be avoided. Add a four year college course for a total of seven years for a minimal professional training. If graduate work leading to the Ph.D. degree is pursued--as is increasingly necessary in the basic sciences especially--another 4-5 years is required. It seems reasonable to conclude that the junior high school years are the critical ones from the standpoint of the supply of newly trained scientific personnel becoming available some 7-12 years later. Implications for Government programs are serious. When such usual automatic controls, as prospective salaries, innate inclination, social status, prestige, etc., become subordinated to Government policy in vocational choice, it becomes especially important that Government exercise the greatest wisdom in its appraisal of requirements a decade in the future.

In concluding these rather disjointed remarks, certainly the scientific and technical occupations have been growing at tremendous rates, although not evenly in all fields. We do not know the extent of present shortage, if one can be said to exist. But perhaps this is not the relevant question. From the standpoint of Government, the critical question relates to the future. It centers on the likely future requirements for scientists and engineers in terms of programs public policy accepts as important. It asks what is likely to be our supply on the basis of demographic trends, training facilities, alternative opportunities, etc. The gaps between these estimates 10 years or more in the future should be the measure of public concern and the basis of remedial Government programs. It is in this area that the Government is currently most concerned and in need of the tools which will permit development of better measurements.

Bibliography

Blank, David M. and Stigler, George J., The Demand and Supply of Scientific Personnel; 1957; National Bureau of Economic Research; New York.

Cole, Charles C., Encouraging Scientific Talent; College Entrance Examination Board; 1956, New York.

Engineers Joint Council, Demand for Engineers 1957; Policy Guidance Task Committee on Surveys of EMC; 1957; New York;

_____, Professional Income

of Engineers - 1958; Engineering Manpower Commission; EJC Report Number 112; January 1959; New York.

Los Alamos Scientific Laboratory of the University of California, National Surveys of Scientific Professional Salaries; 1949-58; Los Alamos, New Mexico.

National Science Foundation, Federal Funds for Science VII; NSF 58-30; 1958, Washington;
_____, Scientific Manpower in the Federal Government, 1954;
_____, Scientists and Engineers in American Industry - January, 1957; NSF 58-38; Washington.

U. S. Department of Labor, Labor Market and Employment Security; Bureau of Employment Security; Washington.

U. S. Department of Health, Education and Welfare, Engineering Enrollments and Degrees; U. S. Office of Education; annual; Washington.

U. S. Civil Service Commission, Occupations of Federal White Collar Workers; Pamphlet 56-1; June 1958; Washington.

THE PROBLEM FROM THE STANDPOINT OF THE SCIENTIST AND TECHNICIAN

By: M. H. Trytten, National Academy of Sciences

In commenting on the topic which is the theme of this symposium from the point of view of the scientist and the engineer, it should perhaps be noted at once that the question of numerical shortages in these professions is somewhat alien to their typical attitude towards their professions. The typical scientist, and to some extent the engineer, seldom thinks in quantitative terms about the members of his profession. In spite of the trend towards team research and towards massive research programs involving highly organized groups, the scientist is even yet fundamentally a lone worker. His job is a highly personal one to him. As a consequence, the differences among scientists are more important than the similarities. These differences relate to personal competence, personal specializations, personal working habits and personal techniques. The scientists is much more inclined to think in terms of individuals and especially the outstanding contributors than in mass or quantitative terms. The scientist knows too well that the top 100 scientists can outproduce the next 100 by a substantial factor. Hence, he is much more conditioned to worrying about quality than he is about quantity, and this is particularly true about the leadership among the scientists.

It seems important to begin this discussion in this way because the question may well be whether one should look at this matter as a statistical problem in which one seeks the most accurate means of relating our needs or our demands for personnel in these categories against supply or whether one examines the evolution of technology as a process of the whole.

In approaching the question in this fashion I assume we are not mainly concerned with the instantaneous relationship between supply and demand as of January 1, 1959, about which one can not affect the emerging discrepancies too greatly, but about the situation which would appear to be confronting us some time in the future when those now at various stages in the educational process emerge to play their parts in the evaluation of science and in the maintenance and development of our technology. The significance of this question lies mostly in its future implications.

To the scientist, the process of the evaluation of scientific discovery is more a natural point of departure. To the economist the ebb and flow of economic forces play a more dominant role in his thinking and no doubt this point of departure leads to the statistical approach. Neither of these is adequate alone in understanding the forward movement of technology and in deducing what the future may hold in respect to the relationship between our future supply and our future needs for scientists and engineers.

One thing seems clear. As a nation we have moved definitely far away from the pre-world War II position in which the beneficent processes of our society were dependent upon to supply adequate personnel resources in these fields. Since World War II we have moved far toward subsidies for individuals and institutions by private and public moneys to stimulate the flow of personnel in these categories. It is becoming more imperative that we somehow shall be able to understand the underlying dynamics of both supply and demand. When we attempt to buy up the difference between supply and demand we are more likely to be concerned about quantitative evaluation of the deficit between the two.

The scientist is, as indicated above, more likely to be aware of the dynamic character of scientific discovery not only in leading to economic exploitation of the new discovery but in the uncovering of new areas of scientific investigation. The rich discoveries of the Elizabethan era probably resulted from no discernible economic motivation but did lead to the evaluation of the basic structure of science. The discoveries of Oersted and the subsequent invention of the motor and the dynamo probably were events in which economic motivations were inconspicuous. Yet modern technology is, to a large extent, based on these two instruments. The electrical engineers and technicians needed in their production, use, and improvement are legion.

Similarly the vacuum tube, the vacuum pump, the processes of fission and fusion, and innumerable others are similarly cases in point. Additional examples from every field of science can be adduced.

None of these developments was unaffected in their rate of development by economic considerations and more recently by political considerations. But each possessed a dynamic impetus intrinsic to itself. The mere existence of a new discovery opens up areas of discovery, on the one hand, and opportunities for application to commerce and industry, on the other. Some of these latter opportunities are so attractive that exploitation occurs whether the economic climate is generally favorable or not.

Scientists are aware, too, of the insistent demands of a growing technology for information which has not yet been uncovered. This, too, is one of the underlying dynamic

forces which contribute to the expansion of technology. In recent years this factor has been of growing importance as a factor. Examples are numerous. The petroleum industry has resorted to more and more scientific methods in petroleum exploration and in the processing of petroleum. Born of necessity the resulting information has had significance beyond the problems which gave rise to the new knowledge. High temperature metallurgy called for by the developments in nuclear technology and jet propulsion is another case in point. Similar cases abound in the mathematics of complex wave structures and supersonic shock waves, in the researches related to the major medical problems, in nuclear chemistry and a host of other fields.

One more consideration is becoming more insistent in its effect on the technological activities of our own country in particular, but on others as well. This is the regenerative phenomenon in which technological developments create technological problems which must be solved if our society and its scientific activities are not to choke themselves to death. Examples of this also abound and are more or less obvious on the face of events. Growth of population and of communities, of transportation, of communications, and all the phenomena of expansion present massive problems. The road network now evolving has made civil engineers a shortage category, but also presents intricate engineering problems in traffic control, in the design and structure of roads and a host of satellite problems of a social, economic and political nature, as well as a technological one. The growth of the information corpus presents new and massive problems in the storage, analysis and retrieval of information which again must rely on technology for their solutions. The logarithmic rate of increase in our demands on food, fiber, minerals, and water are foreshadowing problems in the supply of raw materials for our civilization, which will no doubt also grow logarithmically in their demands for scientific and technical effort and for appropriate personnel. We are even now changing from a "have" to a "have not" nation in respect to certain key commodities and raw materials essential to our industries, such as iron ore, and other metals. These shortages create the need for much research to reduce to the minimum our dependence on lines of supply which are vulnerable and may be in jeopardy at some future critical time.

This is by no means a catalog of the intrinsic factors which are inherent in science and its environment and which contribute to the expansionist tendencies of science. They should, however, indicate that these factors exist. It is quite apparent that they are powerful in the determination of the level of demand which will probably exist in the future. It should also be apparent that as yet there is no good statistical method of measuring them.

Any judgment amounting to a quantitative assessment of future manpower needs must find a way to grapple with these intrinsic dynamisms. They may be said to be those dynamic attributes of science which because scientific information exists, further technological efforts arise to utilize it.

More recently, two new motivations have arisen which lead substantially to increased activities in these fields. Because scientific information is now more generally considered to be of eventual benefit and economically profitable, industry has progressively become more research minded, and in increasing cases is willing to support research and development as a venture. Whether the research supported is basic or applied, programmatic or purely exploratory is less likely to be a matter of concern. The net effect is an increased stimulus to the growth of research laboratories, the creation of new knowledge, and hence the stimulation of the growth of technology as a whole, and consequently greater demands for personnel.

The other development is the extraordinary increase in the use of public money in the support of research and development, some of it obviously to serve specific military ends and purposes, some for the general welfare, some for the augmentation of the personnel resources and the increase in the rate of research itself.

These new developments create a new problem in the measurement of our future requirements because fundamentally the motivations in these two cases arise from policy judgments made by scientific laymen (in boards of directors, in the Government, both in Congress and the executive branch) and may at times be subject to fluctuations in response to gross economic and other changes in the climate of research support.

From what has been said, it might seem that the necessary conclusion is that the prediction of the course of technology and of science is a vain hope. It is true that we do not yet know how to assess the future impact of a given discovery, such as, for example, the practical use of semi-conductors in research and in industry. We know the transistor has been a powerful instrumentality and its applications are in their infancy. But no one in 1840 could have foreseen the manifold uses of the dynamo and the motor and their effect on later technology. No one could have foreseen the development of electronics following Dr. Forest's invention of the three electrode vacuum tube. No one could have foreseen the nuclear powered submarine following the discovery of nuclear disintegration.

And yet the growth of technology has a remarkable consistency when viewed as a whole historic movement. Whether one views one of its indices or another the characteristics are

similar. The growth of technology and its underlying scientific structure is logarithmic. Even as far back as the days of steam power, and before electrical power took over, capital investment in steam power installations approximately doubled every ten years. The rate of employment of scientists and engineers has followed a logarithmic curve. The rate of production of persons trained in these technical specialties follows a similar curve. The production of persons trained to the doctorate in our country has increased roughly seven per cent per year for many decades, with interruptions only for major events such as World War II.

It seems likely that an adequate analysis of these past trend lines might lead to some understanding of the major determinants of growth, and of the limiting factors. These may turn out to be the per cent of the gross national product allocated to supporting certain types of scientific or technical activities. The limiting factor may be personnel. It may indeed be personnel of a special creative type. The role of a particular limiting factor may historically be more dominant at one period than at another.

Perhaps this discussion would not be complete without introducing one somewhat philosophical element. Assuming that in the long run the well-being of our scientific and technological effort will depend mostly on its cutting edge - its creative research scientists and its inventors and innovators - should every effort be made to search out, train and support such persons. Should we have conscious policy that we can never have too many of such persons actively engaged in the exploration on frontiers of science? Certain countries in the world appear to be following such a policy now. In our own country there has been a tendency to move in this direction. Obviously, the more we approximate such an accepted policy the more rapidly we are likely to accelerate the growth of our scientific knowledge and the technology on which it is based.

As one looks at this matter, therefore, in perspective both looking forward and backward in time, it would seem probable that our scientific evaluation and its technological associated development are from their very nature going to expand at least as rapidly as in the past. To a large extent this will be due to the explosive growth of scientific knowledge. It has recently been estimated that this knowledge increases presently at the rate to double every nine years. Not only will this momentum continue, but this knowledge itself will exert continuous pressure to increase activity in both basic and applied science.

The relatively new factors of political, economic and military interest should add impetus to this rate of development as they have since World War II especially.

The two emerging problems which will increasingly preoccupy attention and stimulate technical effort are, first, the problems arising from the increasing needs of our society for energy, raw materials, water and other resources, and second, the self-induced problems of an increasingly massive technology.

With respect to the first, there is in the long run no solution except through research. We shall increasingly be concerned with augmenting water resources, creating new substitute materials, new energy sources, and in more shrewdly using and controlling the use of what resources we do have.

In respect to the second, we shall be more concerned with storage, and retrieval of the massive flow of information arising from our scientific efforts in all fields of investigation. We shall be more concerned with problems of land use, highways and their effects, transportation channels, the effects of new technologies such as radiation hazards and high altitude and other novel environments, with relatively unsolved scientific problems of new areas of the globe, such as the tropics and the high latitudes, and with the new scientific problems associated with supersonic speeds, high temperature and low temperature, high pressure and other new realms of phenomena in which practical technology is now concerned. Most of these new problem areas are not linear but logarithmic in respect to the exploratory efforts which they will require.

One other facet of this matter should be discussed. This relates to the relationship of the individual in our society, to his scientific and technological environment and to his degree of familiarity with and understanding of the society in which he lives. This again points to the kind of education he receives in his school and college days.

It is quite possible for a person to live and be happy with no knowledge of science. But it is also possible to live and be happy with no knowledge of Greek and Latin, of the Renaissance, and even of the nature of modern society. But we generally assume that an intelligent citizen should be educated both for the sake of his own richer life, and so that he can function as a responsible citizen in a democracy. But we are by no means agreed as to the ingredients of his education.

This much is, however, obvious. The world into which young people emerged from their educational experience even as late as a generation ago differs from the present world perhaps more than the world of 1920 differed from that of many generations prior to that time. Assuming that education should be relevant to the world in which the student emerges, it would seem that education should have changed accordingly.

I presume that no one would seriously argue that it has. Most of the difference between life a generation or two ago and today is either technological in nature or caused by technology. Presumably the new importance of this aspect of life should have been apparent in a revised educational program.

There has been much concern over this matter recently as evidenced by massive studies aimed at improved curricula in certain sciences, legislation such as the Defense Education Act and discussions in the press and in other settings. However, most of the concern has been focused on pre-professional training.

The relevance of this point here is that probably no single change in our demands for personnel at home in the sciences could present such massive requirements for properly trained personnel as would a wholly adequate curriculum in the sciences at elementary, secondary, and college levels.

In summary, the scientist and the engineer, while becoming more conscious of the statistical aspects of the growing demands of our society for more persons trained in their disciplines, are much more conscious of the need for quality. They are primarily interested in the functioning of persons in their disciplines as individuals and are aware of the great difference between individuals in the contributions they may make. But they are aware also of the evolution of science and technology as a phenomenon with its own intrinsic pressures for expansion and growth. They are aware of the new and expanding role, and the increasing challenges to science lying just ahead. If science and technology are to continue to grow adequately to meet these challenges, undoubtedly the awareness of the layman of science must continue to grow. This is why scientists generally are becoming much more concerned over education, not only for prospective students in these specialties, but for the layman as well.

ON OBTAINING AN ANSWER TO THE QUESTION: "IS THERE A SHORTAGE OF SCIENTIFIC AND ENGINEERING MANPOWER?"

By: Philip M. Hauser, University of Chicago

The advent of Sputnik about a year ago shocked a United States, unprepared for the event, into a number of frenzied reactions. Two of these reactions bordering on the hysterical, were glaring non-sequiturs in respect of the quality of U.S. education and of the supply of U.S. scientists and engineers. The USSR presumably achieved its triumph in space rocketry because U.S. high school and college training had gone "soft" and did not match Soviet educational standards; and because the U.S. did not have enough, as well as not able enough, scientists and engineers.

The USSR achievement certainly was based on Soviet strength, rather than U.S. weaknesses. The failure of the U.S. to be the first in space rocketry as in many other scientific and engineering developments could conceivably be, and probably was, not an indication of inferior U.S. education or of shortages in scientists and engineers but, rather, a manifestation of the way in which these human resources were allocated. Our scientists and engineers were busy on other matters--including ways and means of cramming more horsepower, more chrome and longer fins into and onto automobiles.

To the pressure for more and better information about scientists and engineers generated by the contrast between U.S. and USSR achievements in rocketry and its implications for national defense, is that produced by the continuing and heightening interest in the role of scientists and engineers in our expanding and increasingly complex economy. The need for information about these highly specialized human resources for "peacetime" purposes is at least as important as that for national defense and, in modern "total war" is in a large measure actually coterminous with national defense needs.

Even if it were demonstrable with available data that there is at present a shortage of scientific and technical manpower, there would still remain large areas of ignorance as to the reason for such a deficiency. Would such a shortage be the result of decreasing, or if not decreasing, in any case, inadequate proportions of students entering the scientific, engineering and technical professions or arts? Of if there were a shortage, would it merely reflect the changes in the age structure of the population? These have in recent years included decreasing proportions and even absolute numbers of persons of labor force entrance age, as a result of the depressed marriage and birth rate of the 1930's. Because of the changing age structure of the U.S. population it has been possible in recent years for the number of new entry scientists, engineers and technicians, as well as new entries to other occupations, to decrease even if the proportion of persons reaching working age entering the occupations increased.

These essayist observations are designed briefly to portray at least parts of the context in which the question was raised: "Is there a shortage of scientific and engineering manpower?" Is there?

The answer is that there are conflicting answers to the question.

In recognition of the inadequacies of the information available about the number and quality of scientists, engineers, and technicians in relation to demand, the President's Committee on Scientists and Engineers, on the one hand, and the Bureau of the Budget through the National Science Foundation, set out to develop a program for the collection of the desired data. An Advisory Panel was jointly established by these agencies to review and evaluate the available information and to make such recommendations as might be indicated for ways and means of improving it.[1] The Advisory Panel was assisted in its work by staff members of the Surveys and Research Corporation which served as its secretariat;[2] and by representatives of several Federal Governmental Agencies who participated as consultants to the Advisory Panel.[3] The results of the labors of the Advisory Panel are reported in A Program for National Information on Scientific and Technical Personnel.[4] The discussion which follows is based on the work of the Advisory Panel, on which the writer was privileged to serve as Chairman.

Available Data

Some data are, of course, available on scientific, engineering and technical personnel. These were inventoried and described in the Report to which reference has been made. They are available as part of the statistical and general informational output of such agencies as the Bureau of the Census, the Bureau of Labor Statistics, the Office of Education, the National Office of Vital Statistics, the National Research Council and of one of the sponsoring agencies of the Advisory Panel—the National Science Foundation.

In addition to these agencies with relatively broad and comprehensive types of information, a number of other agencies have more restricted or spot survey types of information. These include such organizations as the Department of National Defense, the National Education Association, the American Chemical Society, the National Society for Professional Engineers, the National Manpower Council, the National Bureau of Economic Research, and the American Society of Engineering Education.

In general, the available data include more about the supply of scientific and engineering personnel, than about the demand for them; more about their quantity than about their quality; more about their placement than about their actual utilization; and more about them in cross section than in longitudinal flow. The Advisory Panel's review of available information disclosed an especially serious deficiency in information about the demand for scientists and engineers, and, also, serious shortcomings in the techniques of anticipating and measuring future demand. Also badly deficient is the available information on the flow of personnel into the scientific, engineering and technical occupations, on their actual utilization, on their quantity, including their innate capacities and training, and on their attrition. Finally, as a climactic indicator of the ignorance we have accumulated about our scientific and technical human resources, it early

became apparent to the Panel that there was not even available an adequate framework for the identification of scientific and technical manpower. No adequate classification of scientists, engineers, and technical personnel is available with which even to delimit the specific occupations of concern, let alone one which provides a frame which analyzes work content, or describes specific jobs for either statistical or use purposes.

The Advisory Panel, after a fairly intensive review of the information on scientific manpower, concluded, with good reason, that "the data found available are in general painfully inadequate."[5] It also concluded that "In the face of the foreseeable long-term manpower needs, the Nation cannot long afford to leave the gaps in its information unfilled and the deficiencies in data unremedied."[6] It was with this perspective and with the instruction from the sponsoring agencies to create a plan for an adequate body of data, that the following framework for a statistical program was developed.

Statistical Program

As a preliminary to the design of a blueprint for an adequate statistical program, consideration was given to the basic issues which confront the agencies faced with the formulation of national policy on scientific and engineering manpower questions. These turned out to be centered largely around problems of recruiting, training, utilizing (including conscripting) compensating, and conserving these relatively scrace and expensive human assets. The issues practically all involved long run considerations. And, as the issues were explicitly stated and specific questions put, the inadequacies of available data became increasingly apparent.[7] The basic issues formulated and the specific questions which were raised led to the development of a conceptual framework around which the statistical collections and reports were to be organized.

Conceptual Framework

The key elements in the conceptual framework which emerged from the deliberations of the Advisory Panel were "supply", "flow" and "demand".

1. Supply. "Supply" was elaborated into four major categories, each with sub-components, namely: 1. current supply; 2. potential increments to supply-short term; 3. potential increments to supply-long term; and 4. "all other." This categorization of the supply factor was made operational through linkage with the concepts used in the labor force measurement as developed by the Bureau of the Census and cooperating agencies. More specifically the total population 14 years of age and over was dichotomized into those "in the labor force" and those "not in the labor force." These groups were then sub-divided into 8 groups, with some sub-total combinations, as indicated:

Total Population 14 Years Old and Over by Labor Force Categories--

In the Labor Force:
1. Employed or Seeking S & T (Scientific & Technical positions)
2. Trained in S & T
3. Could be trained in S & T
4. Could not be trained in S & T

Not in the Labor Force:
5. Trained in S & T
6. Being trained in S & T
7. Could be trained in S & T
8. Could not be trained in S & T

These labor force groupings were then related to the four major categories of supply as follows:

Total Population 14 Years Old and Over by S & T Supply Categories--

A. Current Supply
   1. Employed and seeking work in S & T
B. Potential Increments to Supply - Short term
   2. In Labor Force trained in S & T
   5. Not in Labor Force, trained in S & T
   6. Not in Labor Force, being trained in S&T
C. Potential Increments to Supply - Long term
   3. In Labor Force, could be trained in S&T
   7. Not in Labor Force, could be trained in S & T
D. All Others
   4. In Labor Force - could not be trained in S & T
   8. Not in Labor Force - could not be trained in S & T

With these eight groupings of the population cross-classified by labor force and scientific and technical status identified, the next step was that of determining the kinds of information needed for each group. In general, it was concluded that five types of information were desired; namely, statistics relating to:

1. economic characteristics, including scientific and technical jobs, industries in which utilized, compensation and fringe benefits, functions performed, working conditions and the like;
2. demographic characteristics--location, sex, age, race, etc.;
3. educational and training characteristics
4. aptitude and skill characteristics, including level of mental ability;
5. community characteristics--the prestige accorded science, engineering and technical vocations, status, attitudes, etc.

2. Flow. Concern with scientific and technical manpower necessarily must take the time dimension into account. Such personnel is the product of relatively prolonged training, and the demand for them is vitally affected by trends in research, technology, and economic and social organization. Any cross section picture of scientific and technical personnel must, therefore, be supplemented by longitudinal data--by inflow and outflow information.

The supply of scientific and technical personnel, of course, is first of all a function of the size and composition and changes in size and composition of the total population. Starting with the measurement of total population, happily available even on a current basis for larger areas of the United States, the inflow of scientific and technical personnel is dependent on the following: 1. the formal educational system; 2. special training courses and apprenticeship; 3. informal training; 4. immigration (of trained personnel); 5. reentry of retirees and others not

in the labor force; 6. transfers from other occupations.

The supply of scientific and technical personnel is subject continuously to outflow as well as to inflow. The chief categories of outflow are: 1. mortality, 2. retirement, 3. emigration (of scientific and technical personnel), 4. transfer to other occupations, 5. departure from the labor force, other than through death or retirement.

The information needed about inflow relates both to those who come into the scientific and technical manpower pool and the characteristics of the source. For example, inflow from the formal educational system requires information both about the trainees and the educational system which trains them. Similarly, information is needed both on the trainees and the institutional provisions which provide scientific and technical persons from sources outside the formal educational system--in-plant training and various types of special training programs. The Panel report spells out in some detail the kinds of information needed about each of the inflow sources to provide the needed longitudinal picture.[8]

To complete the flow picture, statistics are also needed on outflow--on the attrition of the scientific and technical manpower pool through death, retirement, emigration, transfer and departure from the labor force, for reasons other than death or retirement. These types of data are also described in the report.[10] Attrition rates, by source, must necessarily be part of any program designed to measure the pool and changes in the pool of scientists, engineers and technicians.

3. Demand. "Demand" was conceived of as consisting not only of unsatisfied but, also, met requirements for scientific and technical personnel. Thus, persons employed in scientific and technical jobs are part of the demand. Moreover, only "effective demand" was considered --demand that was, or could be met, under actual or obtainable salary schedules and working conditions. Demand, for operational purposes, was regarded as consisting of three components: 1. scientific and technical personnel currently employed; 2. established but unfilled scientific and technical positions; and 3. need for scientific and technical personnel recognized by employers (willing and able to attract personnel) but for which positions were not yet established.

To measure demand in cross section, the actual employment of scientific and technical personnel and current needs for additional personnel must be ascertained. Such information is presumably available from employers and potential employers. To measure future demand, however, is not so simple. Even the willing employer respondent is likely not to know. To assay future demand, knowledge of the factors which produce the demand and the trends in these factors is necessary. In consequence, an effort was made to analyze the "determinants of changes in demand" to provide a frame for anticipating future demand. These elements were classified into 7 groups: 1. population growth, 2. economic growth and change; 3. growth in urbanism; 4. levels, content and quality of education; 5. research and development; 6. requirements of national defense; 7. "other factors" including such basic things as the changing role of the United States in world affairs, governmental policies, organization structure of scientific activity, changing practices in use of scientific and technical personnel and the relevant democratic value system.

Trend analyses of each of these factors in relation to demand for scientific and technical personnel, it was posited, could provide a firmer basis than now exists for the quantification of future demand.

Projects

The conceptual framework which was developed served as the basis for the specific project recommendations which were made. The specific statistical and research operations which are proposed are, with a summary of the available data, presented in the Report in relation to, and in the context of, the concepts described. In all, 15 projects involving some 27 surveys or researches are recommended. They are presented in the Report in three classifications: first, in priority order by functional description;[10] second, in relation to the conceptual framework developed--under supply,[11] flow,[12] and demand;[13] and third, in operational terms, in the "Summary" under headings of: 1. "extensions of existing programs of data-collection and analysis"; 2. "new surveys conducted through establishments"; 3. "new surveys of population"; and 4. "new research in special problem areas."[14]

In the priority category of "most urgent" surveys the following are recommended:[15]

1. Identification of Scientific and Technical Occupations
   Analysis of work content and job descriptions in scientific, technical and related fields, to develop a set of job definitions and a classification, and to develop methods for identifying these occupational categories with adequate precision when conducting such surveys.
2. Pool of Scientific and Technical Personnel
   A periodic survey, using appropriate sampling techniques, of organizations and agencies of all types which are employers of scientific and technical personnel, to obtain information on the numbers and characteristics of such personnel.
   From time to time intensive analyses, through personal interviews, of a sub-sample of present and former scientific and technical personnel, to determine the number, occupation, and economic and personal characteristics of these persons; and to obtain information concerning the nature of movement into and out of the supply.
3. Periodic Study of the Demand Outlook for Various Categories of Scientific and Technical Personnel in Each Major Activity
   The study would include the analysis of: (a) employment and production growth trends; and (b) the changing roles of particular

classes of scientific and technical personnel, to develop current estimates and forecasts of demand for such personnel by occupation and educational levels.

These projects would, among other things, provide a better answer than is now possible to the question of whether there is a shortage of scientific and technical personnel. The first of these projects would make possible good identification of the pool of manpower that is included when the rubric "scientific, engineering and technical" is used. The second would provide a much more adequate measurement than now exists, and on a periodic basis, of the actual pool of employed scientific and technical personnel, on their number and their characteristics, including their mobility or flow. The third project would provide a benchmark measurement of current demand, and, eventually, some projection of demand which, in relation to supply, could provide other than a speculative answer to the question of alleged "shortage."

These projects alone could greatly increase our present knowledge about scientific and technical manpower. But they, by themselves, would fall far short of answering the crucial questions which must be answered if policy and action are to be based on sound knowledge. The other projects proposed were also regarded by the Advisory Panel as "urgent"—even if not as "most urgent." On a functional basis these "urgent" projects are listed as follows:[16]

Studies of Supply and Utilization

4. 1960 Census Survey of College Graduates: Scientific, Technical and Other Specialized Personnel
   A special direct survey of a large sample of the persons recorded in the 1960 Census enumeration as college graduates or as persons currently or last employed in scientific and technical positions, whether college graduates or not, to determine relationships between training and subsequent occupation.
5. Sample Population Survey
   An occasional survey similar to the Current Population Survey of the Bureau of the Census, to provide source data for estimating both the current supply of scientific and technical personnel and potential increments to it, as well as net turnover.
6. Analysis of Data from National Register of Scientific and Technical Personnel
   Continuing analysis of Register data on social and economic characteristics of scientific personnel listed, to determine age, levels of education, functions, type of employer, scientific specialization, income level, etc.

Extension of the Coverage of School Reporting Comparison of Data from Diverse Sources, and Other Educational Data

7. Extension of School Reporting
   Extension of the annual and biennial statistical reporting program of the Office of Education to obtain a more complete view of the nation's educational system, with particular reference to private schools, college enrollments and course enrollments.
8. Comparison of Data from Diverse Sources on School Enrollment and Educational Attainment.
   An intensive statistical analysis to compare and coordinate Office of Education data on enrollments and graduates with Bureau of the Census data on enrollment and highest grade of school completed.
9. Reports of Government Technical Schools
   Initiation of annual reports from the appropriate governmental agencies concerning technical schools for higher education operated by them, the character and extent of science instruction provided and the utilization of personnel trained in science and engineering; to determine the level and extent of such training.
10. Analysis of Factors in School Retention
    Studies of the factors related to drop-outs of students in various fields and of different ability levels, at particular grades in school, seeking particularly to measure the effects of financial factors.

Research into Aptitudes in and Attitudes Toward Science and Engineering

Several related types of inquiry: (a) to determine the distribution of youth in terms of general mental ability and other more specific aptitudes; (b) to measure comprehension of scientific subject matter; (c) to ascertain the attitudes of students toward science and engineering at several educational levels; and (d) to ascertain student's occupational conceptions and choice.

12. Follow-up Surveys of Recent College Graduates
    Surveys of samples of recent college graduates made two, four, and six years after graduation to learn the beginning career patterns of young graduates as these relate to their major fields of study and other personal characteristics.
13. Research into Community Attitudes Toward Scientific and Technical Personnel
    Studies of the attitudes of American communities toward science and engineering and the individuals who are occupied in these fields, to determine their importance in influencing occupational choice.

Surveys of Curricula and Physical Facilities Available for the Teaching of Science and Engineering

14. Surveys of Curricula in Science and Engineering
    Periodic surveys of the courses offered in the science and engineering programs of schools and colleges.
15. Surveys of Physical Facilities for the Teaching of Science
    Periodic surveys to determine the adequacy of laboratory and other equipment used for the teaching of specific science and engineering courses in the nation's schools and colleges.
16. Research into the Quality of Science Instruction

Studies to determine the effectiveness of science and engineering teaching, with the principal purposes of isolating the background factors and personal qualities which characterize the most effective teachers.

Surveys of Training Programs in Science and Engineering Outside the Formal Education System

17. Surveys of Training Programs for Scientific and Engineering Technicians in Industrial and Commercial Establishments

Sample surveys to obtain information on formal courses given to technicians within such establishments, including apprenticeship programs for training technicians, to determine their importance in supplementary technical training.

18. Surveys of Advanced Scientific Training Programs of Industrial and Commercial Establishments

Surveys of the formal, advanced scientific and technological training programs conducted by business establishments for their scientific personnel possessing college degrees or equivalent backgrounds, to determine the level, content, and relative importance of such programs in the nation's scientific training effort.

19. Study of Background of Scientific and Technical Personnel Who Have Had No Formal Training

Identification of selected scientists and technicians who have received no formal training and the analysis by case studies of their background and the means through which their status was obtained; in order to appraise the possible importance of the informal training methods disclosed.

Surveys of Changes in Supply from Immigration and Emigration, Retirement, Other Causes and Death

20. Reports on Immigration and Emigration of Scientific and Technical Personnel

Initiation of annual reports on the scientific and technical personnel of foreign citizenship who immigrate into or emigrate from this country. Study the extent and character of emigration of scientific personnel who are citizens to determine its relationship to the overall supply.

21. Interview Survey of Retirees

Special survey to be repeated periodically of a sample of retired scientific and technical personnel, designed to ascertain their ability and willingness to resume active employment full-time or part-time if needed.

22. Additional Study of Occupational Mortality Rates Among Scientific and Technical Personnel

Intensive study of differential mortality among occupational groups of scientific and technical personnel, to improve estimating methods.

Studies of Demand

23. Correlation Regressions Among Major Determinants of Demand for Scientific and Technical Personnel

Historical studies of the regression relationships between demand series and trends in several factors hypothesized to be major determinants of demand with a view to determining the degree of predictive value which each of them may possess.

24. Surveys of Employers Expectations of Demand for Scientific and Technical Personnel

Periodic surveys of the expectations of present and potential employers of scientific and technical personnel as to the numbers and types of such personnel they will need at specified times in the future, and as to the major factors that are expected to determine their future needs with particular attention to lessons that may be learned for improving methods of projection.

Of these proposed projects, eight are new surveys involving the establishment approach. Two of these (numbers 1 and 2 above) are in the "most urgent" category and the remainder are "urgent" (numbers 9, 17, 18, 19, 21, 24). Three additional surveys also are new ones. These would involve the population approach including a follow up of the 1960 Census (numbers 4, 5, 12). In addition to the surveys, four new researches are included in the list, one in four parts (10, 11a,b,c,d; 13, 23). Finally, nine extensions of existing programs of data collection and analysis are also listed. These would involve addition to programs already under way (3, 6, 7, 8, 14, 15, 16, 20, 22).

With the diversity of projects and the number of different agencies involved, it is clear that the proposed statistical program will necessarily involve provision for implementation and coordination of program. For this reason the Advisory Panel, after pointing to the need for the cooperation of Federal, State and local government agencies and also private organizations, recommended that an appropriate Federal agency be given explicit responsibility for coordinating that part of the program which involves the Federal Government.

Concluding Observations

The situation in respect of knowledge about scientific and technical personnel is reminiscent of the situation in the 1930's in respect of knowledge about unemployment. In the absence of a direct measurement of unemployment various conflicting indirect metrics were obtained which, as the record shows, varied with the interest of the estimator. With the advance of sampling and survey methods utilized in the Monthly Report of the Labor Force, initiated by the W.P.A. and continued by the Bureau of the Census, a direct and reliable measurement of unemployment has been available for almost 20 years. In consequence political squabbles about the facts of unemployment are now mostly a matter of history, and policy and action in respect of employment and unemployment can now be based on fact rather than conjecture.

Adequate facts about the supply of scientific, engineering and technical personnel in relation to demand are not now available. These highly trained and skilled human resources have become of vital moment to national defense and to national economic advance. At the present time considerations of policy and program designed to assure an adequate and capable corps of scientific and technical personnel involve, among other things, debate about the facts themselves. And facts cannot be ascertained by means of polemics and majority vote.

The science and art of statistics has progressed to a point where the needed facts about scientific and technical personnel can be had with relatively little cost and effort. A program has been planned through which the needed information can be collected. It would seem that the course of wisdom is to proceed to obtain them.

Footnotes

1. Members of the Advisory Panel were: Philip M. Hauser, University of Chicago, Chairman; Philip M. Coombs, Fund for the Advancement of Education; Henry David, National Manpower Council; Coleman R. Griffith, American Council on Education; Merriam H. Trytten, National Academy of Sciences and National Research Council; Ralph J. Watkins, Brookings Institution; Dael Wolfle, American Association for the Advancement of Science.

2. The staff who served were: Stuart A. Rice, Libert Ehrman, Robert H. Mugge, Lorand D. Schweng, and Jeremy C. Olin.

3. Agencies represented by the consultants were: National Science Foundation; President's Committee on Scientists and Engineers; U.S. Department of Labor, Bureau of Labor Statistics; U. S. Department of Commerce, Bureau of the Census; U.S. Department of Health, Education and Welfare, Office of Education and National Institutes of Health. The Bureau of the Budget, Office of Statistical Standards, was represented by observers. In addition, the Office of Defense Mobilization and the Interdepartmental Committee on Scientific Research and Development were requested to submit suggestions on the adequacy of and the needs for scientific and technical personnel information on behalf of those agencies not directly represented.

4. Published by the National Science Foundation on behalf of itself and the President's Committee on Scientists and Engineers, NSF-58-28, August, 1958.

5. Ibid., p. 6
6. Ibid., p. 6.
7. Ibid., pp. 15ff.
8. Ibid., pp. 23-24.
9. Ibid., pp. 40-44.
10. Ibid., pp. 7-11.
11. Ibid., pp. 25-28.
12. Ibid., pp. 34-45.
13. Ibid., pp. 58-60.
14. Ibid., pp. 62-63.
15. Ibid., p. 7
16. Ibid., pp. 8-11.

DISCUSSION

William H. Miernyk, Northeastern University

Mr. Blank has indicated some of the difficulties involved in defining a shortage of a particular type of manpower. In his paper, however, he has reiterated the conclusion reached earlier by himself and George J. Stigler that "the record of earnings would suggest that up to at least 1955 there has been no shortage - in fact an increasingly ample supply - of engineers." 1/ This conclusion was based on the finding that up to that year the average salary of engineers had gone up relatively less than the average wage level, and also less than the average increase in earnings of other professional groups.

At least two questions might be raised about the method employed by Mr. Blank and Professor Stigler in reaching this conclusion. First, is it appropriate to apply product market reasoning to the labor market without important qualifications? Second, how relevant is a comparison of the average salaries of engineers and other occupational groups?

It is entirely true that if the demand for a commodity in a competitive market is in excess of the supply there will be an increase in price. And price adjustments occur very rapidly in some types of product markets. Studies of the labor market by Reynolds, Shister, Myers, Shultz and a number of others have indicated rather clearly, however, that there are significant differences in the operation of labor and product markets. For example, wage and salary adjustments occur much more slowly, for a variety of reasons, than do price adjustments in many commodity markets. Even during the capital boom of 1956, when much was heard about a shortage of engineers, the average salary of engineers did not rise sharply. There are, however, anti-pirating conventions among employers; many of long standing. To be sure, these are not entirely adhered to, but even when they listed many engineering job openings employers did not rush to bid up salaries.

One reason for this is that employers are quite aware that aggressive "price competition" of this sort would be a losing game. The supply of engineers cannot be increased quickly by raising the average level of salaries. Various other recruiting tactics were employed during this period in an effort to fill job openings, but it is not surprising that employers failed to engage in an aggressive, and self-defeating, price war in trying to attract more engineers.

To turn to the second question, a comparison of movements of the average salaries of engineers and other professional groups tells us relatively little about the demand for and supply of any of the occupational groups involved. The relevant comparison is

of starting salaries. New entrants into the labor market are typically more mobile than those already established. Hence there is more open bidding for the services of the former than of the latter. During the time when the "engineering shortage" was a topic of widespread discussion, the starting salaries of new engineering graduates did increase substantially more than those of other college graduates. Indeed, this compounded the problem of "salary telescoping" among engineers. Those with long service saw the salaries of new entrants into the profession rising rapidly while their own salaries went up slowly. This encouraged a greater leakage of engineers into non-engineering occupations, further contributing to the imbalance between engineering manpower requirements and the available supply at that time. The leakage of higher-salaried engineers, if they were replaced by younger men at lower rates of pay, also might have exerted downward pressure on the average level of salaries of all engineers.

As Mr. Mills has pointed out there has been much less clamor about an engineering shortage since the beginning of the recession of 1957. At the same time, new engineering graduates have not experienced serious difficulty in finding employment. There are indications, however, that recruiting activities are now being stepped up. While this year's June graduates will undoubtedly not have the range of choice available to those who graduated in June 1956, recruiters are already far more active than they were at this time last year.

Both Mr. Mills and Mr. Trytten have stressed the dynamic, and long-run aspects of the engineering and scientific manpower problem. Mr. Blank, however, has expressed concern that undue stress on the "engineering shortage" will induce too many students to enter the field, and make it difficult for them to find employment upon graduation. It would indeed be unfair to encourage too many students to enter colleges of engineering. But the proportion of engineers and scientists in the labor force has been steadily increasing, and there is a strong likelihood that this proportion will continue to increase in the future. 2/ Given some of the current problems of elementary and secondary education, I doubt that an excessive number of students will be encouraged to seek careers in engineering and science.

Using existing data and the techniques of analysis which have been employed up to the present, the question of an "engineering shortage" could be debated ad infinitum. Professor Hauser's paper goes to the heart of the problem - the lack of reliable data to provide a definitive answer to the question of whether or not there has been a shortage of engineering manpower. It is encouraging that something is to be done about this data deficiency.

The Advisory Panel established by the President's Committee on Scientists and Engineers and the National Science Foundation has done its job well. If the statistical projects proposed by the Panel are carried out, we should know much more about the market for engineering and scientific manpower in the future than we have in the past. A number of important policy issues hinge upon better information about the demand for and supply of scientific, engineering and technical personnel. And as Professor Hauser says, "facts cannot be ascertained by means of polemics and majority vote." If the broad program proposed by the Advisory Panel is carried out with dispatch, it will obviate future meetings of the kind we have had tonight to discuss the controversial question of whether or not there is a shortage of technical manpower.

---

1 /The Demand and Supply of Scientific Personnel, New York: National Bureau of Economic Research, Inc., 1957, p. 29.

2 /Since 1870, for example, employment in science and technology has increased about 17 times as fast as the labor force. See Trends in the Employment and Training of Scientists and Engineers, National Science Foundation Publication 56-11, p. 4.

XI

MEASURING INDUSTRIAL AND POPULATION GROWTH: TOOLS FOR MARKETING AND SMALL AREA RESEARCH

Chairman, Dorothy Swaine Thomas, University of Pennsylvania

Demographic Aspects of Military Statistics—Jacob S. Siegel and Meyer Zitter, Bureau of the Census

A Method of Projecting the Number of Households in Small Areas—William Hodgkinson, Jr., American Telephone and Telegraph Company

Effectiveness Of Our Tools For Estimating Population Change In Small Areas—Carl M. Frisen, California Department of Finance

Discussion—Chester Rapkin, University of Pennsylvania

DEMOGRAPHIC ASPECTS OF MILITARY STATISTICS

by
Jacob S. Siegel and Meyer Zitter
U. S. Bureau of the Census

This paper is intended to focus attention on military personnel statistics as a body of basic statistics important to the field of general demographic analysis. Demographic statisticians, and statisticians in general, have paid little attention to this material although data on military population enter into a number of problems which engage their attention. We shall concern ourselves here with a discussion of some of these problems, the sources of statistics on the military population, the basic reporting systems, and the comparability and consistency of the data.

The military population has represented a numerically significant segment of the population of the United States continuously since 1941, as a result of World War II and the ensuing period of sustained "cold war". It has exceeded one million for all the years since 1941, rising to over 12 million at the climax of World War II; as of July 1958 it numbered 2.6 million. Gross changes in the group are impressive; even in the year 1957-58, when the armed forces declined by about 200,000, there was a turnover involving about 1,200,000 men.

The inclusion or exclusion of armed forces distinguishes the three principal types of population figures for the United States published by the Census Bureau. These are (1) the civilian population, (2) the total population resident in the United States (including armed forces in the United States), and (3) the total population including armed forces overseas. Since the first and second types of estimates are obtained by subtraction from the third type, it is necessary to have figures on the total number of United States military personnel and the number inside or outside United States to prepare these types of estimates. One problem, here, arises from the fact that the armed forces of the United States include persons who were inducted or who enlisted from outside continental United States. An accurate separation of the armed forces into "inside" and "outside", consistent with the census definition of residence inside and outside continental United States, presents its special difficulties.

Of the three types of population figures noted for the United States above, only the first two are ordinarily considered appropriate for states and local areas. Thus, the Census Bureau publishes only the total resident population and the civilian population for states, and excludes overseas armed forces from its official counts and postcensal estimates for states. Similarly, the Census Bureau publishes only total resident population for its projections of state population. Yet, contrariwise, in their censuses the states of Massachusetts and Kansas include absent military persons with their families living in the state and exclude military persons stationed in the state. In general, this definition of population is demographically unsound as well as impractical. The military man's previous household may have been relocated or terminated as a result of his induction or enlistment; he may marry while in military service or on leaving it and establish his own home; for other reasons he may not return to his preservice home. To assign military persons back home would be unrealistic, therefore; it would also be operationally very difficult to enumerate accurately "at home" so large a population "away from home". Furthermore, for most of the purposes for which population figures are used, they should represent the population which public and private facilities in the area must serve; that is, the resident population. On the other hand, for some analytical purposes and particularly as a stage in making certain population estimates, it is useful to consider a hypothetical population representing the sum of the civilian population and persons away in the armed forces. For example, this would seem an appropriate type of figure for certain family studies, for measuring the potential voting population, or as an intermediate element, for deriving projections of local population.

Derivation of census counts and postcensal estimates of the total and civilian population of states and local areas requires data on the number of military persons "resident" in each state or area. Because of the mobile character and special living arrangements of the armed forces, the collection and compilation of data relating to them, particularly data for small areas, present special problems.

If state and local population estimates are made by the component method or if population changes are to be analyzed in terms of components, another type of military statistics is needed—net movement of civilians to the military population. The component method involves the estimation of population by combining separate allowances for births, deaths, net civilian migration, net movement of civilians to the military population, and net change in the number of military persons stationed in the area, with the figure from the last census. The net movement component represents the balance of the number entering military service from the civilian population and the number leaving military service for civilian life. There are two basic sources of figures on "net movement" — data on the "preservice residence" of the armed forces, by states, from the Department of

Defense and data on so-called "net credits", by states, from the Selective Service System. Theoretically, to represent "net movement", the change in these series over the estimate period must first be derived and then increased by the small number of deaths of persons from the area serving in the Armed Forces.

It is clearly important to take military changes into account in the analysis of population changes in states and local areas. An increase in the size of the military station, or more exactly a large net military migration into the area, may account for a substantial portion of the population growth in the area. Net military migration may be derived as the difference between net total migration (the difference between total change and natural increase) and net civilian migration (a by-product of the component estimating procedure). For the Middle Atlantic States, net military migration between 1950 and 1957, for example, was as large a component of population change as net civilian migration; on the other hand, for the Pacific Division, net military migration was negligible in comparison with net civilian migration.

Let us now consider directly the various sources of military personnel statistics in terms of the separate reporting systems producing the data. There are seven separate reporting systems producing military statistics, of which the five branches of military service considered together (Navy, Marine Corps, Army, and Air Force in the Department of Defense and the Coast Guard in the Treasury Department) are the most important. The other principal sources are the Selective Service System and the decennial censuses. Taken in combination, these provide a wide variety of military statistics on a regular basis.

Most of the data useful for demographic analysis from the Defense Department and the Selective Service System come as by-products of the administrative needs of these agencies; very little information is made available directly for so esoteric a purpose as demographic analysis. Two main procedures are used by the military branches to produce their "demographic" data; one is the use of sample surveys taken in the field at various time intervals and the other is the use of the regular central reporting system or central file in which 100 percent of all personnel are covered. The Department of the Army and the Department of the Air Force obtain data on the number and geographic distribution of their military personnel—that is, the number assigned to and located at specific duty locations—from a complete accounting based on "morning reports". They use sample surveys to obtain information on such characteristics as age, marital status, educational status, etc.

The Department of the Navy and the Coast Guard develop most of their information from personnel records kept in the central file in Washington and do not resort to the use of sampling procedures. The compilation of Marine Corps statistics is primarily the responsibility of the Marine Corps itself, although the Marine Corps is administratively part of the Navy Department.

The Army conducts its survey four times a year. This survey is designed to give a five percent return of all Army personnel, the sample of individuals to be included being determined by a random process on the basis of the last two digits of their military service numbers. Questionnaires are sent to all military installations (inside and outside continental United States) with appropriate instructions to the personnel clerk at each installation. The latter generally has overall responsibility for the completion of the survey at each location. The exact procedure has changed over time. At one time, personnel clerks used to fill out the schedules from individual personnel records and transmit them to Washington. Now, however, the serviceman is required to fill out most of the items personally, and, before transmitting the schedules, the personnel clerk checks the serviceman's answers against his personnel folder and reconciles any differences. Although the survey is taken every three months, there is no fixed time schedule for the individual items and their frequency varies in accordance with the needs of the Department. Some demographic items are included frequently. An item like State of pre-service residence, on the other hand, has relatively low priority and is not likely to be included more than once a year.

The sample survey of the Air Force is similar to that of the Army although now it is conducted only three times a year. It, too, is designed for a five percent sample return. The demographic items included are also similar to those in the Army survey, e.g., age, marital status, educational status, etc.

It should be pointed out that no attempt is made to achieve uniformity in the wording of questions common to both the Army and Air Force surveys. In some items, such as age, variations in the wording hardly affect the consistency of the results since the alternative questions call for essentially the same thing (e.g., age last birthday). In an item such as state of pre-service residence, on the other hand, variations such as "legal residence", "voting residence", or "state of induction"

tend to reduce the comparability of the results to some extent. Unfortunately, such differences in wording exist not only between the Army and the Air Force surveys taken at a single date, but may also appear in surveys taken by the same sources from time to time.

To what extent the differences between the branches in procedure of data collection and in the wording of questions affect the comparability of the data is not known. For the most part, the data from the various branches are believed to be additive and the data from each branch are believed to be consistent over time. In combination, these sources present the data needed to secure a picture of the current size, distribution, and composition of our armed forces.

There is within the office of the Secretary of Defense a small statistical staff, which in a very limited way serves as a central source of military personnel statistics and which combines selected types of statistics from the individual military departments. Its principal function, however, is to provide the Secretary of Defense and his staff with the statistics necessary for the overall operations of the Department of Defense. Its functions do not include the role of central statistical office for military statistics. Thus, the Bureau of the Census relies on the individual sources for most of the required personnel statistics. It may be possible in some instances to secure military station data for specific local areas from the local military commander (or local military public information officer). The information available locally may enable one to distinguish between those living on post and those living off post or may even indicate the exact distribution by area of residence—the type of information needed to tie in with the 1950 Census materials.

In addition to the five branches of the military service, there are two other important sources of military statistics—the Selective Service System and the decennial censuses of the United States Bureau of the Census. Of minor importance as a source is the Census Bureau's Current Population Survey which affords limited data on military personnel in continental United States living off post or with their families on post.

The Selective Service System is, of course, the machinery through which young men are drafted into the armed forces. Its records contain much information about registrants and draftees, and provide a guide to the available military manpower pool for the Department of Defense. In addition to maintaining the regular registration and selective service records, the system receives information from the armed forces on those who enter the military outside Selective Service channels. Furthermore, regardless of whether an individual was inducted through Selective Service or enlisted, a record (home address report) is sent to the local Selective Service Board for each individual separated from service. Thus, Selective Service may serve as a source of information on entries into, and separations from, military service; estimates of changes in the overall strength of the armed forces; and figures on the number of persons serving in the armed forces from each state, referred to as "net credits". The latter can be used to develop readily estimates of the net movement of the civilian population into the Armed Forces from each state over a period of time, previously noted as an important component of population change.

These military data from Selective Service sources may be compared with the data obtained from the Department of Defense. With respect to overall strength, recent Selective Service figures indicate an armed forces level on July 1, 1958 some 300,000 lower than the "actual" mid-1958 figure (approximately 2.6 million). The difference between the two totals varies from period to period. The differences appear to be greater during periods of rapid build-up in our armed forces; for example, in July 1951 the Selective Service figure on total strength was about ½ million lower than the 3.3 million reported by the Department of Defense. Apparently, then, there is some "catching-up" during periods of relative stability of the armed forces.

Both for the number serving from each state as of any current date and for the estimated net movement into the armed forces from each state for any postcensal period, there are substantial differences between the Selective Service series and the Defense Department series. Even when the Selective Service data are adjusted to the total U. S. strength figures from the Department of Defense, the net movement estimates for 23 states differ by 20 percent or more in a comparison for 1950-56. Although the Defense Department figures may be taken as the standard here for demographic purposes, they too are subject to error arising primarily from misassignment with respect to residence and from sampling variability. Even though the percentage error in the net movement component may be large, however computed, the resulting error in the population estimates is relatively small, and in most instances substantially smaller than the error involved if this component is ignored completely, as is sometimes done.

For local areas, such as counties and cities, the Selective Service System represents the only source of an actual tabulation of persons currently away in the armed forces, although estimates are possible on the basis of

the Defense Department data. For counties, or cities which have their own selective service boards, these figures may be readily available; otherwise, special tabulations from basic records are necessary.

For its state population estimates the Bureau of the Census employs the Defense Department's preservice residence data, and for its county estimates preservice residence is approximated by taking that proportion of the state total represented by males of military age in the county in 1950. Selective Service data are not used.

As mentioned earlier, military persons living in continental United States are counted in the decennial censuses of population. In addition, in 1950 military persons living abroad were also separately counted. In the 1950 Census, military persons living in barracks on military posts were covered by use of special census forms and included in the population count for the area in which their camp or post was located. These forms were distributed to all such persons through the personnel officer (or his designate) at each camp or post, filled out by the individual, and returned to the Census Bureau through the personnel officer. Military persons living off post or in regular type quarters on post were enumerated on regular schedules by regular enumerators at their place of residence. A similar procedure was carried out in counting the Armed Forces abroad.

The Navy presents a special problem because many seamen are "resident" on ships afloat or in port. According to the census rule, persons assigned to ships in port were to be enumerated at the port and, hence, as part of the port city's population. The chance assignment of ships to ports at the time of the census means that there is some instability in the count of population of cities having ports in which military vessels are berthed.

The 1950 Census reports do not generally show military population directly, but such data can be obtained from the labor force tables as differences between the total and the civilian labor force. The 1950 Census provides information on the size and geographic distribution of the Armed Forces within continental United States, by states, counties, urban places, and urbanized areas, and on the size of the Armed Forces outside U.S. The census also provides information on the age-sex distribution of the military population inside the United States, outside the United States, and in each state, and on a wide variety of other demographic and social characteristics.

The Census figure for the total military population (as best as can be derived) was about 10 percent lower than the Department of Defense figure for April 1, 1950, with the continental figure being about 7½ percent lower. There are some large differences between the two sets of figures by states, a few of which result from differences in basis of allocation by states (e.g., District of Columbia). For its national and state population estimates for both 1950 and later dates, the Bureau of the Census employs the military station figures from the Department of Defense (with the exception of the District of Columbia, Virginia, and Maryland). In this way, comparability is achieved between the military figures for various estimate dates.

In the 1950 census, all military persons were theoretically enumerated as of the place where they lived, regardless of place of assignment or place of preservice residence. This was necessary if the principle of "usual place of residence" was to be followed and if counts of total population by residence were to be obtained for all census tabulation areas. For purposes of population analysis, this type of information is, in a sense, more satisfactory than the type of data provided by the Department of Defense which relate to the number of military persons assigned to, or on actual duty at, the various military posts in the area. In most cases, this number will agree approximately with the number of military persons living in the area. However, for many metropolitan areas, although the overall number of military persons stationed there may be about the same as the number residing there, the distribution within the area may be substantially different on the two bases. For the Washington, D. C., Metropolitan Area, for example, the percent of the military population assigned to the District of Columbia according to the military reports, differs substantially from the proportion residing there according to the census. State population figures may also be affected when military installations are located close to state boundaries or cross state lines. For its occasional work in estimating the population of metropolitan areas and counties, the Bureau of the Census generally employs the military figures from the census for April 1950 and the Department of Defense figures for postcensal dates, adjusted if necessary, for comparability with the census figure.

Census materials are, of course, available only once every 10 years; this limits sharply the usefulness of the census as a source of military statistics. The census can currently serve as a basis for determining the distribution of military personnel by residence within metropolitan areas and for adjusting Department of Defense figures for postcensal dates from a "place of work" to a "place of residence" basis. The plans for the 1960 Census with respect to the military population, so far as they have

been developed, are essentially similar to those for 1950. The "labor force" basis of securing the data suggests again a wide range of tabulation detail.

For lack of time, we have not considered here a number of facets of our general subject, such as, the use of Selective Service registration data and Department of Defense data to evaluate the accuracy of census data; analysis of the demographic and economic characteristics of the military population per se in comparison with those for the general population; problems in the compilation and analysis of migration, marital status, household, labor force, education, and vital statistics data for the general population resulting from the fact of a large military population; and other topics.

A METHOD OF PROJECTING THE NUMBER OF HOUSEHOLDS IN SMALL AREAS

By: William Hodgkinson, Jr., American Telephone and Telegraph Company

The purpose of this paper is to outline a methodology for breaking down a given projection of households for an area into individual household projections for subdivisions of that area. It assumes that population projections are available for each subdivision. Household projections are needed in marketing studies and in various fields of long-term planning. Households, of course, are the natural market for long lists of commodities and services. The procedures to be described do not constitute the basic type of research which is involved, say, for estimates of the future number of households in the United States. Rather, they are such as to make use of such estimates.

A household is defined in the Population Census as one person or a group of not necessarily related persons occupying a dwelling unit. In the Housing Census there are data on occupied dwelling units, and since the same definition of a dwelling unit is used in both Censuses, there is an identity of concept. This identity unfortunately did not carry over into the tabulations for 1950 because of necessarily separate processing, but the differences between the two are relatively small. The critical part of the definition is bound up in the concept of "dwelling unit." In the main a dwelling unit is a house, an apartment, trailer, etc.

The population census in this country has always been taken on a de jure basis, so that wherever they might be enumerated, persons are attributed to their permanent residences. One sees a reflection of this fact in the counts of "non-resident dwelling units" in the housing census.

The Bureau of the Census uses a dichotomy for classifying the population: everybody is assumed to live either in a household or in a quasi-household. The essential point is that a quasi-household is not a household. The population in quasi-households is made up of the institutional population (hospitals, prisons, homes for the aged, etc.) and "other." This latter group covers a wide variety of abodes - those who live in all but the smallest lodging houses, in hotels, those quartered in lumber and mining camps, and those members of the resident armed forces who live in barracks. The following are the figures as of April 1, 1950:

Type of Residence	Population	Percent of Total
Total	150,697,361	100.00
Households	145,030,888	96.24
Quasi-households	5,666,473	3.76
Institutions	1,566,846	1.04
Other	4,099,627	2.72

If estimates are available of the future population in an area, and if in addition one knows for that area the future ratios of population to households (often referred to as "population per household," or "persons per household"), the problem of this paper is obviously solved: these ratios when divided into the population give the number of households. Any procedure which attempts to derive household from population estimates must concern itself with this relation.

Any error in such ratios is of course transmitted to the household estimates as a reciprocal. Thus a 10% underestimate of the ratio produces an 11.1+% overestimate of households, etc. It is significant, therefore, to have some idea of possible variations in the ratio. To this end a distribution was made of all 3,102 counties (or equivalents) in 1950 according to the magnitude of the ratio of total population living in households to the number of households. (These ratios were published to two places of decimals in Table 42 of the individual State parts of Volume II of the 1950 Census of Population.) Data derived from this compilation are presented in Table 1. It is to be noted that the distribution is of counties and not of households, since the emphasis here is on variations in the averages between counties. The mean of the averages for all counties is 3.54 (compared to the mean size of all households in the nation of 3.38). The range in the distribution is from 2.19 persons for Esmeralda County, Nevada, to 5.12 persons for Leslie County, Kentucky, or 2.93 persons. This range is over 80% of the mean and 7.7 times the standard deviation of 0.38. The coefficient of variation is 10.7% for the nation but is naturally less in all divisions except the extensive Mountain division. Time does not permit more than mere mention of the well-known regional variations in the ratio. It is evident, if proof were needed, that account must be taken of area variations in average household size if projections of the number of households are to do more than indicate a gross order of magnitude.

Even if projections are given of average household size for the nation or other large area, it remains a formidable task to make direct projections of average household size for a small area. Population and households, the numerator and denominator of the measure, are influenced by different factors or by the same factors in varying degree.

The problem is simplified by projecting what is termed herein, "relative household size." This is a ratio of two ratios: the numerator is the ratio of total population to the number of house-

holds for the small area, while the denominator is the same ratio for the large area for which projections of households are available. Two characteristics of this ratio are to be noted (see equation (1) below). In the first place, extensive cancellations take place upon its formation; secondly, the ratio is also identical with that obtained by dividing the proportion of the large area's population contained in the small area by the proportion of the large area's households contained in the small area.

The cancellations assist in the analysis of a given situation and hence permit more intelligent forecasting by clearing away extraneous detail which masks the essential point. A notable example of cancellation is the historic decline in average household size which has persisted off and on from the first decennial census in 1790. It is of course cancelled out since it tended to affect all areas. A further point not to be overlooked is that any breaks in the continuity of the Census record due to definitional or procedural changes also tend to cancel, since both large and small areas are similarly affected. The varying treatment of quasi-households is an example.

What is left after the cancellations is a record over time of average household size in the small area as a proportion of that in the large area. This of course is precisely what is needed to estimate average household size in the small area, since by assumption the average size of households in the large area is known. The ratios of relative household size tend to fluctuate in time about unity, since any residual trend represents a divergence in average household size between the two areas. This fact suggests that the widening of such divergencies (the ratio moving further from unity) can be of only limited duration. Plotted as a time series on a chart, the line for unity will, in effect, operate as a magnet. The following are three examples of exceptional deviation from unity which has persisted for a long time. The reasons, however, seem fairly evident:

	Aroostook County, Me.	Dukes County, Mass.	Nantucket County, Mass.
1890	1.24	.73	.72
1900	1.24	.70	.70
1910	1.23	.71	.71
1920	1.23	.78	.72
1930	1.27	.81	.83
1940	1.23	.87	.89
1950	1.19	.85	.85

The measure, relative household size, has served its purpose as such, when, from its past performance and what other considerations come into play, it is projected into the future. Simple graphical projections seem indicated, with readings made from the charts. It is at this point that the second property of this ratio which was mentioned above, becomes of service. For if we divide the proportion of large area population in each of the component small areas by the figure for relative household size, we obtain the proportion of households contained in each small area. The sum of these proportions will be precisely unity at any census date. For projected values, however, this sum will deviate from unity because each projected relative household size was determined independently. Hence a small adjustment is required. Application of the proportions, adjusted to add to unity, is the final step, yielding a projection of households for each component of the large area. The algebraic relations will now be set out, together with a practical computing procedure.

For simplicity of notation a subscript i is suppressed in writing each lower case letter. Thus "q" stands for "q_i". The letter i denotes a particular county, the counties always to be taken in the same sequence, i = 1, 2, . . ., N. N is the number of counties in the state. Summations, denoted by S( ), are all from i = 1 to i = N. In general a different set of the relations below will hold for each point in time.

p = population,	h = households,
P = S(p),	H = S(h),
q = p/P,	r = h/H,
S(q) = 1;	S(r) = 1.

It is assumed that H and the N values of q are given. Relative household size, denoted by s, is defined as:

$$(1)\quad s = \frac{p/h}{P/H} = \frac{p/P}{h/H} = \frac{q}{r}\text{, so that}$$

$$(2)\quad r = q/s = h/H.$$

Summing equation (2) we have:

$$(3)\quad S(r) = S(q/s) = (1/H)S(h) = 1.$$

At any census equation (3) must hold exactly. If for the projection period the values of s are determined independently of one another, however, equation (3) will not hold exactly. Writing as primed characters those quentities subject to adjustment and assuming that the adjustment can be spread proportionally over all counties, we have, with K a constant:

$$(4)\quad (1/K)S(r') = (1/K)S(q/s') = (1/KH)S(h') = 1.$$

Multiplying each member by K we see that

$$(5)\quad K = S(q/s').$$

Equating the third member of (3) to the second member of (4), and also equating the second members of (3) and (4), we obtain:

(6) $h = (H/K)(q/s')$, and

(7) $s = Ks'$.

Practical computation:

a) Read values of s' from charts (or otherwise) and
b) divide them into corresponding (given) values of q.
c) The sum of these quotients is K, the adjustment factor.
d) Divide H by K and multiply this constant quotient into each quotient obtained in b), thus deriving the adjusted number of households in each county.
e) Compute adjusted values of s by multiplying each s' by K. Effect of adjustment can then be gauged.

Reference has been made herein to two different ways of computing average household size in an area: (1) to relate total population to households, and (2) to relate only the population living in households to households. For any one date it is obviously imperative that average household size should be computed on the same basis for both the small areas, for which household estimates are to be made, as well as for the large area for which household estimates are already available. The use of either base (preserving consistency at any date) is permissable if it does not produce distortion in relative household size.

Distortion can be caused by variations in the relative balance between population living in households as opposed to population in quasi-households. The danger is of a disproportionate change in quasi-household population in the small area.

The following table indicates the magnitude of average household size using both bases for different types of area.

Ratios of Population by Type of Residence to Total Households in Area, 1950

Area	In Households	In Quasi-Households: Institutions	Other	Total	TOTAL
Urbanized Areas:					
Central cities	3.186	.025	.138	.163	3.349
Urban fringes	3.400	.021	.070	.091	3.491
Total	3.248	.024	.118	.142	3.390
Other urban	3.224	.036	.107	.142	3.367
Total urban	3.242	.027	.115	.142	3.384
Rural nonfarm	3.454	.092	.088	.180	3.634
Rural farm	3.984	-	.012	.012	3.996
Total rural	3.667	.055	.057	.112	3.780
U. S.	3.384	.037	.096	.132	3.516

Table 2 gives an illustration of the first part of the procedure for estimating the number of households in each geographic division. The data through the next to the last column should be clear. The last column gives relative household size. The United States average in the next to last column becomes the denominator for the corresponding date for each of the nine divisions.

The accompanying charts show fluctuations from 1870 in relative household size, with possible projections indicated for the period after 1950.

Readings from these charts provide all the data needed to produce estimates of future households in these areas except for the required population projections and the national household projection. These charts with their accompanying tables are included to give some idea of the behavior of these ratios. It is also worth noting that the measures of relative household size need not be changed if a different household projection should be substituted for the large area, or if different general population projections should be used.

RELATIVE HOUSEHOLD SIZE

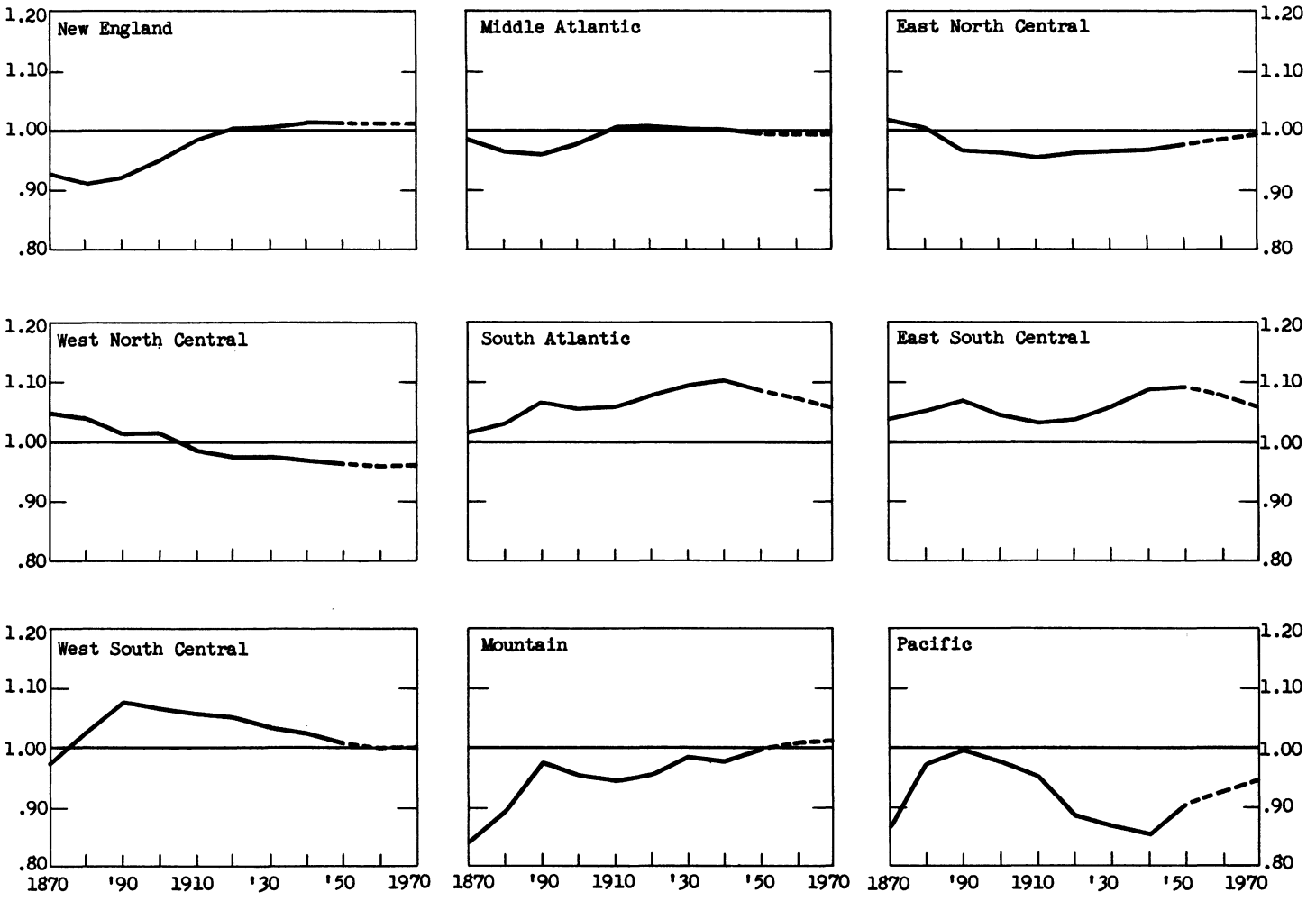


TABLE 1

VARIATIONS IN MEAN HOUSEHOLD SIZE BETWEEN COUNTIES

1950

Area	Number of Counties			Mean*			Standard Deviation			Coefficient of Variation		
	Total	Metro.	Non-Metro.	Total	Metro.	Non-Metro.	Total	Metro.	Non-Metro.	Total	Metro.	Non-Metro.
United States	3,102	284	2,818	3.54	3.51	3.56	.38	.22	.39	10.7	6.3	10.9
Regions												
Northeast	217	77	140	3.41	3.51	3.40	.17	.13	.19	5.1	3.8	5.6
North Central	1,056	82	974	3.36	3.50	3.36	.25	.16	.25	7.3	4.6	7.5
South	1,416	95	1,321	3.76	3.50	3.78	.38	.25	.38	10.0	7.3	10.0
West	413	30	383	3.33	3.49	3.35	.36	.20	.36	10.7	5.8	10.8
Divisions												
New England	67	20	47	3.39	3.51	3.39	.19	.20	.22	5.6	5.7	6.6
Middle Atlantic	150	57	93	3.41	3.51	3.41	.16	.15	.17	4.8	4.2	5.0
East North Central	436	56	380	3.34	3.51	3.34	.17	.14	.17	5.0	4.1	5.1
West North Central	620	26	594	3.37	3.25	3.38	.29	.17	.29	8.5	5.3	8.6
South Atlantic	582	54	528	3.89	3.53	3.92	.37	.25	.36	9.4	7.0	9.1
East South Central	364	17	347	3.87	3.57	3.89	.34	.21	.34	8.8	5.8	8.7
West South Central	470	24	446	3.51	3.39	3.52	.29	.27	.29	8.2	7.9	8.1
Mountain	280	9	271	3.43	3.36	3.44	.38	.16	.38	11.0	4.7	11.1
Pacific	133	21	112	3.13	3.08	3.14	.18	.16	.19	5.8	5.1	5.9

* Unweighted means of county averages of persons per household.

Note: Except for New England, metropolitan counties are those included in Census of Population: 1950, Vol. II, Part 1, Table 26. For the New England States, they are the counties (rather than towns and cities) given in County and City Data Book: 1952, Appendix Table D-1. Berkshire County, Mass., was also regarded as metropolitan.

TABLE 2

RELATIVE HOUSEHOLD SIZE

Illustrative Computation for Geographic Divisions

Area	Census Date	Population	Households	Pop./Households In Area	Percent of U.S.
United States	1870	38,558,371	7,579,363	5.087	100.0
	1880	50,155,783	9,945,916	5.043	100.0
	1890	62,622,250	12,690,152	4.935	100.0
	1900	75,994,575	15,963,965	4.760	100.0
	1910	91,972,266	20,255,555	4.541	100.0
	1920	105,710,620	24,351,676	4.341	100.0
	1930	122,775,046	29,904,663	4.106	100.0
	1940	131,669,275	34,948,666	3.768	100.0
	1950	150,697,361	42,857,335	3.516	100.0
New England	1870	3,487,924	740,271	4.712	92.6
	1880	4,010,529	872,075	4.599	91.2
	1890	4,700,745	1,034,262	4.545	92.1
	1900	5,592,017	1,236,929	4.521	95.0
	1910	6,552,681	1,464,942	4.473	98.5
	1920	7,400,909	1,703,812	4.344	100.1
	1930	8,166,341	1,981,499	4.121	100.4
	1940	8,437,290	2,208,351	3.821	101.4
	1950	9,314,453	2,616,797	3.559	101.2
Middle Atlantic	1870	8,810,806	1,757,223	5.014	98.6
	1880	10,496,878	2,151,666	4.878	96.7
	1890	12,700,800	2,677,980	4.743	96.1
	1900	15,454,678	3,320,337	4.655	97.8
	1910	19,315,892	4,235,675	4.560	100.4
	1920	22,261,144	5,085,080	4.378	100.9
	1930	26,260,750	6,374,380	4.120	100.3
	1940	27,539,487	7,294,488	3.775	100.2
	1950	30,163,533	8,622,808	3.498	99.5
East North Central	1870	9,124,517	1,757,835	5.191	102.0
	1880	11,206,668	2,213,547	5.063	100.4
	1890	13,471,840	2,820,912	4.776	96.8
	1900	15,985,581	3,488,620	4.582	96.3
	1910	18,250,621	4,214,820	4.330	95.4
	1920	21,475,543	5,143,913	4.175	96.2
	1930	25,297,185	6,362,823	3.976	96.8
	1940	26,626,342	7,290,676	3.652	96.9
	1950	30,399,368	8,829,542	3.443	97.9
West North Central	1870	3,856,594	722,476	5.338	104.9
	1880	6,157,443	1,175,470	5.238	103.9
	1890	8,890,439	1,777,693	5.001	101.3
	1900	10,347,423	2,143,928	4.826	101.4
	1910	11,637,921	2,592,069	4.490	98.9
	1920	12,544,249	2,957,849	4.241	97.7
	1930	13,296,915	3,317,881	4.008	97.6
	1940	13,516,990	3,698,161	3.655	97.0
	1950	14,061,394	4,153,167	3.386	96.3

TABLE 2

- 2 -

Area	Census Date	Population	Households	Pop./Households in Area	Percent of U.S.
South Atlantic	1870	5,853,610	1,132,621	5.168	101.6
	1880	7,597,197	1,463,361	5.192	103.0
	1890	8,857,920	1,687,767	5.248	106.3
	1900	10,443,480	2,078,603	5.024	105.5
	1910	12,194,895	2,539,270	4.803	105.8
	1920	13,990,272	2,991,628	4.676	107.7
	1930	15,793,589	3,511,860	4.497	109.5
	1940	17,823,151	4,291,395	4.153	110.2
	1950	21,182,335	5,540,342	3.823	108.7
East South Central	1870	4,404,445	833,694	5.283	103.9
	1880	5,585,151	1,053,186	5.303	105.2
	1890	6,428,770	1,217,097	5.282	107.0
	1900	7,547,757	1,520,339	4.965	104.3
	1910	8,409,901	1,796,832	4.680	103.1
	1920	8,893,307	1,977,381	4.498	103.6
	1930	9,887,214	2,273,359	4.349	105.9
	1940	10,778,225	2,626,791	4.103	108.9
	1950	11,477,181	2,991,927	3.836	109.1
West South Central	1870	2,029,965	408,717	4.967	97.6
	1880	3,334,220	644,364	5.174	102.6
	1890	4,544,123	854,023	5.321	107.8
	1900	6,532,290	1,287,871	5.072	106.6
	1910	8,784,534	1,827,105	4.808	105.9
	1920	10,242,224	2,242,810	4.567	105.2
	1930	12,176,830	2,868,262	4.245	103.4
	1940	13,064,525	3,386,552	3.858	102.4
	1950	14,537,572	4,103,354	3.543	100.8
Mountain	1870	315,385	73,597	4.285	84.2
	1880	653,119	144,891	4.508	89.4
	1890	1,156,326	239,940	4.819	97.6
	1900	1,674,657	367,932	4.552	95.6
	1910	2,633,517	614,656	4.285	94.4
	1920	3,336,101	803,853	4.150	95.6
	1930	3,701,789	914,408	4.048	98.6
	1940	4,150,003	1,126,190	3.685	97.8
	1950	5,074,998	1,446,725	3.508	99.8
Pacific	1870	675,125	152,929	4.415	86.8
	1880	1,114,578	227,356	4.902	97.2
	1890	1,871,287	380,478	4.918	99.7
	1900	2,416,692	519,406	4.653	97.8
	1910	4,192,304	970,186	4.321	95.2
	1920	5,566,871	1,445,350	3.852	88.7
	1930	8,194,433	2,300,191	3.563	86.8
	1940	9,733,262	3,026,062	3.216	85.4
	1950	14,486,527	4,552,673	3.182	90.5

TABLE 2

For a discussion of comparability of data see U.S. Census of Population: 1950, Vol. IV, Special Reports, Part 2, Chapter A, General Characteristics of Families, pp. 8 - 9. The data herein follow Table B, p. 8, of this reference. Table B, however, is for the U.S. as a whole and covers only the period 1890 - 1950.

Household data taken from the following Population volumes:

Data for	Volume for	Reference
1870	1890	Part I, Table 87, p. 914;
1880 - 1900	1900	Vol. II, Part II, Table LXXXVIII,p.clx;
1910 - 1920	1920	Vol. II, cap. XIV, Table 2, p. 1267;
1930	1930	Vol. VI, Table 40, p. 33;
1940	1940	Vol. IV, Part 1, Table 51, pp. 162 - 163;
1950	1950	Vol. II, Part 1, Table 47 for U.S. and Table 22 of each State Part.

Prior to 1950, except for 1900 and 1930, the data labeled "households" are for "families" and necessarily include the relatively small number of quasi families. In 1900 and 1930, however, the data are for "private families", (excluding quasi families) and are generally comparable with "households" in the 1950 Census.

Population data taken from the following volumes:

Data for	Volume for	Reference
1870 - 1950 (except 1890)	1950	Vol. II, Part 1, Table 6;
1890	1900	Vol. II, Part II, Table LXXXVIII,p.clx.

Persons in Indian territory and on Indian reservations were enumerated for the first time in 1890. No "family" data, however, are available and hence the 325,464 persons are excluded.

EFFECTIVENESS OF OUR TOOLS FOR ESTIMATING POPULATION CHANGE IN SMALL AREAS

By: Carl M. Frisen
California Department of Finance

The widespread interest in demographic data relating to small areas is reflected in the number of agencies listed in "Local Population Estimates Prepared by State and City Agencies: 1957-58", published by the United States Bureau of the Census in its Current Population Reports, Series P-25, No. 178, in June 1958. A summary is provided of the estimating work done by 62 state agencies and 33 city agencies. Also included is a useful discussion of the problems and limitations involved in the preparation of small area estimates. From the information included, it is apparent that the users of population data, involving both governmental agencies and private organizations, have been able to stimulate a great deal of activity aimed at providing population information on a current basis. As a result, county population estimates are now generally available and in many states the populations of cities and towns are estimated at regular intervals.

From the standpoint of the producer of demographic materials, problems of method - such as the choice of the most effective tool in a given instance - depend upon the questions posed by the user of the data. For the most part, it is likely that these questions will call for one or more of the following types of information:

1) estimates of total or civilian population, for one or more small areas;
2) measures of the components of change, particularly the volume of net civilian migration;
3) analysis of the composition of population with the concern likely to involve age, sex and color or race;
4) estimates of the number of households, and factors relating to the formation of households and families since the last federal census.

A second consideration relates to the nature of the small areas for which these types of data may be needed. If the State is considered the largest "small area" in this discussion, administrative needs may range from state, counties and cities, through less widely recognized but formally delineated units such as hospital districts, school districts and townships, to such amorphous areas as unincorporated places, marketing areas and the like.

From the standpoint of the producer, then, the administrator's request for demographic information poses problems in terms of the nature of the data that are desired and in terms of the specific areas involved. This means that the demographer must first determine whether symptomatic data are available that can measure the factor with which he is concerned, be it total population, net migration, age distribution or whatever, and secondly he must ascertain whether the symptomatic data are available for the geographic areas designated. Once the availability and accuracy of the indicators are determined, there remains the problem of their applicability in terms of technique.

At this point it is deemed appropriate to suggest two classificatory schemes that may contribute perspective in analyzing the current picture. The first of these is applicable to the symptomatic data used in the estimating process, while the second is concerned with how the estimating technique relates to components of population change.

The indicators selected by various estimators may be classified on the basis of the degree to which they relate to the population that they seek to measure. One class of variables may be termed indirect, since measurements are expressed in units other than persons. Data on utility connections, rural route box holders and occupied or total dwelling units are of this type, for they require a factor for the average number of persons per unit before an estimate of population can be made. A second class of variables includes those that are direct, in that they are measured directly in terms of numbers of people. A distinction may be made in this class between indicators that are partial and total. Partial-direct variables are limited to certain classes or elements of the population and the relationship of the subgroup to the total must be established prior to their use in estimating. Among the variables of this type are school enrollment data, selective service registration, birth statistics and motor registrations. Total-direct variables are complete in their coverage, relating to all segments of the population and include morbidity and mortality records. Illness, accidents and deaths occur at all ages, in both sexes and among all races.

An alternative classification may be made in terms of how the estimating techniques measure population change. Some methods, such as the Census Bureau's methods I and II, estimate only net civilian migration, and other components of change must be obtained separately and combined with the migration estimate to indicate total population change. Other methods provide a measure of total population change only, and net migration may be derived only as residual by subtracting other components. Bogue's vital rates method is of this type, as is the dwelling unit method currently being used in California's city population estimating program. A third category, which may be termed composite, includes estimates based on combinations of methods some of which measure net migration only while others estimate total change. One such combination has been used by Bogue and Duncan in estimating the population of cities in Illinois by age, sex and color. Here, school enrollment is used for the younger ages, fertility ratios for the middle ages, and specific death rates for the older ages.

As modified by the Census Bureau for purposes of a test against the 1950 census, reported on in a paper by Shryock, Siegel and Greenberg read at the Population Association meeting in May 1957, the Census Bureau's method II was used to estimate preschool and school-age migration, while the 18-44 year old change and the 45 year and over change were estimated as total population change.

For the moment, let us direct our attention to the questions facing the demographer who is assigned a specific problem for solution. Many of the analyses of estimating methods fail to give sufficient emphasis to population estimates in terms of their role in the administrative process. The effect of this role may be illustrated in the case of various requests for city population estimates in California. From time to time State agencies ask for current estimates for one or more incorporated places in the State, perhaps for use in making a general evaluation of changes in agency workloads. Ordinarily the Department of Finance will depend upon extrapolations from special censuses or locally published estimates as sufficiently accurate under these circumstances. The problem is of a different order when the question is posed: can satisfactory population estimates for cities be prepared that will serve as the basis for allocating State-collected funds on a per capita basis? This question was raised by the League of California Cities in exploring alternatives to the use of special censuses in officially determining population change. It is readily apparent that in the requests cited we are concerned with two different definitions of "satisfactory" as applied to population estimates for cities, and the demographer must be aware of these differences and what they mean in terms of data gathering and method. In the case of the study for the League of California Cities, an effort was made to furnish a criterion for determining whether a given estimate was satisfactory by approximating the financial effect of consistently underestimating a city's population as compared with the funds received under procedures then currently in use by the city. That is, an estimating technique might be deemed unsatisfactory if a likelihood existed that some cities would receive smaller amounts of allocated funds over a period of time than would be true under the procedure previously in effect.

Before any test of accuracy or acceptance can be applied, however, the producer of population estimates must determine what tools are at hand for the task assigned by the administrator. The first question to arise is: what symptomatic data are available for the specific areas involved? When cities are the object of study, a number of possible indicators may be explored. Records of elementary school enrollment are sometimes available, though in California few cities have school district boundaries that coincide with or even approximate city boundaries. In many instances current information can be obtained from school records relatively easily, but it is not possible to derive data for an earlier census date that must serve as the benchmark. Statistics on births and deaths are tabulated annually for most of the cities in the State, with detailed data, such as deaths by age, sex and race and births by age of mother, compiled for cities of 50,000 and over. Again, local building permit records on residential construction are maintained by most California cities, though there are differences in the completeness of coverage and the time span covered by the records. Finally, information on residential meter connections can be obtained from utility records, though there may be some problems of interpretation of data maintained by some municipal utilities. This brief summary of the major types of indicators found in seeking to establish a broad test of city population estimating techniques shows that it is possible to find some instances of all three classes of symptomatic data, and in turn to utilize them through all three classes of method.

Looking first at partial-direct variables, the present discussion is limited to the use of reported school-age population rather than school enrollment. Special censuses report population by age for cities of 50,000 and over and for a limited number of smaller cities that contract for the detailed information. Since the Census Bureau's methods I and II use the school enrollment to estimate related population age groups, it is feasible to substitute the appropriate population directly in the methods, thus testing their ability to estimate net civilian migration for all ages. Through this procedure, it is possible to avoid problems of availability and accuracy of school enrollment data.

Questions of this order should be explored thoroughly, however, if only to assure that the results will receive proper interpretation. For example, the City of Los Angeles has been included in the published results of a number of tests of method II as applied to large cities in the United States. A proper check in advance would have shown that elementary school enrollment data for the Los Angeles school district cover an area approximately 20 percent larger in population than the total within the corporate limits of Los Angeles. In actuality, it is not possible at the present time to employ only a school enrollment-based method to estimate the population of the City of Los Angeles, and it is preferable that this be recognized before the city is included in the test of the accuracy of such an estimating technique.

Also needed is a series of tests of the reliability of school enrollment as a measure of population of a specific age group, rather than the present implicit assumption in many tests that differences from a census standard

result only from the operation of the estimating method. This is important because errors in the measurement of the population of school age become magnified in the natural-increase-and-net-migration methods.

The use of special censuses by California cities has provided a basis for testing of estimating methods against census standards, but the distribution of special census dates through time operates to limit the strict comparability of results. For most tests, therefore, a distinction is made between estimates for dates in 1955 and later, as compared with tabulations that include estimates for earlier years.

An initial test was made of the Census Bureau's method II as applied to 12 California cities, the total number of cities for which necessary data by age has been published. These estimates differed on the average by 5.8 percent from the published census figures. It was possible to apply method I to a total of 26 cities that had taken 34 special censuses after April 1, 1950. In using this method, an adjustment factor was incorporated into the procedure on the basis of earlier experimental work which showed that greater accuracy should result. For the total of 34 special censuses, the average difference of the estimates was 3.6 percent, while estimates for the 17 special census dates after January 1, 1955 differed from the census results by an average of 4.9 percent. Using the same 12 cities included in the method II test, the average deviation for method I was 4.3 percent. When the method I and II estimates were averaged for the 12 cities, the average difference was reduced to 3.4 percent.

A second test combined a partial-direct variable, live births, and a total-direct variable, deaths, with a method that measures total population change, that is, Bogue's vital rates method. Forty-five cities were included in this test, and the average difference of the estimates was 11.9 percent. Only 13 of the 45 estimates were within 5 percent of the special census results and 23, or more than half, differed by 10 percent or over. As might be expected, the estimates were predominantly lower than the special census results.

It is evident that the vital rates method is at best applicable only to large and very slow-growing cities. This is confirmed by data in Shryock's paper on the "Development of Postcensal Population Estimates for Small Areas", presented at the 1957 Conference on Research in Income and Wealth, showing that the "natural increase method", which makes use of birth and death statistics and assumes that net migration equals zero, showed smaller average deviations than the vital rates method for the 92 cities included in the test.

The third class of estimating technique was termed the composite, because it combines both component-of-change and total-change methods. Some experimental work has been carried out with the age-specific-death-rate method, which may be used to estimate the population aged 15 years and over in combination with either method I or II for estimating the population under 15 years of age. Estimates for cities were prepared, using deaths by sex and by 10-year age groups. For eight cities included in the initial test, estimates of the population aged 15 years and over differed from special census totals by an average of 4.7 percent. Interestingly enough, the greatest differences in age groups appeared in the 15-24 year, 25-34 year and 75 year and over age groups.

In 1954 a test of the age-specific-death-rate method was made, deriving 1950 population estimates for 29 states with relatively small nonwhite populations. These states were used so that 1940-50 change in the age-specific white death rates for the United States could be used in estimating changes in state death rates. The estimates of total population aged 15 years and over differed from the 1950 census by an average of 4.6 percent. For the population aged 15 years and over, the average deviation for the 29 states was 2.4 percent. These results point to the utility of the age-specific-death-rate method for preparing estimates by age, particularly for the older age groups. In this same test, a check of the 1950 population estimates by 10-year age groups for ages 15 through 84 years and 85 years and over for the five largest states, New York, California, Pennsylvania, Illinois and Ohio, showed that the estimate for only one age group -- the 15 to 24 year olds in Pennsylvania -- differed by more than 5 percent from the census figure.

The discussion so far has dealt only with partial-direct and total-direct indicators, involving the use of symptomatic data expressed in numbers of people. The tests of methods and indicators suggested the need for further efforts to improve the accuracy of city estimates. The next step was to examine the availability, accuracy, and applicability of such indirect indicators as building permit records and utility data.

One test endeavored to explore the possibility of estimating change in the number of persons per household, assuming that an accurate measure of change in the number of households could be achieved. That is, the reported number of occupied dwelling units at the time of the special census was used to measure change in dwelling units from 1950, while change in the number of persons per household was estimated by relating each city's average to the national pattern as reported in the Current Population Survey. Population estimates were prepared for 59 California cities, including 66 special censuses in 1955 and 1956. The average deviation of the estimates was 3.63 percent. These cities ranged in size from 1,700 to 2,244,000 at the time of the special censuses, with growth after the 1950 Census ranging up to 536 percent, and with at least 10 cities reporting over 100 percent gains.

A second test dealt with the use of utility records as a basis for estimating change in the number of occupied dwelling units or households. Through the cooperation of the major utilities in the State, information for almost all of the more than 360 cities in California is available for use in tests and estimate preparation. The major problems of the use of electric meters for this purpose are well known: the existence of master meters particularly in public housing and in trailer parks; the tendency to leave meters active although the premises are vacant; the effects of economic change on the relationship of meters to households; and others. In tests to date, estimates of the number of occupied dwelling units have been prepared for 44 cities that have taken 64 special censuses since January 1, 1955. In a test of 26 special censuses taken in the San Francisco Bay Area, the average deviation of the estimates of the number of occupied dwelling units was 2.5 percent from the reported figures, with 5 cases showing differences of 5 percent or more. These five cities are all located in one county and involve areas where war housing and master metering was operative in 1950, with considerable change since then. A check of estimates applied against 38 special censuses in the Los Angeles-Long Beach Metropolitan Area showed an average deviation of 1.5 percent from the reported occupied dwelling units, with the largest deviation under 5 percent. In the Southern California test, the benchmark dates were special censuses in 1952 and 1953 rather than the decennial census of 1950.

These data suggest that it is possible to prepare estimates of the numbers of dwelling units in cities that are sufficiently accurate to stand as estimates of the number of households for those agencies needing such figures. Further, it would appear that satisfactory estimates of population can be derived from these data. One test, applied to 23 cities in the San Francisco Bay Area, related percentage change in the number of occupied dwelling units based on utility data to percentage change in the total population, using an equation based on census data from 32 cities that took special censuses between January and April of 1956. The average difference of the 47 estimates for these 23 cities from the census standards was 4.2 percent, with 6 estimates differing more than 10 percent.

The Census Bureau, in its use of this general type of approach, has included a modification in which the numbers of persons per household under 18 years of age and 18 years and over have been handled separately. This modification has been tested, using the national pattern of change as a guide for estimated local averages. For the same 12 cities used in the test of methods I and II, estimates of total population aged 18 years of age and over showed an average deviation of 2.5 percent from the special censuses. At the present time this procedure is used for all official city estimates where a basis exists for establishing trends of change for the two age groups.

Another modification has been incorporated in the preparation of recent estimates for the cities of Los Angeles and Oakland. With the cooperation of the City Planning Commission staffs, data were developed for four subareas in Los Angeles and five subareas in Oakland, so that allowance could be made for subarea differences in change in the number of persons per household. It is believed that both estimates gained greater accuracy through this intermediate step, because both cities display markedly different patterns between subareas.

While the conclusion must be considered tentative, it is important to observe that the tests of the dwelling unit method suggest that neither the size of a city nor its rate of growth appears to influence the accuracy of an estimate. On the other hand, only the taking of a large number of special censuses has provided assurance that change in the number of persons per household can be traced over the years since the last decennial census. Further, each estimate must be developed individually in order to achieve maximum accuracy.

Do we have effective means for measuring population change in small areas? The answer would appear to be a qualified "yes". A variety of methods exist for translating symptomatic data into population measures, at least for the most widely recognized formally defined small areas such as states, counties and incorporated places. With varying degrees of applicability and accuracy, these methods will describe total populations, characteristics such as age, sex and color, components of change, numbers of households, and so forth. The degree of effectiveness is likely to depend in large measure on the recognition that indicators and methods must be related to specific cases. This is illustrated by the fact that limited testing of method I on California cities led to the introduction of a relatively simple adjustment factor that results in method I giving generally more accurate estimates for California cities than method II. It is quite possible that a further analysis of method II as applied to cities will again reverse the relationship, although there is no inherent reason for one method being more accurate than the other. Their differences arise primarily from the age groups selected as partial-direct indicators of migration.

The comments of the Census Bureau on the very limited extent to which state and city agencies report testing of the methods they use points to the need for a considerable expansion of testing activity before population estimates are accepted as satisfactory in terms of the criteria applicable. Admittedly this type of testing cannot always be done, but it may be that the producers of population estimates must share some blame for this in that they have not convinced the users -- the persons who originate requests and utilize results -- that the lack of sufficient time and adequate facilities to collect and check symptomatic data and to test and modify methods can only reduce the effectiveness of the population estimates and limit their value as administrative tools.

PROJECTING THE DISTRIBUTION OF POPULATION BY SUBAREAS OF THE URBAN COMMUNITY

By: Chester Rapkin, University of Pennsylvania

I think Congress should strike a special medal to be awarded to statisticians and demographers who venture estimates of the population of small areas. Some of us remember past population estimates and projections for the nation as a whole and offer embarrassed retrospective explanations. In small areas where the basic data components of an estimate are less reliable and where the evil hand of chance swings a more uncertain arc, our calculations are even more fraught with uncertainties. As yet, for example, the preparation of estimates of the size and composition of inmigrants to an urban area are still within the realm of art rather than science.

Within the past two decades, an additional difficulty has been added to our already heavy burden. Military service calls some of the young men of the population and by a complicated process redistributes them: some within the nation, and some abroad; some in concentrated accommodations, some in civilian quarters. Before I had the pleasure of reading the work of Siegel and Zitter, I participated in or witnessed many a session devoted to statistical projections in which the size of the military population was estimated in very much the same manner as the young men are selected for duty.

But despite all our misgivings, Frisan tells us that our instruments for estimating population change in small areas are perhaps sharper than we have thought. In any event, we now have available an impressive array of techniques, methods, sources of data, and even some significant theory. On the whole, therefore, aside from those areas that will experience a drastic change in regimen, it would appear that the margins of error may be sufficiently narrowed to permit the projections to serve a useful purpose.

Just as the force of circumstance demands the preparation of population estimates and projections for urban or metropolitan areas, be it for use by a public agency or a private firm, there has developed a growing demand for population data on subareas of the urban community. Although the public and private planners and the market analysts have frequently asserted that they would prefer data by census tract, they readily settle for a larger geographic unit, such as the planning analysis area or health area or a homogeneous group of tracts that they can label "a neighborhood."

It is understandable that efforts in this direction have been distressingly meager. Perhaps the best known are the gravity models prepared by the regional scientists and the distance-density formulae derived by Ernest Jurkat in which the distribution of the anticipated population is taken to be a function of existing population density and the distance from the core of the city, with modifications to allow for unusual features of the terrain or for natural impediments.

I should like to suggest that some future effort be devoted to an exploration of the problem of estimating the future population of subareas within a city through an analysis of the housing sub-markets which comprise these areas. At this point I do not have a formula to propose and my comments therefore are general and in the nature of an initial exploration.

The first step in such a series of calculations is the derivation of a projection of the population of the city or metropolitan area based upon anticipated employment opportunities. Although people may desire to locate in one area rather than another for a multitude of reasons, the existence of acceptable or desirable income earning opportunities is unquestionably the most important. There is no doubt, for example, that the population of areas which exert a locational pull because of natural advantages would be sharply reduced if inmigrants could not find employment.

The population projection must be translated into an estimate of the total housing requirement for the community. This is done by calculating the manner in which the population will arrange itself into households which we take to be equal to occupied dwelling units. The number of households may be derived by the technique suggested by Hodgkinson, or by the techniques developed by Glick or by Winnick, or by any other suitable method. This figure yields an estimate of the total housing requirement for the community. The difference that exists between this figure and the present number of dwelling units represents the likely volume of new construction, after due allowances are made for changes in existing stock and in vacancy ratio.

The second step in this process rests on the assumption that the characteristics of a population in any given area will be closely related to the type of housing accommodations available in that area. Let me pose several situations for

your consideration. These may constitute some of the items in the catalogue of subareas and the types of changes that may occur within them.

(a) Undeveloped or Sparsely Utilized Land

Land of this description is almost invariably found on the outskirts of the city or metropolitan area and, in fact, is even becoming more difficult to find in these locations. If the post-war patterns do not change substantially, we can with confidence anticipate that a major proportion of the new housing requirement will be met by the construction of single family homes in presently undeveloped areas, for since the end of World War II approximately 85 percent of all residential construction in the United States have been in units of this type. In a few scattered instances rental developments have also been erected in these areas. For the most part, these rental units have not been in high elevator structures, but in two- or three-story garden type accommodations. But in either event, the population of the new developments has been singularly homogeneous, a fact which warms the heart of the statistician, but saddens the sociologist.

For example, a recent survey of population and households in the Philadelphia SMA conducted by the Bureau of the Census for the Institute for Urban Studies revealed that the purchasers of recently constructed homes had the following characteristics: approximately three-quarters of the household heads were between 25 and 44 years of age; three-fifths earned over $6,000; and three-fifths of the households consisted of three or four persons. Almost one-half of all the households possessed all three characteristics. These figures apply to all new home purchasers. There was far less variation by subarea and by subdivision.

(b) Built-Up Areas in Which There is Little or No Change Anticipated in the Number or Characteristics of the Structures

The occupants of dwellings in areas of this type alter in number and in composition as a consequence of births and deaths and the aging of the population and of the successive stages through which the family cycle passes. Thus, if we think of an area that was a new subdivision a decade ago, we now find that the toddlers have entered high school, that some of the grandparents have gone to their reward, and that father is at least ten years older. The changes wrought by time are mitigated by the turnover of property in the intervening period, a fact that serves to keep the aggregate for the subarea from altering commensurately with changes in the resident families. But, by and large, we have observed that the newcomers to an area possess characteristics which corresponded quite closely with those of the resident group.

One of the most interesting population profiles is to be found among the occupants of the more distinguished (and more expensive) downtown apartment structures. Most cities of size possess a number of accommodations of this type, some of which are merely comfortable, while others are indeed luxurious. A few recent studies have shown that the population composition of the occupants of these accommodations constituted a readily distinguishable group. In fact, a recent report conducted by the Institute for Urban Studies relating to the city of Philadelphia showed surprisingly few deviations from the group pattern.

Field interviews revealed that 97 percent of the heads of households were over 35 years old; in 94 percent of the cases there were no children in the household; 96 percent enjoyed a family income of more than $5,000 a year. The employment pattern was similarly concentrated. Ninety percent either worked in center city or were not in the labor force. Over 80 percent of the respondents fit into the first three categories. Every respondent had at least one of the four characteristics.

By comparison, in the total metropolitan area 80 percent of the household heads are over 35; 50 percent have annual family incomes of more than $5,000; 50 percent have no children; and 33 percent work in center city or do not work. Only six percent, or 68,000 households, fit into all four categories. Thus, center city provides another illustration of a sub-market that draws a segment of the population with palpably distinct and readily identifyable characteristics.

(c) Existing Areas in Which the Housing Stock Undergoes Major Changes

Some sections of the housing inventory undergo sharp changes in quality within relatively short periods of time. In some cases, deterioration is experienced both in the physical nature of the accommodation and in the surrounding area. Frequently this is accompanied by the conversion of accommodations by subdivisions into quarters of smaller size or into rooming houses. Change of this sort is experienced during a period in which a large and rapid inmigration of population confronts an inflexible supply of housing. If the inmigrant population has a low income earning capacity, it will not be able to pay the rents or prices

commanded by accommodations of the size and quality available on the market. They therefore will be inclined to purchase smaller amounts of space at higher cost per square foot in the subdivided buildings. In a twenty-block area of the west side of Manhattan, for example, the population rose from 33,000 in 1950 to 39,000 in 1956 with virtually no new construction during that interval. In areas of this type, the population pattern alters sharply. Middle income families with children of school age tend to be reduced in number. The incoming population is likely to be a racial or national minority with a distinct demographic structure.

The extent of the inmigration in any particular area can be estimated by analyzing the composition of the housing stock. The large and well-kept units are likely to remain in the same use, for the rents they command (per room) equal or exceed the payments that can be made by the low-income groups. The units that are susceptible to conversion possess few amenities and can only be let at low rents. It is in these that the aggregate rent roll can be increased substantially by conversion to rooming houses. In the area noted above, brownstone and old-law tenement structures constituted virtually this entire category.

Conversion of housing need not necessarily imply deterioration of quality. In many instances physical alterations are undertaken to modernize and improve a structure. These changes range from superficial refurbishing to major reconstruction. Conversions of the latter type usually take place in the more expensive areas of the city in the sections surrounding the central core. It is in these areas that a sector of the upper income group of the population maintains its residence, and the area is therefore sought not only for the convenience and quality of location, but also because it is a "good address." With few variations the characteristics of the occupants of these structures resemble very closely the residents of the expensive downtown apartments described previously.

(d) Demolition of the Existing Stock and Succession of Uses

In our dynamic urban society, the face of the city is constantly changing. Many a residential structure is demolished and replaced by a new building. Very often the demolished structure is of low quality and, in fact, slum clearance may have been the reason for its destruction. On the other hand, the land may be so valuable in an alternative use that the acquisition and demolition of an existing structure, even of high quality, is justified in order to release the site from its previous use. In New York City the Hotel Marguery and the Ritz Carlton Hotel were demolished to make place for office buildings. In other instances, structures of quality were removed in order to provide land for public facilities, such as expressways, bridges, courthouses or schools.

In order to identify those areas that might be subject to demolition, it is necessary for the population analyst to study the real estate market. The demand for office space in New York, for example, provides an index of the rate of construction of new buildings and of the rate of demolition of property presently standing on potential office building sites. If these are residential structures, then we can of course expect a further decline in the resident population as land use in the area continues to change.

It is undoubtedly easier to estimate the changes that are expected to take place in an area slated for public renewal or redevelopment. In these cases the superseding use is stipulated in the plans of the governmental agencies and, in fact, may be dictated by a specific population policy. In other words, the renewal may be undertaken to provide accommodations for middle income families with children or to make available subsidized public accommodations for elderly low-income persons.

Summary

If the statistician is to explore the possibilities of utilizing changes in the housing market to assist him in estimating anticipated variation in the distribution and size of local population, he will of course seek to identify those variables that are strategic and that lend themselves to quantification. In view of the foregoing discussion, it would appear that experimentation with the following possible independent variables may prove fruitful:

1. Quality index of dwelling units
2. Age of dwelling units
3. Persons per room
4. Persons per dwelling unit
5. Dwelling units per structure
6. Persons per acre
7. Distance from center of city
8. Time of travel from center of city
9. Rent or value per dwelling unit
10. Rent or value per room
11. Ratio of the value of land to total of land and improvements

Experimentation with these factors may yield more than suitable techniques for estimating the population of subareas within a city. Without doubt, such endeavor will also serve to enhance our understanding of the nature and processes of neighborhood change.

XII

CONTRIBUTED PAPERS

Chairman, Robert E. Johnson, Western Electric Company

Delineation of Demographic Areas and the Contiguity Ratio—James M. Beshers, Purdue University

A Suggested Method for the Delimitation of Population Clusters—Arthur F. Loeben, Planning Commission, Montgomery County, Pennsylvania

Patterns of Heaping in the Reporting of Numerical Data—Stanley H. Turner, University of Pennsylvania

Fertility Estimates from Birth Statistics—Kuno R. Gabriel, University of North Carolina and Hebrew University, Jerusalem and Ruma Falk, Hebrew University, Jerusalem

Negative Skewness and its Significance in Relation to Distributions of Performance Ratings of Civil Service Employees—James P. George, University of Tennessee

DELINEATION OF DEMOGRAPHIC AREAS AND THE CONTIGUITY RATIO

By: James M. Beshers, Purdue University

In this paper several problems in defining and delineating demographic areas will be discussed, and the potential value of the contiguity ratio in solving these problems will be indicated.[1] Let us regard a demographic area as a bounded area with a set of distinctive demographic characteristics, as compared with neighboring areas. Thus a demographic area is defined in terms of the characteristics of sub-areas which lie within and outside any given area. Therefore every demographic area should be delineated in terms of (1) the distribution of demographic characteristics within a larger area, of which the particular area is a part, and (2) the distribution of demographic characteristics within sub-areas of this area.

Note that hierarchies of areas can exist—the area at one level of discussion may become the sub-area at the next higher level. Census tracts and their combination in urbanized areas may each in turn be examples of demographic areas. However, the implications of this paper are not limited to these areas.

We shall consider the usefulness of a particular delineation of areas for research purposes. Criteria should be developed for the suitability of a given delineation for research purposes. Such criteria would have two functions: first, they would serve as a guide to the delineation of new areas, and second, they would enable us to evaluate existing areas for their appropriateness in research. We may evaluate existing areas without reference to the original purposes for which they were delineated—whether for administrative or for research purposes.

But we must see immediately that there is no general solution to this problem. No single criterion can be posed for a given set of areas. As the purposes of given research projects differ, so must the criteria which determine the usefulness of a set of areas for these purposes. In particular, the variables of importance will differ from project to project, as well as the statistics computed from these variables. The Shevky-Bell[2] literature asserts that there are three types of demographic variables with different spatial distributions. Nevertheless, we can consider the properties of criteria for particular variables and estimates.

Any criterion which we propose will be relative to the number and size of areas into which a larger area is to be subdivided. Assuming equal sized areas (either geographic size or population size), we must consider the specific number of areas, for if the number of areas is increased, then better delineations may become possible. There are two problems—first, determining the best delineation relative to a specific number of areas; second, determining the adequacy of this delineation with respect to the purposes of a specific project. If this delineation is not adequate, then the number of areas may be insufficient.

In order to illustrate these statements, we may consider two of the purposes for which a research project may use observations on areas. The area data may be used to obtain estimates on sub-units, either sub-areas or individuals, or the area data may be related to structural, or aggregate, concepts. In the former case, the appropriate criterion must be a function of the ratio of variation between areas to variation within areas, i.e., a function of the correlation ratio.[3] In the latter case a criterion is not so readily apparent. However, if areas meet a correlation ratio criterion, it seems likely that they meet most criteria appropriate for the latter case as well.

Although a correlation ratio criterion seems appropriate to the discussion of single variables, two problems remain. First, this criterion will vary from project to project as the need for greater or less accuracy differs. Translation of a correlation ratio criterion into a statement specifying the accuracy of conclusions based on certain areas might be a solution. Confidence intervals within areas could be constructed. Second, when two or more variables are considered, and the purpose of research is to estimate the relationship between these variables, then the covariance should be maximized by the delineation of areas.

A correlation ratio criterion may enable us to determine the best delineation of areas relative to a given number of areas. It may also permit us to evaluate this relative optimum in terms of the accuracy needed. Further work must be done on the latter possibility. The contiguity ratio, however, may shed some light upon the former problem (as well as many other demographic problems.)

The contiguity ratio, developed by R. C. Geary,[4] is an appropriate measure of the spatial clustering of characteristics of areas. The contiguity ratio is a two-dimensional generalization of the Von Neuman ratio used in time series analysis. The contiguity ratio compares the sum of squared deviations between the value for each area with the values for its contiguous areas summed over all areas in the numerator with the variance in the denominator. Constants are chosen so that the expected value for a random distribution is unity.

In Geary's notation, let the number of areas be n, the measure of the T-th area Z_T, with number of connections K_T. The contiguity ratio C is given by

$$C = \frac{n-1}{2k} \cdot \frac{\sum_T \sum_{T \neq T'}{}' (Z_T - Z_{T'})^2}{\sum_T (Z_T - \bar{Z})^2}$$

where

$K = \sum' K_T$ — T′ designate the areas contiguous to area T,

$\sum$ = sum over all counties

$\sum'$ = sum over contiguous counties — $Z_{T'}$ is the measure of the contiguous area,

Suppose we have three areas lying end-to-end, with values 5, 4, and 3 for each area. Then we have

$$\frac{2}{4} \cdot \frac{4}{2} = 1$$

revealing no effect of contiguity.

Let us consider the uses of the contiguity ratio. By itself, the ratio tells us whether a single variable has a significant clustering effect. (Significance may be determined either by randomization or by classical normal theory.) Further, for a given area, the degree of clustering between several variables may be compared.

How does this measure of clustering relate to the delineation of demographic areas? Recall our definition of a demographic area as a bounded area with a set of distinctive demographic characteristics, as compared with neighboring areas. The measurement of clustering effects of characteristics therefore has a two-fold significance for the delineation problem. Distribution of characteristics by sub-areas has significance outside and inside a particular area. The fact of clustering itself, as may be demonstrated by the contiguity ratio, must be evidenced before an area may be delineated.

But the contiguity ratio may guide delineation more specifically when used in conjunction with regression analysis. After the existence of a clustering effect has been demonstrated, we may seek an explanation for this effect. Regarding the clustered variable as a dependent variable, we may select independent variables and compute a regression equation. The effects of the independent variables may be removed, and the residuals tested for a clustering effect. If the residuals are not clustered, then the independent variables have "explained" the clustering effect. Subject matter theory must supply the meaning of this "explanation."

If the independent variables are distance measures, then the regression is equivalent to fitting a surface to the original variables with the contiguity ratio employed as a criterion of "goodness of fit." These distance measures may be represented in rectangular co-ordinates or polar co-ordinates. The distance measures in the regression may be supplemented by classifications of areas, which may be introduced into the analysis by covariance methods.

By these means the gradient hypothesis and various zonal hypotheses of urban ecology may be tested. The amount of variation attributable to each effect may be determined. The contiguity ratio may be used to determine whether the clustering effect has been accounted for by these hypotheses.

Thus the contiguity ratio aids us directly in studying the distribution of demographic characteristics, and therefore in the delineation of demographic areas. But these methods may be turned to the evaluation of a given delineation as well.

Recall that we wish to determine the best delineation of areas relative to a given number of areas. The surface representing the distribution of a variable must be considered. If the variation can be represented by a continuous smooth surface, then almost any delineation of areas will be as good as the best delineation. But if the variation is characterized by sharp fluctuations— canyons, deep gorges, and isolated peaks, so to speak— then the delineation must be tailored carefully to these configurations. The contiguity ratio used with a regression surface provides a partial answer to this question. The "goodness of fit" of smooth surfaces can be evaluated by these techniques.

If our variables have smooth surfaces, then we may neglect detailed delineation problems, and concentrate our attention on insuring that a sufficiently large number of areas are used. If our variables have surfaces which are "almost smooth", then we may be able to smooth them out by increasing the number of areas. The choice of smooth surfaces is by no means an easy task.

In conclusion, we need criteria for the usefulness of areas. These criteria should derive from the consequences of using the areas, from the risk involved. These criteria should determine the best delineation relative to a given number of areas, and they should evaluate this "best" delineation. A correlation ratio criterion might be used for both of these problems. An alternative approach using the contiguity ratio calls attention to the smoothness of the surface of distribution of a variable. If the surface is smooth, then more attention should be paid to providing a large enough number of areas, and less attention should be paid to the particular delineation of boundaries for areas.

---

1 This discussion is drawn in part from the author's unpublished Ph.D. thesis, "Census Tract Data and Social Structure: A Methodological Analysis," University of North Carolina, 1957. Daniel O. Price, Rupert B. Vance and James A. Norton have been of assistance in formulating these ideas. This paper was presented at the Annual Meeting of the American Statistical Association, Dec. 1958, under the title, "The Definition of Population Clusters and the Contiguity Ratio,"

2 In particular, Social Area Analysis by Eshref Shevky and Wendell Bell, (Stanford: Stanford University Press, 1955).

3 The coefficient of intraclass correlation, rho, may be the best statistic for this purpose. See Leslie Kish, "Differentiation in Metropolitan Areas", American Sociological Review, 19 (August, 1954).

4 "The Contiguity Ratio and Statistical Mapping," by R. C. Geary, The Incorporated Statistician, Vol. 5, No. 3.

A SUGGESTED METHOD FOR THE DELIMITATION OF POPULATION CLUSTERS

By: Arthur F. Loeben

Planning Commission, Montgomery County, Pennsylvania

A settlement pattern can be imagined in units of people, or in units of the places where people collect. In this paper the latter approach is used and the basic unit of the study is called the "stead." The farmstead and homestead is a familiar and well used notion in this country. It generally means the principal building, related structures, and areas directly related to the use of the stead. That notion has not been changed in its use in this paper. The use of the combining form "stead" has been expanded beyond farmstead to include each single unit of activity such as storestead, schoolstead, theaterstead, filling station stead and so on.

Steads are the locuses, or focal points of the activity of people. Each of us has a set of steads within which we circulate, spending most of our time at one stead or enroute to another. The home, work place, church, store, school, theater, post office and other special steads constitute a normal set of steads. Some busy people have additional steads, such as this and similar type meeting rooms.

The total pattern of steads in a landscape is a very useful index to the distribution of the population. If the steads can be isolated, and mapped, the raw-material or basic unit of the settlement pattern will become available in correct position and ready for further analysis.

Individual steads can be identified on aerial photography at a scale of 1:20 000 and symbolized with a dot. The true shape of a stead could be retained but, it has been found that if the work is to be used at scales smaller than 1:250 000, the shapes will blur and lose their identity.[1] A large dot 140 feet in diameter[2] is used to symbolize each stead, and is shown reduced to one-sixth its linear or one-thirty-sixth its area on Figure 2.

Each dot is a dispersed stead. Very large and prominent steads such as golf clubs, stone quarries, drive-in theaters and others are shown separately at true size and shape by a distinctive tonal pattern. Steads which are in clusters, close together and even piled on top of each other are isolated and delimited in what has been called a nucleated area. It is unnecessary to separately identify each stead within such areas. Therefore, many of the problems of photo interpretation of steads in congested areas are relieved for it is not always possible to single out a stead in such conditions.

The mapping of a nucleated area is the first level in the generalization of steads, and also the first cluster of population to be delimited. The nucleated areas have important characteristics as follows: first, the shape of the area conforms to the shapes of the steads enclosed; second, the areal extent is determined by the number and shape of the steads enclosed; third, the number and classification of steads enclosed must be determined separately and usually by field investigation if it is necessary; fourth, the criteria used in establishing the areas can be varied as changing conditions arise.

A nucleated area is characteristically the nucleus aroundwhich additional growth occurs in an urbanizing population. In quantitative terms a nucleated area is the area enclosing a minimum of 5 steads with the distance between any two steads no more than 250 feet, measured from center to center. When minimum conditions are established, nucleated areas are extended when steads lie within 250 feet of the edge. Notice on Figure 2 that nucleated areas vary both in size and, shape. The largest is the Borough of Gettysburg, and the smallest is often 5 steads compactly spaced at a road junction. Some areas have extended fingers, others seem to jump back and forth across the road. In some cases, only a few more steads located in the gaps would be enough for the nucleated areas to coalesce into a long string.

The criteria of 5 steads and a spacing of 250 feet between steads are based upon previous studies as cited below. Further investigation into each dimension would be most desirable and in one almost a necessity.

The spacing distance of 250 feet is from work by the Census Bureau.[3] In the definition of "urbanized area," certain areas are included if they have a group of 100 dwelling units or more with a density of 500 or more per square mile. If the 100 units (steads) were evenly dispersed as far apart as possible over one-fifth of a square mile they would be at a density equivalent of 500 per square mile. The distance between the units in such a dispersant would be 250 feet.[4] This distance is used in forming nucleated areas and in each area therefore, the steads are at a density equivalent to, or greater than the urban density used by the census.

The minimum number of 5 steads rests on the work of Trewartha at the University of Wisconsin.[5] His study of the Unincorporated Hamlet is an important work on the beginnings of nucleation in the settlement pattern. The work is based on intensive field study in Wisconsin, followed by an organized study over the United States with cooperating scholars and institutions. Trewartha concluded that a hamlet must have at least 5 steads. In addition, he suggested that the

hamlet have some non-farm activity and that the maximum linear distance between outermost buildings be no greater than 1/4 mile. It is interesting to note the close relationship between the 1/4 mile and the 250 feet of the census. If 5 steads are in a string each 250 feet apart the distance from center to center of the end steads would be 1 000 feet. By measuring from outside to outside rather than the center, the distance could easily be 1 300 feet, just 20 feet short of 1/4 mile.

Further work is certainly needed in the problem of the minimum number of units in the smallest of population clusters. While Trewartha found 5 to be significant, further study may find some other number, or a combination of factors to be more significant.

Figure 2 presents the settlement pattern of most of Adams County, Pennsylvania at the first stage of compilation and generaliztion. It was originally compiled at a large scale of 1 inch equals 1 667 feet and subsequently reduced by photography to its present scale. It is really the basic data compiled for further generalization and analysis. Figures 5, 6 and 7 are the end products of this stage in the research and are generalizations of Figure 2.

A major development of the project has been the D-Line Method of analyzing a distribution. The D-Line Method has been developed by Klimm[6] as a tool in the grouping, delimiting and generalizing of dispersants. The technique is fully explained in Technical Report 3 of this project, however, it can be explained here briefly as it is applied to points. Only when it is applied to units having area, do several complications arise.

Figure 4 shows the various types of agglomerations produced by applying the D-Line to a set of points. D is applied to a set of points and lines are drawn connecting each pair of points which lie within D. Three types of agglomerations are formed in this matter; linkages, hollow agglomerations and agglomerated areas.

A linkage is two steads connected by a line D or less in length and, linkages can occur singly or in a series. If a series of linkages closes in on itself like a string of beads a hollow agglomeration is formed. Notice example K on Figure 4. The distance across the center between any 2 steads is more than D, and therefore cannot form an agglomerated area.

An agglomerated area is formed by 3 steads each within D of the other 2. In Figure 4, H is a single agglomerated area, an isosceles triangle exactly D on each side. Groups of agglomerated areas occur as at J or E. In practical use the lines connecting individual agglomerated areas are not shown, and only the outer edges as at N are shown. But, it must be remembered that agglomerated areas of three or more steads are really a compounding of triangles.

While the technique of using D is independent of the length of D, the selection of D to be used depends on several factors. The choice of D is critical and, a search for the D seems futile. Because regions are different, and because study objectives do vary, a flexible D suited to varying demands is probably the best answer. Furthermore, there is the question of scale. The smaller the D the less agglomeration of steads with a resulting large scale generalization. As D increases more steads are agglomerated and a smaller scale of generalization is accomplished. A decision on the length of D must consider the region, the objective, and the scale of work and presentation.

Three different D's were applied to the same data on Figure 2. The D's used were 1 000 feet, 1 900 feet, and 1 mile. The first was arbitrarily selected, the second induced from the data on Figure 2, and the third used because it related to other studies in the project.

The empirical D of 1 900 feet is the distance between the dispersed steads on Figure 2 if all such steads were uniformaly spaced over that area. There are 3 835 dispersed steads on Figure 2, and the gross area of the map is 460. 2 square miles. The steads within the nucleated areas are not included and neither is the area of all such nucleated areas. The total area of the nucleated areas is estimated to be 4. 6 square miles. If the 3 835 steads were uniformaly spaced, each the center of a hexagon, over the 455. 5 square miles, the distance between any two would be 1 900 feet.[7]

The 1 900 foot D used on Figure 6 pulls together into agglomerated areas groups of steads which are closer together than the uniform distance. Many steads lie within D of one stead, but not within D of 2. Such cases appear as linkages. There are a number of dispersed steads more than 1 900 feet from their nearest neighbor. They appear as dots. Finally, there are empty areas without steads.

The nucleated areas are usually associated with, or the center of an agglomerated area. In the actual agglomerating technique each nucleated area is considered to be a mass of steads and the D distance is measured from the edge to the nearest stead. Notice examples M and L on Figure 4. The distance of the stead from the area determines whether a linkage, or an agglomerated area will be formed. In fact, when D becomes a mile in this area, as on Figure 7, the nucleated areas are submerged within the entire agglomerated area.

Figure 7, with its mile D, is really a map of the empty areas. The emptiness of South Mountain on the West, and its foot hills, as well as scattered empty areas south and northeast of Gettysburg are well brought out. The State Park, and the institution on the western edge are oasis of settlement in a sea of emptiness. To the south is a finger of settlement poking up a valley into the mountain. The scalloped edge is due to the limits

of the map. As D is increased a large proportion of the area is agglomerated and the degree of generalization increased so as to exclude only direct and relatively pure areas of opposite characteristics.

The effects of a successively increasing D, are illustrated in Figures 5 to 6 to 7. These maps could be altered by dropping the road net work, the linkages, and the dispersed steads thereby making a clearer map. There is also the possibility of generalizing nucleated, or agglomerated areas by applying a D-Line analysis to them rather than to steads. It must also be remembered that this test area around Gettysburg is well developed. The average density is about 85 persons per square mile. These techniques would have different effects applied to areas of different densities, or to areas with a more urban population.

The work reported on here was supported by the Geography Branch of the Office of Naval Research. The basic research project into principles and methods of generalization in geography was conducted at the University of Pennsylvania, and directed by Professor Lester E. Klimm. This particular section concentrated on the problem of generalizing features in a landscape which could be represented as points. Other sections of the project sought to generalize single features with areal extent such as cropland; and, multi-feature combinations such as a combination of woodland and pasture. The results have been reported to the Office of Naval Research under contract Nonr 551 (01).

FOOT NOTES

1. Arthur F. Loeben, 'Geographic Representation of Rural Settlement," an unpublished paper read before Section E of the 1956 Annual Meeting of the American Association for the Advancement of Science.

2. The dimension of 140 feet was selected after studying the average size of stead, and the visual retention of a dot upon several reductions.

3. Bureau of the Census, U. S. Census of Population: 1950, Vol. 1, Number of Inhabitants, p. XXVII.

4. In such a uniform distribution each stead is assumed to be the center of a hexagon. The distance between steads can be determined by the formula $D = 1.0746\sqrt{\frac{A}{N}}$ where D is the distance, A the area, and N the number of steads. The formula was introduced in: James A Barnes, and Arthur H. Robinson, "A New Method for the Representation of Dispersed Rural Population," Geographical Review, Vol. XXX (1940), pp. 134-37.

5. Glenn T. Trewartha, "The Unincorporated Hamlet," Annals of the Association of American Geographers, Vol. XXXIII (1943), pp. 32-81.

6. Lester E. Klimm, The D-Line Method of Analyzing a Distribution, Technical Report 3, Office of Naval Research, Contract Nonr 551 (01), 1958.

7. Barnes and Robinson, op. cit.

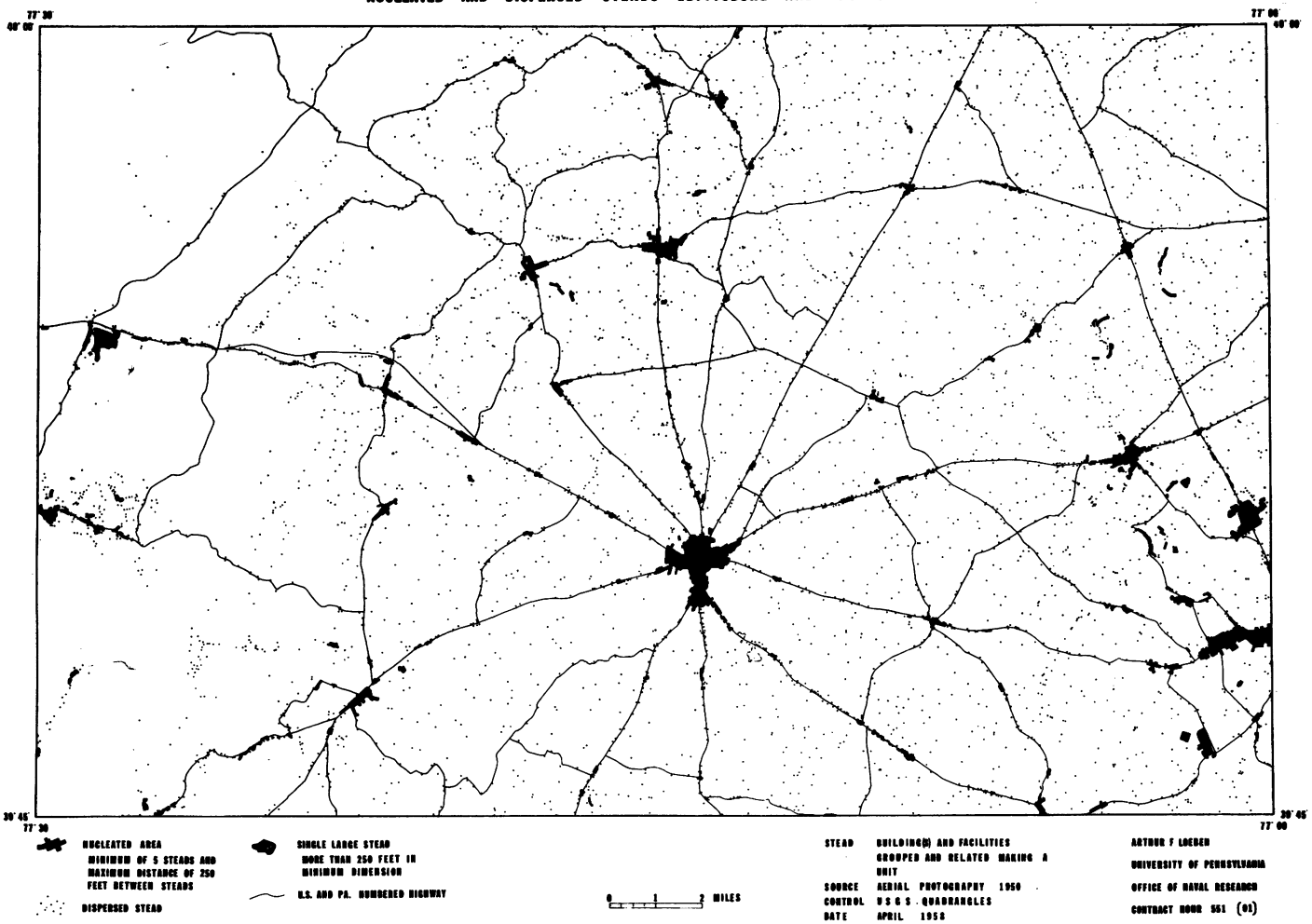

NUCLEATED AND DISPERSED STEADS GETTYSBURG AREA PENNSYLVANIA
77° 30'
40° 00'
77° 00'
40° 00'
39° 45'
77° 30'
39° 45'
77° 00'
NUCLEATED AREA
MINIMUM OF 5 STEADS AND MAXIMUM DISTANCE OF 250 FEET BETWEEN STEADS
DISPERSED STEAD
SINGLE LARGE STEAD
MORE THAN 250 FEET IN MINIMUM DIMENSION
U.S. AND PA. NUMBERED HIGHWAY
0 1 2 MILES
STEAD BUILDING(S) AND FACILITIES GROUPED AND RELATED MAKING A UNIT
SOURCE AERIAL PHOTOGRAPHY 1950
CONTROL U S G S . QUADRANGLES
DATE APRIL 1958
ARTHUR F LOEBEN
UNIVERSITY OF PENNSYLVANIA
OFFICE OF NAVAL RESEARCH
CONTRACT NOOR 551 (01)

Figure 2

TYPES OF AGGLOMERATIONS

Horizontal Scale.... 1 inch = 1 mile

D = 0.5 inches

H	Single Agglomerated Area
J, E	Group of Agglomerated Areas
N	Agglomerated and Nucleated Area
M	Linkage, Exactly D in Length
L	Agglomerated Area Attached to a Nucleated Area
G	Series of Linkages
A, F	Linkages to Agglomerated and Nucleated Areas
B, C	Hollow Agglomerations Attached to Agglomerated and Nucleated Areas
K	Hollow Agglomeration

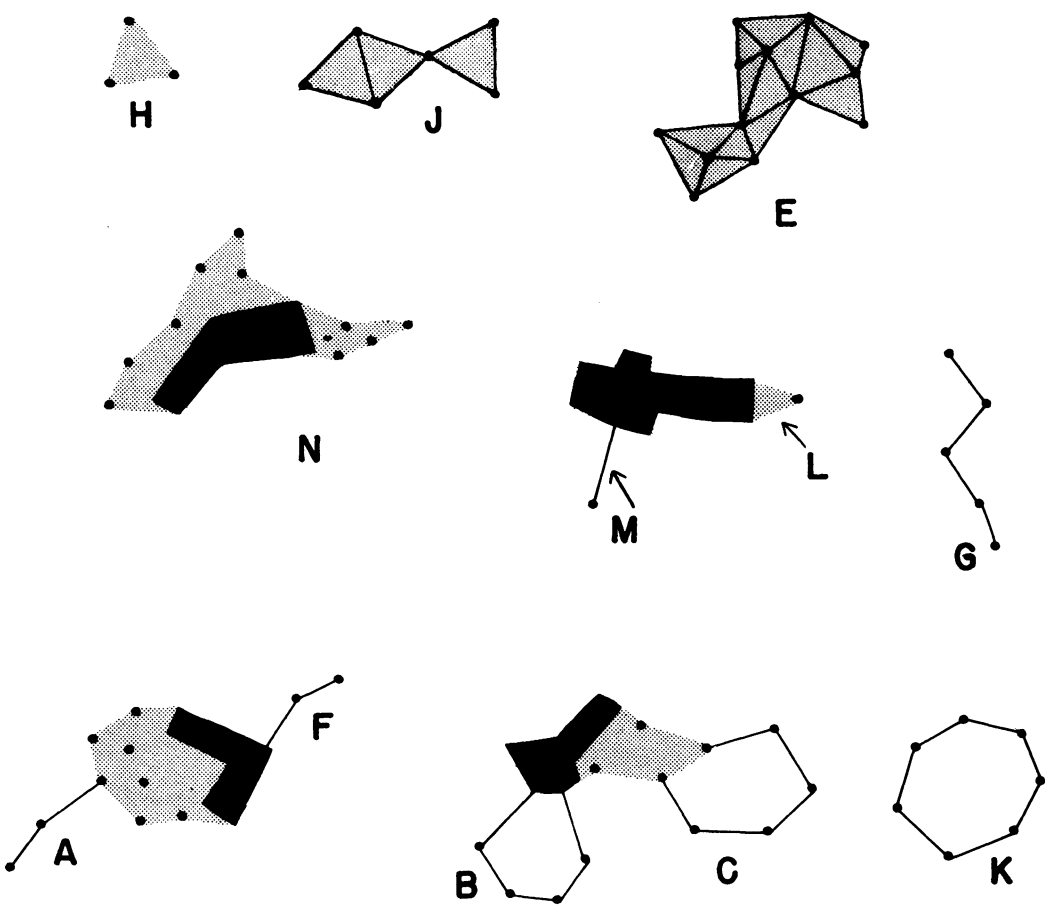
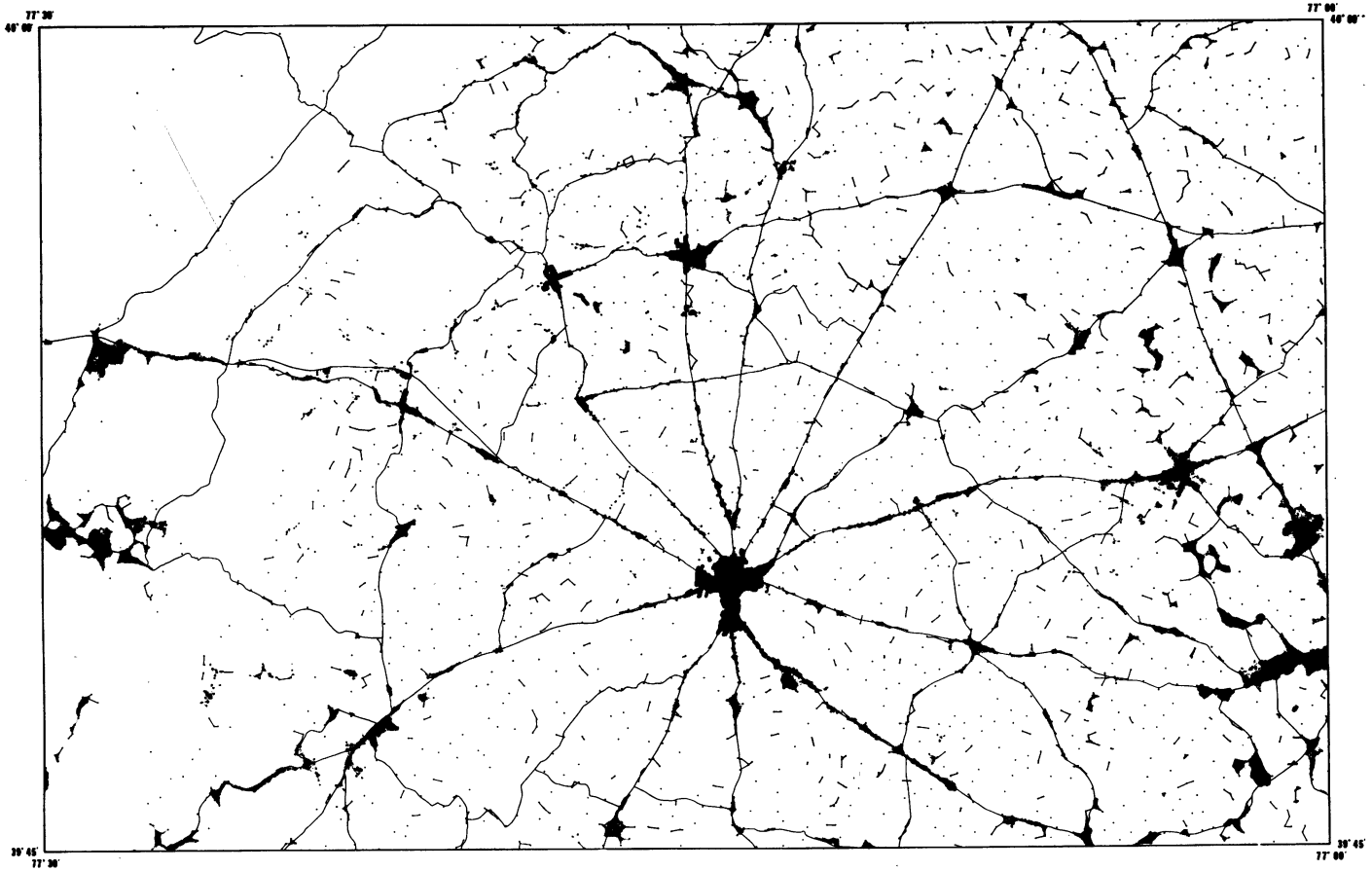


Figure 4

AGGLOMERATED AREAS AND LINKAGES GETTYSBURG AREA, PENNSYLVANIA

D = 1000 FEET



AGGLOMERATED AREA
NUCLEATED AREA
SINGLE LARGE STEAD
SINGLE DISPERSED STEAD
LINKAGE
U.S. AND PA. NUMBERED HIGHWAY

0 1 2 3 MILES

Figure 5

LINKAGE: A LINE CONNECTING 2 STEADS WITHIN D DISTANCE OF EACH OTHER.
AGGLOMERATED AREA: THE AREA ENCLOSED BY LINES CONNECTING 3 STEADS, EACH WITHIN D DISTANCE OF THE OTHER 2.

ARTHUR F. LOEBEN
UNIVERSITY OF PENNSYLVANIA
OFFICE OF NAVAL RESEARCH
CONTRACT NONR 551 (01)
OCTOBER 1958

AGGLOMERATED AREAS AND LINKAGES GETTYSBURG AREA, PENNSYLVANIA

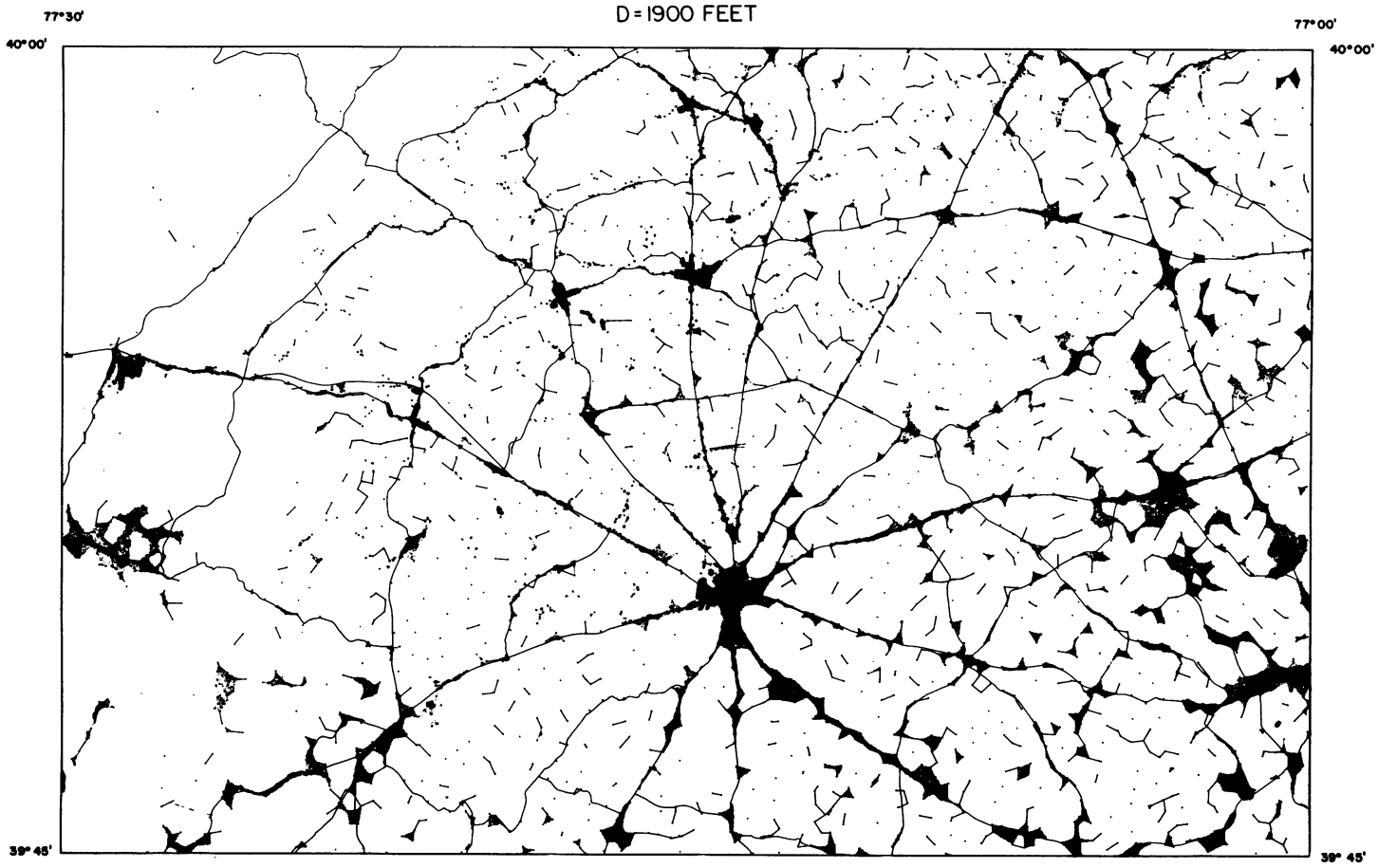


Figure 6

AGGLOMERATED AREAS AND LINKAGES GETTYSBURG AREA, PENNSYLVANIA

D = 1 MILE

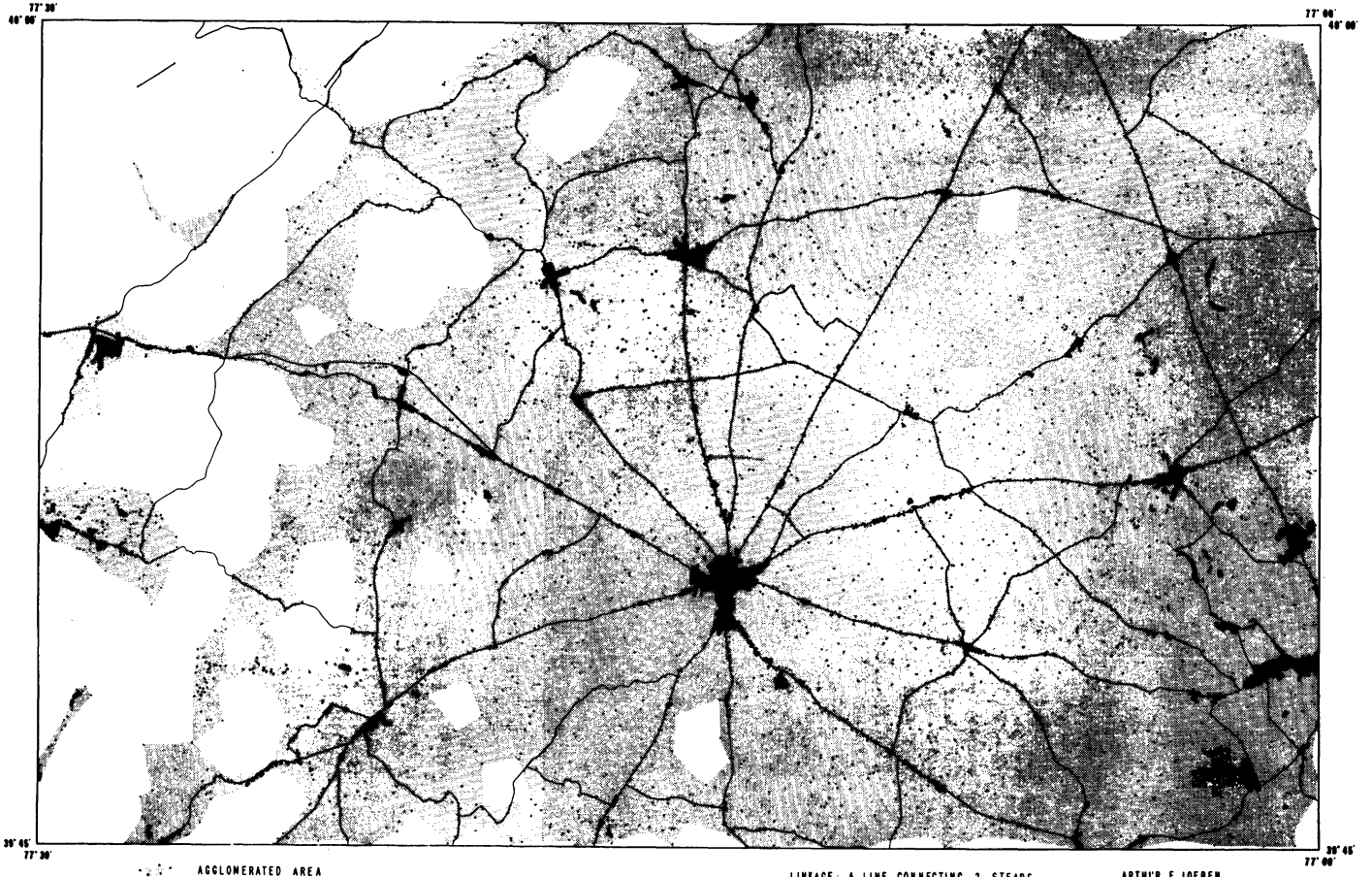


Figure 7

PATTERNS OF HEAPING IN THE REPORTING OF NUMERICAL DATA

By Stanley H. Turner, University of Pennsylvania

When a person is asked to volunteer numerical data, he either gives a precise answer or he does not. That is, he bases his response either on precise information or a somewhat hazy estimate. This paper tries to explain a certain type of pattern that emerges from data based on numerical estimates.

The central idea of this paper is that this pattern that emerges is related to the number system used by the estimator. To put it simply: the way we count, influences the way we estimate. That is, when a person estimates, he should do so in convenient units provided for him by the number system. Specifically, he should tend to over-report digits which are multiples of the divisors of the base of the number system and under-report digits which are not multiples of the divisors of the base of the number system.

As an example, consider the reporting of age. We may be unsure of our age or we may be asked to estimate the ages of other persons. Which digits are we more likely to report? This paper is concerned only with the ending digit of age since it is assumed that the decade of age is known accurately. Therefore, which ending digits of age are we more likely to report?

The most familiar way of counting is with the base ten. The divisors of the base ten are ten, five and two. The hypothesis states that estimates should heap at multiples of these three divisors; but more than that, it states that the most heaping should occur at ages ending in multiples of ten, the next largest at multiples of five, and the next largest at multiples of two. Figure I shows the rank order of heaping for part of the ending digits of age.

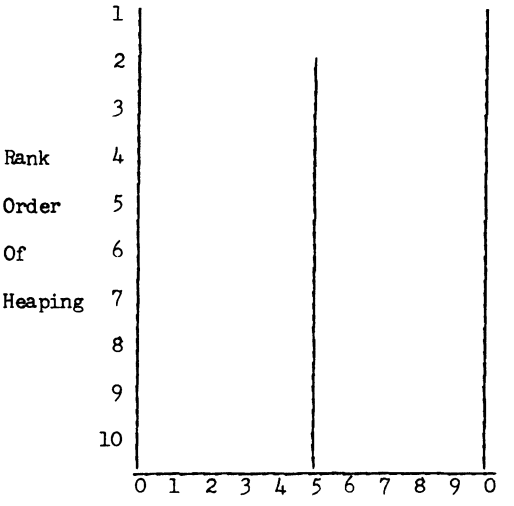


Figure I

Thus far, ages ending in multiples of ten, that is, ages ending in zero, are supposed to receive the most heaping. Ages ending in five, the next largest divisor, are supposed to receive the next largest amount of heaping. Only a single zero is needed in Figure I, but two are shown for symmetry and clarity.

Multiples of the next largest divisor, two, should all come next. That is, ages ending in two, four, six and eight should all receive the next largest amount of heaping. But notice that four and six are right next to five, which is supposed to attract a good deal of heaping. Furthermore, the other two even digits, two and eight, are not next to either zero or five. Therefore, two and eight should be free to attract more heaping than four or six. This line of reasoning implies the following additions to the expected pattern of heaping:

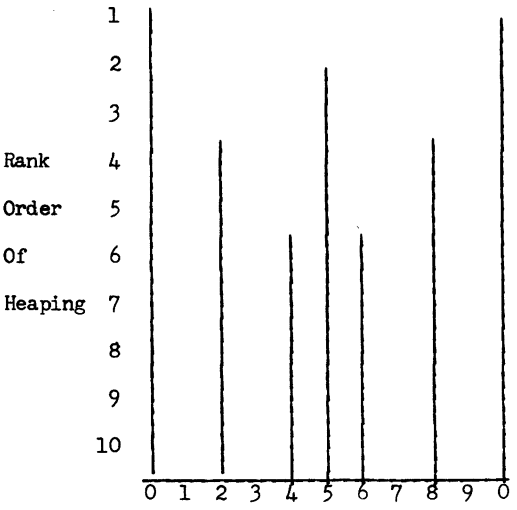


Figure II

All that remains is to fit in the remaining odd digits - one, three, seven and nine. Notice that the digits one and nine are between digits which are ranked as attracting a large amount of heaping. This should put one and nine at a disadvantage compared to three and seven. This enables the ranking to be completed as follows:

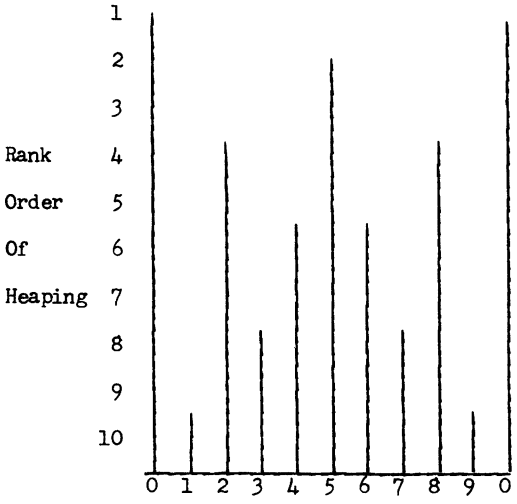


Figure III

From Figure III, the complete rank order prediction for all ending digits of age can be made:

Ending Digit Of Age	Predicted Rank Order
0	1.0
1	9.5
2	3.5
3	7.5
4	5.5
5	2.0
6	5.5
7	7.5
8	3.5
9	9.5

Eight decennial censuses of the United States population, covering the period from 1880 to 1950 were used to test the above expected rank orders. The number of people whose ages ended in each digit in each census was determined. However, a correction suggested by R. J. Myers was needed. If heaping were estimated by adding together all people whose ages ended in one, two, three, etc., a bias would be introduced. Consider the group of people whose ages end in one and those whose ages end in two. The former group is younger and therefore usually more numerous. That is to say, the sum of those aged 10 + 20 + 30 + 40 + 50 + 60 + 70 + 80 + 90 is usually greater than the sum of those age 11 + 21 + 31 + 41 + 51 + 61 + 71 + 81 + 91. In general, starting with any age tends to overstate the heaping at that age. Myers suggested that this bias could be removed by starting at each digit in turn and averaging the results.[1]

Table 1. shows the rank order of heaping for all ending digits of age during the entire period from 1880 to 1950.

Table 1.

RANK ORDER OF HEAPING FOR ENDING DIGITS OF AGE, BOTH SEXES, U. S. CENSUSES FROM 1880 - 1950*

Digit	1880	1890	1900	1910	1920	1930	1940	1950
0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0
2	4.5	4.0	4.0	4.0	4.0	4.0	4.0	3.0
3	7.0	6.0	8.5	8.0	8.0	8.0	9.0	8.5
4	6.0	7.0	6.0	6.5	8.0	6.5	7.0	6.5
5	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
6	4.5	5.0	7.0	5.0	5.0	6.5	6.0	6.5
7	8.0	8.0	8.5	9.0	8.0	9.0	8.0	8.5
8	3.0	3.0	3.0	3.0	3.0	3.0	3.0	4.0
9	9.0	9.0	5.0	6.5	6.0	5.0	5.0	5.0

*Columns 1880 - 1930 are based on data taken from Robert J. Myers (See Bibliography)

Table I shows that zero was the most frequently reported ending digit of age during the entire period. Ending digit five was next most frequently reported and digit one was the least reported of all the ending digits of age. The other digits changed their rank orders somewhat. It is interesting to note that the digit nine was the least stable in its rank order. It ranked ninth in 1880 and 1890 and then rose to fifth rank in 1900. None of the other digits displayed such variability.

It might be helpful to analyze these data separately for males and females since the sexes are known, or at least reputed, to differ in their willingness to report their ages. Indeed, the censuses do provide a breakdown of age reporting for males and females. But such figures can be easily misinterpreted. Remember that the census enumerator does not ask each and every person his or her own age. Rather one person is commonly asked to report the ages of perhaps several other persons. Since females may be reasonably expected to be home more frequently when the enumerator calls, then many of the figures listed in the census as male's ages are actually reported or estimated by females. In view of this observation, the decision was made not to analyze the ages of each sex separately.

Instead, an average rank order of each of the ending digits of age was computed. This was done in order to compare the observed rank order derived from the census data to the expected rank order derived from the hypothesis.

The results are shown in Table 2. The difference between the expected and the observed rank orders is small except for the digit nine. A statistical test showed that the overall pattern of heaping conformed quite closely to the predicted pattern. (Spearman's Rank Correlation Coefficient $r' = 0.96$).

Additional work is being done to test the hypothesis against census materials in other countries. All countries tested so far give similar results.

Table 2.

AVERAGE RANK ORDER OF HEAPING FOR ENDING DIGITS OF AGE, BOTH SEXES, U. S. CENSUSES FROM 1880 - 1950*

Ending Digit of Age	Predicted Rank Order	Observed Rank Order	Difference
0	1.0	1.0	0.0
1	9.5	10.0	0.5
2	3.5	3.9	0.4
3	7.5	7.9	0.4
4	5.5	6.7	1.2
5	2.0	2.0	0.0
6	5.5	5.7	0.2
7	7.5	8.4	0.9
8	3.5	3.1	0.4
9	9.5	6.3	3.2

*The average rank order of each ending digit was the sum of its rank order from 1880 - 1950 divided by eight, the number of censuses. All values were taken from Table 1.

Footnotes:

1. For further discussion of this technique of measuring heaping consult the excellent study by R. J. Myers, "Errors and Bias in the Reporting of Ages in Census Data." in the Handbook of Statistical Methods for Demographers, A. J. Jaffe, U.S. Government Printing Office. Washington, 1951.

Heaping as defined by Myers is equal to:

Heaping At Age Ending In Digit	Number of Persons At Age		Number of Persons At Age
0	=(10 + 20 + 30 + ... + 90)	x 1 +	(20 + 30 + 40 + ... + 100) x 9
1	=(11 + 21 + 31 + ... + 91)	x 2 +	(21 + 31 + 41 + ... + 101) x 8
2	=(12 + 22 + 32 + ... + 92)	x 3 +	(22 + 32 + 42 + ... + 102) x 7
.	.	.	. .
.	.	.	. .
.	.	.	. .
9	=(19 + 29 + 39 + ... + 99)	x10 +	(29 + 39 + 49 + ... + 109) x 0

FERTILITY ESTIMATES FROM BIRTH STATISTICS

By: K. R. Gabriel, Hebrew Univ., Jerusalem and Univ. of N. C. and Ruma Falk, Hebrew Univ., Jerusalem

I. THE NEED FOR MEASURES OF FERTILITY BASED ON BIRTH STATISTICS ONLY, AND THEIR INVESTIGATION

Refined techniques of fertility analysis have received much attention in recent years. They usually require detailed statistics both of births and of the composition of the female population. For many populations and sections for which fertility analysis is of interest -- such as occupational and ethnic groups -- such data are not usually available and simpler, if cruder, methods are still of importance.

Indices of fertility based on statistics of births only, i.e., not requiring data on the population, are of particular usefulness. This is because statistics of births (or confinements) can often be broken down by social, economic, residence, origin or other characteristics of the population for which no census data on age and sex composition are available. Such indices will be specially useful if their sampling errors can be estimated, thereby allowing their use in relatively small sub-groups of the population. A number of such indices are proposed and investigated in this study.

The validity of indices of fertility and other demographic indices is usually inferred from the logical implications of their methods of computation. Empirical validation is not often attempted, largely because statistically measurable ultimate criteria of fertility are rarely available -- except perhaps where cohort analyses have been completed. Indirect validation of indices may be obtained by correlating them with such time-honored measures of that of Total Fertility, i.e., age-of-mother-specific-birth-rates summed for all ages. Total fertility is, however, known to be prone to considerable short term fluctuations, yet neither a precise evaluation of its validity nor an acceptable method of smoothing its fluctuations is available. Critical examination of Total Fertility and similar measures based on age-specific-birth-rates is therefore indicated as a preliminary to the use of such measures as validating criteria.

Empirical investigation should ideally be carried out on a random sample from the universe of populations and dates for which validity is being studied. This is hardly practicable in demography where the sample is usually determined by the availability of statistics. In this study fertility indices have been computed and investigated on data for Australia for the 47 years 1909-1955 (this is apparently the longest available sequence of birth statistics cross classified by age of mother and order of birth). Further data are available for the Jewish population of Israel and its main origin sub-groups, though only for very few years. The Australian data give information about the behavior of the indices in time, whereas the Israeli data add some information about the characteristics of the indices in different populations of very different fertility (i.e., the origin sub-groups). The sources of the data are described in detail in Appendix II to this paper.

II. TOTAL FERTILITY AND TOTAL MATERNAL FERTILITY

Some notation will be introduced, all symbols referring to events in some given year. Denote the number of women of age a (or age group a) in the population by P(a); also the number of births of order i, and of all births, by $B_i(a)$ and B(a), respectively. The age-specific-birth-rate for mothers of age a is denoted by f(a) and defined as

$$f(a) = \frac{B(a)}{P(a)}$$

for all births,
and as

$$f_i(a) = \frac{B_i(a)}{P(a)}$$

for births of order i.
Thus

$$\sum_i f_i(a) = f(a).$$

($\sum_i$ and $\sum_a$ are used to indicate summation over all birth orders or all age groups, respectively). Total Fertility of all births is denoted by F and defined as

$$F = \sum_a f(a).$$

(The Gross Reproduction Rate is 0.485 F, where 0.485 is the proportion of females among births.) Similarly, Total Fertility of i-th births is denoted by F_i and defined as

$$F_i = \sum_a f_i(a),$$

whence

$$F = \sum_i F_i.$$

Total Fertility -- or the Gross Reproduction Rate -- is widely accepted as a fertility measure because it is age standardized and because it would measure the average number of births per mother precisely if there were no variations in age-specific-birth-rates. By the same reasoning total fertility of i-th births -- F_i -- would measure the proportion of women having i-th births. Also since the number of mothers corresponds to the number of first births (though not necessarily every year) F/F_1 would measure the average number of births per mother and thus qualify as a measure of total maternal fertility.

True total maternal fertility is presumably very highly correlated with true total fertility, as the proportion of childless women -- which determines the difference between the two values -- probably varies very closely inversely to fertility. However, year to year fluctuations in birth rates affect both the measure of total fertility F and that of total maternal fertility F/F_1, but not always in the same way, so that the correlation between the two measures is reduced. It is

difficult to judge generally whether F or F/F_1 would be more affected by fluctuations, though it might be thought that in F/F_1 fluctuations in the numerator and denominator might cancel out. As the validity and variability (due to fluctuations) of neither measure has been adequately investigated and as the logical construction of both is much the same, there seems no a priori reason to prefer either one as a measure of fertility.

The sequences of F and F/F_1 for Australia and for Israel are presented in Appendix Tables A-1 and A-2, respectively. The correlation between them is r = 0.542 for Australia and r = 0.883 for Israel with standard errors of estimate of F given F/F_1 being 0.352 for Australia and 0.515 for Israel (Tables 2 and 4). Inspection of the sequences shows that the correlation is due to the correspondence of the trends in Australia and the correspondence of the origin differentials in Israel. Short term fluctuations in the two indices differ markedly in each country, as for instance in the great depression when first births were relatively scarce and F/F_1 increased though F decreased a good deal (Australia). Undoubtedly, F/F_1 cannot be used to indicate changes in birth rates, but this does not necessarily impair its validity as a measure of fertility.

It must be remarked that it is not possible to eliminate these fluctuations by any of the standard statistical techniques. For these fluctuations are clearly neither random nor even of a simple stochastic character. Indeed for Australia it is not judged possible even to estimate the relative variability of the measures as it is not clear to what extent such variations as the fall of birth rates in the early thirties and their subsequent recovery should be considered fluctuations or real changes in fertility. For a period of relative stability in fertility -- though not in birth rates -- in Israel, i.e., for 1938-49, relative variability ($\sigma/\bar{x}$) was found to be 18% for F and 11% for F/F_1.

It is to be hoped that analysis by cohorts may solve some of these problems and provide some sort of final criterion for validation of fertility measures. Until such an analysis is available no final decision about the validity or superiority of either F or F/F_1 seems possible, hence both are used in this study.

III. RATIOS OF ALL BIRTHS TO FIRST BIRTHS

The use of birth statistics for fertility analysis is based on the distribution of births by order. Thus one expects high fertility to be associated with relatively many higher order births. Possible distortions due to the age composition of mothers may be corrected by considering each age group separately and then possibly averaging them with suitable weights. Two approaches are studied here, that using the ratio of all births to first births and that using mean birth order.

Estimation of maternal fertility by calculation of the ratio of all births to first births from statistics of confinements seems to have been first advocated by Gini (1934). This calculation can be relied on to give a correct estimate of fertility only if numbers of births and of first births do not vary much from year to year. Denoting Gini's index by S_g, we have

$$S_g = \frac{\sum_a B(a)}{\sum_a B_1(a)} = \frac{\sum_a P(a)f(a)}{\sum_a P(a)f_1(a)}$$

and this shows the index's dependence on the age composition of the female population. As age composition is irrelevant to fertility, S_g cannot be considered a satisfactory index of fertility.

Ratios of all births to first births of mothers of a specific age are obviously independent of the female population's age composition. Denote them by S(a) so that

$$S(a) = \frac{B(a)}{B_1(a)} = \frac{f(a)}{f_1(a)} .$$

The ratio of all births to first births for all ages can then be expressed as a weighted mean of the age specific ratios. Either as an arithmetic mean

$$\frac{\sum_a w_1(a)\, S(a)}{\sum_a w_1(a)} ,$$

or as a harmonic mean

$$\frac{\sum_a w(a)}{\sum_a w(a)\, S^{-1}(a)} ,$$

where $w_1(a)$ and w(a) are weights.

A special case of these means is Gini's index S_g, being the arithmetic mean with weights $w_1(a) = B_1(a)$ and the harmonic mean with weights $w(a) = B(a)$. Another special case is total maternal fertility

$$\frac{F}{F_1} = \frac{\sum_a f(a)}{\sum_a f_1(a)} .$$

This is given by the arithmetic mean with weights $w_1(a) = f_1(a)$, and by the harmonic mean with weights $w(a) = f(a)$.

These weighted means are of importance when the true weights are unknown and assumed weights are used instead. Of course only the proportional distribution of the weights matters, and small differences are unlikely to affect the weighted means very much. Thus if S(a) can be computed from birth statistics, an estimate of maternal fertility can be obtained by averaging the S(a) with appropriate weights.

Hajnal (1948) proposed using as weights the rates of another population, preferably one with similar fertility patterns to those of the population studied. (Hajnal's study is in terms of ratios and rates specific to marriage duration. Our use of ratios and rates specific to age of mother (Gabriel, 1953) is analogous, though the results may differ). Thus if a bar denotes the rates and ratios from some population used as a standard, the weights would be

$$\bar{w}_1(a) = \bar{f}_1(a)/\bar{F}_1$$

and

$$\bar{w}(a) = \bar{f}(a)/\bar{F} .$$

The two estimates of fertility according to Hajnal thus are: the arithmetic mean

$$S_f = \sum_a \bar{w}_1(a)\, S(a) ,$$

and the harmonic mean

$$S_a = \left\{ \sum_a \bar{w}(a)\, S^{-1}(a) \right\}^{-1} .$$

Evidently if age-specific-birth rates in the standard population are proportional to those in the population studied, i.e.,

if $\bar{w}_1(a) = f_1(a)/F_1$

and $\bar{w}(a) = f(a)/F$,

then $S_f = S_a = F/F_1$.

Ratios of all births to first births increase with age of mother. Hence if the standard weights are concentrated at higher ages than the unknown true weights, then the means will become too large, and conversely for standard weights concentrated at lower ages.

It might be surmised that generally the less fertile a population is, the later the ages at which first births are concentrated, and the earlier the ages of all births. The latter might be explained by the fact that family limitation usually is most marked at later childbearing ages. (Hajnal (1948) has suggested that such a relationship might hold for marriage-duration-specific-birth-rates, i.e., for marriage duration at confinement. It might of course hold for marriage duration and yet not for age, or vice versa). If these surmises are right then when the age-specific-birth-rates of some standard population are used as weights one would find

$$S_a < F/F_1 < S_f$$

if the population studied has higher fertility than the standard population, and

$$S_a > F/F_1 > S_f$$

if the population studied has lower fertility than the standard population. In either case F/F_1 would be bracketed by S_f and S_a. The actual existence of such a pattern is investigated below.

Actual computations were carried out for five year age groups of mothers. For Australian statistics two alternative sets of weights were used: (1) corresponding to high fertility years -- 1909 - 1927 average age-specific-birth-rates, and (2) corresponding to low fertility years -- 1928-1945 average rates. For Israel all computations were weighted corresponding to 1949 age-specific-birth-rates in the entire population.

Choice of weights was found to affect the actual level of S_f and S_a appreciably, but to have practically no effect at all on comparisons of S_f and S_a between times and populations. Thus for Australia the two sets of weights gave S_f values with a difference of 0.11 on the average for the 47 years and a correlation of 0.9994, and S_a values with a difference of 0.33 on the average and a correlation of 0.9992. Few corresponding data are available for Israel and for a few computations similar results were observed.

The relation between the two measures S_f and S_a is much as the relation between two differently weighted estimates of either of them. There exists a difference in level: S_f is on the average 0.12 above S_a for Australia (second set of weights) and 0.34 for Israel. The correlation between the two measures are 0.9867 for Australia and 0.9932 for Israel.

Since choice of weights affects the level of S_f and S_a and since the two measures are strongly correlated it is not surprising to find that the surmised relation between them and F/F_1 does not generally hold. In fact for both Australia and Israel the majority of observations show S_f and S_a both above or both below F/F_1. Evidently the relations between fertility and ages at confinement are not as simple as surmised above.

Though no consistent ordering of the three measures S_f, S_a and F/F_1 can be observed, the three are strongly correlated. Correlations, regressions and standard errors of estimate of F/F_1 given each of S_f and S_a, are shown in Tables 1 and 3. It appears that each of the indices can give a close estimate of F/F_1, and S_f is the better one -- the standard error of estimate of F/F_1 with respect to S_f being 0.10-0.13 for Australia and 0.22 for Israel. The corresponding correlations are about 0.98 in either country. The regression equations are rather different for the two countries, possibly due to the different weights employed.

Correlations with F are considerably lower -- Tables 2 and 4 -- and the standard error of estimate of F is about 0.35 for Australia and 0.50 for Israel. As the measures S_f and S_a were shown to correspond in their construction to F/F_1 rather than to F it is not surprising that they should be more highly correlated with the former. Also, the correlations of S_f and S_a with F are

very nearly the same as those of F/F_1 with F, both for Australian and for Israeli data. One may therefore say that possible shortcomings of S_f and S_a as measures of fertility may lie in their being estimates of F/F_1 rather than of F. It is remarkable that the S_f and S_a measures seem to be practically as good for estimating F as the F/F_1 ratio is, despite the fact that the former are calculated without statistics of the age composition of the population.

IV. MEAN BIRTH ORDER

An alternative approach to the measurement of fertility from birth statistics is based on mean birth order. Though mean birth order is not a function of maternal fertility (i.e., mean number of births per mother) only, it can reasonably be assumed to be closely correlated with it.[1]

Mean birth order, denoted by μ_B, will be defined for a population with uniform age distribution in the reproductive ages (just as Total Fertility is), i.e.,

$$\mu_B = \frac{1}{F} \sum_i i \, F_i \ .$$

Introducing the expression for F_i in terms of births and population this becomes

$$\mu_B = \frac{1}{F} \sum_a \frac{1}{P(a)} \sum_i i \, B(a) \ .$$

Defining mean birth order for mothers of age a as

$$\mu_B(a) = \sum_i i \, B_i(a)/B(a)$$

we obtain

$$\mu_B = \frac{1}{F} \sum_a f(a) \, \mu_B(a) \ .$$

Now when the age-specific-birth rates f(a) are not known for a population one might substitute those of a standard population with similar fertility, $\bar{f}(a)$, say. Defining the substitute weights as $\bar{w}(a) = \bar{f}(a)/\bar{F}$ one would then obtain the estimate

$$\hat{\mu}_B = \sum_a \bar{w}(a) \, \mu_B(a) \ .$$

The weighting is the same as for S_a, and the reasoning that with suitable weights such an estimate be good also is similar to that for S_f and S_a. In this study $\hat{\mu}_B$ was computed only by using five-year age groups.

A possible advantage in the use of $\hat{\mu}_B$ over S_f, S_a could be in its using the information of the entire birth order distribution whereas S_f, S_a use only the proportion of first births. This might reduce the sampling error of $\hat{\mu}_B$ relative to S_f and S_a. An obvious disadvantage of $\hat{\mu}_B$ is, however, that even μ_B which it is supposed to estimate is not strictly the same as fertility. Hence $\hat{\mu}_B$ can at best only serve as an index of fertility.

Finally, some statisticians have used $\mu_B(a_s)$ of some particular age group a_s directly as an index of fertility. (We are indebted for this idea to Prof. R. Bachi of the Hebrew University, Jerusalem.) This seems a justifiable procedure as it is reasonable to suppose that variations in μ_B will generally be reflected in similar variations in $\mu_B(a)$ at all ages -- though exceptions are of course possible. This method has the attraction of easy computation and no need for assumptions about the suitability of weights but it is of course subject to possible distortion if the $\mu_B(a)$'s should behave differently at different ages. The age group chosen is usually about the middle of the reproductive ages, and in this study the use of each of the three age groups 25-29, 30-34 and 35-39 -- denoted by a_1, a_2 and a_3, respectively -- has been investigated.

For Australia $\mu_B(a_s)$ has been computed for the three age groups mentioned above, and the mean $\hat{\mu}_B$ has been calculated with weights proportional to the mean of the 1909-1955 age-of-mother-specific-birth-rates. For Israel only $\mu_B(30\text{-}34)$ is available. (See Appendix Tables A-1 and A-2.)

Correlations of the μ_B indices with F/F_1 and F -- Tables 1 to 4 -- are found to be just slightly lower than those of S_f and S_a. It is not possible to draw reliable conclusions as to the relative standing of the various indices as the observations are too few. Yet, it would seem that $\mu_B(25\text{-}29)$ is practically as good an index of maternal fertility as S_f and S_a. Also its correlation with F is much below its correlation with F/F_1, clearly for the same reasons as were discussed with regard to S_f and S_a.

The finding that $\mu_B(a)$ predicts F/F_1 best for the lowest age group, even better than $\hat{\mu}_B$ does, indicates that in the higher mother age groups mean birth order is less closely related to mean fertility. Perhaps this is because only the more fertile women bear at those ages at all.

From the purely computational point of view the $\mu_B(a)$ measures are definitely preferable to any of the weighted means. Both for the estimate itself and for its standard error (see Appendix I) the computations are greatly reduced. Also if only one age group is used, this may mean considerable saving in sorting and tabulating whenever these are done specially for the purpose of fertility analysis.

V. CONCLUSIONS

It has been shown that indices of fertility based on birth statistics alone are very highly correlated with a measure of maternal fertility. For most purposes it would seem that some of the indices are practically as reliable as the ratio F/F_1 computed from age-specific-birth-rates. The best of the indices seem to be S_f and $\mu_B(25\text{-}29)$, the latter having the additional advantage of requiring only few and simple computations, both for the estimate and for its standard error.

Indices of fertility which are computed from birth statistics are based on the distribution of births by order. They are therefore related directly to maternal fertility rather than to all fertility. This explains why their correlations with Total Fertility are much the same as those of Total Maternal Fertility.

For the lack of any available ultimate statistical criterion of fertility it cannot be said whether F -- which estimates mean number of births per woman -- or F/F_1 -- which estimates mean number of births per mother -- is the more valid measure of fertility. Hence no final evaluation of the measures based on birth statistics only is possible at this stage.

APPENDIX I

SAMPLING ERRORS

The magnitude of sampling variability is presented here by the standard error or its square, the variance. The sampling distributions have not been studied but for sizeable samples they may safely be assumed to be approximately normal.

Mean birth order estimates are clearly unbiased but the bias of the S_f and S_a estimates has not been investigated. For large samples it is likely to be negligible.

Variances were first computed for the individual terms of the various indices, i.e., for proportions of first births to all births, or mean birth order, at a given age of mother. Then they were transformed and summed to give the variances of the fertility indices.

For this derivation we used the two well known theorems that if X_i are independent variables and a_i constant weights

$$\mathrm{Var}\left(\sum_i a_i X_i\right) = \sum_i a_i^2 \,\mathrm{Var}(X_i)$$

and

$$\mathrm{Var}\left(\frac{X_1}{X_2}\right) = E\left(\frac{X_1}{X_2}\right)^2 \left\{\frac{\mathrm{Var}(X_1)}{(E X_1)^2} + \frac{\mathrm{Var}(X_2)}{(E X_2)^2}\right\} \text{ approximately}$$

Further we have used the known expressions for the variance of a proportion and of a mean. However, in order to simplify the presentation we have introduced in the denominator the number of observations instead of the degrees of freedom which would have been one less.

Thus as

$$S(a) = B(a)/B_1(a)$$

$$\mathrm{Var}(S^{-1}(a)) = \frac{(S^{-1}(a))\,(1-S^{-1}(a))}{B(a)}$$

and therefore

$$\mathrm{Var}(S(a)) = \frac{S(a)\,(S(a)-1)}{B_1(a)} .$$

Hence it is found that

$$\mathrm{Var}(S_f) = \sum_a (\bar{w}_1(a))^2 \,\frac{S(a)\,(S(a)-1)}{B_1(a)}$$

and

$$\mathrm{Var}(S_a) = \frac{1}{S_a^4} \sum_a (\bar{w}(a))^2 \,\frac{S^{-1}(a)(1-S^{-1}(a))}{B(a)} .$$

Also as $\mu_B(a) = \sum_i i\, B_i(a)/B(a)$

we obtain $\mathrm{Var}(\mu_B(a)) = \sigma_B^2(a)/B(a)$

where $\sigma_B^2(a) = \sum_i i^2 B_i(a)/B(a) - (\mu_B(a))^2$.

For $\hat{\mu}_B$ it follows that

$$\mathrm{Var}(\hat{\mu}_B) = \sum_a (\bar{w}(a))^2 \,\frac{\sigma_B^2(a)}{B(a)} .$$

It should be quite clear that the above formulae refer to estimates of the standard errors which can be calculated from the sample and not to the population values. All the expressions given here can easily be calculated along with the estimate itself.

APPENDIX II

DATA AND ESTIMATES FROM WHICH FERTILITY ESTIMATES WERE COMPUTED

The measures of fertility investigated in this paper are illustrated by computations from Australian data for each of the years 1909 to 1955. The basic figures were taken from publications of the Australian Commonwealth Bureau of Statistics and supplemented by figures made available to the authors by the courtesy of that Bureau. Where figures were incomplete or did not correspond precisely to the requirements of this study, estimates and adjustments were introduced, as described in this Appendix. These estimates and adjustments were often crude but were deemed adequate for the purpose of calculating and comparing fertility measures in this paper. No other use of these estimates should be made without investigating whether they are reliable enough for such purposes.

Figures of the female population by five year age groups were available for all years since 1921. They are based on the censuses of 1921, 1933, 1947 and 1954, and on intercensal estimates. These figures were used here as published.

For the years 1909-1920 an interpolation-extrapolation procedure was used to estimate the female population. For each five year age group linear interpolation between 1911 and 1921 census

figures gave a first estimate (for 1909 and 1910 the same trend was extrapolated). The interpolated figures were then adjusted proportionally so that their sum for ages 10-54 be equal to the estimated number of women at those ages in the population of that year. Since no such numbers had been published, they were estimated by multiplying the total number of females in each year (published) by the average proportion in age group 10-54 among females in the censuses of 1911 and 1921.

The basic data on births were tables of nuptial confinements by age of mother and previous issue. These figures relate only to live births and to previous issue from the present marriage. For the estimation of fertility it would have been desirable to have figures for all previous issue -- however, no adjustment has been made for this. Furthermore, ex-nuptial confinements were not included in the calculations, except that 15.2 percent of them were added to nuptial first births. (This is the percentage of legitimations to ex-nuptial births in the preceding year, computed from 1924-1955 figures.) These additional 15.2 percent of ex-nuptial live births were distributed among mothers' age groups in proportion to other first births. A small number of births for which age of mother and previous issue were unknown were also added in proportionally to other births.

These adjusted figures for births were used in the computation of birth rates. The other measures of fertility were computed directly from the unadjusted data. Omission of the legitimations (about 1 percent of the births) must thus have slightly increased both mean birth order and ratios of all births to first births. However, since this omission occurred in all years its effect must have been systematic and comparisons of trends and changes should not have been affected.

The data for Israel are taken from the first author's research on fertility in Israel (Gabriel, 1957). The calculations are based on official government statistics of births and of the population, supplemented by some estimates where data were incomplete. For details the reader is referred to Chapter 6 of the above work.

ACKNOWLEDGEMENTS

We are indebted to the Australian Commonwealth Bureau of the Census and Statistics for providing us with information and data on Australian births and populations. The interpolations and adjustments of the data are, however, entirely our responsibility.

We wish to acknowledge the help of Tuvia Blumental and Menahem Naveh who did most of the computing work, and the financial support of the Hebrew University, Jerusalem and the University of North Carolina who supported, respectively, the analysis and the presentation of this study.

REFERENCES

R. Bachi, (personal communication).

K. R. Gabriel (1953) "The Fertility of the Jews in Palestine", Population Studies, VI. pp. 273-305.

K. R. Gabriel (1957), Nuptiality and Fertility in Israel, Ph.D. Thesis at Hebrew University, Jerusalem (in Hebrew with English Summary).

C. Gini (1934) "Sur la fécondité des marriages", Bull. Inst. Intern. Statist. XXVII 2 livr.

J. Hajnal (1948) "The Estimation of Total Family Size of Occupation Groups from the Distribution of Births by Order and Duration of Marriage", Population Studies, II, pp. 305-317.

FOOTNOTES

1. It is easily shown that, if μ_M and σ_M^2 are the mean and variance of the distribution of mothers by number of births and μ_B is mean birth order, then $\mu_B = 1/2\left\{1 + \mu_M \left[1 + \left(\frac{\sigma_M}{\mu_M}\right)^2\right]\right\}$.

Thus, mean birth order is affected by the relative variation of fertility as well as by mean fertility. Yet it would appear that relative variation of fertility does not vary as much as mean fertility and thus the correlation between μ_B and μ_M would be high.

Table 1 Correlations, Regressions and Standard Errors of Estimate of Various Measures with Total Maternal Fertility - F/F_1. (Australia 1909-55; 47 observations, F/F_1-mean 3.345, Standard Deviation .5581).

	Mean	Regression Equation	Standard Error of Estimate	Correlation
S_f(1909-27 weights)	3.509	1.0458 S_f - 0.3248	.1043	.9824
S_a(1909-27 weights)	3.609	1.2445 S_a - 1.1464	.1892	.9408
S_f(1928-45 weights)	3.396	1.1011 S_f - 0.3943	.1296	.9726
S_a(1928-45 weights)	3.281	1.4350 S_a - 1.3633	.2116	.9253
μ_B(25-29)	2.381	2.9642 μ_B - 3.7127	.0997	.9839
μ_B(30-34)	3.301	1.4711 μ_B - 1.5112	.1798	.9467
μ_B(35-39)	4.447	0.8222 μ_B - 0.3112	.2639	.8811
$\hat{\mu}_B$(1909-55 weights)	2.926	1.7767 $\hat{\mu}_B$ - 1.8536	.2109	.9258

Table 2 Correlations, Regressions and Standard Errors of Estimate of Various Measures with Total Fertility - F. (Australia 1909-1955; 47 observations, F-mean 2.659, Standard Deviation .4189)

	Mean	Regression Equation	Standard Error of Estimate	Correlation
F/F_1	3.345	.4066F/F_1 + 1.2988	.3521	.5417
S_f(1909-27 weights)	3.509	.4463 S_f + 1.0931	.3475	.5584
S_a(1909-27 weights)	3.609	.5545 S_a + 0.6577	.3475	.5584
S_f(1928-45 weights)	3.396	.4799 S_f + 1.0292	.3457	.5648
S_a(1928-45 weights)	3.281	.6488 S_a + 0.5302	.3478	.5573
μ_B(25-29)	2.381	1.2039 μ_B - 0.2075	.3546	.5323
μ_B(30-34)	3.301	0.5189 μ_B + 0.9463	.3752	.4448
μ_B(35-39)	4.447	0.2175 μ_B + 1.6917	.3982	.3105
$\hat{\mu}_B$(1909-55 weights)	2.926	0.5315 $\hat{\mu}_B$ + 1.1039	.3894	.3689

TABLE A-1. VARIOUS INDICES OF FERTILITY - AUSTRALIA 1909-1955.

	F	F/F_1	$S_f^{(1)}$	$S_f^{(2)}$	$S_a^{(1)}$	$S_a^{(2)}$	$\mu_B(a_1)$	$\mu_B(a_2)$	$\mu_B(a_3)$	$\hat{\mu}_B$
1909	3.178	4.406	4.425	4.238	4.216	3.780	2.665	3.839	5.412	3.382
1910	3.211	4.407	4.460	4.271	4.245	3.805	2.667	3.873	5.316	3.375
1911	3.249	4.149	4.173	4.008	4.059	3.658	2.634	3.803	5.266	3.330
1912	3.405	4.015	4.109	3.950	4.023	3.620	2.629	3.797	5.232	3.318
1913	3.339	3.864	3.974	3.824	3.939	3.552	2.602	3.738	5.144	3.270
1914	3.301	3.891	3.956	3.815	3.982	3.597	2.599	3.730	5.151	3.271
1915	3.155	3.886	3.963	3.819	3.984	3.599	2.582	3.703	5.110	3.249
1916	3.028	3.911	4.021	3.873	4.006	3.615	2.565	3.675	4.995	3.217
1917	2.950	4.291	4.304	4.149	4.286	3.862	2.625	3.732	5.084	3.271
1918	2.800	4.501	4.526	4.358	4.445	3.993	2.659	3.747	5.010	3.281
1919	2.670	4.282	4.354	4.194	4.222	3.786	2.655	3.746	5.016	3.260
1920	2.911	3.443	3.560	3.424	3.535	3.193	2.476	3.571	4.940	3.122
1921	2.893	3.316	3.387	3.263	3.429	3.107	2.429	3.553	4.935	3.090
1922	2.900	3.586	3.678	3.546	3.727	3.374	2.449	3.555	4.931	3.115
1923	2.815	3.620	3.677	3.552	3.779	3.428	2.451	3.504	4.844	3.091
1924	2.773	3.648	3.731	3.606	3.832	3.474	2.454	3.492	4.829	3.090
1925	2.739	3.637	3.782	3.647	3.838	3.474	2.475	3.513	4.831	3.103
1926	2.638	3.566	3.700	3.578	3.797	3.441	2.483	3.479	4.751	3.079
1927	2.594	3.480	3.673	3.541	3.727	3.374	2.468	3.451	4.741	3.066
1928	2.554	3.406	3.594	3.473	3.684	3.341	2.448	3.474	4.685	3.052
1929	2.437	3.417	3.608	3.490	3.717	3.375	2.448	3.436	4.650	3.038
1930	2.389	3.407	3.592	3.474	3.717	3.377	2.445	3.391	4.638	3.016
1931	2.179	3.464	3.682	3.562	3.800	3.450	2.460	3.443	4.715	3.051
1932	2.019	3.549	3.811	3.683	3.898	3.531	2.475	3.442	4.683	3.061
1933	2.002	3.414	3.689	3.556	3.751	3.399	2.421	3.410	4.598	3.018
1934	1.947	3.233	3.476	3.350	3.552	3.226	2.381	3.344	4.604	2.981
1935	1.958	3.037	3.257	3.138	3.359	3.059	2.332	3.241	4.489	2.917
1936	2.021	2.874	3.026	2.922	3.173	2.901	2.249	3.114	4.354	2.826
1937	2.046	2.826	3.079	2.957	3.158	2.886	2.218	3.083	4.339	2.805
1938	2.048	2.765	2.876	2.786	3.077	2.827	2.180	2.992	4.235	2.747
1939	2.065	2.708	2.798	2.717	3.028	2.789	2.154	2.948	4.096	2.701
1940	2.150	2.699	2.747	2.669	2.980	2.747	2.135	2.871	4.020	2.652
1941	2.213	2.588	2.677	2.603	2.909	2.682	2.101	2.846	3.896	2.603
1942	2.222	2.593	2.703	2.628	2.922	2.687	2.105	2.844	3.838	2.584
1943	2.400	2.539	2.625	2.558	2.853	2.627	2.085	2.797	3.732	2.531
1944	2.459	2.825	2.876	2.805	3.099	2.844	2.143	2.825	3.671	2.540
1945	2.565	2.822	2.883	2.811	3.098	2.841	2.142	2.826	3.619	2.527
1946	2.804	2.725	2.811	2.743	3.039	2.792	2.119	2.798	3.587	2.491
1947	2.871	2.585	2.763	2.696	2.973	2.728	2.103	2.793	3.592	2.477
1948	2.772	2.816	3.036	2.971	3.269	3.003	2.188	2.853	3.594	2.528
1949	2.769	2.897	3.184	3.117	3.410	3.130	2.200	2.915	3.619	2.556
1950	2.845	2.985	3.288	3.219	3.525	3.237	2.229	2.918	3.631	2.573
1951	2.832	3.005	3.365	3.292	3.598	3.299	2.256	2.930	3.669	2.598
1952	2.945	2.963	3.354	3.280	3.581	3.281	2.282	2.973	3.676	2.617
1953	2.955	3.008	3.442	3.368	3.667	3.360	2.314	3.011	3.707	2.648
1954	2.953	3.084	3.558	3.482	3.789	3.469	2.345	3.036	3.764	2.676
1955	3.005	3.147	3.692	3.612	3.912	3.577	2.394	3.077	3.777	2.709

Notes $S_f^{(1)}$, $S_a^{(1)}$ computed with 1909-27 average birth rates as weights, $S_f^{(2)}$, $S_a^{(2)}$ with 1928-45 average rates. a_1, a_2 and a_3 refer to age groups 25-29, 30-34 and 35-39, respectively.

Table A-2 VARIOUS INDICES OF FERTILITY - ISRAEL: ALL WOMEN 1930-1954, WOMEN BY CONTINENT OF BIRTH 1938/40, 1944/45, 1949, 1951-1954.

	F	F/F_1	S_f	S_a	$\mu_B(30\text{-}34)$
All Women					
1938	2.48	2.23	2.24	2.24	
1939	2.23	2.20	2.21	2.22	
1940	2.35	2.17	2.15	2.17	
1941	2.12	2.05	1.98	2.03	
1942	2.38	2.00	1.93	2.01	
1943	3.11	2.13	1.98	2.07	
1944	3.44	2.64	2.38	2.41	2.50
1945	3.53	2.81	2.51	2.56	2.58
1946	3.34	2.72	2.52	2.53	2.63
1947	3.54	2.46	2.37	2.35	2.73
1948	3.08	2.32	2.25	2.27	2.73
1949	3.43	2.45	2.44	2.45	2.76
1950	3.90	3.05	3.09	2.99	3.12
1951	4.01	3.04	3.45	3.25	3.44
1952	3.98	3.29	3.68	3.51	3.61
1953	3.88	3.37	3.88	3.69	3.69
1954	3.59	3.63	4.20	3.97	3.75
Israel-born Women					
1938/40	3.54	3.85	4.15	3.59	4.60
1944/45	3.87	3.95	3.56	3.13	3.69
1949	3.56	3.18	3.22	2.83	3.86
1951	3.55	3.46	3.65	3.12	3.80
1952	3.36	3.50	3.75	3.15	3.61
1953	3.22	3.39	3.61	3.06	3.70
1954	2.89	3.44	3.38	3.10	3.54
Women born in Asia and Africa					
1938/40	4.62	5.25	6.84	5.86	5.61
1944/45	4.99	4.71	5.46	4.50	4.84
1949	4.45	4.01	5.42	4.29	4.72
1951	6.30	4.53	6.12	4.86	5.07
1952	6.23	5.07	6.95	5.45	5.21
1953	6.15	5.13	6.85	5.66	5.27
1954	5.57	5.56	7.32	6.21	5.28
Women born in Europe, etc.					
1938/40	1.85	1.73	1.57	1.72	2.10
1944/45	3.06	2.32	2.00	2.04	2.04
1949	3.21	2.11	2.06	2.10	2.22
1951	3.17	2.33	2.57	2.51	2.25
1952	3.04	2.38	2.58	2.55	2.28
1953	2.87	2.39	2.70	2.61	2.31
1954	2.63	2.48	2.81	2.74	2.39

Table 3 Correlations, Regressions and Standard Errors of Estimate of Various Measures with Total Maternal Fertility - F/F_1. (Israel, total population and origin groups 1938-1954; 38 observations, F/F_1 Mean 3.140, Standard Deviation 1.030).

	Mean	Regression Equation	Standard Error of Estimate	Correlation
S_f	3.488	$.6213\ S_f + .9733$	.2247	.9759
S_a	3.153	$.8511\ S_a + .4571$	.2537	.9692
$\mu_B(30\text{-}34)$*	3.499	$.8985\ \mu + .1862$	.2644	.9655

(* only 32 observations, F/F_1 mean 3.330, F/F_1 St. Dev. 1.015)

Table 4 Correlations, Regressions and Standard Errors of Estimate of Various Measures with Total Fertility - F. (Israel, total population and origin groups 1938-1954; 38 observations, F-mean 3.595, Standard Deviation 1.096).

	Mean	Regression Equation	Standard Error of Estimate	Correlation
F/F_1	3.140	$.9390F/F_1 + .5468$	.5148	.8828
S_f	3.488	$.5053\ S_f + 1.4843$	.4912	.8940
S_a	3.153	$.8301\ S_a + .9786$	.5025	.8837
$\mu_B(30\text{-}34)$*	3.499	$.8044\ \mu + .9970$	.5825	.8333

(* only 32 observations, F-mean 3.811, Standard Deviation 1.054)

NEGATIVE SKEWNESS AND ITS SIGNIFICANCE IN RELATION TO DISTRIBUTIONS OF PERFORMANCE RATINGS OF CIVIL SERVICE EMPLOYEES

By: James P. George, University of Tennessee

In the social and behavioral sciences it has been said that frequency curves skewed to the left or in a negative direction are uncommon, and that data which are characteristically so skewed are so rare, so unusual, or of such infrequent occurrence as to be practically nonexistent. It, furthermore, appears to be the opinion of some authorities that frequency distributions of performance ratings are apt to be either moderately skewed to the right or nearly symmetrical in conformation.

This paper reports the findings resulting from analyses of more than 250 frequency distributions of performance ratings of Civil Service employees, representing the Departments of Navy, Agriculture, and the Veterans Administration. These analyses were carried to the point of determining the criteria of curve type. Of the total number of distributions, coming within the scope of this study, only 12 were skewed in a positive direction. In other words, fewer than five distributions out of a possible 100 tended to be skewed to the right. From the standpoint of size it is important to observe that one hundred distributions of ratings were representative or 500 or more individuals. Over-all distributions in certain instances (See Table 1) were sufficiently large to be regarded as parent populations or universes in and of themselves.

Table 1. Negative Skewness and Degree of Significance as Measured by Normal Deviate of T Test, Performance Ratings of Civil Service Employees, Departments of Navy and Agriculture, and Selected Services, Data for Years Ending March 31, 1942 and 1943.

Governmental agency and service	Measures descriptive of certain aspects of a frequency distribution				
	Total frequency or N	Criterion of curve type	Alpha 3	Standard error of Alpha 3	Normal deviate or T value
Navy Department:					
Mar. 31, 1942	17078	I	-0.1045	0.0187	5.58
Mar. 31, 1943	15178	I	-0.9463	0.0199	47.60
Agriculture, Departmental					
Service: Mar. 31, 1942	12832	I	-0.6540	0.0216	30.25
Mar. 31. 1943	10870	I	-0.3764	0.0235	16.02
Field Service:					
CAF Mar. 31, 1943	15771	I	-0.6162	0.0196	31.44
P Mar. 31, 1943	10900	I	-0.7125	0.0235	30.37

For example, performance ratings covering 17078 employees were reported by the Navy Department as of March 31, 1942. The Clerical, Administrative and Fiscal personnel, comprising the Field Service of the Agriculture Department, numbered 15771 as of the same date. On March 31, 1942 CAF personnel, constituting the Departmental Service of the Department of Agriculture, numbersome 12832 individuals. The Navy Department reported performance ratings on 15178 Civil Service employees for the year, ending March 31, 1943. The Departmental Serivce of the Department of Agriculture reported rating on employees to the number of 10870 on March 31, 1943. Comparable numbers for the Professional personnel in the Field Service of the Agriculture Department amounted to 10900 as of the same period.

Of those 12 distributions of Navy Department personnel, which were skewed to the right or in a positive direction only four such distributions numbered in excess of 500 individuals. (See Table 2 which follows:

Organizational Unit	Total frequency or N	Curve type	Alpha 3	Standard error	Normal deviate
New York Navy Yard Regular employees, 1944	5319	I	0.2601	0.0336	7.74
Bureau of Supplies and Accounts, 1942	2640	I	0.4156	0.0477	8.72
EXOS & OS, 1942	2283	IV	0.0982	0.0513	1.92
Marine Corps	1059	I	0.2235	0.0753	4.71

In the above table those normal deviate of T values marked with an asterisk indicate that the distributions with which they are associated are highly significantly skewed. In other words, three out of four distributions were characteristically skewed to the right or in a positive direction. Performance ratings of EXOS & OS combined gave rise to a frequency curve, conforming to the Karl Pearson Type IV criterion. Skewness in this instance was not so great but what it could be accounted for on the basis of chance. (Consult Table 2.)

Alpha 3 is a coefficient of relative skewness and possesses the desirable quality of being independent of the unit of measurement. Inasmuch as it is not restricted in range, it is a very sensitive measure. For the purpose of this study Alpha 3 has been used.

A deviation from any statistical measure in terms of its standard error is called a normal deviate. The normal deviate is otherwise designated by the letter T. In calculating the T value the deviations are always considered to be positive. For any symmetrical distribution the value of a coefficient of skewness is always zero. To measure the departure from absolute symmetry we divide Alpha 3 by its standard error. For example, the value of Alpha 3 descriptive of the distribution of performance ratings for the Navy Department as a whole, for the year ending March 31, 1943, was -0.1045. The standard error of this measure was 0.0187. Diving the former by the latter we obtain a quotient of 5.58.

The value of the normal deviate or T for the .95 probability level is 1.960, and for the .99 probability level, 2.576. Deviations are said to be significant when T is equivalent to 1.960 or more, and highly significant when T is equivalent to 2.576 or more. Since the computed value of T for this particular frequency distribution is 5.58, the departure from absolute symmetry of form is to be interpreted as highly significant. We conclude that the data from which this measure derives are characteristically skewed in a positive direction.

Certainly insofar as the Departments of Navy and Agriculture, and the Veterans Administration, Branch Area 7, are concerned the evidence at hand fails to support the hypothesis with respect to symmetry, or direction and extent of departure from symmetry, as expressed in the introductory paragraph of this paper. Negative skewness very definitely gives evidence of being a dominant characteristic of such data as does conformity with respect to shape to the Karl Pearson I or III criteria.

The research of this writer leads him to the conclusion that the distributions of raw scores tend to conform to the same pattern as described above. Only when the examination upon which the raw scores are based has been constructed by a person trained in the technicalities of psychometrics have we a right to expect even approximate symmetry of form of these scores. It is likewise true that having passed through a so-called "normalizing" or adjustment process at the hands of a person so trained the distribution of the resulting scores may be roughly summetrical. It is important to remember in this connection, however, that first line supervisors in government agencies and members of the teaching profession have neither the time, the technical "know how," nor the computing machines for such refined analyses and adjustment.

As submitted to the writer for analytical treatment the performance ratings of Civil Service employees of the Navy Department for the year ending March 31, 1943 were coded and in the form of frequency distributions, concerning of nine classes. This statement is equally true of the Department of Agriculture for the same time period. In other words, in both instances the original ratings were in numerical form. However in the case of the Navy Department, Departmental as well as Field Services, for the year ending March 31, 1942, and in the case of the Veterans Administration, Branch Area 7, the original measures of performance were reported as adjective ratings. These data were coded by the writer in order to make statistical manipulation possible. Intervals along the base line were equally or uniformly spaced. Through the use of logarithmic transformations no doubt these Pearson Type I or Type III frequency curves would have assumed in approximation of normality. In this event the intervals along the base line of course would no longer have been uniformly spaced.

AUTHOR INDEX