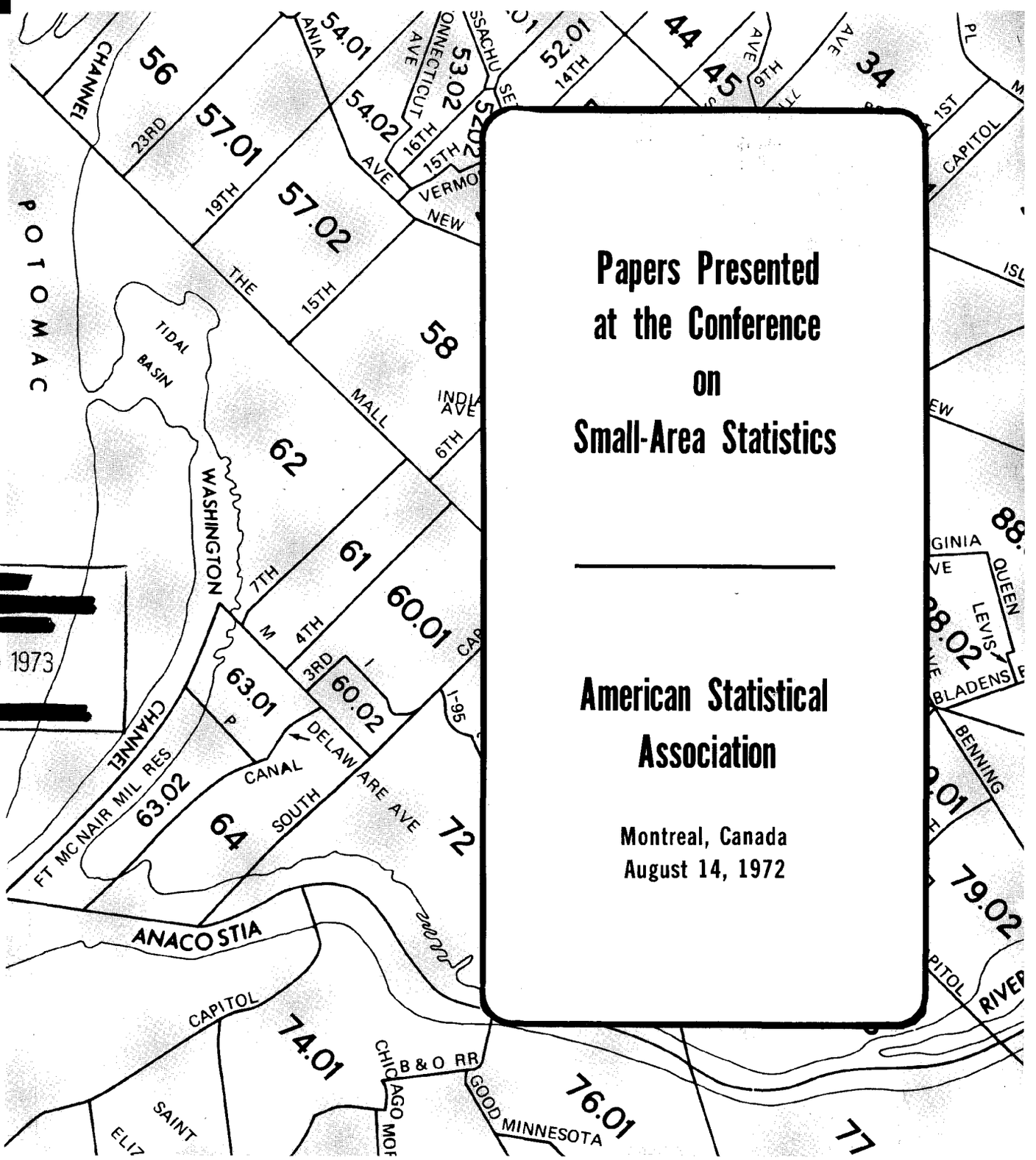
Census Tract Papers Series GE-40, No. 9

A UNITED STATES DEPARTMENT OF COMMERCE PUBLICATION

Social Indicators for Small Areas



Papers Presented at the Conference on Small-Area Statistics

American Statistical Association

Montreal, Canada
August 14, 1972

Census Tract Papers Series GE-40, No. 9

Social Indicators for Small Areas

Papers Presented at the Conference on Small-Area Statistics American Statistical Association

Montreal, Canada August 14, 1972

Issued May 1973

U.S. DEPARTMENT OF COMMERCE **Frederick B. Dent,** Secretary

Social and Economic Statistics Administration
Edward D. Failor, Administrator

BUREAU OF THE CENSUS, Vincent P. Barabba, Acting Director

BUREAU OF THE CENSUS

Vincent P. Barabba, Acting Director
Robert L. Hagan, Deputy Director
Paul R. Squires, Associate Director

GEOGRAPHY DIVISION

Morton A. Meyer, Chief

Library of Congress Card No. A68-7769

SUGGESTED CITATION

U.S. Bureau of the Census, *Census Tract Papers, Series GE-40, No. 9, Social Indicators for Small Areas,* presented at the Conference on Small-Area Statistics, American Statistical Association, Montreal, Canada, August 14, 1972. U.S. Government Printing Office, Washington, D.C., 1973.

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402
Price: $1.00, domestic postpaid; 75 cents, GPO Bookstore. Stock Number 0324-00193

PREFACE

The five papers published in this report were presented at the Small-Area Statistics Conference at Montreal, Canada on August 14, 1972 during two sessions of the annual meeting of the American Statistical Association (ASA). This conference was sponsored by the ASA Committee on Small-Area Statistics and the Social Statistics Section (ASA) with the program planned by Albert Mindlin, Chairman of the Committee.

All papers deal with Social Indicators for Small Areas, a subject that has received increasing attention in recent years from statisticians and social scientists with the greatly expanded availability of small-area social statistics. Albert Mindlin chaired both morning and afternoon sessions and has provided an introductory statement.

This report was organized and prepared under the direction of Robert C. Klove, Geographic Research Advisor, Office of the Associate Director for Statistical Standards and Methodology, Bureau of the Census.

CONTENTS

INTRODUCTION

Albert Mindlin, Chairman
ASA Committee on Small-Area Statistics

There is a long history of research efforts to develop typologies of cities and neighborhoods. Most of this history has been academic, although it has interfaced with market research, and to a much lesser extent public administration. More recently, and independently of this history, there has developed a large body of literature on "social indicators." Interest in social indicators has been stimulated, far more than the earlier history of urban typologies, by problems of effective public administration of vast new social programs. It has been to a considerable extent a product of the urban revolution and various social issues, such as poverty and civil rights. These phenomena greatly accelerated in the sixties and are still in full motion. The expression "social indicators" is clearly an echo of "economic indicators." Economic indicators developed over many years as input to an economic accounts system which underpins much of the control and guidance of our national economy. It was hoped during the sixties that a social accounts system could be developed with analogous purpose, and with various "social indicators" as input. Such a system has not emerged, nor does one receiving general agreement appear likely in the foreseeable future. However, the movement to construct "social indicators" has developed a life of its own, and has absorbed much of the earlier research on urban typologies.

The earlier literature on urban typologies tended to be dominated by sociologists; and more recent, literature on social indicators has been dominated by political scientists, sociologists and economists. There has been relatively little input from professional statisticians or of a practical, concrete character. However, such input is beginning to appear. The purpose of the 1972 ASA Conference on Small-Area Statistics was to bring together in a single extended presentation some of the best known statistical work now going in the field of social indicators for small areas. All of the work described in these proceedings include both theory and application down to concrete product.

The morning session was given over to two papers that describe work based on sophisticated statistical methodology. Bixhorn and Mindlin, in **Composite Social Indicators for Small Areas—Methodology and Results in Washington, D.C.,** describe the effort to develop a practical on-going social indicator system for the District of Columbia. Using census tracts as a geographic unit, the paper focuses on the creation and use of composite indicators, i.e. index numbers based on several individual statistical series. The paper describes the line of reasoning by which individual series are brought together to define a composite indicator. It then describes and illustrates the method of principal components by which an index number is calculated from these separate series. It next describes and illustrates a method of cluster analysis which classifies census tracts on the basis of their index numbers. Finally it examines these clusters and illustrates uses in practical public administration.

Korper, Deshaies, Schuerman, and Crellin, in **Composite Social Indicators for Small Areas—Census Use Study—Recent Developments in Methodology and Uses,** provide an extended overview of the methodological work that the Census Bureau has been carrying on in this field for several years. The Bureau is drawing on a broad range of statistical tools, and the paper touches, at least briefly, on most of them. It may be noted that in the evolution of Census Bureau research it has moved from analysis of single series indicators to construction of composite indicators, and from ranking and related methods of classifying areas to cluster analysis. The Census Bureau paper describes its methodological efforts to construct a valid trend model, a difficult problem.

The afternoon session presented three papers. Weiner, in **Application, Limitation, and Methodology of the Scientific Urban Matrix (SUM),** describes a rather large scale effort in Los Angeles to create and utilize what he calls a Scientific Urban Matrix. He uses a variety of technics to create composite indicators—factor analysis, subjective grouping, cluster analysis. A unique feature of his methodology is to divide the city into "excellent," "good," "endangered," and "blighted" areas according to attitudinal perceptions on a variety of characteristics obtained from a sample of residents. He obtains a score for a composite indicator in a neighborhood by a modified summed ranks procedure, and then regresses these scores on the attitudinal perception rating of the area. His purpose is to rate the various categories of community conditions on a scale of citizen perception of the quality of life in the neighborhood. The improvement of quality of life would focus on those categories that most change the citizens' perceptions. The use of citizen perception as a standard rather than "objective" criteria defined by professionals is a fascinating approach that merits further discussion.

Garn and Flax, in **Indicators and Statistics: Issues in the Generation and the Use of Indicators,** present a summarized description of the work of the Urban Institute. They review briefly some of the definitional and conceptual issues that have emerged in the literature, and then proceed to a short description of each of several reports prepared by the Urban Institute. The Institute's work for the most part avoids construction of composite indicators. It selects several functional areas of social concern. In each area it selects a single measure. For each measure it attempts comparisons between metropolitan areas, between city and suburbs, and over time in central city and suburbs. The Institute has made inter-city comparisons in greater detail in selected functional areas, in particular racial comparisons, and comparisons for a variety of statistical series in education. In the latter case an effort has been made to distinguish between social condition variables, educational

system input variables, and educational condition output variables, and to correlate these, with some interesting results. Having clarified much of its methodology, the Institute is now directing attention to how to make data series immediately useful to program administration.

The Goldsmith and Unger paper, **Social Area Analysis: Procedures and Illustrative Applications Based Upon the Mental Health Demographic Profile System** is essentially taxonomic rather than statistical, and in this sense is a continuation of a substantial body of social indicator literature. The authors examine 1970 census data, create a set of social dimensions, and then select census characteristics which they believe measure each dimension. For each statistical series they establish a set of intervals. They then classify census tracts into these intervals. Thus a census tract is classified on each of a large number of census variables. Since the variables have been grouped into sets that define a social dimension, each census tract can then be described in a profile sense along these social dimensions. The authors do not attempt to create composite indicators. Their taxonomy has been computer programmed and in this form is available to others who might wish to re-create it with their own census data.

Thus, the Montreal Conference covered an array of statistical efforts to bring the social indicator movement to small areas, and also to transform it from theory to practical application. The efforts presented ranged from sociological taxonomy, essentially non-statistical, through increasing degrees of statistical methodology to highly sophisticated technics—but all oriented directly to practical application by program administrators.

COMPOSITE SOCIAL INDICATORS FOR SMALL AREAS—METHODOLOGY AND RESULTS IN WASHINGTON, D.C.

Herbert Bixhorn
and
Albert Mindlin

Government of the District of Columbia

Introduction

A large body of literature has developed on "social indicators." Interest in the subject is to a considerable extent a product of the urban revolution and various social issues, such as poverty and civil rights. These phenomena greatly accelerated in the sixties and are still in full motion. The expression "social indicators" is clearly an echo of "economic indicators." Economic indicators developed over many years as input to an economic accounts system which underpins much of the control and guidance of our national economy. It was hoped during the sixties that a social accounts system could be developed with analogous purpose, and with various social indicators as input. Such a system has not emerged nor does one receiving general agreement appear likely in the foreseeable future. However, the movement to construct social indicators has developed a life of its own.

Several features of the present literature are noteworthy. First, it is dominated by political scientists, sociologists, and economists. There has been very little input from professional statisticians. Second, much of it is of a theoretical, conceptual, or subjective taxonomic character with little direct applicability by public administrators. Third, much of it is national or at least macro-oriented. Relatively little is local or micro-oriented, i.e., dealing with cities or neighborhoods. The purpose of this paper is to fill some gaps in all three of these features. It presents the work being done by mathematical and other professional statisticians in the Government of the District of Columbia, oriented toward the city and its neighborhoods, and intended for direct policy and program application by city administrators. So far the work is very much in the "pilot" stage, but enough has been done to warrant public and professional exposure. The context in which we are working imposes an outlook and constraints. We are not seeking to break new theoretical ground, either conceptually or mathematically, nor are we seeking to develop data that do not currently exist on a neighborhood level. Using known mathematical and statistical methodologies and existing data, we are seeking to develop a social indicator model and system whose results are meaningful and usable for planning and management by local public agencies.

We define a social indicator as follows: *it is a quantitative measure of the quality of community life.* Community life of course exists in many dimensions or functional areas—health, welfare, housing, etc. Consequently there must be more than one social indicator. Furthermore, rarely can a single quantitative variable constitute an adequate measure of the quality of life in any functional area. Almost always there are several quantitative variables measuring various aspects of that area. Indeed there are dozens of statistical series measuring aspects of community life. One can be inundated with statistics. One of the objectives of a social indicator system, as we perceive it, is to illuminate the underlying reality by identifying the separate quantitative variables that measure aspects of a functional area, and then to bring these together rationally into a single composite indicator. Thus the thrust of our work is to develop meaningful and useful composite indicators.

General Outline of a Local Social Indicator System

The system has several objectives. In each dimension or functional area, it seeks to present a comprehensible profile of that dimension over neighborhoods of the city; to determine the trend of the dimension over time; to illuminate geographic areas of needed governmental attention; to determine for a given neighborhood what social conditions seem more in need of attention than others relative to those conditions in other neighborhoods; and if possible to help in evaluating the impact of governmental programs.

In order to accomplish these objectives, the following steps are currently followed:

1. Define specific social dimensions or functional areas of life in broad social fields such as health, welfare, housing, public safety, etc.

2. Determine the available statistical series that constitute quantitative measures of various aspects of the quality of life in each dimension.

3. Construct a methodology to bring these series together into a composite indicator for each dimension.

4. Delineate geo-areas of the city that are relatively homogeneous with respect to the composite indicator.

5. Conduct various analyses of the indicator, comparing it with other indicators in the same neighborhood, comparing neighborhoods, disaggregating it to evaluate its components, seeking insights heretofore unnoticed, and attempting to draw inferences of policy, programmatic, managerial, and budgetary significance.

6. Develop a presentation package that will most effectively convey the subject to agency administrators.

At the present time we have fairly fully developed points 1-4, but have taken only a few steps in points 5 and 6. In this paper we will present the general methodology first, and then illustrate it with actual data from the District of Columbia.

Steps 1 and 2

Defining the Dimensions and Determining the Available Statistical Series

These steps are inextricably intertwined. There are numerous classification schemes of dimensions of social life. Ideally a taxonomy should be constructed or justified by its relevance to sociological theory, or public administration, or some other intellectual framework quite independent of the availability of data. This has been attempted; the social indicator literature is filled with taxonomies. Regrettably the absence of available data makes almost all of them useless, regardless of other virtues or defects they may have. It simply is not practical to delineate dimensions of social life in order to construct social indicators without having available data play an integral part in the delineation; and indeed, every local social indicator effort that has sought practical development that the authors are aware of (such as efforts of the Census Bureau, Urban Institute, and others) have either taken the body of available data as their point of departure, or have determined data and taxonomy in unison.

There are objective mathematical methods for separating a large number of statistical variables into groups of highly correlated variables. The best known of these is factor analysis. We rejected a mathematical approach at this stage for the following reasons: first, factor analysis often produces factors, or variable groupings, either that cannot be understood by administrators ("Just what does this factor measure, as against that factor? What does it mean?") or cannot be used by administrators ("Well, your mathematics may tell you that variables a, b, c, d, and e 'hang together,' but that doesn't help me. There is no government program that deals with a and b, and c is not under my control; d and e are the only ones I can deal with"). Second, we want our system to be used. The best way to assure this is to build into it from the beginning the participation of planning and operating agency staff.

We are proceeding as follows: starting with a very general, more or less arbitrary and fairly obvious classification (health, welfare, public safety, education, etc.) we conduct a series of meetings with operating officials in each of these fields. It is these officials who hammer out the dimensions of their field and the statistical series that go into each dimension to create a composite indicator. Their deliberations take place under the painful constraint of available data, and with the participation of the statisticians, data processors and other hangers-on. The comments and questions raised are such as:

1. "Is this classification meaningful in terms of administration?"
2. "that's not a measurable variable"
3. "no information exists on that"
4. "yes, we have records on that variable, but they aren't mechanized. Do you have the resources to go through 50,000 records by hand every year?"
5. "that's only available from the census, so you can't make a time series"
6. "that's only available for the whole city, not by census tract, so you can't get a geo-distribution unless you add a coding operation"
7. "the quality of that file is so poor, you could prove black is white from it"
8. "that variable requires a cross-tabulation of two files, and they're incompatible"
9. "this is a community condition variable while that is a program workload variable. It may not be meaningful to put them together in the same composite indicator," etc.

These meetings not only engage planning and operating officials in active planning of the system, they also are of considerable educational value to the officials. The officials sometimes lead themselves to recognition of the importance to their operations of statistical series generated elsewhere, and thus themselves create indicators that cross administrative lines. We have no *a priori* hang-up on taxonomy, but the officials themselves create a practical one. Also they learn some consequences of past decisions to mechanize or not to mechanize specific data items, and of the consequences of using geo-codes which are compatible with other files (census tract, city block, etc.); or of unique geo-codes unrelatable to anything else (a significant feedback of the sessions we have held has been an increased appreciation of an integrated management information system), and gain insight into some of the types, uses, and possibilities of statistical analysis.

Step 3

A Methodology to Construct a Composite Indicator

We now have a defined dimension and several available statistical series to measure it. It could be argued that we have an operational definition—it is the series we use that quantitatively defines the dimension. In D.C. the census tracts are our geo-units. We have a value for each of several variables on each census tract for approximately the same time period. The variables are converted to rates, so that for each variable taken separately, the values are comparable from one census tract to another. At this point we cop-out of a theoretically important issue. Should weights be assigned to the variables, reflecting differential importance in the composite? Differential importance is a matter of social values and public policy. It is not a matter to be decided by statisticians. We laid this matter before the program administrators, with a result surprising in the event, not so surprising although in hindsight. They refused to pass judgment on the relative social importance of the variables. For example, one functional area hammered out by health officials was "health conditions surrounding birth." The variables they settled on were birth weight under 5½ lbs., infant mortality, no or inadequate prenatal care, out of wedlock, and age of mother under 20 years. The administrators refused to rank these in any way in social or health importance. So the variables entered the mathematical maelstrom weightless, all treated equally. However, the mathematics we use creates automatic weights. These weights reflect

the degree of correlation between the variables. A variable which correlates poorly with the others plays a lesser role in the composite, and may even be rejected. This point will be elaborated later.

How to bring the variables together into a single composite? Perhaps the simplest way is to rank each variable over the census tracts and then add the ranks. This "summed ranks" method has decided advantages, such as simplicity of calculation and easy comprehension by non-technicians. But it has a monumental flaw: it throws away too much important information. For a given variable its unique geo-distribution, where it piles up, where it stretches out, is lost; the ability to perceive natural groupings of tracts is lost; poorly correlated variables are not detected and can produce anomalies. Some of these disadvantages can be mitigated by modifying the method, for example converting the values to standard scores within each variable and then summing the standard scores. But modification still leaves too much to be desired. We rejected this method and all modifications of it that we tried

It is obvious to think of straightforward multiple regression. This, however, cannot be applied because it requires specification of a dependent variable. There is no dependent variable in this work. However, multiple regression has many desirable properties, and it leads us to the method that we currently employ to create an index number representing a composite indicator—the method of principal components. The easiest way to explain this method nontechnically is to describe it graphically in 2-dimensional space.

Consider figure 1. There are two variables, x_1, x_2. The points are census tracts. In ordinary regression x_2 is denoted the dependent variable, and the regression line is so chosen as to minimize the sum of squared deviations from the points to the line along the x_2 axis. In the method of principal components the line is so chosen that it minimizes the sum of squared deviations along the perpendicular from the points to the line. Thus variables x_1, x_2 are treated equally. There is no dependent variable.

Figure 1

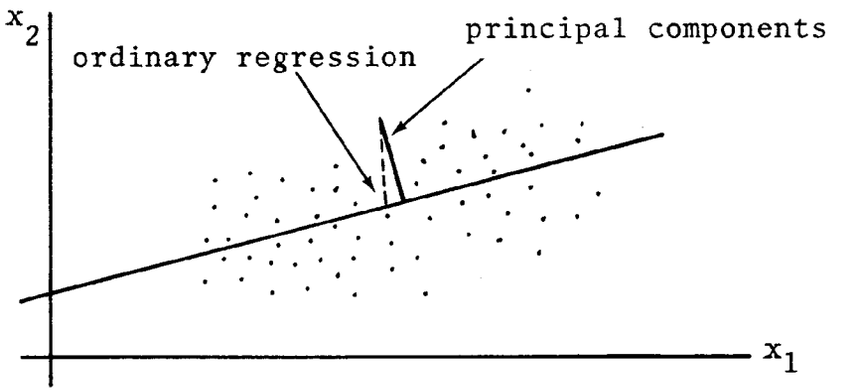
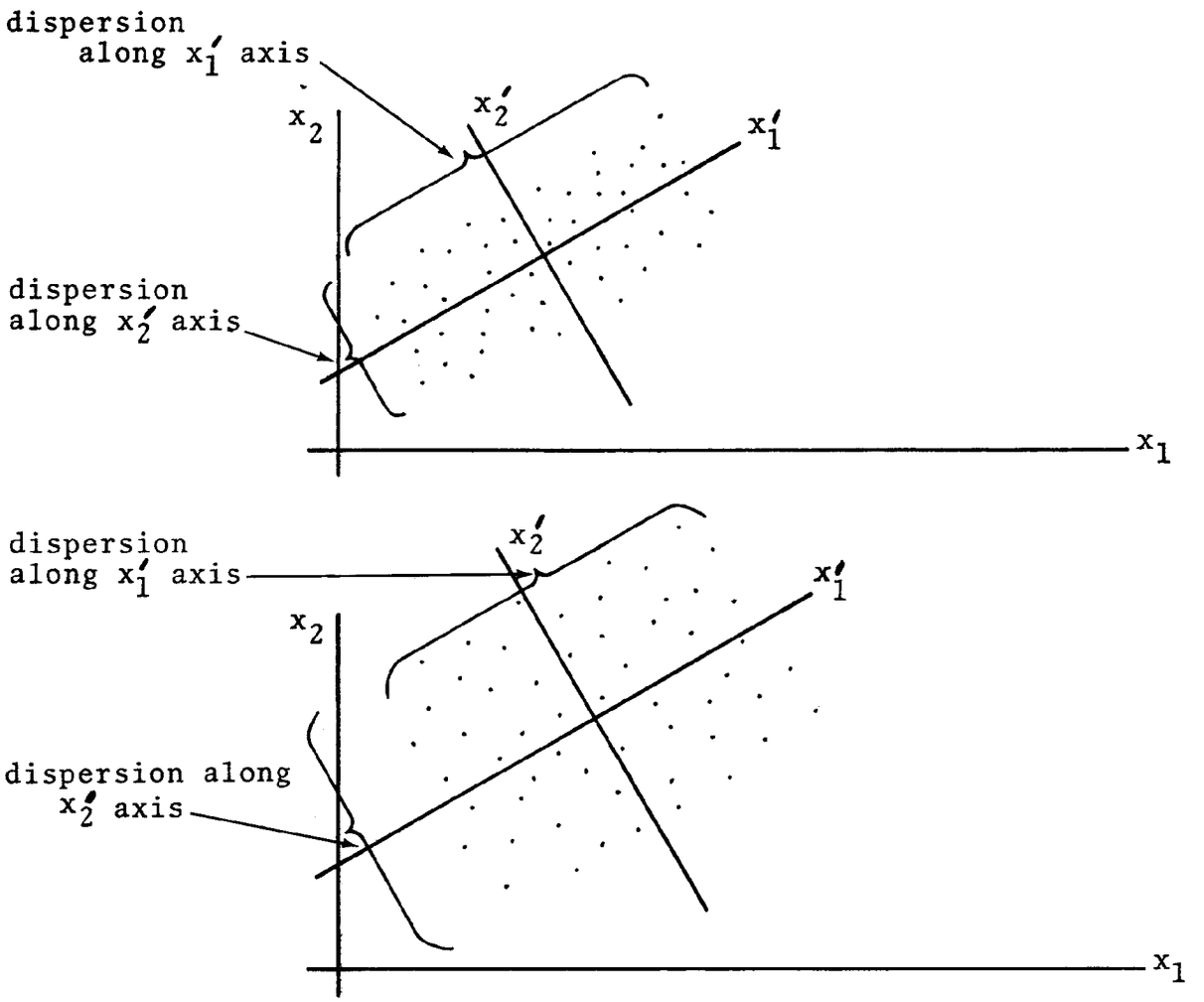


Figure 2



Consider figure 2. The method first standardizes the variable $\frac{x - \bar{x}}{\sigma_x}$. This makes all variables directly comparable with each other. It also places the origin at $\bar{x}_1, \bar{x}_2$. The method then rotates the original x_1, x_2 axes into other axes denoted by x'_1, x'_2. The formula for each of these new axes is a linear combination of the original variables; that is, $x'_1 = ax_1 + bx_2, x'_2 = cx_1 + dx_2$. Each axis in this new coordinate system is called a "principal component." The regression line which minimizes the perpendicular distance from points to line is the x'_1 axis. It is called the "first principal component." The x'_2 axis is called the "second principal component." There is a dispersion of the points along each axis, measured by the variance along that axis. Thus associated with the x'_1 axis is a "first principal component variance"; associated with the x'_2 axis is a "second principal component variance." The sum of the variances is called the "total variation of the points." Compare the two graphs in figure 2. It is clear that in the top graph the first principal component variance constitutes by far the largest proportion of the total variation of the points; while in the bottom graph the two variances contribute about equal proportions to the total variation of the points. In the top graph the first principal component "explains" most of the total variation in the points. In the bottom graph the first principal component "explains" much less of the total variation. All of this can be generalized to n dimensions, with n principal components. (See [1] and [2].)

We now apply the method of principal components to the construction of a composite social indicator. The formula for the combination of the separate variables into an index number is the equation of the first principal component. If there were only two separate statistical series in the composite, the indicator could be shown graphically as in figure 2. For each point, i.e., each census tract, the index number would be the value of x'_1 : $x'_1 = ax_1 + bx_2$. It should be noted that this definition of the composite indicator, i.e., this index number, is more meaningful if the first principal component accounts for a high proportion of the total variation of the points (i.e., the points are strung out along the first principal component axis much more than along any other principal component axis), and less meaningful if the first principal component does not account for a high proportion of the total variation. Consider again figure 2. Clearly the desired condition obtains when the variables are highly correlated, and drains away as the correlation between the variables decreases. This conforms to intuition. If several variables are highly correlated, it is reasonable to infer that their combination is measuring some meaningful underlying reality, each variable making a contribution to the index number. But if some variable does not correlate highly with the others then it is reflecting some other underlying functional area. Its presence in this composite distorts the index number and makes it hard to evaluate and less useful in analysis. The actual mathematical expressions of the coefficients, or weights, attached to each variable in the first principal component formula reflect, in an indirect way, the correlation of that variable with the other variables. In practice we adopt the following rule of thumb: we drop a variable from the composite if (a) its weight (i.e., its coefficient) is less than one-half the maximum weight; or (b) the first principal component takes up less than 65 percent of the total variation of the points. Dropping a variable under this rule results in a reduced composite that usually is mathematically satisfactory. This has considerable significance in analysis. Since the subject experts placed that variable with the others, dropping it is equivalent to pointing out something special about it of programmatic significance. This feature is one of the advantages of the principal component method. It will be illustrated later with D.C. data.

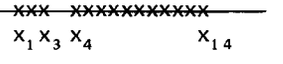
Step 4

Delineating Geo-areas of the City that are Relatively Homogeneous with Respect to the Composite Indicator

We now have a composite indicator with an index number for every census tract. We wish to see how this index number classifies the tracts over the city—i.e., we wish to cluster the tracts so that those that are relatively homogeneous fall into the same cluster. In the first instance we could merely order the tracts by the index number and arbitrarily cut them into, say quintiles or deciles. This has the advantage of simplicity, but it has the disadvantage of utter arbitrariness. What is magical about five groups, or ten—or four or seven or any other number? What is magical about having the same number of tracts in each cluster? Furthermore the range of values in each group may be very dissimilar. For example, in the top group the lowest index number may be only 5 percent less than the highest index number, while in the next group the difference may be, say 30 percent. The means of the two groups might actually be very close together, and there may actually be more within-variability than between-variability. In our judgment ordered cuts into an arbitrary number of groups with an equal number of tracts in each has little intrinsic merit.

Alternately, we could plot the index number on a scale and look for "natural gaps." There are two problems here. First, there may not be any obvious "gaps." The plot may look like this:

Second, "natural gaps" can be deceptive. For example, suppose the plot looks like this:



Visually, there would appear to be two "natural" groups—the left three, and the remaining eleven. But in fact tract 4 is closer to tract 1 than it is to tract 14, and is closer to the mean of the left group than it is to the mean of the right group. So "natural gaps" can produce anomalies. We have searched for a more objective way to delineate the city by means of the index number. We believe we have found it in the mathematics of cluster analysis. There are many types of cluster analysis. We employ a technique proposed by Rubin and Friedman [3], primarily because it has a simple and precise criterion for clustering the tracts which in addition has several desirable properties. The technique has been reduced to a canned computer routine which prints out on the side a number of secondary results, such as all of the two-way correlations between the variables, and symbolic plots of the tracts along a line according to their index numbers [4]. The technique may be simply explained as follows: suppose there are six census tracts altogether, each having a composite indicator index number. We wish to group the tracts into two clusters such that the tracts in each cluster are relatively homogeneous with respect to their index numbers—that is, they are more like each other than they are like the tracts in the other cluster. N tracts identified by an index number and ordered from low number to high number can be separated into two ordered groups in N-1 ways. For six tracts, let x_1 be the lowest index

number and x_6 the highest. Then the tracts can be grouped into $(x_1, x_2x_3x_4x_5x_6)$, $(x_1x_2, x_3x_4x_5x_6)$, $(x_1x_2x_3, x_4x_5x_6)$, $(x_1x_2x_3x_4, x_5x_6)$, $(x_1x_2x_3x_4x_5, x_6)$. Which is the best clustering? The Rubin-Friedman paper first standardizes the variables, as does principal components. It then sets up the following criterion for selecting the "best" clustering. Denote the total variance of the standardized index numbers over all six tracts by T. For a single clustering, e.g., $(x_1x_2, x_3x_4x_5x_6)$, there are two "within" variances, one for each cluster. Denote the sum of these two "within" variances by W. Form the ratio $\frac{T}{W}$. This ratio is called the criterion measure. If six tracts are to be divided into two clusters, there are five such ratios. If heterogeneous tracts are placed in the same cluster, W will be large and, since T is a constant, $\frac{T}{W}$ will be small. The technique examines all five ratios and selects the clustering with the largest ratio, i.e., it finds the grouping that maximizes the criterion measure. The best clustering may be e.g., $(x_1x_2, x_3x_4x_5x_6)$ or $(x_1x_2x_3, x_4x_5x_6)$. The number of tracts in each cluster under the best clustering may vary. This is one advantage of the technique—it does not force an equal number of tracts in each cluster.

The technique is not restricted to a single dimension, i.e., a single value for each tract, although we currently employ it this way. It is a general technique. It could be employed even if we had not first used the method of principal components to reduce all of the variables to a single index number. If each tract is represented by several values, then the criterion measure becomes the ratio of two covariance determinants $\frac{T}{W}$. Thus, this method could be used independently to cluster areas of the city with respect to several variables considered simultaneously. Indeed probably that is its more common use.

The user must specify in advance the number of clusters to be created. If he wishes two, the technique finds the "best" two as defined by the criterion measure. It also yields the value of the criterion measure for its "best" selection. If the user wishes three clusters, the technique finds the "best" three and yields the value of the criterion measure. As the number of clusters goes up, the value of the criterion measure also goes up. Hence we use the technique iteratively, calculating each time the difference between successive criterion measure values. We stop the iteration when the difference markedly decreases. To cluster beyond that does not seem justified, for further clustering does not appear to increase homogeneity substantially. This ability to indicate an optimum number of clusters is another advantage of the technique.

As stated, this technique can be employed to cluster the tracts when each tract has several values, such as the value of the separate variables making up the composite indicator, and there are distinct advantages to doing this. There was a definite line of reasoning that led us to employ principal components first, and then to employ cluster analysis on the single index number so obtained, which is a simplified use of the technique. The cluster analysis technique by itself clusters the tracts but does not yield an actual numerical value of the composite indicator, i.e., an index number. Having an index number to represent the composite indicator is valuable for a variety of purposes, e.g., to see its trend over time, to compare the value for a single tract with the value for the whole city, and other purposes. However, if the cluster analysis is made first and the principal component technique applied independently afterwards, seeming contradictions between the results of the techniques can and do appear. In particular, let us assume that low index numbers are possessed by the "best" tracts. The cluster analysis used independently may place a tract in the "best" cluster, while principal components may give that tract an index number that is higher than the index number of a tract that the cluster analysis places in a "poorer" cluster. This is because the two methodologies are different, and an off-beat collection of values in a tract could produce this anomaly. The anomaly cannot occur if the cluster analysis is applied to the single index number produced by the principal component technique.

Application of the Methodology to Census Tract Data in the District of Columbia

We now show the application of the above methodology to the creation of social indicators in the District of Columbia. We held several meetings with health officials in our Department of Human Resources, during which they hammered out the composite indicators that seemed most useful to them as public health administrators. After much discussion, a consensus was reached that indicators based on age seemed the closest related to the general classification of their programs. Hence they arrived at the following five indicators: conditions surrounding birth, childhood, adolescence, maturity, and old age. For each indicator they arrived at a consensus as to the most useful statistical series among those available—i.e., existing in mechanized form on a census tract level and reasonably recurrent. For birth and old age the series consisted entirely of data developed by health officials. For the other ages the officials recognized the importance of some non-health variables. For pilot project purposes we have concentrated on the indicator, "health conditions surrounding birth." As mentioned earlier, this indicator initially contained five variables:

a. births to mothers under age 20
b. no or inadequate prenatal care
c. birthweight under 5½lbs.
d. births out of wedlock
e. infant mortality

Discussions with other public officials produced several other composite indicators that became part of the pilot project—dependency, housing, and general socio-economic conditions. The initial variables in each are shown in table 1.

Table 1. Composite Indicators

1. **Health Conditions Surrounding Birth**

   a. Births to mothers under age 20
   b. No or inadequate prenatal care
   c. Birthweight under 5½ lbs.
   d. Births out of wedlock
   e. Infant mortality

   Rates: The denominator for all variables is number of live births

2. **Dependency**

   a. Persons under age 18 on Aid to Families with Dependent Children (AFDC)
   b. Persons on Old Age Assistance (OAA)

c. Persons on General Public Assistance (GPA), Aid to the Blind (AB), and Aid to the Permanently and Totally Disabled (APTD)
d. Persons receiving food stamps
e. Persons registered for medicaid, and persons on D.C. medical care

Denominators for:

a. total persons under age 18
b. total persons age 65+
c. total persons age 18-64
d. total population
e. total population

3. **Housing**

a. Overcrowded households
b. Incomplete plumbing
c. Low rent
d. Value of owner-occupied one-unit structures
e. Vacancies

Denominators for:

a. occupied housing units
b. total housing units
c. renter-occupied units
d. owner-occupied one-unit structures
e. all housing units

4. **General Socio-economic Conditions**

a. Overcrowded households
b. Incomplete plumbing
c. Median family income
d. Matriarchy index (female-headed families with children under age 18 ÷ families with children under age 18)

Denominators for:

a. occupied housing units
b. all housing units

Indicators 1 and 2 include only locally produced data which can be updated at least annually. Indicators 3 and 4 contain only 1970 census data.

Each of the indicators was put through steps 3 and 4. We shall illustrate the results of the procedure with the composite indicator "Health Conditions Surrounding Birth: 1969." Principal components produced the following first principal component, i.e., the expression that provided the initial index number:

$$x_1' = .52x_1 + .47x_2 + .38x_3 + .54x_4 + .28x_5$$

x_1 = Births to mothers under age 20

x_2 = No or inadequate prenatal care

x_3 = Birthweight under 5½lbs.

x_4 = Births out of wedlock

x_5 = Infant mortality

The proportion of total variation "explained" by this expression was 59 percent. Applying the rule discussed in step 3 for acceptance or rejection of variables, we rejected "infant mortality." The lack of close relationship between infant mortality and the other variables was confirmed by the matrix of 2-way correlations, shown in table 2.

Table 2. Health Conditions Surrounding Birth: 1969

Pearsonian Correlation Matrix

	x_1	x_2	x_3	x_4	x_5
x_1	1	.67	.40	.87	.27
x_2		1	.39	.68	.21
x_3			1	.48	.34
x_4				1	.32
x_5					1

For definitions of variables see text.

Eliminating "infant mortality" and rerunning the other variables produced the following first principal component expression:

$$x_1' = .54x_1 + .50x_2 + .38x_3 + .56x_4$$

The proportion of total variation explained by this expression was 70 percent. We refer to this as the "reduced composite."

The census tract index numbers produced by the expression above were not used directly. They were first converted into numbers ranging from 1 to 146, the number of noninstitutional census tracts in 1970. Standardizing the range of values this way expedites understanding and analysis by making each index number of each indicator easily relatable to the range of tracts and allowing comparison of index numbers for a single tract. The index numbers of the "reduced composite" were then put through step 4. The criterion measure values for various numbers of clusters were as follows:

	Criterion measure	Difference
2 clusters	1.16	–
3 clusters	2.24	1.08
4 clusters	2.79	.55

The optimum number was taken to be three clusters.

The cluster analysis method does not inherently make a judgment that one cluster of tracts is "better" than another. The mathematics simply seeks to put together tracts that are "like each other" when all of the variables are treated collectively. For the purpose of public policy the tracts are judged "better" or "worse" on each variable, e.g., tracts with a low proportion of "birthweight under 5½lbs." are "better" than tracts with a high proportion. All of the variables are structured to run in the same direction, i.e., low values are "better" than high values. Thus, the clusters are designated "B(best)," "M(medium)," "W(worst)." This does not mean that the value of all variables in all tracts in the "best" cluster are lower than those in the "medium" cluster.

The clustering is based on the index number. It puts together tracts that are most like each other in the composite. The irregularities of the number could easily produce a tract in the "best" cluster with a value on some variable that is worse than some values of that variable in the "medium" cluster. But taking all of the variables in that tract together, it is more like tracts in the "best" cluster than tracts in the "medium" cluster. The mean values of the variables in each cluster in "health conditions surrounding birth: 1969," are shown in table 3.

Table 3. Health Conditions Surrounding Birth: 1969

Mean Values in Each Cluster

Reduced Composite	Best	Medium	Worst
x_1 Births to mothers under age 20	6.4	23.5	38.6
x_2 No or inadequate prenatal care	9.1	24.2	33.3
x_3 Birthweight under 5½lbs.	8.3	12.8	16.6
x_4 Births out of wedlock	10.1	32.5	55.3
Number of tracts	30	66	50

On this basis tracts are labeled B, M, W.

The analysis of rejected variables is important. Hence "infant mortality" was separately put through the same cluster analysis to form its own clusters. The "best" and other clusters here are not of course composed of the identical tracts as the reduced composite.

The total procedure summarized in the paragraphs above was followed for all four of the composite indicators in the pilot project.

Step 5

Using the Indicators

We are at an early stage of community analysis by means of composite social indicators. We have not yet developed an integrated analytic package. Our studies so far have been largely restricted to preliminary results of the cluster analysis. We have not yet explored the uses of the index numbers. Consequently this paper does not include the index numbers. We present here some initial findings and paths we are exploring.

1. **Rejected Variables**

a. **Infant mortality**—This was rejected from the composite health indicator. For many years health administrators, faced with a mountain of health data and a chaotic statistical picture, have accepted infant mortality as a key general measure of community health, using it to judge that one neighborhood is better than another in community health and to point to geo-areas that require special health program emphasis. Results of the composite health indicator challenge this classic role of infant mortality. The analysis shows that infant mortality correlates poorly with other health conditions surrounding birth. Whatever its role historically, in D.C. today it no longer appears to distinguish good from bad neighborhoods with respect to other health conditions surrounding birth.

This can be seen clearly in Appendix, figure 2a. This is a cluster map of the birth health composite indicator, 1969. The colors represent the distribution of best, medium, and worst tracts based on the reduced composite. The patterns represent the distribution for infant mortality. It is clear that infant mortality is distributed somewhat haphazardly with respect to the composite indicator. Thus the analysis suggests that judging the health of a community by infant mortality and using infant mortality as a critical variable in directing birth health programs over the city are not for the most part supportable. A side comment: The central mathematical statisticians discovered the low relevance of infant mortality, but it was the research and statistics group in the Health Services Administration that recognized the significance of this discovery. This was a good example of the assertion that effective use of a social indicator system requires the team participation of statisticians and subject experts. Nonsubject oriented statisticians do not have the expertise to draw from the system all that it can yield. The subject experts must be actively involved.

b. **Incomplete plumbing**—This was rejected from the housing indicator. As with infant mortality, it has a low correlation with other variables in its indicator. However, unlike infant mortality in which high rates are dispersed through the city, high rates of incomplete plumbing are concentrated in a few tracts. This is clearly seen in figure 3a. The vast bulk of census tracts fall into the best cluster on incomplete plumbing. Only three tracts are classified as worst in incomplete plumbing, all in the heart of the inner city.

c. **Matriarchy index**—This was rejected from the general socio-economic indicator. It is similar to infant mortality in being spread through the city more haphazardly than the other variables in its indicator, although less markedly than infant mortality. However, a different phenomenon appears to be at work here—the failure of nonspecialized intuition. The socio-economic indicator is the only one which was put together by the central statisticians alone. It was constructed from selected items from the 1970 census. We assumed, without testing, that the selected items "hang together" geographically as typical city ecology.

We were wrong. The two-way Pearsonian correlation matrix of these items is shown in table 4.

Table 4. Socio-Economic Composite Indicator (1970 Census)

Pearsonian Correlation Matrix

		a	b	c	d
Overcrowded households ...	a	1	.16	-.69	-.11
Incomplete plumbing	b		1	-.31	-.04
Median family income	c			1	-.16
Matriarchy index	d				1

The poor correlation of matriarchy index with the other variables is clear. However, the other variables also correlate poorly with each other. The rejected variable was simply the one whose correlations were poorest. The only reasonably good correlation is between family income and overcrowding. The remaining correlations involve incomplete plumbing. As indicated earlier, incomplete plumbing is a highly polarized variable which apparently correlates poorly with other community variables. Furthermore, overcrowded households as the single best measure of social deterioration may not be as tenable a generalization as many believe.

2. **Some Inferences Suggested by the Cluster Maps**

We now turn to the reduced composite indicators, whose clusters are shown by colors on the maps. We take up the maps in turn, comparing the census tracts. However, the discussion inevitably will involve comparing maps, i.e., comparing different indicators for a given census tract. Since this paper is being written for our own administrators as well as for presentation here, certain details of primarily intramural interest will be pointed out. The large numbers on the maps refer to the nine "service areas" (SA) into which the city has been divided by the District Government. SA 6 is the "model cities" area, the heart of the inner city.

a. **Figure 2a, Health conditions at birth, 1969**

First we note certain "islands." In SA 5, ct 66 is a best tract surrounded by poor and worst tracts. This is the heart of the Capitol Hill renovation effort. It would appear that this effort has had some success. In SA 9, ct 54.2 is downtown. It has a small population and few births. In SA 4, ct's 76.2 and 76.3 are of more programmatic significance. These have always been good residential areas, but the increasing threat to them of deteriorating health-at-birth is clear. Census tract 76.1, which has been a good residential area in the past has deteriorated, and infant mortality, the rejected variable, is already getting a grip on 76.2. If we look at these tracts on figure 5a, dependency, 3a, housing, and 4a, socio-economic, we see that they have relatively few welfare families and other dependency characteristics, and that their general socio-economic condition is still good, but that their housing, while still not bad, is going downhill. These tracts have a good prognosis for improvement because of their seeming basic strengths, provided preventive public action is focused on them. But without public attention they will probably go the way of their neighbors.

Let us pursue further the effort to find possible areas of favorable prognosis if public attention is focused. In SA 5, ct's 79.2 and 68.3 are bad in the health-at-birth indicator. However, the general socio-economic condition (fig. 4a) and housing (fig. 3a) are fair, and they have little dependency (fig. 5a). Why is the health indicator so bad? The other indicators show strength. Special attention seems in order. Another such area is SA 9, ct 54.1. This tract also is bad in health-at-birth. Yet it has good housing (3a), fair socio-economic condition (4a), and relatively little dependency (5a). Certainly on the suggestive evidence of the maps, ct 54.1 has the strength to benefit substantially from concentrated official attention.

The reverse of this type of search can also be seen. Note SA 3, ct's 78.1, 78.4, 78.7, 78.8. In the far northeast section of Washington these are worst in health-at-birth. Prognosis for improvement does not seem good according to the maps, for they are already being overwhelmed in other dimensions, even though their general socio-economic condition is fair. The matriarchy index is very high there (4a); bad housing exists and has already overwhelmed ct 78.3 (3a) and dependency has made inroads (5a). However, this inference is only suggestive and should lead to closer analysis. There is considerable public housing in ct 78.1 and 78.4 which may account for the poor showing of the other indicators in those tracts. If the poor health condition can be identified to public housing occupants to a considerable extent, then ameliorative steps may be more feasible and the prognosis improved.

Certain ct's stand out as "peninsulas" that may be the "cutting edge" of serious deterioration, e.g., SA 1, ct 24. This is worst in health-at-birth, and seems to be the herald of the expanding deterioration behind it. But it has relatively low infant mortality and is in fair shape on the other indicators. Can concerted public attention to this area hold back the advancing deterioration? Another area in a similar position is SA 7, ct 87.

One must beware of possible artifacts. For example, SA 8, ct 4 shows best on the reduced composite (and best on all other indicators) but worst on infant mortality. However, in this tract there were only 24 births throughout 1968 and 1969, and two cases of infant mortality. This is a very high rate of infant mortality on a very small base. It may be statistically unimportant—or it may indicate a serious bad health pocket in an otherwise excellent tract. The same remarks apply to SA 1, ct 26, and SA 9, ct 57.1.

b. **Figure 3a, Housing**

The first thing to bear in mind about this indicator is that the reduced composite has no variable that directly measures the physical condition of housing. All of the variables, including the rejected one, are 1970 census variables. The rejected variable, incomplete plumbing, is the only one dealing directly with physical characteristics. The remaining variables are overcrowding, low rent, vacancy rate, and value of owner-occupied one-unit structures. They are variables usually associated with geographic areas of poor housing, but are not themselves direct measures of that.

This caution is appropriate because the thing that strikes the eye of a person familiar with Washington is the concentration of "worst" housing in SA 4, which contrasts with a considerably better showing of that area on the other indicators. Some of the tracts in SA 4 have a good deal of public housing as table 5 indicates. Almost all of the tracts with substantial public housing are worst on the reduced housing indicator. Does this reflect primarily low rent? Or does it also reflect severe overcrowding? Abnormal vacancy rate? Comparing tracts with and without public housing in the same general areas, table 5 suggests that there is indeed excessive overcrowding and generally above-average vacancy rate in public housing tracts.

The usefulness of comparing indicators is indicated by SA 4, ct 76.2. On the housing map this tract is inconspicuous. But we saw earlier that, compared to the other indicators, housing seems to be going downhill. Public attention might save this area. In contrast, a tract that appears about to be overwhelmed is SA 7, ct 29. Indeed,

the housing map is dismaying when compared to the other maps, for it shows the advance of the demographic aspects of bad housing in almost every direction from the core of the inner city, SA 6.

It may be noted that almost the entire inner city core, SA 6, is in the worst cluster in housing.

c. **Figure 5a, Dependency**

The medium and worst clusters on dependency are clearly not as widespread as on the other indicators. Outside the inner city core, the worst areas tend to be areas with substantial public housing as might be expected, e.g., SA 4, ct's 74.1, 74.4, 97; SA 5, ct's 71, 72; SA 9, ct 60.2. An anomaly seems to be SA 7, ct 36 which is a medium tract almost surrounded by worst tracts. Its social characteristics are generally believed to be not markedly different from its neighbors. This should be investigated. Whatever the explanation, the socially dependent population is clearly moving up this general corridor.

Table 5. Some Aspects of Housing in Selected Census Tracts: 1970

SA	CT	Percent population in public housing	Reduced housing indicator	Percent over-crowded	Vacancy rate percent
4	73.4	0	W	23	8
	74.1	49	W	31	9
	74.2	8	W	13	3
	74.4	82	W	35	6
	74.5	0	M	18	3
	75.1	0	W	22	7
	75.2	39	M	13	3
	97	46	W	37	13
	98	29	W	29	6
3	78.1	39	W	26	5
	78.3	0	W	19	3
	78.4	46	W	20	7
	78.5	0	M	22	3
	78.7	0	M	20	2
	78.8	51	W	27	6
5	70	28	M	11	7
	71	33	W	22	7
	72	88	W	28	12
9	60.2	99	W	54	0

W = worst M = medium

These notes are intended to indicate some of the ways that cluster analysis can be used to alert administrators to what is going on in the city in various functional areas. At the present stage of analysis they are hypotheses suggested by the cluster maps. We don't know if the hypotheses are correct, but they illuminate the paths of deeper analysis and possible administrative action. Further analysis will involve the index numbers, which provide more precise information, as well as disaggregation of the variables to determine whether certain variables are playing heavier roles than others. There are other uses to which the indicators can be put. For example, the index numbers and the clusters provide rational stratification devices for sample surveys, and similarity groupings for experimental–control experimental designs. In the D.C. Government each service area has its own inter-agency committee, which concentrates on governmental services in the service area of its responsibility. The indicators can be used by these committees seeking program guidance in their own service area. We are only on the threshold of a social indicator system for neighborhoods. We expect to develop more indicators, for example, in public safety and education. We expect to develop time trends (there are methodological problems here). We expect to improve the methodology. For example, we believe that there is a conceptual difference between program variables, such as those that make up the dependency indicator, and community condition variables such as those that make up the housing indicator; this should be explored. We hope to develop indicators based entirely on locally produced and recurrent data in order to construct useful time trends. From another perspective, the system is impelling a deeper use of the data by operating staff than has occurred in the past.

A final note of re-emphasis. Such a system cannot be effectively developed and used by the statisticians alone. They don't know enough. It must be embedded in developmental and analytic teamwork of statisticians, subject experts, and data processors; and it must be used in government planning. Its success as an operation of local government will depend ultimately on how useful it proves to be for concrete governmental action.

REFERENCES

[1] Anderson, T. W., **An Introduction to Multivariate Statistical Analysis,** John Wiley and Sons, Inc., 1958

[2] Morrison, D. F., **Multivariate Statistical Methods**, McGraw Hill, Inc., 1967

[3] Rubin, J. and Friedman, H. P., "On Some Invariant Criteria for Grouping Data," **Journal of the American Statistical Association,** December 1967

[4] Rubin, J. and Friedman, H. P., "A Cluster Analysis and Taxonomy System for Grouping and Classifying Data," IBM Program Library, August 1967

APPENDIX FIGURES

2a. Health Conditions at Birth: 1969
3a. Housing: 1970
4a. Socio-economic Conditions: 1970
5a. Dependency: 1971

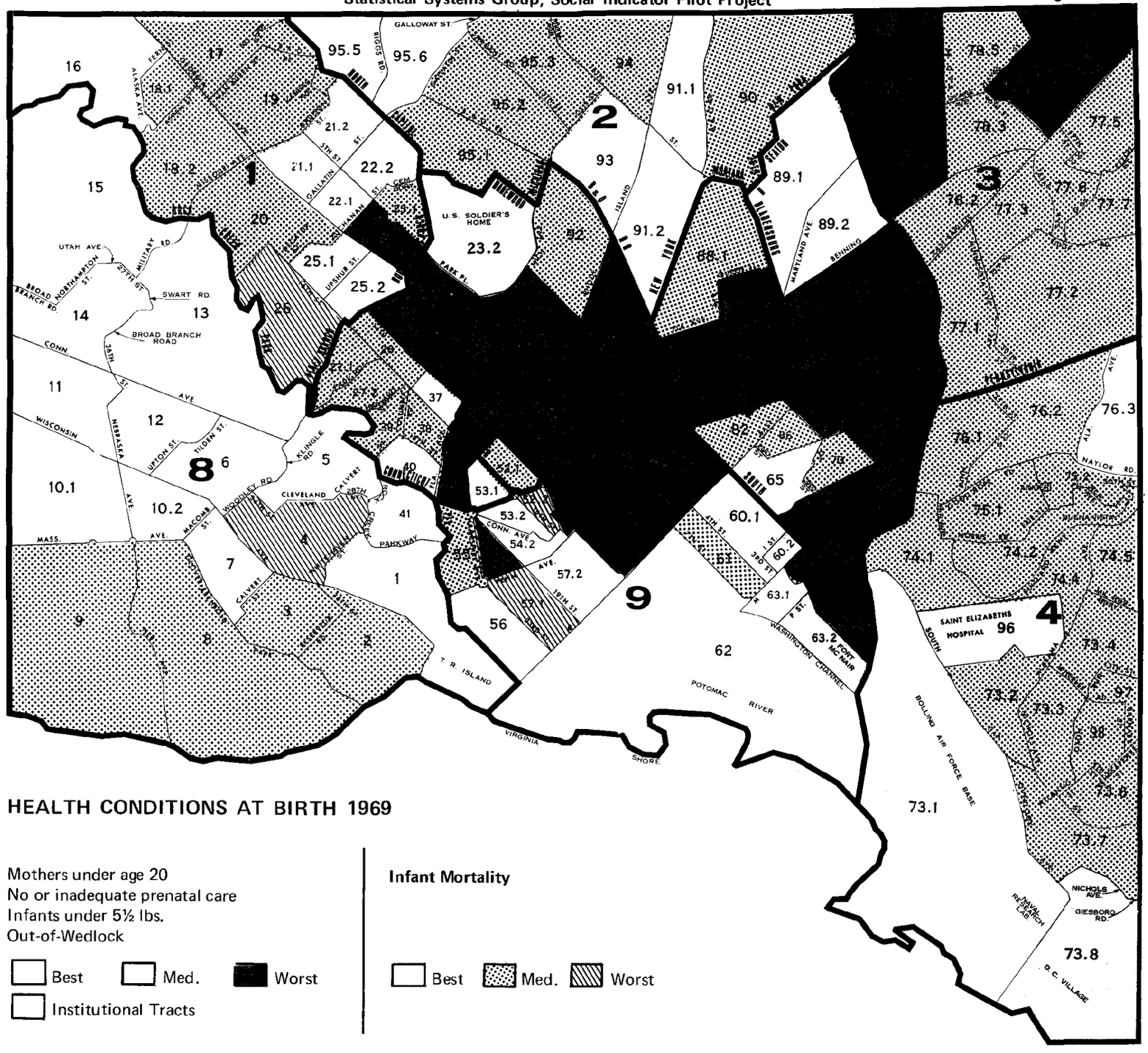
2b. Health Conditions at Birth: 1969
3b. Housing Indicator Group Means: 1970
4b. Socio-economic Variables Group Means: 1970
5b. Dependency Indicator Group Means: 1971

Figs. 1a and 1b not referred to and not included in this paper.

DISTRICT OF COLUMBIA, EXECUTIVE OFFICE

Statistical Systems Group, Social Indicator Pilot Project

Figure 2-a



HEALTH CONDITIONS AT BIRTH 1969

Mothers under age 20
No or inadequate prenatal care
Infants under 5½ lbs.
Out-of-Wedlock

Best Med. Worst

Institutional Tracts

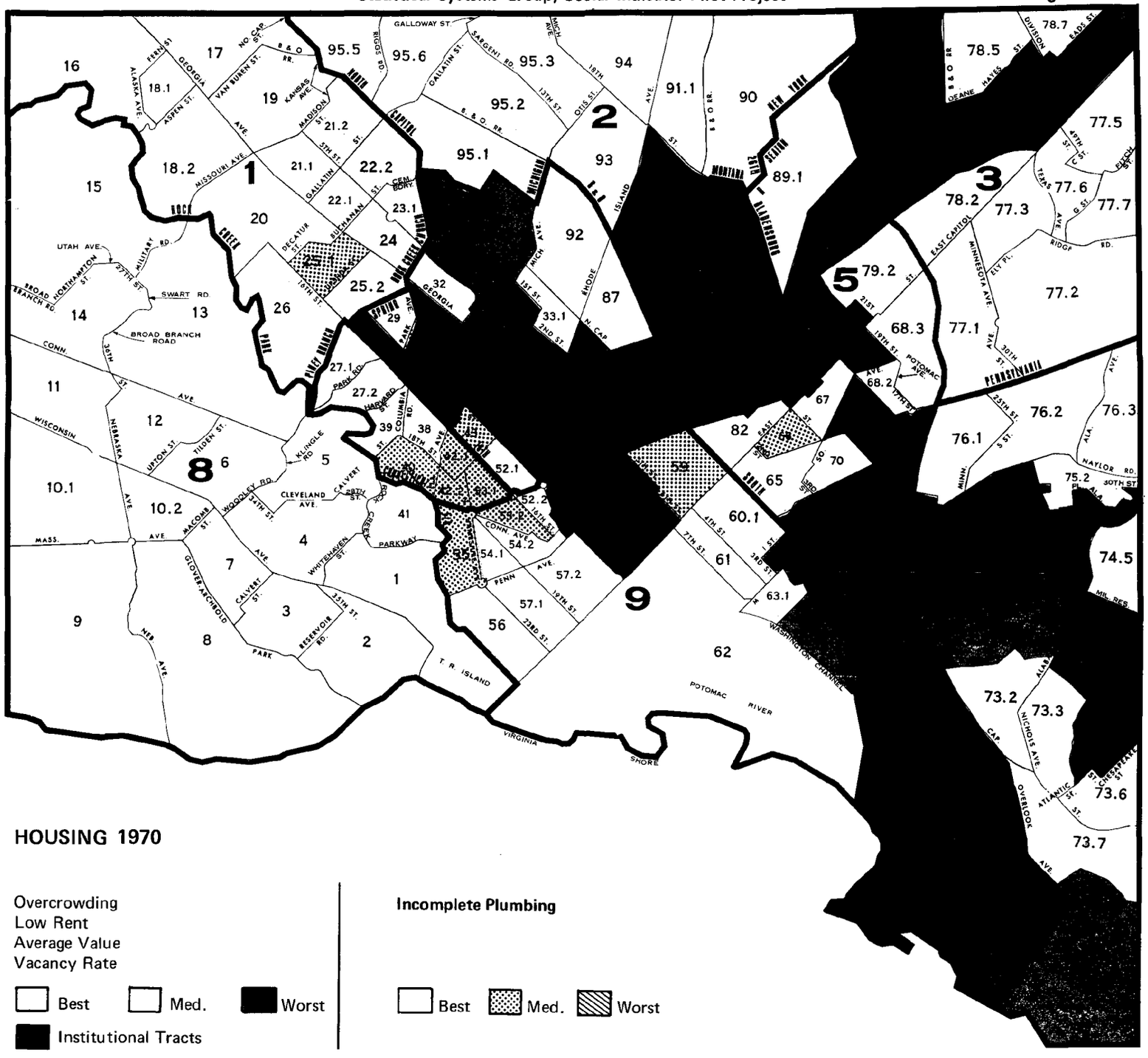
Infant Mortality

Best Med. Worst

DISTRICT OF COLUMBIA, EXECUTIVE OFFICE

Statistical Systems Group, Social Indicator Pilot Project

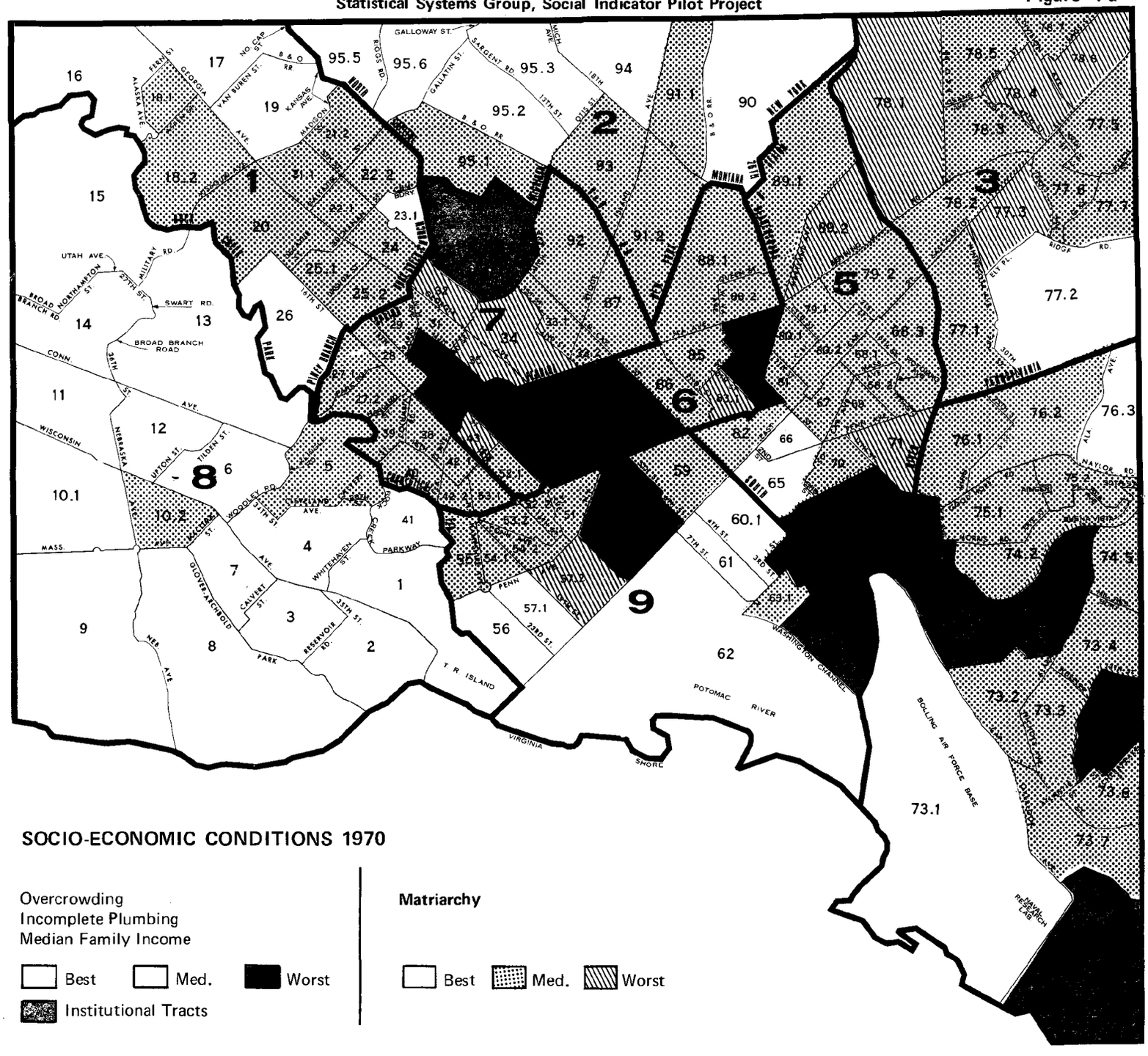
Figure 3-a



DISTRICT OF COLUMBIA, EXECUTIVE OFFICE

Statistical Systems Group, Social Indicator Pilot Project

Figure 4-a



SOCIO-ECONOMIC CONDITIONS 1970

Overcrowding
Incomplete Plumbing
Median Family Income

Best | Med. | Worst

Institutional Tracts

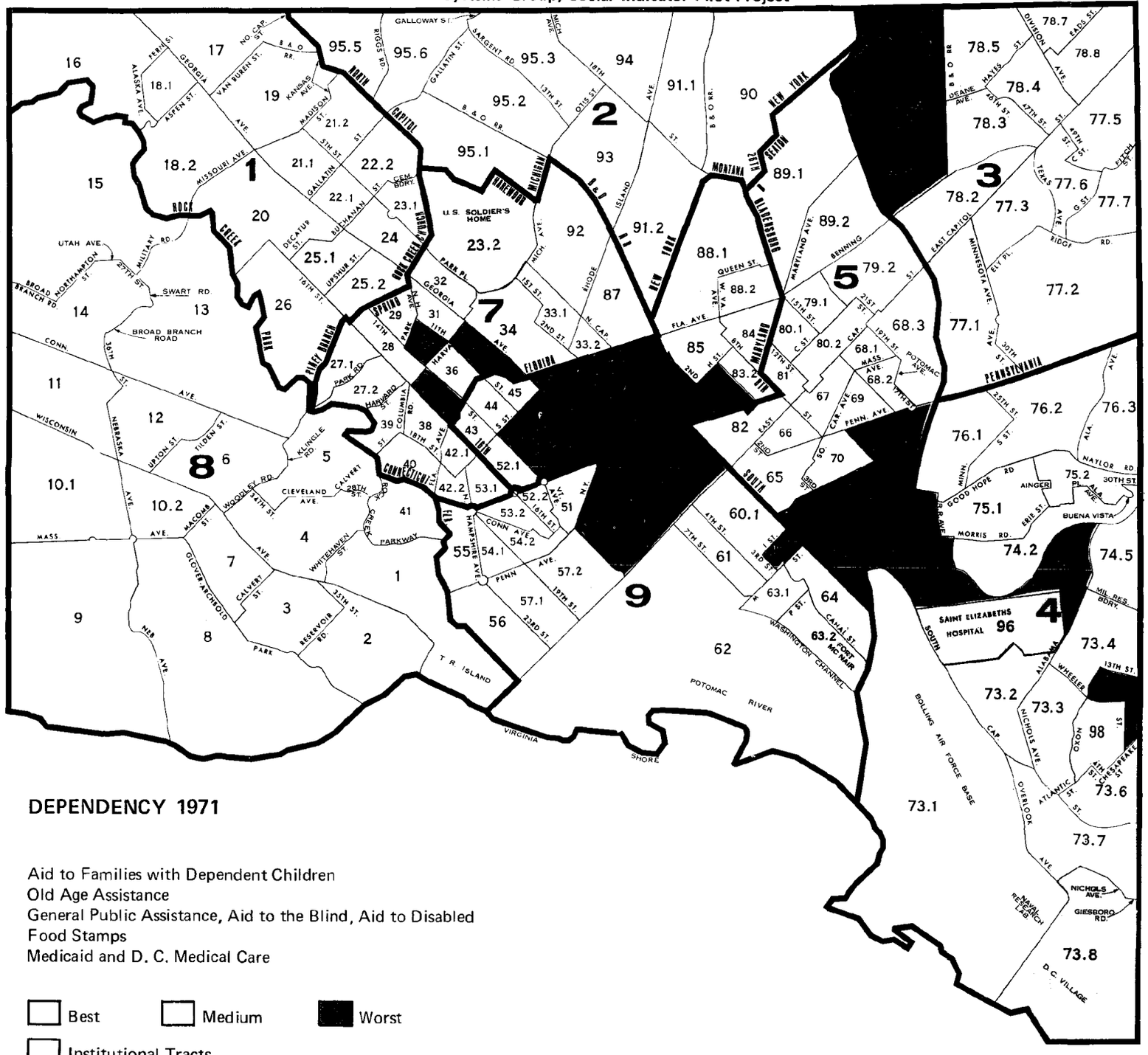
Matriarchy

Best | Med. | Worst

DISTRICT OF COLUMBIA, EXECUTIVE OFFICE

Statistical Systems Group, Social Indicator Pilot Project

Figure 5-a



Figures

2b. Health Conditions at Birth: 1969

(Group means in percent)

	Best	Medium	Worst	City-wide
Mothers Under Age 20.....	6.4	23.5	38.6	26.0
No or Inadequate Prenatal Care.....	9.1	24.2	33.3	25.0
Birth Weight Under 5 1/2 Lbs.....	8.3	12.8	16.6	13.4
Out-of-Wedlock.....	10.1	32.5	55.3	36.7
NUMBER OF TRACTS.....	28	54	40	122
Infant Mortality.....	1.7	3.4	6.4	3.0
NUMBER OF TRACTS.....	41	73	8	122

OPM - Statistical Systems Pilot Project

3b. Housing Indicator Group Means: 1970

(Group means in percent)

	Best	Medium	Worst	Citywide
Overcrowding....	3.5	12.7	22.2	12.2
Low Rent.....	4.2	16.9	39.0	19.9
Average Value of Owner-Occupied One-Housing Unit Structures.....	$37,000	$20,300	$17,400	$26,600
Vacancy Rate...,	2.0	3.7	7.0	4.1
NUMBER OF TRACTS	41	55	50	146
Incomplete Plumbing.....	1.4	6.2	21.6	2.3
NUMBER OF TRACTS	121	22	3	146

OPM - Statistical Systems Pilot Project

4b. Selected Socio-Economic Variables Group Means: 1970

(Group means in percent)

	Best	Medium	Worst	Citywide
Overcrowding Index.....	3.5	14.5	24.7	12.2
Incomplete Plumbing.....	1.0	2.1	7.1	2.3
Median family income.....	$16,300	$8,400	$5,900	$9,600
NUMBER OF TRACTS.....	40	84	22	146
Matriarchy index.....	13.6	30.4	45.6	30.1
NUMBER OF TRACTS.....	36	82	28	146

OPM - Statistical Systems Pilot Project

5b. Dependency Indicator Group Means: 1971

(Group means in percent)

	Best	Medium	Worst	City-wide
Aid to families with dependent children.....	7.0	26.1	42.8	22.0
Old Age Assistance.....	1.4	7.8	16.5	4.7
General Public Assistance; aid to blind; disabled.....	0.8	3.0	7.2	2.3
Food stamps.....	3.1	15.6	29.4	11.3
Medicaid and D.C. Medical Care.....	8.9	34.4	57.1	25.0
NUMBER OF TRACTS.....	66	59	21	146

OPM - Statistical Systems Pilot Project

COMPOSITE SOCIAL INDICATORS FOR SMALL AREAS – CENSUS USE STUDY– RECENT DEVELOPMENTS IN METHODOLOGY AND USES

Samuel P. Korper, John C. Deshaies, Leo Schuerman, and Ronald Crellin
U.S. Bureau of the Census

Introduction

To lend proper perspective to this paper, let us begin by briefly indicating the general Census Use Study program areas and developments we would like to cover. The earliest of these, the New Haven Health Information System, was begun in 1967 under the dual sponsorship of the Bureau of the Census' Census Use Study and the Maternal and Child Health Division of the Connecticut State Department of Health, with major support from the then Children's Bureau of the Department of Health, Education, and Welfare. Since its early formative years (1967-69), the Health Information System program has been evolving and expanding in a number of ways.

First, the subject matter which was originally focused on maternal and child health has been expanded to not only include many other components of health but also other subject matter areas such as welfare, education resource economics and social pathology.

Second, there has been a movement from a focus on small-area analysis of a one-time data base to one which could be tracked over time to discern changes in health and social status, the level of delivery of health and social services, the level of community resources, and so forth.

The third expansion represents the analysis we have been undertaking in composite social indicators. Our emphasis has always been away from one-dimensional social indicators. Even during the early New Haven days, we were working on composite indicators, and in those early days, many of the methods we worked with were hit and miss. Later, we shall discuss refinements in the methodology dealing with composite indicators.

Fourth, is the continuing development of technology by the Census Use Study which has been utilized by the Health Information System program. We believe that new technological developments can not be overemphasized. Preliminary to active concern about "social indicators" or "composite methods," there is a vast range of processing problems and details that must be dealt with. These include problems of data access, preparation and summarization; but most important, and often most vexing, is the problem of geocoding data to small areas (the issue of geographic base files) and ways and means to disaggregate local data files so that they are useful and meaningful for analyses.

The fifth area of expansion is inherent in the fact that our efforts are no longer limited to New Haven. Under a new program the Census Use Study has undertaken for the Office of Economic Opportunity, we are currently transferring our methods and technology to eight new sites–both urban and rural. Social and resource indicators programs are being implemented in Los Angeles, Chicago, Atlanta, Phoenix, Providence, R.I., and Mound Bayou, Miss. Initial forays are also being made into Lowndes County, Ala. and Baldwin, Mich.

A basic thrust of the Health Information System program, and one emphasized throughout this report of our activities, is the effort directed toward developing local data files. While we would be denying our heritage by failing to recognize the necessity for utilizing census data, we feel it is unwise to become too dependent on this resource. Census data are only available once every 10 years. Local data are generally to be had on a continuous basis. The problem, which we in no way intend to minimize, of course, is ferreting them out, and making them useful input to our analyses.

New Haven Health Information System

The major elements, tasks and results of the New Haven Information System program may now be briefly indicated. The fundamental ingredients of the program are:

1. The geographic base file (a computerized source map containing a listing of all possible address ranges and geocodes pertaining to these address ranges. A geographic base file, or the DIME (an acronym for Dual Independent Map Encoding) system, is imperative for supporting the local and census data contained in an information system).

2. Address Matching Computer Programs (a series of computer programs that will break down an unformatted address from a local data file into a number of components, screen a geographic base file to locate these components, and retrieve from the geographic base files the geocodes pertaining to that street address).

3. Data pertaining to public health and related phenomena that could be linked into a system of information through geocoding.

The original data set was derived from five sources: census information consisting of the results of the dress rehearsal census in New Haven (April 1967), both 100 percent and sample data; the vital records system; a special purpose health survey; and hospital obstetrical records obtained from the Yale-New Haven Hospital.

Having these ingredients, a number of subsequent tasks were required in the design of the health information system. These were:

1. to geocode all records in each of the five data files to block groups and census tracts of the City of New Haven.

2. to derive more than 300 data summaries or social indicators for each block group and census tract in the City of New Haven.

3. to delineate neighborhoods (on the basis of scaling and ranking procedures) ascertaining that the characteristics of the population contained therein are relatively homogeneous.

4. to examine so-called "high risk" health characteristics within and between neighborhoods.

5. to perform multi-variate analyses to determine relationships between sets of variables and previously formulated research hypotheses.

These tasks have been completed with a 1967 data base and are documented in Census Use Study Report No. 12.[1] We have organized a data base in New Haven for 1970 so that comparative time series (temporal trends) analyses may be performed.

The 1970 data base which is presently undergoing analysis in New Haven will consist of the following data files:

1. Census Summary Tapes—containing both head count data and sample census data. Social indicators, representing various dimensions of demographic, socio-economic status, migration, and the like, will be derived from these data sets. The emphasis will be on deriving social indicators previously used with the 1967 data base.

2. The Vital Records System—will serve as the source for deriving indicators of fertility, illegitimacy, deaths, and other vital events. The emphasis is again placed on deriving indicators previously used with the 1967 data base.

3. Other sources of information are contained in the 1970 data base which were not available in 1967. These will consist of additional data summaries obtained from agencies such as the Connecticut Mental Health Center, and the mental retardation registry made available through the Connecticut State Office of Mental Retardation. The Mental Health Center data summaries include social and economic information concerning the patient population of the mental health center. Additionally, information will be available for analysis on specific health variables such as days on service, hospital inpatient and outpatient days, admission unit, first treatment unit, interval since last admission, alcohol problems, drug problems, reason for discharge, recidivism, and diagnostic classification.

The mental retardation registry includes small-area data on the distribution of retardation by level of retardation. (There are two retardation levels of interest. The first level might better be labeled cultural deprivation since the retardation is not innate. The second level is severe retardation.)

In addition to these files, the 1970 information base includes further sources which are seen as useful adjuncts to our analysis, and which, it is felt, contribute significantly to a broadly based health planning effort. These are:

a. fire department data, including variables such as fire calls and emergency service calls, with further differentiations into health problem areas such as seizures, maternity, suicide, drug overdose, etc.

b. circuit court statistics, to include such variables as drug offenses, family disruption, and possible relationships to mental disorder.

c. juvenile court statistics, including drug offenses, cases related to conduct, and several indicators of the need for the appropriate health resources.

Several additional files including those of police, transportation, family relocation, and housing code violation authorities are currently being processed for inclusion in the analysis.

The processes for analyzing the 1970 data base, with one major revision and these several expansions, is patterned on the 1967 methodology, as documented in Census Use Study Report No. 12. The data files containing the obviously sizable number of social indicators will be structured through correlation and factor analysis. However, a major departure from the earlier procedures is that neighborhood delineation will be attempted through cluster analysis rather than the composite scaling method used in the 1967 analysis.

The analysis is designed to do two basic things. First, a comparative temporal analysis is being undertaken to identify the changes taking place between 1967 and 1970. The hope is to identify multi-dimensional or configurational changes; the assumption being that one-dimensional social indicators inadequately measure a dysfunctional situation. From 1967 data, some 20 research hypotheses were formulated and will be reconstructed from 1970 data. The rationale for reconstructing research hypotheses with 1970 data is to test the hypotheses over time, and to determine to what extent, if any, they continue to hold. Another aspect of this comparative analysis will be to examine specific neighborhoods to determine the direction of change, if any, of dysfunctional trends originally discerned in the 1967 analysis.

The second thrust of the analysis is to formulate new sets of hypotheses that can be tested on a subsequent data base. Exemplifying the actual and ongoing projects related to this overall work plan is that we have developed in close cooperation with the Connecticut Mental Health Center which serves the South Central Connecticut area. This study is an areal analysis of the utilization of mental health and related services in New Haven, exploring a conceptual area which may best be termed "expressed demand."

[1]U.S. Bureau of the Census, Report No. 12.

It is obvious that a major goal of the community mental health centers movement is the provision of services that are relevant to the needs of the population served, but the determination of the extent to which this aim is realized is no easy task. While general agreement on the best combination of strategies for the evaluation of community mental health programs has yet to be achieved, there is general agreement that an essential component of evaluation is that of "diagnosis" of the community. To this end, something has to be known about the characteristics of the population at risk (potential consumers of services) and of populations currently being served by existing agencies (current consumers of services) as indicators of demand. Clusters of variables (approximately 300) associated generally with socio-economic class are being analyzed as the best available predictors not only of case outcome, but also as pre-hospitalization determinants of who receives attention, the source of treatment, psychiatric diagnosis, and eventual reentry into treatment.

The preliminary analytical components of this study are: (1) the delineation of New Haven neighborhoods (block groups or census tracts) on the basis of factor analytic, scaling, and ranking procedures, and (2) ascertaining that the characteristics of the populations contained therein are relatively homogeneous.

The second analytical thrust of the study will be the development and distillation of data from 10 agencies which reflect a wealth of health and social information beyond that available from, but not limited to, census-type data. Two approaches are taken to further explicate the measurement and operationalization of the term "utilization" for this study. First is the recognition and analysis and description of the a priori separation of services which may be termed specifically mental health as opposed to those that are not. The second major approach is the development of "categories" of service utilized, through the identification and designation of common clusters or modules of service.

The third emphasis of this study is the assessment of utilization of mental health services and facilities by defined population group from small areas and relating their help-seeking behavior to the utilization of other services which call for or reflect need of psychiatric intervention.

The emphasis, then, of this and other studies undertaken in conjunction with the New Haven Health Information System, will be on analyzing non-census data because of the likelihood that the next census will not take place until 1980. What is needed in the interim, therefore, is an information system not completely dependent on census data, hence our conscious emphasis on local records and data sources. An additional important aspect of the analysis will be its capacity for replication in several additional sites throughout the country. Such activities are reviewed below, emphazing research into composite social indicator development and the programmatic role such research plays.

Social and Resource Indicators Program

As we indicated earlier in this paper, the New Haven Health Information System has long since expanded beyond New Haven and is no longer concerned solely with health or analyses utilizing a one-time data base. The O.E.O. Social and Resource Indicators Program alluded to earlier is now in the data collection stages in Los Angeles; Mound Bayou, Miss.; Atlanta; Providence, R.I.; and Phoenix. Basically, the program is being undertaken in three stages. Stage I consists of identifying and assessing data sources and files, access problems, and estimating necessary initial operations and costs to ready the data to be collected for input to the indicator system to be developed in Stage II. Stage III involves the use of operational data of the O.E.O. Health Program in conjunction with the indicator system to assess certain implications of these programs to the populations they service. Stage III also entails structuring the indicator system to be turned over to local groups to be responsible for operating and using the indicator system on a regular, continually updated basis from 1973 into the future. The study design at the present stage of development entails tapping numerous local data files to obtain health, social, and resource indicators for small geographic areas on an annual basis from 1965 through 1970. Without going into detail, the indicator matrix will include, but not be limited to, data summarized from the vital record system; reportable disease registries; public health clinics; health and allied health manpower; facilities such as hospitals, nursing homes, and extended care facilities; financial data on health department operations; welfare data on the five categorical programs; medicaid and food stamps; mental health; education; employment; public and private housing; taxation; land use and access; and deviant behavior—crime and juvenile delinquency statistics.

Basically, the idea is to start with a master matrix that is composed of what would obviously be an unwieldy number of cells dealing with a broad range of subject matter areas and ending up with a structured indicator matrix reduced through multi-variate analytic approaches (such as regression and discriminant analysis) to fewer and more meaningful composite social indicators that are sufficiently viable to be monitored over time. Through a "before-and-after" research design our intention is to establish this indicator system on an on-going basis; tie it into structured operational data to be provided by O.E.O. health centers or health networks; and finally, to select control areas without health programs and attempt to discern what influence, if any, these health programs have or may exert on the health and social status, delivery of health and social services, and level of resources of the communities they service.

The composite indicators, as was the situation with earlier work in New Haven, will consist primarily of logical statements of interrelationships between a number of variables in the system derived and examined through multi-variate analyses. In addition to the hypothetical generalizations we have already mentioned in connection with the early work in New Haven, we will be concerned with a large number of issues which are of some importance to the O.E.O. health programs. We would like to present here two examples of issues we intend to examine in this program.

The first concerns health status. We will formulate such questions as: What combination of factors effected by the O.E.O. health programs contribute in what ways to improving the health status of populations served by those programs? Can we define the effects of the health programs themselves? Does health status change over time? If so, what other characteristics associated with health status also change? For example, if an improvement in health status on a certain category or morbidity is discernible, is or was there also a concomitant change in such areas as level of delivery and quality of health services, social pathology, social status, and so forth?

Second, what combinations of factors effected by the O.E.O. health programs contribute (and in what manner) to more efficient delivery of health and social services? Example: Medicaid, which pays for many of the health services received by the poor, is tied inextricably to the welfare system. In some places, a

prerequisite to the receipt of medicaid benefits is that the person or family be first determined eligible to receive welfare benefits under one of the categorical programs. However, there are many barriers placed in the path of a potentially eligible person who applies for welfare. Unless he is completely disabled, he must usually apply in person to a welfare office which may be difficult to reach. Once he gets to the welfare office, he must fill in complicated application forms which he may have difficulty in completing. Then he must go through an investigation, a series of interviews and a waiting period before the determination is finally made. Once the determination is made, it is subject to review once or twice a year.

The point we want to make here is that one must look at a far-reaching complex of issues to begin to investiage a question such as whether medicaid is, or is not, providing benefits which characterize efficient delivery of health services to the poor. O.E.O. health programs can effect this component of the health delivery system in a number of ways, including, as in this case, the assisting of recipients through the cumbersome welfare process. Through the positing of an hypothesis, or, if you will, composite indicators which aid in the modelling of this whole process, we can begin to assess this aspect of delivery of health services. In the analysis of this situation over time, we must attempt to discern the changes in all the components of the configuration.

We have indicated two examples of the hypotheses to be formulated with the information base being developed by the Census Use/O.E.O. indicators study. There are, obviously, many more dealing not only with health status and delivery of health services but also with related issues such as improved "levels of living," attrition in the educational system, access to recreational and social services, resource allocation in the community, mental health and deviant behavior, and a plethora of other issues which affect personal and public well-being.

At this point in time, we have more questions than we have answers. However, this is what the research and development of composite social indicators is all about. Before any meaningful quantum leaps are forthcoming, it is, we feel, necessary to go through this process of research design and hypothesis formulation, attempting to sharpen the definition of intended output as well as underlying theoretical constructs. Obviously crucial in this endeavor is the derivation of a reasonably acceptable definition and interpretation of terms such as "health" and "social indicator," to mention but two. While having reiterated and vowed allegience to such awesome definitions of health as those held by agencies such as the World Health Organization, operationalization dictated a more limited (read realistic) definition. To this end, sights were lowered, and analyses were refocused on "intermediate outputs" and approaches to examination of causal linkages between programmatic inputs and output in terms of measurable health states or outcomes (prenatal visits and birth weight). These quantified states could never have been approached utilizing the original broad definitions of preferred health status at once so moving and elusive. Without reduction to and construction of testable theory, meaningful evaluation becomes impossible.

The development of social indicators and the mysticism associated thereto presents a problem certainly as complex. Varying interpretations were assessed and often found wanting. It is obviously too simplistic, for example, to suggest that measurable change in social indicators (given their selection) should result merely from health program impact, without simultaneously controlling for a host of social and economic variables. Reasonable estimation of the effects of the health programs on health states must be normalized for social and economic variables both cross sectionally and prospectively in order to have sound evaluative or analytic purpose. The Los Angeles study provides an example of Census Use Study approaches to these concerns.

Composite Indicators of Change—the Los Angeles Case

To place the Los Angeles study in perspective, it is worthwhile to briefly recapitulate the framework for the social indicators program research design. In outline, it includes the following broad steps:

1. Obtain and combine numerous local administrative files on an annual basis from 1965 to the present;

2. Generate relevant social and health indicators for small geographic sub-areas;

3. Formulate research questions and hypotheses focused on community health;

4. Develop an analytical perspective that includes a "before-and-after" design around the emergence of O.E.O. health programs, as well as examine the concomitant relationship between program service, social condition and health status;

5. Compare the community service and health trends in the O.E.O. service areas to appropriate control areas without health programs.

Through a current arrangement with the Los Angeles County Health Care Department, the plan is to assist in the establishment of a countywide system which will continue the monitoring of these data once the initial stages are complete.

In Los Angeles, many of the identified relevant local agency data files, such as those alluded to earlier, are available from 1965 to the present. The files are currently being prepared to cover a common universe of territory corresponding to three designated O.E.O. service areas, with the basic areal unit of analysis being the census tract. For the three Los Angeles centers, there are approximately 200 census tracts identified and data from various governmental, social and voluntary agencies are being aggregated for these sub-areas.

In the main, we are interested in the health center service areas as a whole, but realizing the existence of an uneven texture within these areas, it becomes important to indentify the changing trends and "hills and valleys" of health and health related characteristics in the areal cohort. Such changes are typically approached by examining some rate or ratio of difference (e.g. $x^2/x^1 - 1 \times 100$ = "percent change") for each derived index, for each year being analyzed. These change measures are often displayed in tables or in figures either as a single line or as a multiple of lines, each line representing a different item over time. The figures can, of course, also be displayed in the form of maps drawn by hand, or by computer graphic techniques. The consequence of the relationships between variables might be analyzed in the form of a descriptive discussion or more elegantly presented in statistical summaries.

With regard to the former, the diversity in number of variables, cases and temporal observations involved for the three Los Angeles service areas often makes such an approach voluminous, confusing and, at best, a judgmental summary of the relationships observed. On the other hand, typically derived rates of change mentioned above are often processed utilizing statistical summary measures. When this is done, many troublesome statistical issues are raised, such as "auto-correlation" and "part-whole correlation" influences. Also, variable universes, when extended over time, may be confounded by problems of measurement which include variable decay, inflation, loss of timeliness, comparative reliability, and so on.

The analytical scheme to be implemented in the Los Angeles O.E.O. indicators project is being designed to overcome, insofar as possible, many of the issues that are conceptually, and, ultimately, statistically prevalent in the analysis of multi-dimensional indicators that extend through time. In the main, the procedures described below avoid the creation of ad hoc change variables from the original indexes gathered on a timely basis. Rather, the individual or composite measures are generated out of variant themes using standard regression and correlation techniques. This approach has particular relevance to our study of small areas because it permits the use of various parametric statistics that describe the "fit" and variance of selected indicators.

After the formalization of statements that describe the relationship between indicators reflecting the condition of such variables as community health, social structure, and health services, it is important to obtain measures of ratio of change and the concomitant relationship concerning the velocity of a sub-area's rate of change. For example, where specific health services increased (or decreased) rapidly over time, there would be an association with accelerated changes in community health and social well-being. This is significant in understanding the association between dramatic or subtle shifts in the patterns of health indicators and rapidity of change for health services.

The development of composite indicators for the O.E.O. health service areas has two facets of multi-dimensionality: (1) each index extends through time and is measured on an annual basis; and (2) a composite of indexes can be a conceptually and statistically related set of independent variables to a target or dependent index.[2] The first type of "composite" is, thus, a single item extended through time; and a composite value for each geographic sub-area is not only useful by itself, but properly constructed, it can be used to determine the degree of relative change over time for each areal cohorts' progression or deterioration on the single indicator. Furthermore, it would seem important to know what segments of the distribution change more or less rapidly over the period of observation. These measures can be obtained on an "inter-annual" basis through the extraction of multiple regression residuals, summarized by regression coefficients, and examined for fit by standard error of estimates and multiple coefficient of determination approaches. Inter-annual regressions are calculated for the same variable over two or more points in time. For example, the inter-annual regression for data between 1965 and 1968 is represented as:

$$X_4 = a_{4.123} + b_{41.23}x_1 + b_{42.13}x_2 + b_{43.12}x_3$$

where $x_1 = 1965, x_2 = 1966, x_3 = 1967, x_4'' = 1968$

The individual residuals for each observation (residual = $x_{nth} - x^1_{nth}$) can be interpreted as an inter-annual composite measure of change relative to the universe of cases extended through time for a particular variable. Furthermore, since the same variable for more than one point in time is projected on a hyperplane, a net regression coefficient of unity (i.e. ') is interpreted to mean that any change that took place from time one to time two was, on the average, equal along the entire distribution after all other time measures have been excluded. (A special case is a bivariate distribution along an X and Y axis represented by a linear slope.) It would also follow that coefficients' less than unity indicate that geographic sub-areas low at the earlier time increased, or decreased—whichever is the case—faster between the respective time points than did the higher initial values in the distribution. A similar but reciprocal interpretation is made for coefficients that exceed unity. Thus, there now exists a descriptive summary of the relative average degree of change (in terms of place and degree for an areal cohort for each period of observation).

Complimentary to the inter-annual regression coefficient is the inter-annual multiple coefficient of determination. Here we can summarize the degree of stability (or change) for a given indicator. The higher the coefficient of determination, the more relatively stable the particular health or social phenomena are over time. Finally, standard error of estimates gauge the overall variance that might be expected for the specific indicator through time.

Thus far, the measures and statistical summaries have centered around a time-related "composite index" for a single variable. Yet, earlier it was suggested that to assess the social or health status of a community, it would be advantageous to develop component measures that isolate the concomitant relationship between the velocity of the rate of change in social or health composition for small areas and the rate of change in, say, neighborhood health services.

One procedure that measures the relationship between rates of change for a composite of independent variables to a dependent variable is to calculate coefficients of "multiple deviational" change. This type of correlation provides a statistical description of the degree of association of the velocity of change in each respective measure of health with rates of change for a composite of selected community social or health indicators. Such a statistical procedure also provides a means whereby the residuals can be used as indicators of individual case deviation (away from a relative position) to the universe of cases extended over time.

Specifically, the multiple deviational correlation is a multiple correlation wherein both the dependent and independent variables are the residuals of least squares multiple "inter-annual" regressions. Thus the data used in the multiple correlation are

[2]The statistical arguments presented in the remainder of the paper are derived in part from the following sources: Otis D. Duncan and Ray P. Cuzzort, "Regional Differentiation and Socio-economic Change," Papers and Proceedings of the Regional Science Association, IV (1958); Otis Dudley Duncan, Ray P. Cuzzort and Beverly Duncan, Statistical Geography Glencoe, Ill.: The Free Press, 1961; George C. Myers, "Variations in Urban Population Structure," Demography, I (1964), 156-163; Maurice D. Van Arsdol and Leo A. Schuerman, "Redistribution and Assimilation of Ethnic Populations: The Los Angeles Case," Demography, VIII (November 1971), 459-480.

essentially obtained by first calculating multiple inter-annual regressions based on data with "n" time points of observation for each of the variables to be used in the multiple deviational correlation. To review the inter-annual procedure above, each needed inter-annual regression uses the latest set of n^{th} observations as the dependent variable and each preceding temporal set of cases are treated in turn as independent variables. The deviation of the final temporal observation for each variable from its expected values (X') are the data used in the multiple correlation of deviational change. Therefore, the degree of fit is calculated utilizing these sets of transgenerated forms of the raw data called multiple deviational change measures. The multiple deviational coefficient of determination is particularly suited for the present problem because it reflects the velocity of concomitant change between times of observation after "removing" the influence of "natural" relative stability, or change, for each selected variable. Furthermore (and similar to the logic in the case of residuals obtained from calculating inter-annual regressions), residuals obtained from deviational regressions can now be viewed as relative measures that are a composite of both multi-dimensional social indicators as well as accounting for these indicators extending through time.

Summary

We have tried to present an overview of Census Use Study research and programmatic incorporation of the concepts related to composite social indicators. Various levels of grappling with this concept have been evidenced throughout this work, extending from the relatively fundamental work of data acquisition to the more heady reaches of conceptualization and theory testing. Operationalization on a day-to-day basis, working in the real world of local area data files and all that that implies, has stripped much of this work of its glamour, replacing it instead with the basic concerns and compromises necessary to "get on with it."

Generally, we are vexed by rather primordial questions such as choice of component indicators and the most appropriate methods of combining them. Working with such large data sets which represents the efforts, programs and services of so many agencies requires vigilance in the avoidance of cross-contamination and almost certain tautological traps. Despite resistance to premature conceptual and theoretical closure, we have found that while it is often quite fruitful and/or popular to use several different indicators to represent a single underlying concept, that in addition to the combined use of multiple indicators, there is still much room for the recording and observation of individual associations between each indicators of the independent and dependent concepts under study. Finally, we feel that it is not an essential characteristic of our indices that they be free from conceptual or operational defects or limitations. Rather, we concur with Bickner, who has said "...all that is necessary for such indices is that they be: reasonably interesting, objectively measurable, commonly reported, and reducible to a common number not exceeding three significant digits."[3]

[3]Bickner, Robert E., "Measurements and Indices of Health," Outcomes Conference I - II, Methodology of Identifying, Measuring and Evaluating Outcomes of Health Service Programs, Systems and Sub-Systems, Carl E. Hopkins, ed., National Center for Health Services Research and Development, Department of Health, Education, and Welfare, 1969.

Selected Bibliography

Ball, G.H. and Hall, D.J. *Applications of Cluster Analysis to Bureau of the Census Data,* Menlo Park, California, Stanford Research Institute, 1970.

Bauer, Raymond A., ed. *Social Indicators.* Boston: Massachusetts Institute of Technology Press, 1966.

Bickner, Robert E. "Measurements and Indices of Health," Outcome Conference I-II, *Methodology of Identifying, Measuring and Evaluating Outcomes of Health Service Programs, Systems and Sub-Systems,* Carl E. Hopkins, ed., National Center for Health Services Research and Development, DHEW, 1969.

Bisco, R.L. *Data Bases, Computers and the Social Sciences.* New York: J. Wiley and Sons, 1970.

Bodian, C., Gardner, E.A., Willis, EM. and Bahn, A.K. *Socioeconomic Indicators from Census Tract Data Related to Rates of Mental Illness.* U.S. Department of Commerce, Bureau of the Census, Washington, D.C.

Crellin, Ronald. "Los Angeles County Data Files Geocoded with ADMATCH and Their Applications." American Institute of Planning Annual Conference, St. Paul, Minn., October 1970.

Deshaies, John C. and Korper, Samuel P. "The Use of Census Data and Local Files for Health Planning." Presented to the Urban Regional Information System Annual Conference, Louisville, Ky. September 5, 1970.

Deshaies, John C. and Seidman, David R. "Health Information Systems," *Socio-Economic Planning Sciences,* Vol. 5, pp. 515-533 (1971)

Duncan, Otis D., Cuzzort, R., and Duncan, B. *Statistical Geography,* the Free Press, Glencoe, Ill., 1961.

Ezekial, Mordecai and Fox, K. *Methods of Correlation and Regression Analysis,* London: Wiley and Sons, 1967.

Greenberg, B.G. "Evaluation of Social Programs," *Review of Int. Stat. Institute,* Vol. 36, No. 3, 260-278, 1968.

Jackson, Alton F. and Curtis, R.F. "Conceptualization and Measurement in the Study of Social Stratification," in Blalock, H.M. and Blalock, A.B., eds., *Methodology in Social Research,* New York: McGraw-Hill Book Co., 1968.

Myers, George C., "Variations in Urban Population Structure" *Demography,* I, 1964.

National Institute of Mental Health (Redick, R.W., Goldsmith, H.F. and Unger, E.L.) *1970 Census Data Used to Indicate Areas with Different Potentials for Mental Health and Related Problems,* USPHS, Publ. No. 2171, Washington, D.C. Government Printing Office, 1971.

Reinke, W.A., and Baker, T.D. "Measuring Effects of Demographic Variables on Health Services Utilization" *Health Service Report,* 2:61-75. Spring, 1967.

Tischler, G.L., Henisz, J., Myers, J.K. and Garrison, V.K. *The Use of Mental Health Services by Catchmented and Non-Catchmented Populations.* Mimeographed report for NIMH, 1971.

U.S. Bureau of the Census. Census Use Study. Deshaies, Korper and Siker, *Health Information System II,* Report No. 12, Washington, D.C. 1971.

U.S. Bureau of the Census. Deshaies, *Health Information System* Report No. 7, Washington, D.C. 1969.

U.S. Bureau of the Census. Census Use Study: "ADMATCH: An Address Matching System," Washington, D.C.

U.S. Bureau of the Census. Census Use Study: "DIME: A Geographic Base File System," Washington, D.C.

U.S. Bureau of the Census. Census Use Study: "GRIDS: A Computer Mapping System," Washington, D.C.

U.S. Bureau of the Census, Wallach, Harold and Deshaies, John D. "Development of Social and Health Indicators to Profile and Monitor Populations Served by the Office of Economic Opportunity Health Programs, Atlanta, Ga.—Stage I" Washington, D.C. 1972.

U.S. Bureau of the Census, Deshaies, John C. and Wallach, Harold C. "Development of Social and Health Indicators to Profile and Monitor Populations Served by the Office of Economic Opportunity Health Programs, Chicago, Ill.—Stage I" Washington, D.C. 1972.

U.S. Bureau of the Census, Deshaies, John C. and Korper, Samuel P. "Development of Social and Health Indicators to Profile and Monitor Populations Served by the Office of Economic Opportunity Health Programs, Mound Bayou, Miss.—Stage I" Washington, D.C. 1972.

U.S. Bureau of the Census, Deshaies, J.C., Humphrey, A.B., Korper, S.P. and Wallach, H.C. "Development of Social and Health Indicators to Profile and Monitor Populations Served by the Office of Economic Opportunity Health Programs, Phoenix, Ariz.—Stage I" Washington, D.C. June 1972.

U.S. Bureau of the Census, Korper, Samuel P., Deshaies, John C., and Organic, Harold N. "Development of Social and Health Indicators to Profile and Monitor Populations Served by the Office of Economic Opportunity Health Programs, Providence, R.I.—Stage I" Washington, D.C. June 30, 1972.

Van Arsdol, Maurice D. and Schuerman, L. "Redistribution and Assimilation of Ethnic Populations: the Los Angeles Case," Demography, VIII, 1971.

APPLICATION, LIMITATION, AND METHODOLOGY OF THE SCIENTIFIC URBAN MATRIX (SUM)

(A Case Study of the City of Los Angeles)

Alan J. Weiner, *President
Mitus Corporation, Los Angeles

Context

Due to the vigorous competition for limited resources between the virtually unlimited needs of housing, education, employment, public safety, etc., it is imperative to examine quite critically the entire range of needs that would be affected by given policy and executive decisions.

Organization

The Los Angeles Community Analysis Program has been funded in order to determine the relative importance and extent of the many problems throughout the City in order to establish priorities for Citywide and community programs to correct existing obsolescence and prevent further inroads of urban deterioration. The Community Analysis Bureau (CAB), an independent City department—in fact, a staff function organized as a line department—was established for this purpose by ordinance of the City Council.

Function

The CAB study was designed to serve the evaluation of blight, i.e., deficiencies in the quality of living and usually the failure of one or more of the delivery systems that exist for the provision of goods or services necessary to urban living.

Resources

Critical to an urban system analysis is the flow of pertinent, relevant, and available information from public and private sources. Systematically organized information enables the identification, definition, and analysis of urban problems, the rational recommendation of well-conceived remedial action programs, and the monitoring of the environment to see whether those programs have produced the desired effects. The flow of information, as organized into the Urban Information System, enables the storage and retrieval from a continually increasing number of voluminous and diverse data files that are being acquired from a broad spectrum of public and private sources.

These include data on morbidity, mortality, births and other vital statistics; probations; crimes and arrests; business taxes; mental health incidents; parcel valuation, use, zoning, and appurtenances; building permits, inspections, and occupancy certificates, employment, net worth, gross sales, and credit ratings of establishments by Standard Industrial Classification (SIC) and situs; public school enrollment, attrition, transiency, and test scores; welfare cases, eligibility, and payments; and various special survey files such as the 1940 through 1970 censuses, transportation surveys, industrial surveys, etc.

Scientific Urban Matrix (SUM)

A major step forward—from data to information—toward the solution of the problems of evaluation of the urban system based on descriptive statistics was accomplished by the development of a statistical social trade-off matrix named the Scientific Urban Matrix or SUM. In 1970 and again in 1971, this technique of "pre-mathematics" facilitated this operation by not only describing a number of significant elements in the urban system but, even more so, by suggesting common linkages between these elements with the goal of creating a comprehensive theory of urban blight. An important benefit of this approach is that the linkages in the urban system described by the matrix suggest trade-offs in considering various remedial actions. A vehicle is provided for simplistic, static preplan testing of concepts that lie behind the remedial action programs.

Capabilities

The Scientific Urban Matrix approach subsumes specific procedures and results which both support and follow from it:

1. It provides a means to coordinate subjective qualitative evaluations with objective quantitative data.
2. It provides a means to identify objectively the areas of the City that were blighted and the criteria by which they are classified as blighted.
3. It provides a means to synthesize identified urban threats and requirements.

*Formerly Manager, Data Analysis Division, Community Analysis Bureau, City of Los Angeles.

4. It provides, through the linkages seen within the urban system, a means to evaluate the indirect as well as the direct ramifications of remedial actions prior to their implementations.
5. It provides direction for the monitoring of changes in the City that could result from the implementation of remedial action programs.

Implementation of the Scientific Urban Matrix requires the application of techniques that can deal with complex relationships among many factors. Once these relationships are known, the effects of fluctuations among the characteristics of remedial action programs on the measures of program effectiveness are readily determined through relatively simple linear equations. These, in turn, allow the development of elasticities of change—i.e., measures of the sensitivity of the total urban system to induced changes that derive from remedial actions. Furthermore, through trial and error manipulation of the independent variables, those that represent direct program actions, near optimum solution to urban problems through realistic programs, can be derived.

Structure

The Scientific Urban Matrix performs its analytical function by logically organizing the universe of available urban data around these assumptions:

1. That the City is a system composed of various delivery subsystems.
2. That blight is the symptom of the failure of these delivery subsystems to adequately provide for the needs of their users.
3. That each delivery system is composed of various dimensions that are at least partially measurable by various techniques of social and behavioral science.
4. That the data that are available are sufficient in terms of extensiveness of scope and intensiveness of quality to approximate significant attributes of the various delivery system dimensions after appropriate manipulation for purification and factoring.
5. That the overall urban system performance is best measured by its users rather than by "experts," provided that the criterion of system performance is presented as a user judgment that is simple and that does not require vast quantities of data and techniques of analysis that would normally not be available to the users.

SUM is applied by describing the urban system as a series of delivery systems—health care, education, income production, law enforcement and justice, etc. (Planning, Programming, and Budgeting System (PPBS) categories)—with each having a potential for a set of behavioral dimensions defined as follows:

1. The action of the cultural-attitudinal system (of a given society) is to motivate individuals towards the achievement of specific values or goals. Communications are the instrumental vehicles through which culture fulfills its function. (The process of acculturation or acquisition of new values is a learning process wherein an individual adapts to his human environment.) Cultural learning may be viewed as a process of conditioning to group norms. The function of the cultural system, which is a structure of values, is to motivate individuals towards "very long run" behavior.
2. The societal system functions to provide the aggregate behavior or group roles required to reach the image of the "ideal state" embodied in cultural values (in the long run) The structure of the social system is given by the pattern of interrelationships between the members of the group such as the degree of segregation, the number of roles played, etc.
3. The action of the political system is to modify group goals and conflicts and for mobilizing groups for the attainment of intermediate-run goals. The structure of the political system is found in the government and party institutions regulating political life and in the political participation, political pressure, legal constraints and their enforcement, and other governmental responses, particularly in terms of the level of services rendered to a community or group.
4. The economic system consists of the provision of goods and services for the want satisfaction of consumers in the short run (resource allocation behavior). The structure of the economic system is found in production, distribution, and consumption units. Indicators of this dimension include measures of the effectiveness of the market as a resource allocation process, costs of services rendered, and the costs of failures of the delivery systems.
5. The performance dimension includes the elements most symptomatic of the success or failure of the various delivery systems. It is the set of impacts directly felt by changes within the systems. Hence, the variables included in it are those that seemed to best measure the overall effectiveness of each functional system.

All available urban data, corresponding to these delivery systems, are retrieved and aggregated to common, reasonably homogenous geographic units, the communities. Residents of the City, the system users, are sampled and polled to rate the communities of which they have knowledge to obtain a subjective evaluation of the overall performance of the system.

The condition and services descriptive data or social indicators are statistically evaluated for the degree that they contribute to the overall performance for each delivery system, and they are factored to discern composite indicators that relate to a delivery system. Blight thresholds are found by looking at values of indicators that discriminate a blighted or endangered community from safe communities, and the various delivery systems are weighted based upon the relationships that exist between their data and the overall community ambiance scores, the reactions of the polled residents—this last providing the derivation of system sensitivities or elasticities.

SUM is designed to provide a needs analysis; hence, it is essential to translate the threats that are investigated by the above technique into requirements for action. In order to do this post-mathematics, it is essential to examine the threats by community in an integrated fashion, to translate them into units of response, and to examine on-going programs for their relevance and applicability toward fulfilling those needs. The Scientific Urban Matrix serves this purpose by dividing the communities into blighted, endangered, and safe areas and examining them for threats, requirements, and on-going programs.

In the context of the SUM, the urban system is defined as the conduct of various activities or delivery systems in a multidimensional network where each dimension of the delivery

systems has an impact—direct or indirect—upon every other dimension of the delivery systems. These levels of activity and of the ensuing experiences are treated as subsystems of the total system with inputs, outputs, and transfer processes pertaining to a given level, and appropriate linkages between these levels. Each level has an action—denoted by the type of goal that is set—and also an observable structure.

Thus, the Scientific Urban Matrix (SUM) is initially based upon the assumptions that the urban phenomenon is composed of a series of delivery systems which, in turn, disaggregate into various common dimensions that are at least partially measurable by various techniques of social and behavioral science utilizing data that are generated by each of those systems in their implementations. Therefore, for a variable to be included in the matrix, it is essential that data measuring it be available, that they be sufficiently extensive to cover all areas of the City, and that they be of a sufficient quality to utilize them as surrogates of the various delivery system dimensions.

Variables

Selection criteria for the variables included the relevance to each of the 40 cells—i.e., attitudinal, societal, political, economic, and physical performance manifestations arrayed along the horizontal axis; and accessibility, law enforcement, fire protection, recreation, health care, housing and neighborhood, income production, and education functional subsystems arrayed along the vertical; the ability of the variables to be stated operationally in terms of requirements for action where indicated; and the availability of data that could readily be developed into statistical area aggregations. Seventy-two candidate variables resulted for which data collection and data development efforts were undertaken. Although most data resided on magnetic tape files of the Urban Information System, in some cases the data did have to be hand crafted.

The following table summarizes the variables and their sources.

Summary of Variables and their Sources

ACCESSIBILITY	LAW ENFORCEMENT	FIRE PROTECTION	HEALTH CARE	RECREATION	EDUCATION	HOUSING AND NEIGHBORHOOD	INCOME PRODUCTION
Attitudinal							
Modal preference of transportation d = % Bus - Income - Transit Miles in neighborhood (LARTS67, SCRTD)	Juvenile probations/ Elem. school enrollment (Los Angeles Probation Dept., 1970 Census) Juvenile delinquents elem. school enrollment (LAPD, 1970 Census)	Malicious false alarms/capita (FIREINCD, POPEST) △Malicious false alarms 1965-1968 (FIREINCD, POPEST)	Inoculable diseases of those under 13 yrs./Children age 0-13 yrs. (HEALTH, 1970 Census)	Vandalism losses ($) per park acre (Rec. & Parks Dept.)	High school drop out rates (EDUCATION)	Elem. school enrollment/Population (EDUCATION, POPEST) Single dwelling unit/Total dwelling units (ASSESSOR, POPEST) % Owner-occupied single dwelling units (ASSESSOR)	% White collar (males 25+ yrs. with less than 4 yrs. of college)
Societal							
Traffic arrests per street mile (LAPD) % Using public transportation % households with 2 or more autos (LARTS 67) % Households with 0 autos (LARTS67)	Juvenile dependencies/Elementary school enrollment (LAPD, 1970 Census)	Arsons + Suspected arsons per 100 people (FIREINCD, POPEST)	Suicides/100,000 population (HEALTH, POPEST)	% of population age 20-59 % of population age 0-19 % of population age 60+ years (1970 Census)	Largest ethnic percentage (EDUCATION) % Nonwhite enrollment (EDUCATION)	Elementary school transiency rates (EDUCATION) Elementary school transiency out rates Elementary school transiency in rates	% Households with working wives (LARTS67)
Political							
Deficient Select System/Total Select System Streets (Bureau of Public Works) Public transit miles/Total street miles (SCRTD, LAPD)	Total arrests/100 population (LAPD, POPEST)	Fire engine companies/10,000 population (FIREINCD, POPEST) Fire engine companies/$1,000,000 Improvement value (FIREINCD, ASSESSOR) Fire engine companies/ 1000 Improved Parcels	% of public hospital care (California Hospital Discharge Study)	5 yrs. proposed parks and recreation Capital expenditures per community acre (Rec. & Parks Dept., POPEST) Proposed expenditures to community population	Voter participation rates (City Clerk, LARTS, POPEST)	% Nonconforming uses on residential parcels (ASSESSOR) % Subsidized & public housing per 1000 occupied dwelling units Use 02-03 where zoning is higher or lower Residential zoned parcels with non-residential uses Parcels with use 01-04 which are zoned higher or lower Single family parcels with higher zoning to total parcels	% of 4-5 yrs. children on welfare (DPSS) Income foregone due to ethnicity (LARTS67)
Economic							
% Bypassed employment due to lack of transportation (LARTS67) Motor vehicle accidents per street mile	Burglary and robbery losses ($) per 100 population (LAPD, POPEST) Police costs per capita	Number of fires with greater than $1000 loss per 100 structures (FIREINCD, ASSESSOR)	Death of adults aged 25-44 to total people aged 25-44 (HEALTH, 1970 Census)	Private recreational investment per 100 population (Rec. & Parks, POPEST)	% of people 25+ yrs. who have completed college (LARTS67)	Median imputed rent (1970 Census) % dwelling units with 1.51 or more persons per room Median imputed rent/median household income (1970 Census, LARTS67)	Unemployment rate (LARTS67)
Performance							
Mean work trip time by private transportation modes (miles/hr.) (LARTS67)	Part I felonies committed per 100 population (LAPD, POPEST)	Structural fires per 100 structures (FIREINCD, ASSESSOR)	Infant mortality rates (L.A. County HD)	Number of types of facilities (Rec. & Parks, ASSESSOR) Park acres per 1000 population (Rec. & Parks Dept., POPEST)	Median 6th grade reading achievement scores (L.A. city schools)	% Sound housing 1970 (Current Housing Study) % Dwelling units lacking one or more facilities (1970 Census)	Median household income (LARTS67) Median per capita income (LARTS67)

LARTS67 Los Angeles Rapid Transit Study, 1967. SCRTD Southern California Rapid Transit District. LAPD Los Angeles Police Department. FIREINCD Fire Incidence File, Los Angeles Fire Department. POPEST Population Estimate, Los Angeles City Planning Department. HEALTH Los Angeles Health Department. EDUCATION Enrollment File, Los Angeles City Schools. ASSESSOR Los Angeles County Assessor's Office. DPSS Los Angeles County Public Social Services.

Mapping and arraying of the 72 variables indicated the need for additional items such as the following list:

1. 1960 Sound Housing Units
2. Environmental Quality
3. Housing Units 20 Years or Older, 1970
4. Percent Negro Population
5. Percent Spanish Surname
6. Park Site Deficiencies per Statistical Area
7. Population Density per Park Site
8. Park Acreage Deficiency
9. Population Density per Square Mile
10. Square Miles per Park Site
11. Percent Households in Poverty
12. Estimated (1970) Household Income
13. Estimated (1970) per Capita Income
14. Labor Force Participation Rates
15. 1970 Estimated Unemployment Rate
16. Percent Families Receiving Welfare Payments
17. Percent Families Receiving Aid to the Disabled
18. Percent Families Receiving Medical Assistance
19. Percent Families Receiving Old Age and Survivors
20. Percent Families Receiving AFDC (Aid to Families with Dependent Children)
21. Percent Families Receiving Aid to the Blind
22. Time series for the 72 original variables where available.

The statistical techniques used in the Urban Matrix require that the data used consist of normally distributed random variables. On observing the frequency distributions of the matrix variables, it may be seen that many of the distributions are skewed from the normal. One can correct for skewness by using some function of the observed variable X such that f(X) is a normally distributed random variable. Thirty-three of the 72 original variables required such a transformation prior to their use in the analysis.

Community Ambiance

The various communities comprising the City were rated for their ambiance, or general subjective attractiveness, and their viability in objective terms of housing and neighborhood, fire protection, law enforcement, recreation, accessibility, education, income production, and public health. According to each standard, the communities were categorized as either blighted, endangered, or safe.

The 65 City Statistical Areas were rated by a poll in order to assess their ambiance or degree of general atttractiveness. This qualitatively determined variable was developed to serve the purposes of blight definition by both function and location, and the purposes of synthesis of the plethora of blight indicators into a comprehensive "overall blight indicator," and of providing a means to quantify both the primary and secondary effects of remedial action. The respondents rated each community of which they had knowledge on a scale of 0 to 9, provided certain demographic data on themselves, and after performing the ratings expressed their principal areas of concern.

The respondent was asked to rate each factor as to how important it was to him in rating a community. A score of 5 meant that the factor was very important to the respondent, and a score of 0 meant it has no importance. Following are the 17 qualities in the overall rank order that they scored:

Factor	Mean	Standard Deviation	Coefficient of Variation
Crime	4.56	.870	0.19
Neighborhood Cleanliness	4.53	.774	0.17
Housing	4.45	.843	0.19
Schools	4.26	1.021	0.24
Neighborhood Stability	4.19	1.035	0.25
Smog	4.14	1.150	0.28
Fire	3.98	1.141	0.29
Noise	3.86	1.137	0.29
Congestion	3.84	1.085	0.28
Shopping	3.81	1.037	0.27
Employment	3.76	1.230	0.33
Health Care	3.63	1.287	0.35
Scenery	3.53	1.157	0.33
Transportation	3.49	1.279	0.37
Recreation	3.46	1.159	0.33
Climate	3.46	1.242	0.36
Neighborhood Age	3.21	1.230	0.38

Note that as the mean score decreases, the standard deviation tends to increase. This could suggest that on the "less important" factors, there is a larger range of opinions or that the factors become more ambiguous to the respondents. It is recognized that more precise definitions of the factors should be investigated. However, even this level of definition provides some information on how people believe they react to a neighborhood.

Threat Integration

Various statistical techniques were employed to bridge the gap between the analysis of the qualitative "gut" reactions elicited from the residents and the succeeding eight analyses and reductions of the approximately 100 operationally generated variables on the delivery systems. As such, there are no hard and fast techniques that must be employed. Rather, a variety of techniques, such as, factor, cluster, discriminant, and regression analysis, are available to perform the twin tasks of (1) identifying factors and determining the dependence of one factor upon another and (2) discriminating and evaluating the differences between the communities. The objective of the analysis was examination of the structure of the data for purposes of:

a. Identification of logical groupings of the variables (factors).
b. Simplification of the model by reducing the number of variables.
c. Study of the relationship between variables which measure characteristics of the community and the results of the survey which rated communities with regard to desirability

The factor analysis generated five significant factors, all of which are discussed here. The largest and most cohesive factor was "affluence."

Factor 1: Affluence

The variables most highly correlated with the first factor, affluence (or actually lack of affluence since the signs are all reversed), were:

Variables	Factor Loading	Original Dimension Assignment
Mean Sixth Grade Reading Score	−.952	Performance Education
Median Rent	−.923	Economic Housing
Percent Minority Enrollment	+.900	Societal Education
Percent Households w/1.51 or More Persons per Room	+.897	Performance Housing
Percent Income Foregone	+.881	Political Income Production
Percent White Collar Workers Among Males Not Completed College	−.854	Attitudinal Income Production
Percent Sound Housing	−.838	Performance Housing
Median per Capita Income	−.831	Performance Income Production
Median Household Income	−.820	Performance Income Production

Whereas neither the dimensions nor the variables are discrete, there is almost no overlap of the factors. The variables that are highly correlated with a given factor are rarely highly correlated with another factor. It is interesting to note which variables are included in the large affluence factor; i.e., both at the 0.8 plus loadings (above), the variables that are highly correlated with the factor, and a second set with a cutoff of 0.7 that is only moderately strongly correlated with it. In the former set, only income production, housing, and education are represented. When both sets are considered, only the above three categories and accessibility have more than one variable each represented. This is confirmation of how highly related are the functional areas of housing, income, and education although not necessarily to each other.

The second set—the less highly correlated variables with the affluence factor—contain at least one variable from each of the other categories. However, the frequency of the representation and the degrees of correlation of the other functional areas would indicate that they are not so totally tied to "affluence." Furthermore, it is worthwhile to note that the variable most highly related to the affluence factor is reading achievement scores.

The nine variables, most highly correlated with the affluence factor, clustered around the initial assignment of variables to the performance dimension. Five of these nine were initially assigned to the performance dimension while the remaining four were initially assigned, one each, to the cultural-attitudinal, societal, political, and economic dimensions. Thus, there appears to be some similarity between this generated factor, affluence, and the economic and performance dimensions initially assigned.

Factor 2, was named "Urban Stresses and Urban Services" due to the presence of transit miles per street mile which in Los Angeles is concentrated in the Central City areas and the variables of juvenile delinquencies and total arrests. The first variable was initially assigned to the political dimension of accessibility and the second two to the attitudinal and political dimensions of law enforcement, respectively. Again, as in the case of the first factor, there is some similarity between this factor and the initially assigned political dimension as it was defined (political organization, pressure, and response). However, it contains only the law enforcement and accessibility categories and, hence, is less complete in terms of the PPBS categories as a political dimension than the "affluence" (economic/performance) factor.

If the threshold factor loadings are reduced to 0.7, suicide rates and the incidence of inoculable diseases among children also enter the factor. This expands its coverage of functional categories by including two public health variables along with the two law enforcement and one accessibility variables. From the standpoint of horizontal dimensions, it adds one initially assigned attitudinal and one societal variable to the previous initially assigned two political and one attitudinal variables. The lowering of the threshold produces a factor that is even more characteristic of urban stresses and services and not much more mixed in terms of the initial dimension assignments with its political-attitudinal combination.

Factor 2: Urban Stresses and Urban Services

Variables	Factor Loading	Original Dimension Assignment
Total Arrests/Population	+.916	Political Law Enforcement
Transit Miles per Street Mile	+.896	Political Accessibility
Juvenile Delinquents/ Elementary Enrollment	+.824	Attitudinal Law Enforcement
Childhood Illnesses	+.745	Attitudinal Health
Suicides	+.707	Societal Health

The third factor to be generated by the factor analysis was named "resident age." The variables included in it are elementary school enrollment, percent population 0-19, and percent population 60 plus years of age. It is particularly significant that both the very young and the elderly residents (variables) were included in the factor—with opposing signs—indicating the importance of resident age in describing the communities, regardless of whether that age refers to a high frequency of children or senior citizens.

The original assignment of these variables to functional categories of the urban matrix of housing and recreation was particularly weak, for they were just as applicable to health, accessibility, income production and economic satisfaction, education, law enforcement, and fire protection. The horizontal assignments, especially the societal dimension, stood the test of the factor analysis better than did the vertical category assignments. The uniformity of variables included in the factor references the societal dimension previously assigned.

Other factors that had less cohesiveness were "Neighborhood Instability" and "Police Resources Allocation." The analysis did not result in a discrete, inclusive law enforcement delivery system factor, but the police resources allocation factor was the closest approximation to one. The factor included burglary and robbery losses, police costs per capita, and the rate of auto accidents; however, it did not include Part One Person Felonies, adult arrests, juvenile delinquencies, etc. This factor, however, provided some minimal reinforcement of the initial dimensional assignments in the urban matrix. The nature of the factor and the variables included in it take on a politico-economic tone although again crossing dimensional lines as originally stated.

The Neighborhood Instability factor, which is composed of transfers into and transfers out of elementary schools, provides some highly significant revelations about community dynamics. This factor indicates that the City does not divide into growing and declining communities so much as it does into stable and unstable communities. In other words, those areas that experience high rates of transfers in are also likely to experience high rates of transfers out, hence, instability rather than growth.

As is the case of "resident age," the initial matrix assignment suffered from these variables not being discrete at least in terms of their vertical assignments to functional categories. Although the two variables were assigned to the housing and neighborhood environment sector, they could have been assigned appropriately to education or to other sectors.

The factor analysis did provide valuable insights to the urban system although not as conclusive as anticipated.

1. All of the variables that are most highly correlated with "affluence" are from the categories of education and housing as well as income production.

2. The factor analysis also demonstrates the relatively moderate or poor correlation of variables that characterize a breakdown of law enforcement with those of the "affluence" factor.

3. The "resident age" factor, containing variables measuring the relative frequency of the young and the old but entering the factor with different signs, demonstrates a pattern of segregation by age among the communities of the City.

4. The factor, "urban stresses," shows the association of public health problems and the degree of law enforcement activities as reflected in arrests in the Central City areas, but it failed to demonstrate the assumed locational association between the incidence of crime and the residence of the criminal (if arrests indicate guilt and if arrests occur most frequently in the community of residence). This, of course, does not support the contention that crimes are committed close to home.

5. The "police resources allocation" factor would suggest that the LAPD allocates its costs, and by inference its resources, primarily on the basis of the dollar values of burglary and robbery losses and secondarily on the incidence of auto accidents. Neither Part One Person Felonies (violent crimes) nor arrests or juvenile delinquencies show any significant correlation with the "police resources allocation" factor.

6. The two "neighborhood instability" variables, elementary school transfers in and transfers out both correlated at 0.7 or higher and with the same sign, formed a factor that indicates to a large extent a syndrome of community instability. However, this conclusion should be qualified by stating that neighborhood instability explains 50 to 70 percent of the variation in the two variables; up to 30 percent of the variation in transfers in could be attributed to growing communities and up to 50 percent of the variation in transfers out could be attributed to declining communities.

The factor values for each community (statistical areas of Los Angeles city) were calculated. The communities and their factor scores were arrayed in rank order and were mapped for further intrepretation and study. The clusters, which are built from a distance matrix, are based on sets of variables with the smallest distances between variables. This means that the clusters will contain only variables which are positively correlated. The clusters yielded from the analysis of the 93 SUM variables break into four large groups:

1. Affluence clusters
2. Resource Allocation clusters
3. Urban Problems clusters
4. Variables that do not cluster or form trivial clusters, such as Elementary School Enrollment clustering with percent of the population aged 0-19 years.

Some of the variables that do not cluster would have if the distance calculations used the absolute values of the correlation coefficients. Unfortunately, that option is not available in the BIOMED cluster analysis program.[1] An example of this deficiency is the variable, "Rent/Income," which does not go into any cluster but does have a high (but negative) correlation with median income.

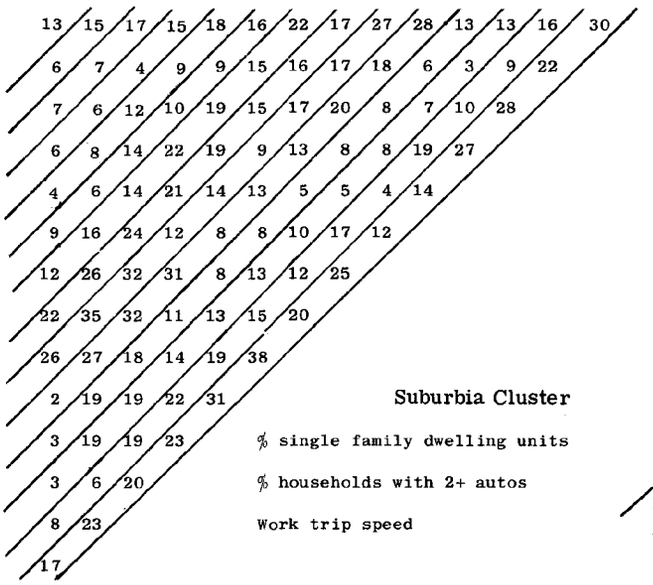
Two cluster trees make up the group of affluence clusters. One might be called "community ambiance," and the second affluence cluster describes suburban characteristics. The "Urban Problems" clusters consisted of an "Alienation" cluster, an "Urban Stresses" cluster, a "Ghetto" cluster, a "Premature Deaths" cluster, a "Poverty" cluster, a "Neighborhood Instability" cluster, a "Non-Conforming Land Uses" cluster, an "Aging and Obsolescence" cluster, and, finally, a cluster entitled, "Economic Effects of Immobility." Five "Resource Allocation" clusters were formed from the 94 variables. These were an "Allocation of Public Housing" cluster, a "Fire Engine Company Allocation" cluster, a "Park Allocation" cluster, a "Private Recreational Investment Allocation" cluster, and a "Law Enforcement Allocation" cluster.

[1] **A statistical package program prepared by UCLA School of Public Health for statistical analysis in biological and medical series.**

Affluence Clusters

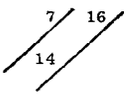
Community Ambiance Cluster

% completed college
6th grade reading score
% white collar employ.
1966 per capita income
Rent
1966 household income
1970 percent sound housing
% Owner-occupied single family dwelling units
% Voters
1970 per capita income
1970 household income
Environmental quality
Ambiance
1960 sound housing



Suburbia Cluster

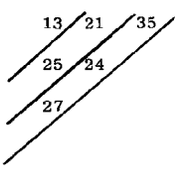
% single family dwelling units
% households with 2+ autos
Work trip speed



Resource Allocation Clusters

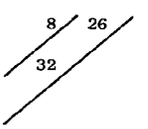
Law Enforcement Resource Allocation Cluster

Police costs
Burglary and robbery losses
Auto accidents per street mile



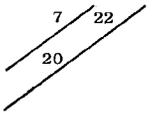
Fire Protection Resource Allocation Cluster

Structures
Population
Improvement value



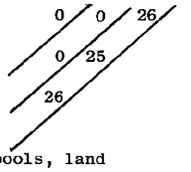
Park Allocation Cluster

Park acreage deficiency
Park acres per population
Park expenditures per population



Private Recreational Investment Allocation Cluster

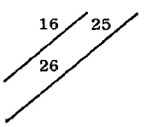
Priv. rec. assessed valuation*
Priv. rec. assessed valuation**
Priv. rec. assessed valuation***
% pop. aged 20-59 years



*Includes improvements, swimming pools, land other than golf courses.
**Includes above plus golf courses.
***Excludes all land.

Public Housing Allocation Cluster

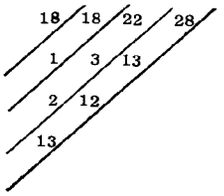
1970 unemployment
% households in poverty
% Public Housing



Urban Problems Clusters

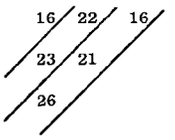
Urban Stresses Cluster

Suicides
Transit miles
Total arrests
Juvenile delinquents
Childhood diseases



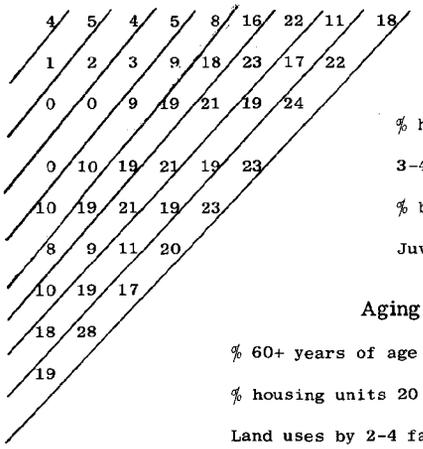
Ghetto Cluster

Juvenile probations
% Negro population
Population density per square mile
Mortality of persons 25-44 years



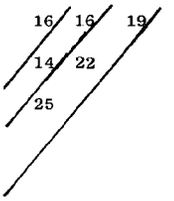
Poverty Cluster

% families receiving welfare
% families receiving Aid to Disabled
% families receiving Aid to Blind
% families receiving old age and survivors' benefits
% families on medical assistance
% use of public transportation
1971 unemployment rate
% housing lacking facilities
% families receiving Aid to Families with Dependent Children
1967 unemployment rate



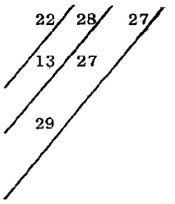
Economic Effects of Immobility Cluster

% households with "0" autos
3-4 years old on welfare
% bypass employment
Juvenile dependents



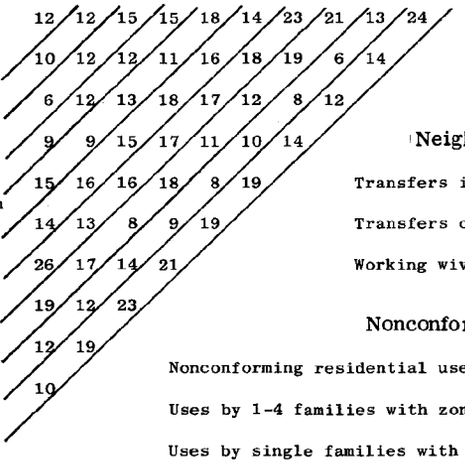
Aging and Obsolescence Cluster

% 60+ years of age
% housing units 20 years and older
Land uses by 2-4 families with zoning variance
Auto accidents per street mile



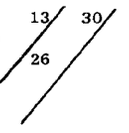
Alienation Cluster

Structural fires
% using public hospitals
Malicious false alarms
Minority enrollment
Felonies
% households with 1.51+ persons per room
Traffic arrests
High school drop-out rate
Income foregone
% households in poverty
% housing unit 20 years and older



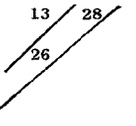
Premature Deaths Cluster

Mortality 25-44
Fire losses > $1,000
Neo-natal mortality



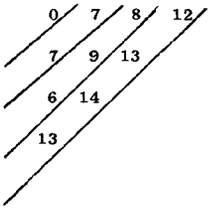
Neighborhood Instability Cluster

Transfers into elementary school
Transfers out of elementary school
Working wives



Nonconforming Land Use Cluster

Nonconforming residential use
Uses by 1-4 families with zoning variance
Uses by single families with zoning variance
Uses by 2-4 families with zoning variance
Housing 20 years and older



For the discriminant analysis, the communities were divided into four groups according to their ambiance ratings. Since a score of 5 was defined in the survey as an average community, any community with a score of less than 5 was considered below average and, thus, either "endangered" or "blighted." If a community had a score of 5 or above, it was above average and, therefore, "good" or "excellent."

The division points of ambiance rating for the four groups of communities were as follows:

Group Name	Ambiance Ratings
Blighted	0 - 3.9
Endangered	4.0 - 4.9
Good	5.0 - 6.5
Excellent	6.6 - 10.0

Using all variables, the program chose 10 variables that best discriminated the communities. Using these variables, the discriminant function depicted only three communities as misclassified ("misclassified" meaning the discriminant function classified the communities in a different category than their original classification by their ambiance rating). It was felt that eliminating service-oriented variables from the discriminant function would perhaps give a set of variables whose meaning was more apparent in relating conditions of blight which distinguish communities. This assumed that service variables are more symptomatic than causative. Also eliminated were variables that only described ethnic composition. Using the remaining variables, a new discriminant function was derived which misclassified three communities.

None of the three communities misclassified by the first discriminant equation were misclassified by the second equation. In the second equation, however, 10 different communities were misclassified. The two equations have only one variable in common, that being Sixth Grade reading scores. This variable was probably the most important single variable in both equations for classifying the communities because of its explanation of more of the variation between communities than any other variable. In the first equation, the variables after reading scores relate to categories, such as:

1. Segregation (percent largest ethnic group)
2. Income (percent hospital discharges which are from public hospitals and percent who have bypassed employment due to lack of transportation)
3. Family size (percent population aged 0-19 years)
4. Urbanization (percent trips using public transportation)
5. Incompatible land uses (parcels in use 02-04 whose zoning is higher or lower)

Other variables are included that are harder to classify, such as, percent of wives who work, value of private recreational investment, and modal preference. Several variables in this equation indicate presence or absence of services or use of services rather than problems or lack of problems which differentiate communities. The equation using the service variables gave much better discrimination than the second equation in which they were deleted. Evidently, then, the service variables give a better description of the communities as far as overall desirability, even though the two equations seem to be approximating much the same functional categories. This is probably partly due to the fact that at least some of the service variables differentiate life in the Central City and the suburbs better than the nonservice variables.

It is interesting that income as a variable is in neither equation directly but only indirectly. This points up the fact that the income of the residents of a community is not the decisive factor in the desirability of a community. Several variables, however, in the equations did come from the "affluence" factor determined in the factor analysis. It would seem, then, that it is the amenities commonly, but not necessarily, associated with higher income that are important in a community. It appears that the communities which were misclassified are borderline communities in many respects and that a high (or low) ranking on one or two variables can shift the community into a new classification.

Since the discriminant analysis program that was used does not give the actual discriminant function in a form that determines the weights of the variables in the equation, it is not altogether clear which variable is most responsible for a community being reclassified. Only through resort to other analytical techniques can the contributing conditions be identified. After interpretation of the results of each discriminant analysis, an assessment is made of each borderline community's problems and problems and probable trends.

In order to determine the relative effectiveness of alternative remedial program strategies to upgrade the quality of life in a community, it is vital at least to know the relative effectiveness of the various urban functions if not the specific programs which would be most effective. This determination of the sensitivity of the overall urban system to deficiencies in each functional area is analogous to the economist's measures of elasticity. In other words, it seeks to weigh the primary and secondary impacts that would derive from programmed changes in alternative functional areas so as to indicate in orders of magnitude the relative effectiveness of remedial action.

The standard against which a category was measured for its contribution to the vitality of a community was the ambiance ratings. Each category was quantified in a community by its composite score as determined in previous analysis. The composite scores were obtained by summing the ranks of a community by the variables in the category where the ranks were weighted by the correlation of that variable with ambiance.

Using the composite score as the observation of each functional category in a community, a regression equation was generated with ambiance as the dependent variable and the categories as the independent variables. From the regression coefficients of each of the categories in the equation can be obtained a standardized regression coefficient. The standardized regression coefficient can be used to rank the categories as to their ability to explain ambiance rating in a community. The rank order of the categories found from this study and their relative weights in the equation are listed below:

Rank	Functional Category	Relative Weight
1	Education	1.000
2	Housing and Neighborhood	.442
3	Income Production	.438
4	Health	.224
5	Law Enforcement	.197
6	Accessibility	.196
7	Recreation	.144
8	Fire Protection	.002

Education has over twice the weight of any other category. This is not surprising since the preceding analyses have shown that the education variables correlate highly with ambiance; that they and environmental quality are good discriminates between communities; and that they are associated with many indicators of affluence and well being. Education, housing, and income production are, in fact, highly associated as can be seen in the factor analysis and again in this study of blight elasticities. In this regression equation, the three above categories explain 88 percent of the variability in ambiance ratings of the communities. The other five categories, together, only explain an **additional** 1 percent of the variance in ambiance. It is interesting to note that law enforcement is so low on the scale of influences, particularly when the group surveyed to obtain the ambiance ratings indicated that neighborhood crime was a very important aspect in their ratings. It may be that there is a threshold of crime in a neighborhood below which the residents are complacent about the problem and above which they find the situation intolerable.

Limitations

Technical

Four major technical deficiencies characterize the SUM process as applied herein. As expected, they derive from and impair the urban system description. These problems include the lack of discrete classification structures for either delivery systems or the social dimensions that describe them, inadequate attitudinal data, occasionally heterogeneous geographic communities, and difficulties with the overall measure of system effectiveness, community ambiance, both as to survey structure and cost and, logically, the scalar nature of the ratings.

Unfortunately, aggregation for statistical reporting and analysis can obscure important differences. The definition of the standard statistical areas as communities fails in some cases; for, within each community, characteristics may vary considerably. The concept of community indices demands homogeneity, a relative quality not always available even at the level of the census tract. Naturally, the larger the geographic unit, the more likely it is to have "pocket" conditions contained within it.

The Statistical Areas used in this report were tested for the degree of dispersion among the block groups (statistical areas average about 50 actual block groups) comprising them for the following characteristics from the 1970 census:

Overcrowded housing (percent units with more than 1.5 persons per room)

Housing condition (percent units lacking plumbing facilities)

Structural type (percent single-family units)

Occupancy type (owner occupied units/single family units)

Age of population (percents 0-19, 29-59, 60 plus)

Housing value (percent rental units under $100 monthly contract rent and percent owner occupied units under $20,000 estimated market value)

Each of these series were examined for their means, standard deviations, and ranges in order to determine if they were relatively homogenous and, if not, if their heterogeneity was due more to "pocket" conditions or to differences across large areas. Only one Statistical Area was considered heterogeneous in four or more of the nine series; and two Areas, in three of the series.

It had been expected that the factor analysis would have yielded more large factors that corresponded at least to the functional categories, e.g., housing, law enforcement, recreation, etc.—and, hopefully, factors that more clearly reflected the behavioral dimensions, e.g., cultural-attitudinal, societal, political, etc. However, limitations in the number of observations, i.e., 65 communities—and the likelihood that those observations were not all relatively homogenous aggregations, undoubtedly, contributed to this deficiency. However, the direction of the findings does indicate that these traditional categories are not discrete and throws open to question the use of urban PPBS categories that so characterize the last 5 years work in the field of urban and regional science.

With the development of the Digital Character Data Conversion System (a computerized system to convert through linear interpolation nonuniform polygon data to point data and, then, reaggregate those points to uniform polygons) being quite imminent, succeeding SUM analyses could be conducted on census tracts which would increase the number of observations more than eleven-fold. Furthermore, the census tracts or aggregations of them to identifiable neighborhoods would be more homogenous units than are the larger Statistical Areas; so, a factor analysis based on the 741 census tracts should yield more details of the interrelationships of the variables.

Although the ambiance ratings are continuous or scalar, rather than discrete or categorical, class limits have been assigned. This, of course, raises certain logical objections. However, although the establishment of certain ranges as discrete classes is arbitrary, it is somewhat mitigated by the imprecise nature of the problem. However, it is recognized that the sample selection desperately needs improvement—both as to size and structure. Efforts are currently underway to "piggyback" the Ambiance Survey onto media marketing surveys.

Political

Although the technical difficulties can be resolved without too much difficulty, there are some far more serious problems that limit the effectiveness of the approach. The generally low level of statistical competence in local government dictates a small audience for full documentation of the Urban Matrix project. The first volume alone contains about 800 pages which, regardless of content, will discourage widespread use in City government. Any substantial reduction in volume would first affect the documentation of the data and their analysis and would degrade the credibility of the findings. The approach to the resolution of this problem that was selected was to produce an analytically sound document for the technical staff of the City Administrative Office and for selected staff of the City Council, Mayor's Office, Planning Department, Redevelopment Agency, etc.

However, experience demonstrated that even with more astute management science types, it was not sufficient merely to hand out the document. Seminars on the document findings and methodology have been and will continue to be conducted in order to guide the reader in his exploration and use of the material. We have found this approach particularly effective in communicating the material to the press.

Since the seminar approach is available for only a relatively limited audience—they typically consume 3 to 4 hours—a printing of highly abridged summaries of the community threats,

requirements, and programs—i.e., two of the three volumes of the State of the City Series—and allied work also performed by the Bureau under other projects will be more widely distributed. However, it must be recognized that the summary will almost inevitably oversimplify and under-document.

It is also quite clear that the main report, "Conditions of Blight and Obsolescence," is of relatively little value to line operating departments. At best, perhpas it serves as a compendium of statistics; but, due to its holistic character, it is probably too broad-brush for its purposes. Furthermore, a principal goal of the report and its methodology is to provide a means to prioritize functions for purposes of allocating resources. Hence, the department managers responsible for those functions which are assigned low priorities will be disinclined to use the findings, and those whose functions are assigned high priorities will tend to abuse the material.

Conclusions

The Scientific Urban Matrix, as it is applied here, basically simulates the research portion of a community renewal planning program that begins with problem identification, continues through its definition and analysis, and culminates in recommendations for remedial action. This paper details the problem identification, definition, and analysis. In recognition of the difficulty of defining the problem, urban blight, a result of its being a composite of many highly complex and interrelated phenomena, the standards that were explored included qualitative composite standards, objective functional standards, and composite statistical standards. It was felt that the integration and synthesis of these into an overall blight standard incorporated the most meaningful and realistic aspects of the parts and produced an effective yardstick by which to identify and measure the extent of urban blight.

Intrinsic to this Scientific Urban Matrix approach is the hypothesis that, even within a discrete functional area, blight is probably multifaceted. The statistical analysis confirmed this hypothesis by demonstrating close relationships between various functional measures of performance and socio-attitudinal, socio-political, and socio-economic manifestations of those functions. Consequently, the cause and effect linkages between the underlying factors and their manifestations in functional performance become far more apparent. This application demonstrates the primacy of social, political, and economic forces, values, attitudes, and goals in the determination of functional performance, along with the futility of proposed remedial actions that avoid these dimensions.

The Scientific Urban Matrix approach, as performed here, admits to definite gaps:

1. The analysis of the data at gross geographical aggregations that are not sufficiently homogenous.
2. Variation in the time period covered by the data.
3. Subjective forcing of data collected on incompatible geographic limits into statistical area aggregations.
4. Uncontrollable inaccuracies in the data collected by operating departments.

Many of these deficiencies will be corrected in succeeding iterations. Yet, in support of the project's ultimate goal of improving the City's environment by optimizing ameliorative programs, this second cycle of the Scientific Urban Matrix Project fulfilled the objectives enumerated in its study design:

1. It synthesized the threats of the functional categories that were studied.
2. By having coordinated qualitatively derived community status indices with conditions and service variables measured by data collected routinely in the City, it expanded the scope of the urban blight phenomenon that can be comprehended and controlled.
3. It provided a firm technical base upon which program recommendations can be founded.
4. It evaluated the relevance of a portion of the Data Base by measuring the sensitivity of the composite phenomenon of urban blight to variations in the state descriptive data contained in and reported from the Data Base.

Hence, the second cycle of this project laid the foundation for the Community Analysis Bureau to meet its goals by developing the fundamental capabilities necessary to preplan evaluation of alternative remedial action programs, based on the significant status and trends in areas of the City, and to post-implementation monitoring of the effects of selected action programs for their effectiveness in improving the function of the urban system.

INDICATORS AND STATISTICS: ISSUES IN THE GENERATION AND THE USE OF INDICATORS*

Harvey A. Garn and Michael J. Flax,
The Urban Institute, Washington, D.C.

This paper is about a special class of statistics—social indicators—and the experimental program in developing indicators, particularly urban indicators, at The Urban Institute. We will discuss:

(1) some definitional issues in the indicator field,

(2) some critical conceptual problems in developing indicators, and

(3) some of our past and current work at the Institute.

Defining Indicators

This is not the place to engage in an extensive discussion of what indicators are. Some comments, however, are in order. As is characteristic of many relatively new and growing fields, there is not a uniform or completely acceptable definition to all the people doing indicators.[1] Most simply and broadly, **indicators are measurements of aspects of life and social conditions related (or believed to be related) to human well-being and satisfaction.** The definitional problems revolve around alternative ways of formulating appropriate reference points for the indicators—means of attaching significance to the measures. Some have argued, e.g., Mancur Olson, that indicators derive their significance from being direct measures of social welfare.[2] That is, they are measures of variables which appear in a social welfare function. Insofar as one can derive such measures, the argument is that the movement of the variables in the appropriate direction represents an unambiguous improvement in social welfare, other things remaining equal. This definition leaves open the questions of how one is to tell what variables should be in the social welfare function, how one is to aggregate differing (and often conflicting) individual welfare functions in determining the social welfare functions, and what to say in cases where other things do not remain equal.

For these and other reasons, many have found such a definition too restrictive. Some of these, e.g., Sheldon and Freeman, have argued that what is important in attaching significance to indicators is the performance of measures of social conditions over time, both to monitor change and to develop reasonable expectations about the future movements of the measures.[3] This definition leaves open the question of what variables should be measured.

Others, e.g., Kenneth Land, see the means of attaching significance to indicators to be their appearance in a social system model in which any individual indicator can be related through the model to other variables in the model.[4] This definition, of course, leaves open the question of how extensive the model must be to qualify, how to choose the appropriate model outputs, and how to assess (in value terms) alternative arrays of model outputs.

Each of these definitions calls attention to important (and possibly unresolvable) issues in the generation of indicators. They are mentioned in the context of this paper for two reasons:

(1) They explicitly raise some of the problems which permeate indicator work but are frequently left implicit, and

(2) they have implications for our work at The Urban Institute and our approach.

[1] It is interesting to report, in this connection, some of the history of the definitions of statistics. Yule and Kendall point out in their **Introduction to the Theory of Statistics** that, "The earliest occurrence of the word 'statistics' yet noted is in **The Elements of Universal Erudition,** by Baron J. F. Von Biefield. ...One of its chapters is entitled **Statistics** and contains a definition of the subject as 'The Science that teaches us what is the political arrangement of all the modern states of the known world.' " E. A. W. Zimmerman said in 1787, "It is about 40 years ago that the branch of political knowledge (statistics) which has for its object the actual and relative power of the several modern states, the power arising from their natural advantages, the industry and civilization of their inhabitants, and the wisdom of their governments has been formed...into a separate science." At this time statistics was largely a verbal, rather than numerical, subject. It evolved towards increasing emphasis on numbers and mathematical concerns and away from the emphasis on "what was noteworthy about the state."

Indicators return to the earlier conception of statistics in the sense that they are about matters of social concern and the actual and relative status of population groups.

*This research was supported by funds from the National Science Foundation under Grant No. GS-29710X1 to The Urban Institute. Opinions expressed are those of the authors and do not necessarily represent the views of The Urban Institute. We wish to thank Jeremy B. Taylor and Herrington J. Bryce for their assistance and many helpful comments.

[2] Therefore,..."statistics on the number of doctors or policemen would not be social indicators, whereas figures on health and crime rates could be." U.S. Department of Health, Education, and Welfare, **Toward a Social Report,** Washington, D.C.: U.S. Government Printing Office, January 11, 1969, p. 97.

[3] E.B. Sheldon and H.E. Freeman, "Notes on Social Indicators: Promises and Potential," Policy Sciences, Vol. 1, 1970, pp. 97-111.

[4] Kenneth E. Land. "On the Definition of Social Indicators," The American Sociologist, November 1971, p. 323.

Some Conceptual Issues

Many of the major problems raised by these alternative definitions and attempts to carry them through in the generation of indicators are conceptual and value-related rather than being primarily statistical or resolvable by statistical techniques, however sophisticated. These problems for those generating indicators include (to summarize some of the implications of the above definitions):

(1) Consideration of just how the measurements to be made are related to human well-being and satisfaction, e.g., how can we identify principal components of the "quality of life?"

(2) A determination of what measures to take and what units of analysis to use, e.g., our "problems" are often defined by what we measure. What are the best measures in each functional area? How should they be disaggregated? Should they be combined into indices?

(3) Since the above are affected by both their intended uses and users, consideration should be given to whether the primary uses of the indicators are likely to be normative, descriptive, predictive, evaluative, or designed to improve understanding of causal relationships. Since the possible uses usually cannot be controlled (or often even anticipated!), what are the probabilities of misuse? Do the anticipated benefits of the particular indicator outweigh the possible costs?[5]

(4) Since the indicators, once generated and presented, may be used for any or all of these purposes, some attempt should be made to consider what inferences and assumptions the users are likely to make. Inferences that are supported by adequate research sould be reinforced; where alternate inferences are equally valid they should be encouraged rather than hidden.

These problems are not likely to be resolved once and for all. We view the development of indicators, therefore, as an inherently iterative process rather than an activity with clear boundaries, fixed purposes, and an accepted theoretical base.

General Comments on Our Work

In our work, we have attempted to develop and use measures which are related to human well-being and satisfaction in the sense of Olson's definition, i.e., direct measures of social outputs, although we have not found it possible or desirable to do so in all cases. We have attempted to use measures which can be taken over time to facilitate monitoring of social conditions, although we have not attempted to use these time series for predictive purposes in our work to date. Finally, we have relied on knowledge in specialized fields both to guide our selection of indicators and in commenting on possible relationships among our indicators. In most of our work we have not found it possible to follow the prescription implied in the third definition—using only measures which appear in established social system models. This has not been possible because in many of the fields which are interesting (in the sense that they relate to areas of pressing social concern), there are few extant comprehensive models which can be used to isolate appropriate variables and their relationships.

Given the large number of unresolved conceptual and modeling issues and our sense that progress in the indicator field is likely to be iterative in any case, we have opted for an evolutionary, experimental program. We started with an attempt to describe important social conditions and recent changes in those conditions across metropolitan areas. The absence of a set of generally accepted value-related or objective standards and established system models has led us to emphasize comparisons as an important means of attaching significance to the indicators developed.

Clearly there is need for a wide variety of indicator research approaches. We have chosen to confront many of the theoretical and measurement problems just mentioned in a descriptive context by generating and presenting indicators and improving them incrementally. We sense a strong and present demand for indicators that can help certain interested users, e.g., government officials and the informed public, obtain a better perception of the present state of affairs as well as some quantitative estimate of recent changes. While responding to this need (and we feel that **some** of those working on indicators should respond) we have had the opportunity to confront many theoretical issues. The experimental, iterative, nature of our work permits us to improve both the conceptual adequacy of the indicators and their sharpness as descriptors of important phenomena. In short, we have chosen to develop descriptive indicators as a means of raising conceptual and methodological questions in specific contexts.

As stated above we have been concerned with the possible uses and users of indicators. How indicators acquire meaning to the user is central to this concern. This is accomplished in at least two ways:

(1) Some form of reference or comparison is needed (as mentioned above), and the indicator should be structured so that this comparison process is facilitated for the user. Comparisons can be made with governmental standards, e.g., does pollution exceed accepted safety levels; goals, e.g., is unemployment below 4 percent; comparisons over time, e.g., is the infant mortality rate higher than last year; comparisons across geographical areas, e.g., are the reading scores for New York City higher than those for San Francisco; comparisons of different population groups, e.g., is unemployment higher in the city than in the suburbs. The important point to be made here is that **some form** of comparison is valuable to the user in attaching meaning to the indicator.

(2) Indicators can be said to **stand for more than they measure.**[6] The inference structure associated with measures varies with the individual and is not subject to precise measurement or prediction. However, the entire process of extracting information from indicators depends on **the inferences users make concerning them, and these inferences may not coincide with those used by the indicator generators.** Indicators may be designed to be primarily descriptive, evaluative, or causal, but their ultimate use is often decided by their users. However,

[5] The question of who receives the benefits and who pays the costs is usually crucial here.

[6] A frown stands for more than the movements of facial muscles—different observers infer various implications concerning a frowning individual.

indicators can be designed to facilitate certain inferences, and that choice remains under the control of those who generate them.

A Summary of Our Reports

Our work so far has resulted in several reports which were, in a sense, "experiments." Each tended to expand our work along a different dimension of indicator problems. Our first study was documented in Paper 136-1, **The Quality of Life in Metropolitan Washington (D.C.): Some Statistical Benchmarks**, by Martin Jones and Michael Flax. This study had several interesting aspects and has guided the work which followed. In the first place, 136-1 reflected our interest in developing indicators over a broad range of areas of social concern. After extensive research and consultation we selected 14 areas of such concern rather than concentrating on one or two. We presented a single indicator for each of these areas. In each case an attempt was made to present a genuinely representative indicator of the area of social concern with appropriate caveats about the possibility of a single indicator providing a proxy for a whole area of social concern.

Next we sought some comparative standards to give the indicators more meaning for our audience. We found that in none of the areas was it possible to adopt either an objective or subjective standard on which most would agree. Therefore, we decided to make our initial comparisons across metropolitan areas. This was done to provide sufficient variation in the measures and because no one had recently attempted to make such inter-city comparisons over such a broad range of social concerns.

Further, comparisons were made of the recent changes that had occurred with respect to each indicator. This broad approach enabled the reader to examine both the characteristics of and recent changes in 18 different metropolitan areas over a wide range of social and economic considerations. It highlighted the fact that there are great variations in the comparative conditions that prevail in any one metropolitan area. For example, Washington ranked from second to eighteenth with respect to the different quality areas. This approach underscored the inadequate state of the available, comparative, recent social indicator data. Comparative data for cities (as opposed to metropolitan areas) were even less available.

In producing this first document we were faced with many of the issues raised earlier. How were we to relate our variables to human satisfaction and well-being? What kinds of inferences and assumptions did we assume on the part of our readers? Since we chose as our audience a wide spectrum of the population (executives, administrators, and the informed public), we tried to choose variables that were of general interest and enjoyed a wide consensus as to "which way was up." We did not attempt to forecast or guide the inferences drawn from the data by this audience. We did, however, present caveats concerning the data used in as clear a manner as possible.

We made no attempt to present an aggregative index of these indicators. We did not feel that such an index would serve any useful purpose, given the state of knowledge, except for the convenience of having a single number. At this stage of our work we feel that the search for a single social indicator or social indicator index would obscure the meaning of the indicators for most users. Eventually it may be possible, as is true of some economic indicators, to formulate composite measures that satisfy both the accuracy requirements of the theoretician and the credibility requirements of most users.

Our subsequent work has consisted of further "experiments" representing incremental improvements in some aspects of our initial paper, 136-1. These are listed below:

1. Experimenting with Levels of Aggregation;
2. Experimenting with Multiple Measures;
3. Experimenting with Racial Comparisons;
4. Experimenting with Improved Classification Systems;
5. Experimenting with Aids to Urban Decision Makers.

1. EXPERIMENTING WITH LEVELS OF AGGREGATION

It is fairly clear that metropolitan areas are not the only, nor necessarily the most interesting, population aggregates on which to make comparisons. We have, therefore, developed additional work in which the objective was to use different population aggregates. The example of this kind of indicator work which we have completed is **"A Study in Comparative Urban Indicators: Conditions in 18 Large Metropolitan Areas,"** Report 1206-4, by Michael Flax, which is an updated and expanded version of our 136-1 report. As part of this report we displayed data for five (data were available for only five) of the 14 indicators in 136-1, broken down into central city and suburban values for the same 18 metropolitan areas as in 136-1.[7] The point again, in addition to the intrinsic importance of central city/suburban relationships and their recent rates of change, was to display a different kind of geographical disaggregation.[8] We intend to experiment with other forms of aggregation for population units in our future work. These may include aggregations of census tracts within a given city, neighborhoods within a city, administrative subdivisions such as school districts or police precincts within the city, or inter-metropolitan comparisons across all of these units. A major reason for the experimental nature of this effort is that serious analytic difficulties exist in establishing criteria for homogeneity of comparison units. We hope that additional insight into these analytic problems will result from experimenting with different forms of aggregation.

2. EXPERIMENTING WITH MULTIPLE MEASURES

The next paper, "Selected Education Indicators for 21 Major Cities: Some Statistical Benchmarks," Working Paper 136-4, by the authors, builds on the work which was begun in our initial paper, 136-1, in a different way. As indicated previously, we recognize that a single indicator is an inadequate representation of any given area of social concern. Working Paper 136-4 expands both the number of indicators and the conceptual framework for considering the indicators in the area of education. The paper presents recent data for 12 different educational indicators for the central cities (21) of each of the metropolitan areas (18) examined in 136-1. The development of additional indicators in a given area of social concern requires that attention be paid to the problem of providing a conceptual framework for the indicators. In this case we utilized essentially an input-output framework and then examined the degree to which input indicators are useful in explaining the variation in output measures across the 21 cities.[9]

[7] In addition to updating the indicators (1968 to 1970), we also expanded our time frame over a 5-year period (instead of the 1-year period in 136-1), and substituted new measurement series for four of the 14 quality areas (education, transportation, narcotics addiction, and pollution). (See Exhibits 1-3 for the variables used, a sample data table and page of caveats, attached.)

[8] For an example of this form of disaggregation, see Exhibit 4 attached.

[9] For definitions of the indicators used, see Exhibit 5 attached.

The indicators themselves were divided into four categories. The first of these consisted of measures which tend to be used as output indicators, e.g., standardized reading scores, draft rejection rates, and the like. The second category of indicators contained essentially measures of resources applied in the educational process directly through the educational system, e.g., per-pupil expenditure, percentage of teachers with advanced degrees, pupils per teacher, and so forth. The third category of indicators contained measures thought of as indirect inputs to education, e.g., those related to educational climate and the socioeconomic characteristics of the populations of the cities. The fourth category contained measures of the relative willingness of people in the cities to choose educational expenditures over other possible uses of public funds.[10]

Since we chose an input-output framework for the indicators, we then tested, using two simple statistical models, the degree to which it is legitimate to use the input measures to explain variations in outputs across the cities. The variation in selected outputs cannot be explained significantly by these input indicators in these city-wide data. For example, it is not true that a lower pupil/teacher ratio tends to be associated across these cities with improved reading scores. Similarly, it is not the case that variations in high school dropout rates can be explained by the proportion of public school students who were either Negro or from low-income families. If nothing more, this work demonstrates the difficulties of using input indicators as proxies for output indicators when making comparisons.

3. EXPERIMENTING WITH RACIAL COMPARISONS

Our next paper, **"Blacks and Whites—An Experiment in Racial Indicators"** 85-136-5, by the authors, represents a different kind of adaptation of the work begun in our initial paper, 136-1. In this case, the relative comparison groups are not, in general, comparisons by geographical aggregates, although we do point out that the nonwhite population is becoming increasingly concentrated in central cities and at a more rapid rate than is true in the white population. In this paper national data for 1960 and 1968, relating to white and nonwhite performance in 16 socioeconomic areas, are presented in a way that highlights the possibility of alternate valid interpretations of the data. In fact, inspection of our charts reveals that for 10 of the 16 functional areas examined, different conclusions as to the relative progress of blacks in our society could be made.[11] By discussing implications of both the indicators and ongoing Urban Institute research in assessing the prospects for future income equality between whites and nonwhites, we provide an example of how research results can be used in explaining the implications of relatively simple indicator data.

4. EXPERIMENTING WITH IMPROVED CLASSIFICATION SYSTEMS

Our working paper, "Income and General Welfare: An Identification of the Socio-Economic Gaps Between Low and High Income Regions," by Herrington Bryce, represents an experiment in identifying indicator structures which enable policy makers to identify key socioeconomic differences among sets of metropolitan regions. In this paper, metropolitan areas were divided into two classes (high and low income regions) and then four discriminant functions emphasizing demographic, employment, housing, and education variables, respectively, were estimated using 1960 census data.[12] The discriminant functions identify linear combinations of the listed classes of variables which most effectively distinguish between high and low income regions. The variables used reflect reasonable theoretical connections among the variables, are clearly defined, and can possibly be affected by public policy.[13] An important advantage of the technique, therefore, is that it calls attention to critical differences in a wide variety of indicators, other than income, which tend to distinguish low and high income metropolitan areas from each other in addition to providing directly a basis for selecting an indicator set for comparative purposes. The technique has the additional advantage of identifying those metropolitan regions which have a typical socioeconomic structures for their income class.

5. EXPERIMENTING WITH AIDS TO URBAN DECISION MAKERS

We are presently preparing an inter- and intra-city (city network) indicator report in cooperation with six cities. We are gathering yearly intra-city data on five functional areas—poverty/welfare, health, education, public safety, and housing—in an effort to answer the following four questions:

(1) Given the limitations of available data, can we present a set of relatively few data series which will prove useful to urban decision makers?

We expect that the presentation of comparative, understandable, yearly data over a large range of concerns will provide a broader perspective and encourage a higher level of discourse concerning certain urban problems.

(2) Given the lack of tested or accepted causal models, can we devise indicators which will provide useful insights that will improve urban decision making?

We have constructed a general conceptual schema in order to classify phenomena which are often thought to have some causal relationships. We expect that discussions of this schema will aid us in devising measures which will provide useful insights for urban decision makers. We would also hope to "adjust" some of our indicators so that data from different areas might be more comparable, e.g., age-adjusted death rates would permit comparisons between areas with different age structures. In a sense this will be a form of evaluation falling somewhere between a thorough analytical study and a purely descriptive catalogue of output measures. We would try to structure and present our indicators in a form useful to the user.

(3) Given the lack of absolute standards or well-defined goals, can we present national, State, city, and intracity data in a comparable format so as to increase understanding and insights concerning urban problems?

Here we hope to locate potentially comparable data from different levels of government and attempt to present them so that users can implement these comparisons to highlight large or unexplained variations wherever they occur. Regression analysis to compute "expected" values controlling for supposedly causal variables such as age, sex, or income may also be used.

[10] For a summary of these data, see Exhibit 6 attached.

[11] See Exhibit 7 attached.

[12] The study can be replicated using 1970 census data, of course.

[13] The variables used are listed in Exhibit 8 attached.

(4) Given the conceptual and methodological difficulties in obtaining attitudinal data, can we combine objective and subjective data so as to be of aid to urban decision makers?

We hope to present existing attitudinal and objective data on the same subjects for the same geographic areas in a way that can be useful to decision makers. We believe that unexpected and possibly unexplainable variations will be brought to light by using objective and subjective techniques to describe the same phenomena. We would expect that this descriptive information might provide added insights to decision makers and provide direction, as well as data, for future research efforts.

Summary

We have discussed several different definitions of indicators. For our own work we have defined them, simply and broadly, as measurements of aspects of life and social conditions related (or believed to be related) to human well-being and satisfaction.

Some of the critical conceptual issues that confront generators of indicators are:

a) how to relate human well-being and satisfaction to measurable variables,
b) how these variables should be defined, aggregated or disaggregated,
c) what is the primary use of these indicators, what misuses can be expected, and
d) what inferences are likely and which should be facilitated by the data presented.

In our work, we are confronting many of these issues in a descriptive context. As a result of our concern about users and uses, two aspects of indicators deserve special mention in this summary:

(1) indicators stand for more than they measure; we advocate explicitly considering the inferences likely to be drawn from indicators in their design; and

(2) some form of comparison is useful in aiding the user to extract meaning from an indicator.

The completed work described, as well as that being planned, are attempts to incrementally address some of the issues raised earlier in this paper.

Some Areas for Future Indicator Experimentation

1. A "Use-Users" Study to define in a systematic way indicators that would be most useful, by different classes of users.

2. Incorporating reasonably well-demonstrated, causal effects into evaluative indicators. Adjusting and normalizing data so as to facilitate the use of comparative data in ways which account for known sources of variation.

3. The uses of comparative techniques for combining national, State, and local data in ways that will highlight large variations from expected performance.

4. The design of combined objective/subjective data gathering procedures so as to facilitate comparing objective and subjective data of the same phenomena, and the making of comparisons among similar geographic, demographic or other areas.

5. The application of indicator techniques to additional functional areas, e.g., racial differences, the criminal justice system, so as to improve understanding and aid decision makers.

6. Application of known statistical techniques to better indicator generation, presentation and use, e.g., validating data sources and leading indicators; adjusting for known or assumed causal variables; clustering or segmenting variables for more meaningful comparisons; correlation techniques for establishing, verifying, or testing relationships.

EXHIBIT 1

Quality Categories and Selected Indicators

QUALITY CATEGORIES	Latest Year Data	INDICATORS USED
Unemployment	1970	% of labor force unemployed[a]
Poverty	1970	% of households with incomes less than $3,000 per year
Income	1969	*Per capita money income adjusted for cost of living differences
Housing	1969	Cost of housing a moderate income family of four
Health	1969	Infant (under 1 year) deaths per 1,000 live births[a]
Mental Health	1969	Reported suicides per 100,000 pop.[a]
Public Order	1970	Reported robberies per 100,000 pop.[a]
Racial Equality	1970	Ratio between nonwhite and white unemployment rates
Community Concern	1970	*Per capita contributions to United Fund appeals
Citizens Participation[b]	1968	*% of voting age population that voted in recent presidential elections
REVISED INDICATORS		
Educational Attainment[c]	1969	*Median school years completed by adults
Transportation[d]	1969	Cost of transportation for a moderate income family of four
Air Quality[e]	1969	Average yearly concentrations of three air pollution components, and change in the concentration of suspended particulates
Social Disintegration[f]	1969	Estimates of number of narcotics addicts per 10,000 population

*An increase in the absolute value of these indicators is assumed to represent an improvement in the quality of life. The reverse is true of all the others.

[a]Data is also provided in central city, suburban, and city/surburban ratio levels and their rates of change. [b]This indicator did not require revision. [c]Selective Service Mental Test rejection rate was used previously. [d]Deaths from auto accidents per 100,000 population (an indicator of traffic safety) was used previously. [e]A composite index of pollutants was used previously. [f]A new method of estimating addiction rates is used.

Source: A Study in Comparative Urban Indicators: Conditions in 18 Large Metropolitan Areas, Michael J. Flax, Urban Institute Paper 1206-4, April 1972.

EXHIBIT 2 An Example of Data Caveats

Infant Mortality Rates—Explanatory Notes

Data Sources

Vital Statistics of the United States, National Center for Health Statistics, U.S. Department of Health, Education, and Welfare (U.S. Government Printing Office), Volume II, part B, table 7-4 and Volume I, table 1-53, from 1962, 1966, 1967, and 1968.

Method of Calculation

The infant mortality rates were calculated by dividing area deaths under 1 year by total area live births for the year in question. The rates of change were calculated as follows:

$$\text{Annual Average Rate of Change} = \frac{\text{Infant Mortality Rate (latest)} - \text{Infant Mortality Rate (earlier)}}{\text{Infant Mortality Rate (earlier)}} \div \text{Number of years between latest and earlier data}$$

Caveats

1. Infant mortality is a commonly accepted measure of community health. It does not explicitly measure the health of the adult population, but it does relate to maternal health, and has been found to correlate with income which is associated with many other health factors. Other indicators available from the above source include neo-natal mortality (including late-period miscarriages, still births, and deaths during the first 28 days of life), and mortality data for 60 ailments. Data are also available from an annual HEW Health Survey on the number of doctors, dentists, and hospital beds in each metropolitan area. Data on the cost of medical care in each metropolitan area are available from the BLS standard budgets, and changes in medical costs can be obtained from the Consumer Price Index.

2. The data used were reported by residence of mother rather than by community where birth occurred. The data were obtained from a 50% sample of microfilmed birth certificates.

3. These data do not include "still births" (babies born dead), late-period (6-9 month) miscarriages, and child deaths occurring later than 1 year after birth.

4. Despite a revision of the International Lists of Causes of Death, the 1967 and 1968 rates are comparable since all infant deaths were reported in each version.

Source: A Study in Comparative Urban Indicators: Conditions in 18 Large Metropolitan Areas, Michael J. Flax, Urban Institute Paper 1206-4, April 1972.

EXHIBIT 3 Indicator Presentation

Metropolitan Infant Mortality Rates*

(A Health Indicator)

Metropolitan Area	(Recent Levels)							
	1962	(Rank)[a]	1966[b]	(Rank)	1967[b]	(Rank)	1968	(Rank)
New York	25.7	(14)	23.2	(11)	22.7	(13)	21.6	(9)
L.A./Long Beach	22.9	(7)	21.0	(5)	20.0	(4)	18.9	(2-3)
Chicago	24.7	(13)	26.3	(18)	24.4	(17)	24.4	(18)
Philadelphia	26.1	(16)	23.9	(12)	24.9	(18)	23.4	(17)
Detroit	24.3	(10)	23.0	(10)	22.0	(11)	22.2	(12)
Boston	22.1	(5)	20.3	(2)	19.4	(2)	20.2	(6)
S.F./Oakland	22.3	(6)	20.5	(3)	19.3	(1)	18.9	(2-3)
WASHINGTON	28.3	(18)	24.9	(17)	20.7	(7)	19.8	(5)
Pittsburgh	23.3	(8)	21.5	(6)	21.4	(9)	21.1	(11)
St. Louis	24.6	(12)	24.1	(14)	24.1	(16)	22.8	(10)
Cleveland	24.4	(11)	20.7	(4)	21.1	(8)	21.3	(8)
Baltimore	27.3	(17)	24.3	(15)	23.2	(14)	23.0	(15)
Houston	25.7	(15)	24.6	(16)	21.9	(10)	23.1	(16)
Minn./St. Paul	20.6	(1)	19.5	(1)	19.7	(3)	18.6	(1)
Dallas	23.8	(9)	24.0	(13)	23.2	(15)	22.8	(14)
Cincinnati	21.1	(2)	22.3	(9)	20.5	(6)	20.3	(7)
Milwaukee	21.3	(3)	21.5	(8)	20.2	(5)	19.2	(4)
Buffalo	21.9	(4)	21.5	(7)	22.3	(12)	22.3	(13)
AVERAGE	23.9		22.6		21.7		21.3	
	(Recent Annual Average Changes)							
	62-67	(Rank)[c]	62-68	(Rank)	66-67[b]	(Rank)	67-68	(Rank)
New York	-2.3%	(8)	-2.7%	(3)	-1.9%	(12)	-4.9%	(6)
L.A./Long Beach	-2.5	(6)	-2.9	(2)	-5.1	(7)	-5.4	(4)
Chicago	-0.2	(17)	-0.2	(17)	-6.9	(4)	-0.3	(13)
Philadelphia	-0.9	(12)	-1.7	(8)	+4.2	(18)	-5.8	(2)
Detroit	-1.9	(9)	-1.4	(13)	-4.1	(10)	+1.1	(14)
Boston	-2.5	(7)	-1.4	(12)	-4.6	(9)	+4.3	(16)
S.F./Oakland	-2.7	(5)	-2.5	(5)	-5.6	(6)	-2.3	(8)
WASHINGTON	-5.4	(1)	-5.0	(1)	-16.7	(1)	-4.3	(7)
Pittsburgh	-1.6	(10)	-0.9	(14)	-0.2	(13)	+3.0	(17)
St. Louis	-0.4	(16)	-1.9	(7)	-0.2	(14)	-9.6	(1)
Cleveland	-2.8	(4)	-2.1	(6)	+2.2	(16)	+1.1	(15)
Baltimore	-3.0	(2)	-2.6	(4)	-4.7	(8)	-0.7	(11)
Houston	-3.0	(3)	-1.7	(9)	-10.8	(2)	+5.5	(18)
Minn/St. Paul	-0.9	(13)	-1.6	(11)	+0.9	(15)	-5.5	(3)
Dallas	-0.5	(15)	-0.7	(15)	-3.4	(11)	-1.9	(9)
Cincinnati	-0.5	(14)	-0.6	(16)	-7.9	(3)	-0.8	(10)
Milwaukee	-1.1	(11)	-1.6	(10)	-6.4	(5)	-5.0	(5)
Buffalo	+0.4	(18)	+0.3	(18)	+4.0	(17)	-0.3	(12)
AVERAGE	-1.7%		-1.7%		-3.7%		-1.3%	

*Infant deaths (under 1 year) per 1,000 live births.

[a]The lower rankings are assigned to the lower infant mortality rates.

[b]These data appeared in the initial version of this report.

[c]The lower rankings are assigned to the higher negative rates of change.

Source: A Study in Comparative Urban Indicators: Conditions in 18 Large Metropolitan Areas, Michael J. Flax, Urban Institute Paper 1206-4, April 1972.

EXHIBIT 4 An Example of City/Suburban Indicator Presentation

1967 City/Suburban Infant Mortality Rates*

(A Health Indicator)

CENTRAL CITY 1967 Infant Mortality Rate[a]			SUBURBAN 1967 Infant Mortality Rate[a]			CITY/SUBURBAN RATIO 1967[a]	
Cincinnati.....	19.9	(1)	Milwaukee.....	14.2	(1)	Cincinnati.....	0.94 (1)
S.F./Oakland.....	21.4	(2)	Minn./St. Paul....	17.4	(2)	Dallas.....	1.00 (2)
L.A./Long Beach...	22.0	(3)	Cleveland.....	17.6	(3)	L.A./Long Beach...	1.17 (3-4)
Minn./St. Paul....	22.5	(4)	WASHINGTON.....	17.8	(4)	S.F./Oakland.....	1.17 (3-4)
Houston.....	22.7	(5)	Boston.....	17.9	(5)	Houston.....	1.18 (5)
New York.....	23.8	(6)	S.F./Oakland.....	18.3	(6)	New York.....	1.20 (6)
Dallas.....	24.0	(7)	Detroit.....	18.4	(7)	Buffalo.....	1.27 (7)
Milwaukee.....	24.3	(8)	L.A./Long Beach...	18.8	(8)	Minn./St. Paul....	1.29 (8)
Boston.....	24.9	(9)					
Cleveland.....	25.0	(10)	18 AREA AVERAGE...	19.1		18 AREA AVERAGE...	1.33
18 CITY AVERAGE...	25.3		Pittsburgh.....	19.2	(9)	Baltimore.....	1.37 (9)
			Houston.....	19.2	(10)	St. Louis.....	1.39 (10-11)
Buffalo.....	25.7	(11)	Baltimore.....	19.4	(11)	Boston.....	1.39 (10-11)
Baltimore.....	26.6	(12)	New York.....	19.9	(12)	Chicago.....	1.40 (12)
WASHINGTON.....	27.3	(13)				Cleveland.....	1.42 (13)
Detroit.....	27.5	(14)	Chicago.....	20.1	(13)		
Chicago.....	28.1	(15)	Buffalo.....	20.3	(14)	Philadelphia.....	1.50 (14-16)
			Philadelphia.....	20.4	(15)	Detroit.....	1.50 (14-16)
Pittsburgh.....	28.8	(16)	Cincinnati.....	21.1	(16)	Pittsburgh.....	1.50 (14-16)
St. Louis.....	29.5	(17)	St. Louis.....	21.2	(17)	WASHINGTON.....	1.53 (17)
Philadelphia.....	30.6	(18)	Dallas.....	22.1	(18)	Milwaukee.....	1.71 (18)

*Infant deaths (under 1 year) per 1,000 live births.

[a]This analysis was completed with 1967 data before the 1968 data became available.

Source: A Study in Comparative Urban Indicators: Conditions in 18 Large Metropolitan Areas, Michael J. Flax, Urban Institute Paper 1206-4, April 1972.

EXHIBIT 5

Definition of Indicators

Indicator number	Indicator
	EDUCATIONAL PERFORMANCE
1	Average 6th grade reading scores--months below equivalent grade level as determined by national norms (1968-69).
2	Percent of Selective Service-draftees that failed the Armed Forces mental test (1969).
3	Percent of high school students not graduating based on pupil loss, grades 10-12 (September 1960-June 1963).
	DIRECT EDUCATIONAL INPUT RELATED INDICATORS
4	Estimated annual current expenditures per pupil in average daily membership (1967-68)--adjusted for cost of living.
5	Median annual salaries for classroom teachers (1968-69)--adjusted for cost of living.
6	Annual Salary Range (difference between first and third quartiles) for classroom teachers (1968-69)--adjusted for cost of living.
7	Percent of teachers having at least a Master's Degree, i.e., Master's, Master's plus 30 hours, or a Ph.D. (Fall 1968).
8	Ratio of pupils to classroom teachers in all elementary and secondary schools (Fall 1968)--pupil/teacher ratio.
	INDIRECT EDUCATIONAL INPUT RELATED INDICATORS
9	Median school years completed by persons 25 years old and over (1967).
10	Percent of adults (age 18 or over) who are taking some form of adult education under public school auspices (1966-67).
	EDUCATIONAL INTEREST AND SUPPORT
11	Per capita total expenditures on education per $1,000 per capita income (1966-67).
12	Educational expenditures as a percentage of total general expenditures (1966-67).

Source: Selected Education Indicators for 21 Major Cities: Some Statistical Benchmarks, Michael J. Flax, Preliminary Draft Working Paper 136-4, May 1971.

EXHIBIT 6 An Example of Indicator Presentation

A 21-City Tabulation of 12 Educational Indicators

CITY	Performance Related			Direct Input Related					Indirect Input Related		Interest and Support	
	1	2	3	4	5	6	7	8	9	10	11	12
	6th Grade Reading Scores: Months Behind National Norms (1968-69)	% Draftees Failing Mental Tests ('69)	High School Dropouts (1960-68)	Adjusted Current Expenditures Per Pupil (1967-68)	Adjusted Median Teacher Salary (1968-69)	Adjusted Salary Range (1968-69)	% Teachers with Graduate Training (Fall 1968)	Pupil/Teacher Ratio (Fall '68)	Median Education of Adult Population, years (1967)	% Adults in Adult Education Programs (1966-67)	Per Capita Expenditure on Education Per $1,000 Per Capita Income (1966-67)	Education Expenditure as a % of Total Expenditures (1966-67)
New York.....	-3 mo.	19%	37%	$865	$8,682	$4,172	34%	21	11.6	1.6%	(NA)	35.5%
Chicago.....	-14	11	34	640	9,141	3,308	26	27	11.4	1.8	$43	38.7
Los Angeles....	-11	9	23	612	9,930	3,793	21	27	12.4	4.5	49	27.7
Philadelphia...	-13	12	47	670	9,247	3,354	37	24	11.2	3.8	32	23.3
Detroit.....	-10	8	35	670	10,543	3,206	34	27	10.9	1.2	31	27.8
Houston.....	-3	11	21	545	7,770	1,784	29	28	12.3	2.1	30	23.0
Baltimore.....	-10	10	35	633	8,588	4,232	31	24	9.9	1.9	33	30.2
Washington.....	-13	14	30	723	8,581	2,907	24	21	12.0	1.5	(NA)	38.1
Dallas.....	(NA)	9	28	467	7,625	1,811	29	27	12.4	2.1	26	24.1
Cleveland.....	(NA)	7	31	561	7,559	2,534	18	28	10.3	3.4	44	24.8
San Francisco..	-16	4	33	639	9,868	3,483	24	22	12.3	4.5	31	37.1
Milwaukee.....	-15	3	26	533	9,132	3,326	32	28	11.5	1.1	34	27.5
St. Louis.....	-5	6	25	616	8,565	3,896	28	29	9.6	1.0	38	22.2
Boston.....	-10	7	36	618	8,711	3,695	43	24	12.1	1.8	42	52.4
Pittsburgh.....	(NA)	4	25	755	(NA)	(NA)	26	24	12.0	1.0	25	30.7
Cincinnati.....	(NA)	5	35	634	8,357	3,540	27	30	11.8	6.6	45	22.9
Buffalo.....	(NA)	5	31	757	8,003	3,075	23	23	10.5	2.9	44	30.6
Minneapolis....	(NA)	2	21	611	8,623	4,016	35	23	12.2	(NA)	(NA)	32.7
Oakland.....	-15	(NA)	25	632	9,348	3,894	38	24	(NA)	2.2	41	(NA)
Long Beach.....	(NA)	(NA)	16	636	10,009	3,369	36	26	(NA)	2.1	42	(NA)
St. Paul.....	(NA)	(NA)	(NA)	623	9,534	3,503	16	25	(NA)	8.3	35	(NA)
WEIGHTED AVERAGE.....	-11 mo.	8%	30%	$640	$9,347	$3,345	29%	25	11.5	2.8%	$37	30.5%

NA--Data not available.

Source: Selected Education Indicators for 21 Major Cities: Some Statistical Benchmarks, Michael J. Flax, Preliminary Draft Working Paper 136-4, May 1971.

EXHIBIT 7 An Example of Indicator Presentation

Rates of Change and Size of Gaps: 1960 to 1968

W-White; NW-Nonwhite; B-Black. About 92% of nonwhites are black.[a]

Rates tend to: △ increase difference ▽ decrease difference Size of gap: ▲ increasing ▼ decreasing

Indicator	Race	Data 1960	Data 1968	Imputed annual rate of change 1960-1968[b]	Analysis: Size of white/nonwhite gap, 1960[c]	Analysis: Size of white/nonwhite gap, 1968[c]	Evaluation: Was the nonwhite rate of improvement greater than the white?	Evaluation: Was the size of the white/nonwhite gap decreasing?
LIVING CONDITIONS AND HEALTH								
Infant mortality (per 1,000 population)[d]	W	17.2	[e]14.7	-2.0% △	9.7%	8.7% ▼	no	yes
	NW	26.9	[e]23.4	-1.8%				
Life expectancy at 35 years[f]	W	73.8 yrs.	[e]74.4 yrs.	+0.10% ▽	4.5 yrs.	4.4 yrs. ▼	yes	yes
	NW	69.3 yrs.	[e]70.0 yrs.	+0.12%				
HOUSING								
% Housing that is substandard	W	13%	6%	-9.2% △	31%	18% ▼	no	yes
	NW	44%	24%	-7.3%				
FAMILY								
% Female-headed families	W	8.7%	8.9%	+0.3% △	13.7%	17.5% ▲	[g]no	no
	NW	22.4%	26.4%	+2.1%				
% Children living with two parents	W	92%	92%	0.0 △	17%	23% ▲	[h]no	no
	NW	75%	69%	-1.1%				
Fertility rates (live birth/1,000 women, age 15-44 years)[i]	W	113	[e]79.4	-4.3% △	41%	38.4% ▼	no	yes
	NW	154	[e]115.3	-3.5%				
% Illegitimate births[j]	W	2.3%	[e]5.5%	+11.4% ▽	19.3%	25.2% ▲	[k]yes	no
	NW	21.6%	[e]30.7%	+4.5%				
EDUCATION								
% Men (age 25-29) completing high school	W	[l]63%	76%	+2.3% ▽	27%	18% ▼	yes	yes
	B	[l]36%	58%	+6.1%				
% Completing at least 4 years of college (25-34 years)	W	11.7%	15.7%	+3.8% ▽	7.4%	9.0% ▲	yes	no
	B	4.3%	6.3%	+4.9%				
EMPLOYMENT								
% Unemployed	W	4.9%	3.2%	-5.2% △	5.3%	3.5% ▼	no	yes
	NW	10.2%	8.7%	-5.1%				
% Teenagers unemployed	W	19.1%	11%	-6.7% △	15.1%	14% ▼	no	yes
	NW	34.2%	25%	-3.9%				
% In clerical occupations	W	15.8%	[e]17.5%	+1.3% ▽	8.5%	5.4% ▼	yes	yes
	NW	7.3%	[e]12.1%	+6.6%				
% In professional and technical occupations	W	12.1%	[e]14.2%	+2.0% ▽	7.3%	6.4% ▼	yes	yes
	NW	4.8%	[e]7.8%	+6.2%				
INCOME AND POVERTY								
Median family income (in 1968 dollars)	W	$6857	$8937	+3.4% ▽	$3063	$3347 ▲	yes	no
	NW	$3794	$5590	+4.9%				
% Persons below poverty level	W	18%	10%	-7.1% △	37%	25% ▼	no	yes
	B	55%	35%	-5.3%				
% Families with incomes greater than $8000 (in 1968 dollars)	W	39%	58%	+4.8% ▽	24%	26% ▲	yes	no
	NW	15%	32%	+10.0%				

Source of Data: The Social and Economic Status of the Negroes in the United States, 1968. Blacks and Whites - An Experiment in Racial Indicators. Michael J. Flax. Urban Institute Report 85-136-5, 1971. [a]Black data are used whenever available. [b]This is the annual percentage increase that would give the 1960-1968 change if compounded annually from 1960-1968. Many of these figures are percentage increases in percentages. [c]Many of these gaps are differences between white and nonwhite percentage figures. [d]Infants dying at less than 1 month of age. [e]1968 values were extrapolated from 1960-1967 or 1960-1969 rates of change. [f]Data on life expectancy look more favorable as the age group rises. [g]The number of female-headed families is increasing for both whites and nonwhites, but the number of female-headed nonwhite families is increasing at a faster rate. [h]The percent of nonwhite children living with two parents is decreasing, while the percent of white children remains the same. [i]A lowering of the fertility rate is assumed desirable. Where income is limited a decrease would tend to encourage desirable trends in many of the other categories measured. [j]The white illegitimacy rate is increasing at double the nonwhite rate, therefore the slower nonwhite change may be expressed as a relatively greater "rate of improvement." [k]The percent of both white and nonwhite illegitimate births is increasing, but that of whites is increasing at a faster rate. [l]1960 data are for nonwhites.

EXHIBIT 8

Variables Included in Discriminant Functions[1]

1. Demography

   a. Percent nonwhite
   b. Crude death rate per thousand
   c. Outmigration rate 1950-60
   d. Crude white fertility rate: white population age 0-4 as percent of white females age 15-44
   e. Nonwhite dependency ratio: nonwhite population 0-4 and 65+ as percent of nonwhite population 20-64

2. Economy (Employment)

   a. Index of the centralization of employment (lack of industrial diversification)
   b. Index of the concentration of nonwhites in low-paying industries
   c. Index of the nonfarm occupational mix

3. Housing

   a. Percent of rental units vacant
   b. Percent of households with two cars
   c. Percent of houses with adequate plumbing
   d. Percent of persons in group quarters

4. Education

   a. Percent enrolled in private elementary and high school
   b. Enrollment rate of those aged 15-19
   c. Percent white population 25 years and over which completed high school
   d. Percent nonwhite population 25 and over which completed high school
   e. Percent of total population 25 and over which completed at least 4 years of college

[1]The variables used in the respective discriminant functions estimated in "Identifying Socio-Economic Differences Between High and Low Income Metropolitan Areas," are listed. Other variables were considered. Urban Institute Working Paper 1206-6, Herrington J. Bryce, July 1972.

SOCIAL AREA ANALYSIS: PROCEDURES AND ILLUSTRATIVE APPLICATIONS BASED UPON THE MENTAL HEALTH DEMOGRAPHIC PROFILE SYSTEM

Harold F. Goldsmith
and
Elizabeth L. Unger
National Institute of Mental Health

Introduction

The National Institute of Mental Health has designed a small area demographic profile system (MHDP) that provides the requisite data for the delineation of meaningful "Social Areas" within mental health service areas, cities, standard metropolitan statistical areas or other designated census areas. In an earlier paper, we discussed the kinds of data comprising the profile and suggested operational procedures to be used to construct social areas (areas with common demographic characteristics) from the data provided (see Goldsmith and Unger, 1972a). Unfortunately, we found that those procedures presented worked best for racially homogeneous husband-wife/family areas, that is, for suburban-like family areas rather than for inner city residential areas. The inner city tends to be characterized by heavy concentrations of poor people, both white and Negro, large proportions of primary individuals who are household heads (nonfamily households), and/or disproportionate numbers of households headed by females. Thus, what is required are standardized procedures to be used in the construction of social areas that are applicable to a wide range of disparate residential areas. Our primary concern in this paper will be to suggest such procedures. We will use the data items in the MHDP, since we are also concerned with familiarizing the reader with the MHDP. Poor white and Negro census tracts of Baltimore city will be used to exemplify the procedures.

Social Area Dimensions

Before presenting the operational procedures, we will comment briefly on the social area characteristics that need to be measured. "Proponents of social area analysis state that much of residence related behavior can be understood and accounted for in terms of three types of society-wide population characteristics or dimensions: social rank, life style or urbanization, and ethnicity" (Goldsmith and Unger, 1970). In our view, neither the social rank dimension nor the urbanization dimension should be treated as single factors (see Goldsmith and Unger, 1970 and 1972b). The standard components of social rank—economic class (ability to consume goods and services), social class (prestige, position) and, perhaps, educational (information) status—should be considered separately. The life style or urbanization dimension is seen as a composite dimension including the following components: family status (family/nonfamily), family life cycle stage, residential life style, and familism (small family/large family). The ethnicity dimension is maintained without modification. In addition to the traditional social area variables, residential stability and area homogeneity are viewed as significant dimensions requiring identification.

Mental Health Small-Area Demographic Profile System (MHDP)

Census measures which serve as indicators of the various social area and related dimensions are presented in appendix tables 1-4 and appendix figure 1. Item selection was based primarily on prior use of specific census measures in demographic and ecological analysis and, secondly, on modifications suggested by subject matter experts. In addition to the data items indexing social area dimensions, we selected (on the basis of examination of epidemiological studies) 14 other census items judged most useful in the identification of subpopulations seen as "high risk" in terms of probable need for mental health services and other related social services (see appendix table 4 and Redick, Goldsmith and Unger, 1971). We feel that the data items in the profile index those demographic and ecological dimensions (1) that best differentiate among residential subareas of American cities, according to the current literature, and (2) that can be measured using census data.

Appendix tables 1 and 2 utilize second count summary tapes which contain complete-count population and housing data at the level of census tracts and larger census areas. Appendix tables 3 and 4 utilize not only second count tapes but also fourth count summary tapes which contain such data as income, occupation and education for the 20-, 15-, and 5- percent sample populations. The MHDP provides age-sex population pyramids for Negro, white and total populations and for rural and urban areas (see appendix figure 1). The MHDP has been designed so that appendix tables 1 and 3 contain overall information about an area's population (i.e., the "dominant" characteristics of that population), and appendix tables 2 and 4 contain specific information about the distribution of an area's population (i.e., indicators of area homogeneity). Profiles can be produced for census tracts, minor civil divisions, counties, and aggregations of these units. Further, the MHDP system allows for statistical manipulation of the area units. The profile items are stored on disks for rapid retrieval for description or statistical analysis. Intermediate tapes (tapes that are derived from census tapes and that contain selected second and fourth count tabulations) were used to build the disk data base.

Procedures For Classifying Areas By Social Area Dimensions

In this section, we will discuss procedures for classifying areas by social area dimensions.[1] We will demonstrate procedures that can be used outside of nonintegrated/husband-wife family areas, i.e., that have general applicability to different types of areas.

The items selected for use here are from the second count summary tapes since the fourth count data items were not available at the time the paper was being prepared. However, where appropriate, we do refer to the fourth count data items since they are now available.

Ethnicity

The ethnic populations that are considered most significant in contemporary social area studies are Negroes and persons of Spanish heritage. Since the second count data do not allow identification of persons of Spanish background, our measure of ethnicity in the poor areas of Baltimore city is the proportion of an area population that is Negro. The classification system used is presented below:

PERCENT NEGRO

(Appendix table 1, Item 9)

0	0.0% - 4.9%
1	5.0% - 9.9%
2	10.0% - 29.9%
3	30.0% - 49.9%
4	50.0% - 74.9%
5	75.0% - 89.9%
6	90.0% - 94.9%
7	95.0% or more

The code was designed to allow differentiation of areas on the basis of the relative visibility of Negro resident populations. Where other ethnic (non-Negro) populations appear in sufficient number in an area, they may be similarly classified. The second count disk data also identifies the proportion of an area population that is neither white nor Negro (see appendix table 1, item 10). With the fourth count summary tape the percent of the population of Spanish heritage (appendix table 4, item 44), foreign stock (appendix table 3, item 14) and southern or eastern European stock (appendix table 4, item 45) can be identified. Programs that will provide data items in appendix tables 3 and 4 for Negro and Spanish heritage populations are being developed. These programs will use the data stored on our intermediate tapes.

Social Rank: Economic Class

Only a crude estimate of area economic class can be made using the second count summary tape data (more precise information will be available when the fourth count summary tape data are incorporated in the system). We suggest that median rent and median house value (appendix table 1, items 7 and 8) be used to estimate **total area** economic class (Goldsmith and Unger, 1972a). In order to estimate the economic status of non-Negro and Negro populations of an area separately, the lower quartile values for owner occupied and renter occupied dwelling units by race should be utilized (see appendix table 2, items 1, 2, 3, 4). The decision to use any of the above items as indicators of area economic class entails the implicit assumption that the higher the rent paid or the higher the house value, the higher the economic class or income.

Since we are concerned specifically in this paper with identifying poor populations in areas by race, we have used the lower quartile values for Negroes of Baltimore city in classifying and designating a population as poor. First, we looked separately at non-Negro headed and Negro headed owner occupied housing units in a census tract and classified them as "poor" if the lower quartile house value was below $7,000; and we examined non-Negro headed and Negro headed renter occupied housing units and classified them as "poor" if the lower quartile monthly rental was less than $65. Then we looked at the distribution of the above groupings, and designated an entire area (census tract) as "poor" if the subgroups classified "poor" comprised 80 percent or more of the households in the area. Only four types of "poor" areas will be examined in this paper. They are specified by race and by renter/owner status as follows:

1. Predominant population Negro renters: "poor" areas in which 80 percent or more of the housing units are rented by Negroes.
2. Predominant population Negro renters or owners: "poor" areas in which 80 percent or more of the housing units are rented or owned by Negroes (less than 80 percent of the units are rented by Negroes).
3. Predominant population non-Negro renters: "poor" areas in which 80 percent or more of the housing units are rented by non-Negroes.
4. Predominant population non-Negro renters or owners: "poor" areas in which 80 percent or more of the housing units are rented or owned by non-Negroes (less than 80 percent of the units are rented by non-Negroes).

In our exploratory analysis of Baltimore city, the first two types are referred to as poor Negro census tracts and the last two as poor white census tracts.

When the fourth count summary data is incorporated into the system, data items for areas such as percent of families and persons in poverty by race (see appendix tables 3 and 4), will provide more precise indicators of the extent to which the populations of areas are "poor." Other estimates of economic class can be arrived at by using such items as median income of families or of unrelated individuals by race (appendix table 4, items 1-5), and unemployment and under-employment rates by race and sex (appendix table 4, items 16-20).

Social Rank: Social Class and Information Status

The system, when it incorporates the fourth count data, will also provide indicators of area social class and area information status. Some examples of appropriate indicators of social status are: percent of males in high status occupations, percent of males in low status occupations (appendix table 3, items 9 and 10), percent of females in low status occupations, etc. (appendix table 4, items 29-32). One indicator of information status would be median school years completed by persons 25 and over (appendix table 3, item 11); and appendix table 4 provides other detailed data pertaining to information status by sex and age (items 34-43).

[1] **The authors have omitted from the present article, due to the limitations of space, two appendix tables and an appendix chart. The reader can receive this supplementary appendix by writing directly to the authors, care of the Mental Health Study Center, 2340 University Blvd. E., Adelphi, Md. 20783. The supplementary appendix tables illustrate in detail the application of the various coding procedures discussed in the text to each of the "poor" census tracts in Baltimore city; the supplementary appendix chart details the 4-item and 3-item disk area family life cycle codes mentioned later in the text.**

Family Status

In an earlier paper on social area identification procedures, our classification of family status was based solely on the percent of husband-wife households in an area: a tract (or other area) was considered a family area if at least 70 percent of the households in the area were husband-wife households. If an area does contain a high proportion of husband-wife households, there is no question that one is dealing with a family area, representative of what many would describe as "normal family" areas. However, this classification does not tell the whole story. For there remain many households (30.6 percent in the United States) that are **not husband-wife** households and many households (19.7 percent in the United States) that are **not family** households, i.e., where the head of household is a primary individual (see U.S. Bureau of the Census, PC(1)C1, June 1972).

A classification procedure that will more accurately describe the family status of an area must distinguish family areas and nonfamily areas and, within family areas, husband-wife family areas and other types of family areas. With these requirements in mind, we re-examined the available disk data items and selected the following data items: (1) percent of households with heads who are primary individuals (appendix table 2, item 18); (2) percent of all households husband-wife (appendix table 1, item 11); and (3) percent of households with female heads (appendix table 2, item 23). We were able to formulate at least a preliminary area family status typology using this set of data items, a typology that would meet the requirements of differentiation and specification mentioned above. Thus, we are able to distinguish (1) nonfamily areas (high percent of households headed by primary individuals), (2) mixed family/nonfamily areas (moderate percent of households headed by primary individuals), and (3) family areas (low percent of primary individuals). We can further distinguish within family areas husband-wife family areas (high percent of husband-wife households) and female-headed family areas (high percent of female-headed households). It should be recognized that this typology is applicable to largely uni-racial areas since the data items refer to total populations; however, appendix tables 3 and 4 do present two of the data items by race which will enable us to make estimates of area family status for integrated areas.

Extensive empirical examination of census tracts in the Washington, D.C., Wilmington, Del., and Baltimore city areas resulted in the following coding system for each of the family status variables:

PERCENT IN CATEGORY

1	Under 15% (very low)
2	15.0% - 29.9%
3	30.0% - 49.9%
4	50.0% - 69.9%
5	70.0% - 84.9%
6	85% and above (very high)

We will give examples of some of the area family status types that we found in Baltimore city later on in the paper.

Area Family Life Cycle

In this section, we will present classification procedures of areas by family life cycle utilizing (a) data stored on the disk and (b) the age-sex population pyramids. The procedures using disk data will be discussed first.

Four-Item Disk Code: A classification system based on four disk items is presented in the paper, **Social Areas: Identification Procedures Using 1970 Census Data** (Goldsmith and Unger, 1972a). Because the 4-item disk code is most meaningful when applied to predominantly husband-wife family areas, a revised classification procedure that would also be applicable to residential areas that are either not family areas or not husband-wife family areas is needed. In response to this need, we have developed a revised (3-item) code which we will present in the next section, following a brief discussion of the 4-item code. (The 4-item code is detailed in supplementary appendix chart 1.)

The development of the 4-item disk code was an outgrowth of a prior analysis of the characteristics of families and areas at different stages in the life cycle (Foote, *et al.*, 1960:96-121; Goldsmith, Stockwell and Unger, 1970). A first step was to assign designations to the various stages in the family life cycle (i.e., pre-family, childbearing, childrearing, childlaunching, post-childlaunching), labels that we hoped would describe families as they age in terms of child-related functions and associated adaptation problems. Then we listed a few related standard and nonstandard demographic characteristics (defining life style characteristics) for area family life cycle stages (see chart 1). This demographic characterization of families proceeds from the general observation that the "average" woman marries by age 20 or 21, has two or three children, and bears her last child before she reaches age 30 (Glick, 1957).

We then selected items from the disk data that we felt would reflect and specify the general characteristics outlined in chart 1. The criteria to be applied to these items in classifying areas by family life cycle stage are presented in chart 2.

Chart 2 illustrates that a husband-wife family area with a low proportion of children of all ages and young household heads would be classified as a pre-family population; whereas the same sort of area with **old** household heads would be described as a post-childlaunching population. A husband-wife family area with a high proportion of children in the under 6 only category (other categories low) and young household heads would be classified as a childbearing population; whereas the same sort of area with a high proportion of children in the **6-17 only** category (other categories low) would be identified as a late childrearing (or childlaunching) population.

Three-Item Disk Code: The theoretical basis for classifying areas by family life cycle stage included consideration of parental age; however, we were obliged to infer age of parents from the disk item "age of household head." Obviously, we are justified in making such an inference only for husband-wife areas (70 percent or more of the households husband-wife) or for family areas (less than 30 percent of household heads primary individuals). Because of this limitation, we decided to delete the "age of household head" item. The remaining 3-item code would still provide information about the families in an area even if that area had a sizeable nonfamily population. We then compared these 2-disk codes to determine the consequences of the revision on the adequacy of the classification system (see supplementary appendix tables 1 and 2). In our judgment, the 3-item code, although less sensitive than the 4-item scale in identifying specific area family life cycle stages, does produce a more reliable index for the **families** in nonfamily areas. (Both codes are presented in supplementary appendix chart 1.)

Chart 1.—DEFINING LIFE CYCLE CHARACTERISTICS FOR AREA FAMILY LIFE CYCLE STAGES

Pre-family	Families with no children and single persons	Young household heads
Childbearing	Families with children under 5 only	Relatively young household heads
Early childrearing	Families with children under 5 and 5-9	Young to middle-aged household heads
Middle childrearing	Families with children 5-9 and 10-14	Middle-aged household heads
Late childrearing	Families with children 10-14 and 15-19	Middle-aged and slightly older household heads
Childlaunching	Families with children 15 and over	Middle-aged and slightly older household heads
Post-childlaunching	Families with no children under 18	Old household heads

Chart 2.—DISK DATA CRITERIA FOR AREA FAMILY LIFE CYCLE STAGES

Family life cycle stage of area	Husband-wife households	Age of household head	Families: With children under 6 only	With children under 6 and 6-17	With children 6-17 only
Pre-family	High	Young	Low	Low	Low
Childbearing	High	Young	High	Low	Low
Early childrearing	High	Young	Low	High	Low
Middle childrearing	High	Middle-aged	Low	Low	High
Late childrearing (or childlaunching)	High	Old	Low	Low	High
Post-childlaunching	High	Old	Low	Low	Low

Age-Sex Population Pyramid Codes: Our exploration of two central cities, Washington, D.C. and Baltimore city, and of suburban Prince George's County made us aware that it was necessary to provide estimates of family life cycle of area populations by race, primarily because the Negro and white populations of a given area may be at different life cycle stages. Appendix figure 1 gives a dramatic illustration of very different age-sex distributions of the white population and the Negro population of a single census tract; we assume that such a discrepancy between the two population pyramids is a reflection of different life cycle stages characterizing the populations. The data items comprising the disk code are not presently available by race (although we do contemplate adding these items by race to the intermediate tape); in addition, the disk codes do not provide a sensitive index for nonfamily areas. Since the age-sex population pyramids are provided by race (see appendix figure 1), and since a classification system of area family life cycle based on the age-sex pyramids would not be contingent upon either area family status or median age of household head; we decided to develop a classification system using the age-sex population pyramids to indicate the family life cycle stages of an area.

In an earlier paper (Goldsmith, Stockwell and Unger, 1970), we developed a procedure using the age-sex pyramids to estimate area family life cycle. Our present procedure, though related to the earlier one, is not concerned with the dominant adult age grouping which was an integral part of the 1970 procedure (see Goldsmith, Stockwell and Unger, 1970). The present procedure, like the disk code, follows logically from charts 1 and 2, and, like the revised disk code, ignores parental age and family status. The 1972 age code primarily relies on the presence (or absence) of children in specified 5-year age categories to estimate area family life cycle, as detailed in chart 3. Attention is paid to adult age grouping only for two life cycle stage designations: pre-family and post childlaunching. These exceptions will be discussed later.

An area often can be classified in more than one stage of the family life cycle. This is because designations are made generally on the basis of Highs always taking into account the adjacent categories. If one or both of adjacent categories are High, then the area receives appropriate designations as indicated in chart 3. Thus, if an area's pattern was H H L H, it would be classified as early childrearing (on the basis of the two Highs in the "under 5" and "5-9" categories) **and** childlaunching (on the basis of the Low and High in the "10-14" and "15-19" categories). Determination as to the primary (first mentioned) designation is based on which High age category is highest. Area family life cycle designations can be preceded by "with some." This wording indicates Mediums rather than Highs in the defining categories (see supplementary appendix tables 1 and 2).

Chart 3.—REVISED AGE CATEGORY CRITERIA FOR DETERMINING FAMILY LIFE CYCLE STAGES OF SUBAREA POPULATION

Family life cycle stage	Type	Percentage of population in specified age category[1]			
		Under 5	5-9	10-14	15-19
Pre-family and/or post-child-launching	1	Low	Low	Low	Low
Childbearing	2	High	Low	Low	Low
Childbearing with early child-rearing	3	High	Medium	Low	Low
Early childrearing	4	High	High	Low	Low
	5	Medium	High	Low	Low
	6	Low	High	Low	Low
Early childrearing with middle childrearing	7	Low	High	Medium	Low
Middle childrearing	8	Low	High	High	Low
	9	Low	Medium	High	Low
	10	Low	Low	High	Low
Middle childrearing with late childrearing	11	Low	Low	High	Medium
Late childrearing	12	Low	Low	High	High
	13	Low	Low	Medium	High
Childlaunching	14	Low	Low	Low	High

[1] Specific criteria for high, medium and low classifications:

High	12.5% or more of total population
Medium	7.5% to 12.4% of total population
Low	below 7.5% of total population

An area population is designated as including a **post-childlaunching** sub-population if 25 percent or more of the population is age 50 or over; the designation "with some post-childlaunching" is used to indicate that from 20 to 24 percent of the population is 50 or over. Designation of a population as **prefamily** depends on the distribution of population in two adult age groupings, "20-24" and "25-29" in combination with the distribution in two child age groupings, "under 5" and "5-9." Thus, an area is designated as "prefamily" if either of the adult age groupings accounts for 12.5 percent or more of the population **and** neither of the child age groupings accounts for more than 7.5 percent of the population. The above designation is qualified and referred to as "with some prefamily" under the following conditions: when the proportion of the population in either of the adult age groupings is high (12.5 percent or more), but the proportion in one or both of the child age groupings is medium (7.5 - 12.5 percent); **or** when the proportion of the population in either of the adult age groupings is medium (7.5 - 12.5 percent), provided the proportions in both child categories are low (less than 7.5 percent).

In addition to the 1972 age code, we have included a visual classification of the age-sex pyramids (in the supplementary appendix tables) because we think that it serves as an additional measure against which the other codes can be checked for reliability. Our previous experience has provided us with considerable skill in coding the age-sex pyramids according to family life cycle (Goldsmith, Stockwell and Unger, 1970). Though admittedly subjective, the visual code incorporates the logic in chart 1 and also takes into account how selective a population is in terms of child and adult age categories.

The designations assigned to each census tract according to the different codes (detailed in the supplementary appendix tables) illustrate the similarities and dissimilarities among the systems. These criteria are based on earlier work and recent exploration of the Washington and Baltimore Standard Metropolitan Statistical Areas.

Further, we have summarized the comparison of the 1972 age codes with the visual codes (table 1) and with the revised disk codes (table 2). One can see that there is considerable agreement among the designations resulting from the different systems, and that the similarity is most marked with respect to classification of family areas. Therefore, we suggest that the disk codes (since the items can be retrieved rapidly) be used to identify family life cycle in family areas, and that the 1972 age code be used in mixed family/nonfamily areas and nonfamily areas.

Residential Life Style

Residential life style indicates the day-to-day activities conducted in the immediate vicinity of the residence. It appears that residential life style of an area is related to the type and condition of available housing units (Goldsmith and Unger, 1970:13-17). We use the item "percent of occupied and vacant year-round housing units that are one unit structures (one unit at an address)" to index type of housing; and the item "percent of persons in households with 1.01 or more persons per room" to index condition of housing (overcrowding). Both of these items are from the second count census tapes. Our codes for these two items are as follows:

TYPE OF HOUSING: PERCENT ONE UNIT STRUCTURES

(Appendix table 1, item 15)

1	80% or more	(single dwelling units)
2	50%-79%	(mixed: single dwelling units with apartments)
3	Less than 50%	(predominantly apartments)

CONDITION OF HOUSING: PERCENT NON-NEGRO OR NEGRO POPULATIONS IN OVERCROWDED HOUSING

(Appendix table 2, items 32 and 33)

1 Less than 5% (not overcrowded)
2 5%-29% (slightly overcrowded)
3 30% or more (overcrowded)

When the data from the fourth count summary tapes are incorporated into the MHDP System, we will be able to identify type of housing by race (percent single dwelling detached units) and will have access to additional useful indicators of residential life style such as percent of housing units that are mobile homes, in large apartment complexes or in highrise apartments (see appendix tables 3 and 4).

Other Relevant Area Characteristics

Area residential instability and area subpopulations with high potential need for health and related services have not been considered in this paper because indicators are derived from the fourth count summary tapes; however, the relevant items are contained in appendix tables 3 and 4. The following census measures are used as indicators of an area's residential instability: percent living in different house 5 years earlier, percent living in different county 5 years earlier, and percent moved into present residence within the last year (see appendix tables 3 and 4). Items identifying "high risk populations" (shown in appendix table 4) encompass data on school dropouts, working mothers, aged persons living alone, disabled populations, and specific poverty groups.

Social Area Characteristics of "Poor" Areas of Baltimore City: 1970

In this section, we describe the kinds of "poor social areas" that we discovered when the procedures described above were applied to Baltimore city (see supplementary appendix tables). These social area types provide a preliminary typology of "poor" residential areas.

Poor Negro Census Tracts

The 23 poor Negro tracts are grouped into 6 basic types according to shared social area characteristics. The social area characteristics of each of the 23 census tracts (presented in supplementary appendix table 1) are summarized by type below:

Set A—Family areas; Negro female-headed families with many school-aged children; overcrowded housing conditions. The family status patterns (224, 225, 124 and 134) for these tracts give credence to their being described as female-headed family areas: in each tract at least 50 percent of the households are headed by females—female headed—and less than 30 percent of the households are headed by primary individuals—family areas. The predominant family life cycle classification for these area families is middle childrearing and early childrearing, with some late childrearing. The age-sex pyramids for the tracts comprising Set A are quite distinctive (e.g., see figure 1), indicating the presence of adult women with many children and the absence of adult males. All census tracts in this set are classified as overcrowded. There is no uniformity in set A in respect to type of housing.

Set B—Mixed family/nonfamily areas; Negro female-headed families in the middle childrearing stage of the family life cycle and some post-childlaunching households; mixed (single dwelling units with apartments) housing areas. The main difference between set B and set A is that set B contains more primary individuals. The family status pattern for this set of 324 describes mixed family/nonfamily areas with many (50 to 70 percent of all households) female-headed households and families. In each tract 30 to 50 percent of all households are headed by primary individuals (i.e., nonfamilies); on the other hand, at least 50 percent are **not** headed by primary individuals (i.e., families). The predominant family life cycle classification for these areas is middle childrearing or post-childlaunching: specifically, tracts 1002 and 1501 were classified as middle childrearing with some early and late childrearing and some post-childlaunching, and tracts 1702 and 1801 as post-childlaunching with some middle childrearing, late childrearing and early childrearing. The proportion of children for this set of tracts is not as large as that for set A (see figure 2 below). The age-sex pyramid also shows more adult females than males between the ages of 20 and 50, indicating the presence of many female-headed families. Set B tracts are classified as mixed housing areas (single dwelling units with apartments); moreover, three of the four tracts in this set are classified as overcrowded.

Set C—Family areas; Negro families husband-wife and female-headed in the middle childrearing family life cycle stage, also some post- childlaunching, late childrearing and early childrearing households. The family status pattern of 233 for set C tracts describes predominantly family areas (no more than 30 percent of all households are headed by primary individuals). Husband-wife families account for from 30 to 50 percent of all households; the same proportion obtains for female-headed households. In terms of family life cycle classification, all but one of the set C tracts are predominantly middle childrearing with some post-childlaunching, early childrearing and late childrearing. For the exception (tract 1502), the post-childlaunching stage was dominant with some middle, late and early childrearing also present. None of the five tracts in set C are classified as predominantly apartment areas, and only one of the tracts is a single dwelling unit area. Four tracts are classified as uncrowded. (See figure 3.)

Set D—Family/nonfamily areas; mixture of Negro nonfamily households, husband-wife families and female-headed families or households; area households in post-childlaunching family life cycle stage, some middle, early and late childrearing households also present. The social area characteristics of this set of tracts are similar to those of set C. The sets differ primarily in the presence of more households headed by primary individuals and in the predominance of the post-childlaunching family life cycle stage. The family status pattern of 333 indicates a mixture of family and nonfamily households, and among families many husband-wife families with a moderate proportion of female-headed families. The predominant area family life cycle stage for these tracts is post-childlaunching with some middle, early and late childrearing families also present. The sex ratios appear to be fairly balanced. None of the set D tracts are classified as single dwelling unit areas. Six of the seven tracts in this set are not overcrowded. (See figure 4.)

Set E—Predominantly nonfamily areas; mixture of Negro and white nonfamily households, and Negro husband-wife and female-headed households; area households in post-childlaunching family life cycle stage, some middle, early and late childrearing households also present; uncrowded, predominantly apartment areas. This set (tract 1402 only) has a family status

Fig. 1. **AN EXAMPLE OF A SET A AGE-SEX PYRAMID**

NEGRO POPULATION OF TRACT 402, COUNTY 510, MARYLAND, INNER CITY CMHC
TOTAL POPULATION 3509 MALES 1503 FEMALES 2006

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
11	0.31	85+		F	0.48	17
7	0.20	80-4		F	0.37	13
20	0.57	75-9	M	F	0.63	22
33	0.94	70-4	MM	FF	0.85	30
35	1.00	65-9	MM	FFF	1.17	41
48	1.37	60-4	MMMM	FFF	1.11	39
50	1.42	55-9	MMMM	FFFFF	1.85	65
45	1.28	50-4	MMM	FFFFFF	2.17	76
39	1.11	45-9	MMM	FFFFFFF	2.48	87
47	1.34	40-4	MMMM	FFFFFFF	2.34	82
46	1.31	35-9	MMM	FFFFFFFFF	3.08	108
32	0.91	30-4	MM	FFFFFFFFFFF	3.70	130
45	1.28	25-9	MMM	FFFFFFFFFF	3.48	122
70	1.99	20-4	MMMMM	FFFFFFFFFFFF	4.02	141
160	4.56	15-9	MMMMMMMMMMMMM	FFFFFFFFFFFFFFF	5.56	195
298	8.49	10-4	MMMMMMMMMMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFFFFFFFFF	8.52	299
295	8.41	5-9	MMMMMMMMMMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFFFFFFFFF	8.75	307
222	6.33	0-4	MMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFF	6.61	232

Fig. 2. **AN EXAMPLE OF A SET B AGE-SEX PYRAMID**

NEGRO POPULATION OF TRACT 1501, COUNTY 510, MARYLAND, PROVIDENT HOSPITAL CMHC
TOTAL POPULATION 5174 MALES 2291 FEMALES 2883

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
12	0.23	85+		F	0.46	24
19	0.37	80-4	M	FF	0.77	40
58	1.12	75-9	MMM	FFF	1.06	55
64	1.24	70-4	MMM	FFFFF	1.89	98
94	1.82	65-9	MMMMM	FFFFFFFF	2.94	152
105	2.03	60-4	MMMMMM	FFFFF	1.78	92
85	1.64	55-9	MMMM	FFFFFFF	2.42	125
104	2.01	50-4	MMMMMM	FFFFFFFF	2.80	145
134	2.59	45-9	MMMMMMM	FFFFFFFF	2.86	148
109	2.11	40-4	MMMMMM	FFFFFFFFF	3.02	156
81	1.57	35-9	MMMM	FFFFFFFFF	3.02	156
70	1.35	30-4	MMMM	FFFFFFF	2.47	128
77	1.49	25-9	MMMM	FFFFFFFF	3.00	155
116	2.24	20-4	MMMMM	FFFFFFFFF	3.27	169
260	5.03	15-9	MMMMMMMMMMMMMM	FFFFFFFFFFFFFFFF	5.76	298
330	6.38	10-4	MMMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFFFF	7.02	363
320	6.18	5-9	MMMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFF	6.44	333
253	4.89	0-4	MMMMMMMMMMMMM	FFFFFFFFFFFFF	4.75	246

Fig. 3. **AN EXAMPLE OF A SET C AGE-SEX PYRAMID**

NEGRO POPULATION OF TRACT 1001, COUNTY 510, MARYLAND, JOHNS HOPKINS HOSPITAL
TOTAL POPULATION 3725 MALES 1800 FEMALES 1925

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
10	0.27	85+			0.13	5
10	0.27	80-4			0.19	7
17	0.46	75-9	M	FF	0.75	28
26	0.70	70-4	MM	FF	0.97	36
57	1.53	65-9	MMMM	FFFFF	1.77	66
78	2.09	60-4	MMMMM	FFFFFF	2.12	79
95	2.55	55-9	MMMMMM	FFFFFFFF	2.68	100
97	2.60	50-4	MMMMMM	FFFFFFF	2.74	102
92	2.47	45-9	MMMMMM	FFFFFFF	2.95	110
93	2.50	40-4	MMMMMM	FFFFFFF	2.90	108
78	2.09	35-9	MMMMM	FFFFFF	2.39	89
86	2.31	30-4	MMMMM	FFFFFF	2.66	99
68	1.83	25-9	MMMM	FFFFFFF	2.74	102
103	2.77	20-4	MMMMMM	FFFFFFFF	3.46	129
210	5.64	15-9	MMMMMMMMMMMMM	FFFFFFFFFFFFFF	5.91	220
252	6.77	10-4	MMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFF	6.71	250
251	6.74	5-9	MMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFF	6.31	235
177	4.75	0-4	MMMMMMMMMMMM	FFFFFFFFFFF	4.30	160

Fig. 4. **AN EXAMPLE OF A SET D AGE-SEX PYRAMID**

NEGRO POPULATION OF TRACT 1601, COUNTY 510, MARYLAND, PROVIDENT HOSPITAL CMHC
TOTAL POPULATION 6103 MALES 2817 FEMALES 3286

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
17	0.28	85+		FF	0.69	42
31	0.51	80-4	M	FF	0.88	54
77	1.26	75-9	MMM	FFFF	1.64	100
119	1.95	70-4	MMMMM	FFFFFFF	2.56	156
171	2.80	65-9	MMMMMMMM	FFFFFFFFF	3.26	199
178	2.92	60-4	MMMMMMMM	FFFFFFFFFF	3.54	216
194	3.18	55-9	MMMMMMMMM	FFFFFFFFFFF	3.83	234
168	2.75	50-4	MMMMMMMM	FFFFFFFFFF	3.65	223
152	2.49	45-9	MMMMMMM	FFFFFFF	2.64	161
150	2.46	40-4	MMMMMMM	FFFFFFF	2.59	158
126	2.06	35-9	MMMMMM	FFFFFF	2.33	142
90	1.47	30-4	MMMM	FFFFFF	2.31	141
141	2.31	25-9	MMMMMM	FFFFFFF	2.65	162
188	3.08	20-4	MMMMMMMMM	FFFFFFFFFF	3.80	232
221	3.62	15-9	MMMMMMMMMM	FFFFFFFFFFFF	4.42	270
254	4.16	10-4	MMMMMMMMMMM	FFFFFFFFFFF	4.31	263
279	4.57	5-9	MMMMMMMMMMMM	FFFFFFFFFFFFF	4.83	295
261	4.28	0-4	MMMMMMMMMMM	FFFFFFFFFF	3.90	238

Fig. 5. **EXAMPLES OF A SET E AGE-SEX PYRAMID**

WHITE POPULATION OF TRACT 1402, COUNTY 510, MARYLAND, PROVIDENT HOSPITAL CMHC
TOTAL POPULATION 349 MALES 121 FEMALES 228

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
9	2.58	85+	MMMMMM	FFFFFF	2.01	7
5	1.43	80-4	MMMM	FFFFFFFFFFFFFFFFFFF	6.88	24
10	2.87	75-9	MMMMMMM	FFFFFFFFFFFFFFFFFFFFFFFFF*	10.60	37
14	4.01	70-4	MMMMMMMMMM	FFFFFFFFFFFFFFFFFFFFFFFFF*	16.91	59
24	6.88	65-9	MMMMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFFFFFFFFF*	13.47	47
6	1.72	60-4	MMMMM	FFFFFFFFFFF	4.30	15
3	0.86	55-9	MM	FFFF	1.43	5
6	1.72	50-4	MMMMM	FFF	1.15	4
5	1.43	45-9	MMMM		0.29	1
5	1.43	40-4	MMMM	FF	0.86	3
2	0.57	35-9	M		0.29	1
2	0.57	30-4	M	FFFF	1.43	5
2	0.57	25-9	M	FFF	1.15	4
13	3.72	20-4	MMMMMMMMMM	FFFFFFF	2.58	9
2	0.57	15-9	M	FFF	1.15	4
4	1.15	10-4	MMM		0.29	1
3	0.86	5-9	MM		0.0	0
6	1.72	0-4	MMMMM	F	0.57	2

NEGRO POPULATION OF TRACT 1402, COUNTY 510, MARYLAND, PROVIDENT HOSPITAL CMHC
TOTAL POPULATION 5223 MALES 2471 FEMALES 2752

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
12	0.23	85+		F	0.48	25
17	0.33	80-4		FF	0.80	42
44	0.84	75-9	MM	FF	1.00	52
94	1.80	70-4	MMMMM	FFFFF	1.91	100
125	2.39	65-9	MMMMMMM	FFFFFFF	2.51	131
108	2.07	60-4	MMMMMM	FFFFFFFF	2.80	146
155	2.97	55-9	MMMMMMMM	FFFFFFFFF	3.18	166
156	2.99	50-4	MMMMMMMM	FFFFFFFFF	3.24	169
161	3.08	45-9	MMMMMMMMM	FFFFFFFFF	3.08	161
167	3.20	40-4	MMMMMMMMM	FFFFFFFF	2.89	151
120	2.30	35-9	MMMMMM	FFFFFFF	2.37	124
116	2.22	30-4	MMMMMM	FFFFFFFF	2.78	145
154	2.95	25-9	MMMMMMM	FFFFFFFF	2.89	151
164	3.14	20-4	MMMMMMMMM	FFFFFFFFFFFF	4.35	227
183	3.50	15-9	MMMMMMMMMM	FFFFFFFFFFF	4.23	221
220	4.21	10-4	MMMMMMMMMMM	FFFFFFFFFFFFF	4.75	248
248	4.75	5-9	MMMMMMMMMMMM	FFFFFFFFFFFFF	5.00	261
227	4.35	0-4	MMMMMMMMMMMM	FFFFFFFFFFFF	4.44	232

pattern of 423 which indicates that most of the households of this tract are headed by primary individuals. This characteristic is unique to tract 1402; for that reason we have designated the tract as a separate type rather than placing it in set D. In terms of the Negro age-sex pyramid, the tract could have been a subtype of set D (see figure 5). Part of the difficulty in classification may relate to the fact that from 5 to 10 percent of the total population in the tract is white, predominantly female and over 65 (see figure 5). Even though the relative size of the white population is not large, it is likely that most of these individuals are female household heads living alone. The disproportionate number of such white households might distort the total figures so that they do not represent the characteristics of the dominant Negro population. Accordingly, in order to make more accurate estimates of family status for such tracts we need data by race. Such data will, of course, be available from the fourth count census tapes. The area family life cycle and residential life style characteristics are like those of set D.

Set F—Family/nonfamily areas; predominantly Negro nonfamily households, husband-wife families and female-headed families or households; area family life cycle stage is middle childrearing with some early and late childrearing households; uncrowded, mixed (single dwelling units with apartments) housing areas. This set is represented by tract 1204. The family status pattern of the total population is 334. Quite likely, the family status pattern of the Negro population is 333 or 233. As in Set E, the difference between the total population pattern and the Negro pattern may be due to the small aged predominantly white population. The family life cycle of this tract is middle childrearing with some early and late childrearing households. It is an uncrowded, mixed (single dwelling units with apartments) housing area. (See Figure 6.)

Poor White (Non-Negro) Census Tracts

The 27 census tracts classified as "poor" and white are grouped into 6 sets (G through L). Together, Sets G and H contain 78 percent of the white tracts. The social area characteristics of each tract (presented in supplementary Appendix table 2) are summarized by type below.

Sets G and H—Family areas; white husband-wife families with some female-headed households in the post-childlaunching family life cycle stage, other life cycle stages also present; uncrowded housing conditions. The 21 census tracts comprising sets G and H are described by one of two family status patterns: 242 or, less often, 243. These patterns describe family areas, specifically characterized by a preponderance of husband-wife families with some female-headed households. The post-childlaunching stage of the family life cycle is the primary classification for all tracts in sets G and H. In each instance, post childlaunching is accompanied by at least one (often several) of the following stages: Middle childrearing, late childrearing, early childrearing, prefamily, childlaunching, childbearing. None of the tracts in sets G or H are overcrowded. Further, none of the tracts are predominantly apartment areas; i.e., at least 50 percent, and often 80 percent or more, of the housing units are single dwelling units. Set H tracts differ from set G tracts only in terms of the proportion of the population Negro: 10 to 30 percent in set H and less than 5 percent in set G. (See figures 7 and 8.)

Set I—Family/nonfamily areas; marked predominance of white husband-wife families with some female-headed households; many households in post-childlaunching family life cycle stage, some middle and early childrearing households also present; uncrowded, mixed (single dwelling units with apartments) housing areas. The family status pattern for set I of 342 describes family/nonfamily areas and, among families, mostly husband-wife families with some female-headed families. Set I (containing only tract 105) resembles set G in all but one respect, i.e., the larger number of primary individuals who are household heads. The predominant family life cycle stage is post-childlaunching with some middle and early childrearing. This is an uncrowded area, with single dwelling units accounting for from 50 to 80 percent of all housing units. (See figure 9.)

Set J—Family/nonfamily areas; white husband-wife families and female-headed households in post-childlaunching family life cycle stage, some late and middle childrearing households also present; uncrowded, mixed (single dwelling units with apartments) housing areas. The family status pattern for set J of 333 describes family/nonfamily areas with families predominating (50 to 70 percent of households are families). Among families, husband-wife families predominate over female-headed families, but to a lesser extent than in set I areas. Post-childlaunching is the predominant family life cycle stage with some middle and late childrearing (also some early childrearing in one tract). Set J has the same residential life style characteristics as set I. (See figure 10.)

Set K—Family areas; predominantly white husband-wife families and female-headed households in the late and middle childrearing family life cycle stages; overcrowded, single dwelling units areas. The family status pattern for predominantly white Set K (tract 2603.03 only) of 233 delineates a family area (70 to 84 percent families) with many husband-wife families and female-headed families. The family life cycle classification for the white population of this tract is late and middle childrearing with some early childrearing and post-childlaunching. With respect to residential life style, over 80 percent of the housing units are single dwelling units and the area is overcrowded. (See Figure 11.)

Set L—Nonfamily areas; predominantly white prefamily and post-childlaunching households; uncrowded, predominantly apartment areas. The family status pattern for predominantly white set L (tract 1206 only) of 423 describes a nonfamily area (50 to 70 percent of the households are headed by primary individuals) with many female-headed households (from 30 to 50 percent of all households). Consistent with this nonfamily designation, set L receives a family life cycle classification of prefamily and post-childlaunching. The area is uncrowded, and over 50 percent of the housing units are apartments. (See figure 12.)

Summary And Conclusions

We have presented the Mental Health Small Area Demographic Profile System (MHDP) and described procedures (using that system) for identifying the major social area axes—economic class, social class, information status, family status, family life cycle and residential life style. Specific items to be used in indexing these dimensions (and other social area dimensions such as familism and community stability) are presented in appendix tables 1 through 4. The social area procedures as presented have general applicability to central city populations as well as to populations of suburban or rural areas.

We have illustrated the use of these procedures with selected poor areas of Baltimore city. This descriptive analysis can be seen as yielding a preliminary typology (made possible by the MHDP system) of "poor" areas that merits refinement and further work. The system allows identification not only of types of poor

Fig. 6. **EXAMPLES OF A SET F AGE-SEX PYRAMID**

WHITE POPULATION OF TRACT 1204, COUNTY 510, MARYLAND, NORTH CENTRAL CMHC
TOTAL POPULATION 235 MALES 106 FEMALES 129

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
0	0.0	85+		F	0.43	1
2	0.85	80-4	MM	FFF	1.28	3
3	1.28	75-9	MMM	FFFFFFFFFF	3.83	9
9	3.83	70-4	MMMMMMMMMM	FFFFFFFFFFFFFFF	5.53	13
6	2.55	65-9	MMMMMM	FFFFFFFFFF	4.26	10
11	4.68	60-4	MMMMMMMMMMMM	FFFFFFFFFFFFFF	5.11	12
10	4.26	55-9	MMMMMMMMMMM	FFFFFFFFFFFFFFFFFFFFF	7.66	18
9	3.83	50-4	MMMMMMMMMM	FFFFFFFFFF	3.83	9
9	3.83	45-9	MMMMMMMMMM	FFFFFFFFFFFFFF	5.11	12
8	3.40	40-4	MMMMMMMMM	FFFFFFF	2.55	6
7	2.98	35-9	MMMMMMM		0.0	0
4	1.70	30-4	MMMM	FF	0.85	2
3	1.28	25-9	MMM	FFFFF	1.70	4
13	5.53	20-4	MMMMMMMMMMMMM	FFFFFFFFFFFFFFF	5.96	14
3	1.28	15-9	MMM	FFFFF	1.70	4
2	0.85	10-4	MM	F	0.43	1
0	0.0	5-9		FFF	1.28	3
7	2.98	0-4	MMMMMMM	FFFFFFFFF	3.40	8

NEGRO POPULATION OF TRACT 1204, COUNTY 510, MARYLAND, NORTH CENTRAL CMHC
TOTAL POPULATION 3222 MALES 1533 FEMALES 1689

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
3	0.09	85+			0.22	7
7	0.22	80-4		F	0.50	16
17	0.53	75-9	M	F	0.50	16
23	0.71	70-4	MM	FFF	1.06	34
35	1.09	65-9	MMM	FFFF	1.40	45
46	1.43	60-4	MMMM	FFFF	1.58	51
64	1.99	55-9	MMMMM	FFFFF	1.92	62
84	2.61	50-4	MMMMMMM	FFFFFFF	2.64	85
96	2.98	45-9	MMMMMMMM	FFFFFFFF	2.82	91
83	2.58	40-4	MMMMMMM	FFFFFFFF	2.95	95
70	2.17	35-9	MMMMMM	FFFFFFFF	2.89	93
71	2.20	30-4	MMMMMM	FFFFFFFFF	3.26	105
67	2.08	25-9	MMMMMM	FFFFFFFF	2.86	92
99	3.07	20-4	MMMMMMMMM	FFFFFFFFFF	3.63	117
177	5.49	15-9	MMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFF	6.58	212
203	6.30	10-4	MMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFFF	6.70	216
202	6.27	5-9	MMMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFF	5.71	184
186	5.77	0-4	MMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFF	5.21	168

Fig. 7. **AN EXAMPLE OF A SET G AGE-SEX PYRAMID**

WHITE POPULATION OF TRACT 101, COUNTY 510, MARYLAND, BALTIMORE CITY HOSPITALS
TOTAL POPULATION 3944 MALES 1867 FEMALES 2077

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
7	0.18	85+		F	0.41	16
12	0.30	80-4		FF	0.76	30
32	0.81	75-9	MM	FFFF	1.50	59
51	1.29	70-4	MMM	FFFFFFF	2.38	94
84	2.13	65-9	MMMMMM	FFFFFFFF	2.84	112
98	2.48	60-4	MMMMMMM	FFFFFFFFF	3.04	120
154	3.90	55-9	MMMMMMMMMM	FFFFFFFFFF	3.68	145
130	3.30	50-4	MMMMMMMMM	FFFFFFFFFFF	4.01	158
135	3.42	45-9	MMMMMMMMM	FFFFFFFFFF	3.73	147
102	2.59	40-4	MMMMMMM	FFFFFFFFF	3.32	131
99	2.51	35-9	MMMMMM	FFFFF	1.95	77
73	1.85	30-4	MMMMM	FFFFFFF	2.54	100
108	2.74	25-9	MMMMMMM	FFFFFF	2.56	101
138	3.50	20-4	MMMMMMMMM	FFFFFFFFFFF	4.41	174
158	4.01	15-9	MMMMMMMMMM	FFFFFFFFFFFF	4.46	176
181	4.59	10-4	MMMMMMMMMMM	FFFFFFFFFFF	4.13	163
165	4.18	5-9	MMMMMMMMMMM	FFFFFFFFFFF	4.11	162
140	3.55	0-4	MMMMMMMMM	FFFFFFFF	2.84	112

Fig. 8. **AN EXAMPLE OF A SET H AGE-SEX PYRAMID**

WHITE POPULATION OF TRACT 1903, COUNTY 510, MARYLAND, INNER CITY CMHC
TOTAL POPULATION 3735 MALES 1796 FEMALES 1939

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
11	0.29	85+			0.29	11
11	0.29	80-4		F F	0.70	26
9	0.24	75-9		F F	0.99	37
36	0.96	70-4	M M	F F F F	1.39	52
72	1.93	65-9	M M M M M	F F F F F F	2.09	78
106	2.84	60-4	M M M M M M M M	F F F F F F	2.01	75
105	2.81	55-9	M M M M M M M M	F F F F F F F F	2.92	109
106	2.84	50-4	M M M M M M M M	F F F F F F F F	3.00	112
99	2.65	45-9	M M M M M M M M	F F F F F F F F	3.24	121
105	2.81	40-4	M M M M M M M M	F F F F F F F F	2.70	101
84	2.25	35-9	M M M M M M	F F F F F F	2.17	81
85	2.28	30-4	M M M M M M	F F F F F F	2.36	88
119	3.19	25-9	M M M M M M M M M	F F F F F F F F F	3.37	126
117	3.13	20-4	M M M M M M M M M	F F F F F F F F F F	4.31	161
151	4.04	15-9	M M M M M M M M M M	F F F F F F F F F F	4.07	152
179	4.79	10-4	M M M M M M M M M M M M M M	F F F F F F F F F F F F F F	5.19	194
186	4.98	5-9	M M M M M M M M M M M M M M	F F F F F F F F F F F F F F	5.01	187
215	5.76	0-4	M M M M M M M M M M M M M M M M M	F F F F F F F F F F F F F F F F F	6.10	228

Fig. 9. **AN EXAMPLE OF A SET I AGE-SEX PYRAMID**

WHITE POPULATION OF TRACT 105, COUNTY 510, MARYLAND, BALTIMORE CITY HOSPITALS
TOTAL POPULATION 2754 MALES 1362 FEMALES 1392

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
5	0.18	85+			0.25	7
20	0.73	80-4	M M	F F	0.80	22
23	0.84	75-9	M M	F F F F	1.63	45
33	1.20	70-4	M M M	F F F F F	1.85	51
50	1.82	65-9	M M M M M	F F F F F F F	2.54	70
81	2.94	60-4	M M M M M M M M	F F F F F F F F F	3.12	86
86	3.12	55-9	M M M M M M M M	F F F F F F F F F	3.30	91
92	3.34	50-4	M M M M M M M M M	F F F F F F F F F	3.01	83
87	3.16	45-9	M M M M M M M M M	F F F F F F F F F	3.45	95
85	3.09	40-4	M M M M M M M M M	F F F F F F F F F	3.23	89
81	2.94	35-9	M M M M M M M M	F F F F F F	2.29	63
73	2.65	30-4	M M M M M M M	F F F F F F	2.32	64
77	2.80	25-9	M M M M M M M	F F F F F F F	2.58	71
131	4.76	20-4	M M M M M M M M M M M M M M	F F F F F F F F F F F F F F	4.90	135
97	3.52	15-9	M M M M M M M M M M M	F F F F F F F F F F	3.52	97
115	4.18	10-4	M M M M M M M M M M M M	F F F F F F F F F F F F	3.89	107
116	4.21	5-9	M M M M M M M M M M M M	F F F F F F F F F F F F	3.78	104
110	3.99	0-4	M M M M M M M M M M M M	F F F F F F F F F F F F	4.07	112

Fig. 10. **AN EXAMPLE OF A SET J AGE-SEX PYRAMID**

WHITE POPULATION OF TRACT 202, COUNTY 510, MARYLAND, BALTIMORE CITY HOSPITALS
TOTAL POPULATION 2763 MALES 1399 FEMALES 1364

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
7	0.25	85+			0.22	6
11	0.40	80-4	M	F F	0.80	22
28	1.01	75-9	M M M	F F F	1.05	29
55	1.99	70-4	M M M M M	F F F F F F	2.06	57
57	2.06	65-9	M M M M M M	F F F F F F	2.14	59
107	3.87	60-4	M M M M M M M M M M M M	F F F F F F F	2.50	69
97	3.51	55-9	M M M M M M M M M M M	F F F F F F F F F F F	3.73	103
137	4.96	50-4	M M M M M M M M M M M M M M M	F F F F F F F F F	3.08	85
110	3.98	45-9	M M M M M M M M M M M M	F F F F F F F F F F F	3.87	107
92	3.33	40-4	M M M M M M M M M	F F F F F F F F F	3.11	86
64	2.32	35-9	M M M M M M	F F F F F F F F	2.68	74
44	1.59	30-4	M M M M	F F F F F	2.14	59
59	2.14	25-9	M M M M M M	F F F F F F F	2.97	82
121	4.38	20-4	M M M M M M M M M M M	F F F F F F F F	3.08	85
96	3.47	15-9	M M M M M M M M M M	F F F F F F F F F F F F F	4.74	131
103	3.73	10-4	M M M M M M M M M M M M	F F F F F F F F F F F F	4.05	112
106	3.84	5-9	M M M M M M M M M M M M	F F F F F F F F F F F F	3.76	104
105	3.80	0-4	M M M M M M M M M M M M	F F F F F F F F F F	3.40	94

Fig. 11. **AN EXAMPLE OF A SET K AGE-SEX PYRAMID**

WHITE POPULATION OF TRACT 2603.03, COUNTY 510, MARYLAND, AREA VII
TOTAL POPULATION 1368 MALES 619 FEMALES 749

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
2	0.15	85+		F	0.37	5
6	0.44	80-4	M	FF	0.73	10
3	0.22	75-9		FF	0.88	12
12	0.88	70-4	MM	FFFF	1.39	19
15	1.10	65-9	MMM	FFF	1.17	16
22	1.61	60-4	MMMM	FFFFF	1.75	24
19	1.39	55-9	MMMM	FFFFFFFFFF	3.36	46
26	1.90	50-4	MMMMM	FFFFFFFF	2.78	38
37	2.70	45-9	MMMMMMMM	FFFFFFFF	2.70	37
23	1.68	40-4	MMMMM	FFFFFFFF	3.00	41
25	1.83	35-9	MMMMM	FFFFFFFFF	3.22	44
21	1.54	30-4	MMMM	FFFFFFFFF	3.14	43
30	2.19	25-9	MMMMMM	FFFFFFFF	3.00	41
39	2.85	20-4	MMMMMMMM	FFFFFFFFFF	3.44	47
92	6.73	15-9	MMMMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFF	6.14	84
96	7.02	10-4	MMMMMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFFFF	7.24	99
90	6.58	5-9	MMMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFF	5.85	80
61	4.46	0-4	MMMMMMMMMMMM	FFFFFFFFFFFF	4.61	63

Fig. 12. **AN EXAMPLE OF A SET L AGE-SEX PYRAMID**

WHITE POPULATION OF TRACT 1206, COUNTY 510, MARYLAND, NORTH CENTRAL CMHC
TOTAL POPULATION 2415 MALES 1226 FEMALES 1189

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
7	0.29	85+		FF	0.99	24
22	0.91	80-4	MM	FFFF	1.49	36
37	1.53	75-9	MMMM	FFFFFF	2.11	51
40	1.66	70-4	MMMM	FFFFFFFF	2.90	70
65	2.69	65-9	MMMMMMMM	FFFFFFFFF	3.23	78
95	3.93	60-4	MMMMMMMMMM	FFFFFFFFFF	3.64	88
79	3.27	55-9	MMMMMMMMM	FFFFFFFFFFF	3.89	94
80	3.31	50-4	MMMMMMMMM	FFFFFFFFFFFF	4.02	97
75	3.11	45-9	MMMMMMMMM	FFFFFFFF	2.77	67
65	2.69	40-4	MMMMMMMM	FFFFFFF	2.44	59
57	2.36	35-9	MMMMMMM	FFFFF	1.95	47
67	2.77	30-4	MMMMMMMM	FFF	1.20	29
118	4.89	25-9	MMMMMMMMMMMM	FFFFFFFFFF	3.69	89
197	8.16	20-4	MMMMMMMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFF	6.00	145
78	3.23	15-9	MMMMMMMMM	FFFFFFFF	3.31	80
46	1.90	10-4	MMMMM	FFFF	1.49	36
46	1.90	5-9	MMMMM	FFFF	1.57	38
52	2.15	0-4	MMMMMM	FFFFFF	2.53	61

populations but also of other subpopulations (such as the aged, female-headed households with children, and populations living in overcrowded quarters) that may run a high risk of health and related problems.

In conclusion, we see the identification of distinctive residential subareas, which in turn have distinctive problems, as a first step in and an integral part of dealing preventively with social problems. In our view, the application of the social area procedures that we have outlined here will result in the identification of distinctive subareas. Thus, with access to the programs that provide the NIMH small area profiles, one can obtain relevant census data rapidly, at a low cost, and in a useful format for effective planning. Given the profiles, one can begin to analyze local census data by applying the procedures suggested in this paper.

References

Foote, Nelson N., Janet L. Abu-Lughod, Mary M. Foley, and Louis Winnick.
1960 Housing Choices and Housing Constraints. New York: McGraw-Hill.

Glick, Paul C.
1957 American Families. New York: John Wiley.

Goldsmith, Harold F., Edward G. Stockwell, and Elizabeth Unger.
1970 A Technique for Classifying Population Age Profiles. Laboratory Paper No. 33, Adelphi, Maryland: Mental Health Study Center, National Institute of Mental Health.

Goldsmith, Harold F., and Elizabeth L. Unger.
1970 Differentiation of Urban Subareas: A Re-Examination of Social Area Dimensions. Laboratory Paper No. 35, Adelphi, Maryland: Mental Health Study Center, National Institute of Mental Health.

1972a Social Areas: Identification Procedures Using 1970 Census Data. Laboratory Paper No. 37, Adelphi, Maryland: Mental Health Study Center, National Institute of Mental Health.

1972b Social Rank and Family Life Cycle: An Ecological Analysis. Laboratory Paper No. 43, Adelphi, Maryland: Mental Health Study Center, National Institute of Mental Health.

Redick, Richard W., Harold F. Goldsmith, and Elizabeth L. Unger.
1971 1970 Census Data Used to Indicate Areas with Different Potentials for Mental Health and Related Problems. National Institute of Mental Health Statistics Methodology Reports, Series C., No. 3. Washington, D.C.: U.S. Government Printing Office.

U.S. Bureau of the Census.
1972 Census of Population: 1970, General Social and Economic Characteristics, Final Report PC(1)-C1, United States Summary. Washington, D.C.: U.S. Government Printing Office.

Table 1. Similarity and Disparity Between Family Life Cycle Classifications of Census Tracts Produced by Visual Code and 1972 Age Code (Specific Measure of Comparison: First Two Visual Code Designations) by Ethnic Status and Family Status: Poor Census Tracts-Baltimore City, 1970

Selected poor census tracts[1] by ethnic status and family status[2]	Number of tracts	First 2 visual code designations occur as first two 1972 age code designations	One of first 2 visual code designations occurs in first two 1972 age designations		First two visual code designations do not occur among 1972 age code designations
			Remaining visual code designation occurs after first two age code designations	Remaining visual code designation does not occur among 1972 age code designations	
Selected poor tracts	50	42	6	2	-
Poor Negro tracts	23	20	3	-	-
Family areas	10	10	-	-	-
Family/nonfamily areas	12	9	3	-	-
Nonfamily areas	1	1	-	-	-
Poor white tracts	27	22	3	2	-
Family areas	22	19	2	1	-
Family/nonfamily areas	4	2	1	1	-
Nonfamily areas	1	1	-	-	-

[1]See text for designations.
[2]Family status classification:
Family areas--less than 30% of households headed by primary individuals.
Family/Nonfamily areas--30-49% of households headed by primary individuals.
Nonfamily areas--50% or more of households headed by primary individuals.

SOURCE: Supplementary appendix tables 1 and 2.

Table 2. Similarity and Disparity Between Family Life Cycle Classifications of Census Tracts Produced by 4-Item Disk Code and 1972 Age Code (Specific Measure of Comparison: First Two Four-Item Disk Code Designations) by Ethnic Status and Family Status: Poor Census Tracts-Baltimore City, 1970

Selected poor census tracts[1] by ethnic status and family status[2]	Number of tracts	First 2-disk code designations occur as first two 1972 age code designations	One of first two 4-item disk code designations occurs in first two 1972 age designations		First 2-disk code designations do not occur among 1972 age code designations
			Remaining disk code designation occurs after first two age code designations	Remaining disk code designation does not occur among 1972 age code designations	
Selected poor tracts	50	11	32	6	1
Poor Negro tracts	23	6	17	-	-
Family areas	10	5	5	-	-
Family/nonfamily areas	12	1	11	-	-
Nonfamily areas	1	-	1	-	-
Poor white tracts	27	5	15	6	1
Family areas	22	4	13	5	-
Family/nonfamily areas	4	1	2	1	-
Nonfamily areas	1	-	-	-	1

[1]See text for designations.
[2]Family status classification:
Family areas--less than 30% of households headed by primary individuals.
Family/Nonfamily areas--30-49% of households headed by primary individuals.
Nonfamily areas--50% or more of households headed by primary individuals.

SOURCE: Supplementary appendix tables 1 and 2.

APPENDIX TABLE 1
PRINCE GEORGE'S COUNTY, MARYLAND
SELECTED STATISTICS FROM 1970 CENSUS OF POPULATION AND HOUSING, SECOND COUNT (100% SAMPLE) DATA FILES

STATISTIC DESCRIPTION	NUMBER, PERCENT, RATIO OR VALUE FOR THE TOTAL AREA	DENOMINATOR FOR PERCENTS AND RATIOS OR POPULATION FOR MEDIANS AND AVERAGES	COMPARISONS WITH OTHER AREAS: SMSA RATIO OR VALUE (A)	STATE RATIO OR VALUE	U.S. RATIO OR VALUE	PERCENT OF VALUES LESS THAN THE COUNTY (C)	MEDIAN VALUE ALL COUNTIES IN MARYLAND (C)
GENERAL POPULATION DATA							
(1) TOTAL POPULATION	660567	...	2861123	3922399	203211905	92.0	53764
(2) NUMBER OF MALES (IN HOUSEHOLDS)	316987	...	1331512	1851019	95456663	92.0	24804
(3) NUMBER OF FEMALES (IN HOUSEHOLDS)	327069	...	1443391	1966563	101943250	92.0	26225
(4) POPULATION IN GROUP QUARTERS	16511	...	86220	104817	5812013	88.0	1589
(5) POPULATION WHITE	561476	...	2124903	3194888	177748975	92.0	46423
(6) POPULATION NEGRO	91808	...	703745	699479	22580289	92.0	7424
SOCIOECONOMIC STATUS							
(7) HOUSE VALUE: MEDIAN VALUE OF OWNER-OCCUPIED AND VACANT FOR SALE HOUSING UNITS.	$23753	91656	$28542	$18776	$17450	88.0	$15849
(8) RENT: MEDIAN RENT OF RENTER-OCCUPIED AND VACANT FOR RENT HOUSING UNITS.	$143	97629	$140	$110	$ 90 (B)	92.0	$ 74
ETHNIC COMPOSITION							
(9) NEGRO: PERCENT OF POPULATION NEGRO (IN HOUSEHOLDS)	13.9	644056	24.6	17.7	11.1	46.0	15.9
(10) OTHER NONWHITE RACES: PERCENT OF POPULATION NEITHER WHITE NOR NEGRO (IN HOUSEHOLDS)	1.1	644056	1.1	0.7	1.4	92.0	0.2
HOUSEHOLD COMPOSITION AND FAMILY STRUCTURE							
(11) HUSBAND-WIFE HOUSEHOLDS: PERCENT OF ALL HOUSEHOLDS HUSBAND-WIFE	75.1	192962	64.8	70.9	69.4	54.0	74.3
(12) AGE OF HOUSEHOLD HEADS: MEDIAN AGE OF HOUSEHOLD HEAD	39.2	192962	45.0	45.7	48.2	4.0	48.7
(13) YOUTH DEPENDENCY RATIO: PERSONS UNDER 18 PER 100 PERSONS 18-64 (IN HOUSEHOLDS)	65.3	374295	60.1	63.4	63.4	63.0	64.1
(14) AGED DEPENDENCY RATIO: PERSONS 65 AND OVER PER 100 PERSONS 18-64 (IN HOUSEHOLDS)	6.8	374295	9.8	13.1	17.4	0.0	17.2
TYPE OF HOUSING (URBANIZATION)							
(15) ONE UNIT STRUCTURES: PERCENT OF OCCUPIED AND VACANT YEAR-ROUND HOUSING UNITS WITH ONE UNIT AT ADDRESS	55.6	200211	53.9	69.0	69.4	0.0	0.0
CONDITION OF HOUSING							
(16) OVERCROWDING: PERCENT OF PERSONS IN HOUSEHOLDS WITH 1.01 OR MORE PERSONS PER ROOM	12.5	644056	13.3	13.6	16.9	25.0	14.5
(17) STANDARD HOUSING: PERCENT OF OCCUPIED HOUSING UNITS WITH DIRECT ACCESS AND WITH COMPLETE PLUMBING AND KITCHEN FACILITIES FOR EXCLUSIVE USE	98.2	192962	97.8	95.4	93.4	92.0	88.9

Symbols apply to this and all subsequent tables.

... Category not applicable.
NA Not available.

(A) Medians estimated from SMSA county medians and county populations.
(B) Occupied housing units.
(C) Based on 23 counties and Baltimore city.

APPENDIX TABLE 2
PRINCE GEORGE'S COUNTY, MARYLAND
ADDITIONAL STATISTICS FROM 1970 CENSUS OF POPULATION AND HOUSING, SECOND COUNT (100% SAMPLE) DATA FILES

STATISTIC DESCRIPTION	NUMBER, PERCENT, RATIO OR VALUE FOR THE TOTAL AREA	DENOMINATOR FOR PERCENTS AND RATIOS OR POPULATION FOR MEDIANS AND AVERAGES	COMPARISONS WITH OTHER AREAS: SMSA RATIO OR VALUE (A)	STATE RATIO OR VALUE	U.S. RATIO OR VALUE	PERCENT OF VALUES LESS THAN THE COUNTY (B)	MEDIAN VALUE ALL COUNTIES IN MARYLAND (B)
SOCIOECONOMIC STATUS							
(1) LOW HOUSE VALUE, NON-NEGRO: LOWER QUARTILE HOUSE VALUE OF NON-NEGRO HEADED, OWNER-OCCUPIED HOUSING UNITS	$19455	79799	$20015	$12581	$11265	88.0	$11004
(2) LOW HOUSE VALUE, NEGRO: LOWER QUARTILE HOUSE VALUE OF NEGRO HEADED, OWNER-OCCUPIED HOUSING UNITS	$15991	10816	$15590	$ 7433	$6330	92.0	$ 5651
(3) LOW RENT, NON-NEGRO: LOWER QUARTILE MONTHLY RENTAL OF NON-NEGRO HEADED, RENTER-OCCUPIED HOUSING UNITS	$122	82107	$111	$ 85	$65	92.0	$ 58
(4) LOW RENT, NEGRO: LOWER QUARTILE MONTHLY RENTAL OF NEGRO HEADED, RENTER-OCCUPIED HOUSING UNITS	$103	10740	$ 78	$ 62	$49	92.0	$ 36
HOUSEHOLD COMPOSITION AND FAMILY STRUCTURE							
GENERAL CHARACTERISTICS							
(5) MEDIAN HOUSEHOLD SIZE	3.1	192962	2.9	2.9	2.7	58.0	2.9
(6) SMALL HOUSEHOLDS: PERCENT OF HOUSEHOLDS WITH ONE PERSON	12.0	192962	19.4	14.9	17.6	29.0	13.4
(7) LARGE HOUSEHOLDS: PERCENT OF HOUSEHOLDS WITH 6 OR MORE PERSONS	11.1	192962	9.9	11.3	10.4	50.0	11.0
(8) SEX RATIO: MALES PER 100 FEMALES (POPULATION IN HOUSEHOLDS)	96.9	327069	92.3	94.1	93.6	58.0	95.5
(9) FERTILITY RATIO: CHILDREN UNDER 5 PER 1000 FEMALES 15-44 YEARS (HOUSEHOLD POPULATION)	443.4	155706	387.7	412.6	415.7	71.0	420.9
(10) CHILDREN LIVING WITH BOTH PARENTS: PERCENT OF PERSONS UNDER 18 IN HOUSEHOLDS WHO ARE OWN CHILDREN IN HUSBAND-WIFE HOUSEHOLDS	86.0	244446	81.2	81.0	82.6	75.0	83.8
FAMILY LIFE CYCLE							
(12) FAMILIES WITH CHILDREN: PERCENT OF FAMILIES WITH OWN CHILDREN UNDER 18	63.5	163379	59.7	57.4	54.9	88.0	55.3
(13) CHILDBEARING ONLY FAMILIES: PERCENT OF FAMILIES WITH OWN CHILDREN UNDER 6, NO CHILDREN 6-17	18.4	163379	15.3	13.5	12.7	92.0	11.9
(14) CHILDBEARING AND REARING FAMILIES: PERCENT OF FAMILIES WITH OWN CHILDREN BOTH UNDER 6 AND 6-17	15.7	163379	14.1	14.1	13.4	71.0	13.6
(15) CHILDREARING ONLY FAMILIES: PERCENT	29.3	163379	30.3	29.8	28.7	67.0	28.3
(16) CHILDREARING COMPLETED: PERCENT OF HUSBAND-WIFE FAMILIES WITH HUSBAND 45 OR OVER AND WITH NO CHILDREN PRESENT	20.7	144880	26.0	30.7	34.7	4.0	35.5
PERSONS NOT IN FAMILIES							
(17) GROUP QUARTERS: PERCENT OF TOTAL POPULATION LIVING IN GROUP QUARTERS	2.5	660567	3.0	2.7	2.9	54.0	2.3
(18) HOUSEHOLDS WITH PRIMARY INDIVIDUALS: PERCENT OF HOUSEHOLDS WITH HEADS WHO ARE PRIMARY INDIVIDUALS	15.3	192962	23.3	17.3	19.7	46.0	15.5
(19) NON-RELATIVES: PERCENT OF PERSONS IN HOUSEHOLDS WHO ARE NON-RELATIVES OF THE HOUSEHOLD HEAD	2.2	644056	2.9	2.0	1.6	79.0	1.7
DISRUPTED FAMILIES							
(20) DIVORCED OR SEPARATED MALES: PERCENT OF MALES 14 AND OVER WHO ARE DIVORCED OR SEPARATED	4.0	224616	5.5	5.0	4.3	38.0	4.1
(21) DIVORCED OR SEPARATED FEMALES: PERCENT OF FEMALES 14 AND OVER WHO ARE DIVORCED OR SEPARATED	6.3	237020	8.2	7.0	6.2	79.0	5.2
(22) WIDOWED FEMALES: PERCENT OF FEMALES 14 AND OVER WHO ARE WIDOWS	7.5	237020	9.6	11.1	12.4	0.0	12.0
(23) FEMALE HEADED HOUSEHOLDS: PERCENT OF HOUSEHOLDS WITH FEMALE HEAD	15.9	192962	23.4	20.0	21.0	33.0	17.1
(24) FEMALE HEADED HOUSEHOLDS WITH OWN CHILDREN: PERCENT OF HOUSEHOLDS WITH OWN CHILDREN THAT ARE HEADED BY FEMALES	9.0	103756	12.4	11.2	10.7	50.0	8.7
SINGLE ADULTS							
(25) SINGLE MALES: PERCENT OF MALES 25 AND OVER WHO HAVE NEVER MARRIED	7.8	156744	11.3	9.1	8.9	25.0	8.8
(26) SINGLE FEMALES: PERCENT OF FEMALES 25 AND OVER WHO HAVE NEVER MARRIED	5.6	163183	10.2	7.1	7.0	21.0	6.7
TYPE OF HOUSING (URBANIZATION)							
(27) LOW OWNER-OCCUPANCY: PERCENT OF OCCUPIED HOUSING UNITS WHICH ARE RENTER-OCCUPIED	49.9	192962	54.0	41.2	37.1	92.0	32.1
(28) TRAILERS: PERCENT OF ALL HOUSING UNITS THAT ARE MOBILE HOMES OR TRAILERS	0.7	200211	0.7	1.5	2.7	8.0	3.5

SEE FOOTNOTES AT END OF TABLE.

APPENDIX TABLE 2--CONTINUED
PRINCE GEORGE'S COUNTY, MARYLAND
ADDITIONAL STATISTICS FROM 1970 CENSUS OF POPULATION AND HOUSING, SECOND COUNT (100% SAMPLE) DATA FILES

STATISTIC DESCRIPTION	NUMBER, PERCENT, RATIO OR VALUE FOR THE TOTAL AREA	DENOMINATOR FOR PERCENTS AND RATIOS OR POPULATION FOR MEDIANS AND AVERAGES	COMPARISONS WITH OTHER AREAS: SMSA RATIO OR VALUE (A)	STATE RATIO OR VALUE	U.S. RATIO OR VALUE	PERCENT OF VALUES LESS THAN THE COUNTY (B)	MEDIAN VALUE ALL COUNTIES IN MARYLAND (B)
CONDITION OF HOUSING							
HOUSING							
(29) VACANCY INDEX: PERCENT OF TOTAL HOUSING UNITS VACANT YEAR-ROUND	3.6	200211	4.2	4.8	6.2	13.0	6.2
(30) STANDARD HOUSING, NON-NEGRO: PERCENT OF NON-NEGRO HEADED, OCCUPIED HOUSING UNITS WITH DIRECT ACCESS AND WITH COMPLETE PLUMBING AND KITCHEN FACILITIES FOR EXCLUSIVE USE	98.8	169792	98.3	96.6	94.6	92.0	92.4
(31) STANDARD HOUSING, NEGRO: PERCENT OF NEGRO HEADED, OCCUPIED HOUSING UNITS WITH DIRECT ACCESS AND WITH COMPLETE PLUMBING AND KITCHEN FACILITIES FOR EXCLUSIVE USE	93.5	23170	95.8	89.2	82.0	88.0	58.5
DENSITY							
(32) NON-NEGRO POPULATION IN OVERCROWDED HOUSING: PERCENT OF THE NON-NEGRO HEADED HOUSEHOLD POPULATION IN HOUSING WITH 1.01 OR MORE PERSONS PER ROOM	10.0	554364	7.5	9.7	14.3	42.0	10.3
(33) NEGRO POPULATION IN OVERCROWDED HOUSING: PERCENT OF THE NEGRO HEADED HOUSEHOLD POPULATION IN HOUSING WITH 1.01 OR MORE PERSONS PER ROOM	27.9	89692	31.1	31.3	38.1	17.0	35.6
(34) OVERCROWDED HOUSING: PERCENT OF OCCUPIED UNITS WITH 1.01 OR MORE PERSONS PER ROOM	6.4	192962	6.9	6.6	8.2	33.0	7.0
(35) PERSONS IN HIGHLY OVERCROWDED HOUSING: PERCENT OF THE HOUSEHOLD POPULATION IN HOUSING WITH 1.51 OR MORE PERSONS PER ROOM	2.7	644056	3.9	3.0	5.0	25.0	3.8
GROUP QUARTERS							
(36) INMATES OF INSTITUTIONS: PERCENT OF POPULATION IN GROUP QUARTERS WHO ARE INMATES OF INSTITUTIONS	14.5	16511	23.2	37.9	36.5	13.0	42.0

(A) Medians estimated from SMSA county medians and county populations. (B) Based on 23 counties and Baltimore city.

APPENDIX TABLE 3
PRINCE GEORGE'S COUNTY, MARYLAND
SELECTED STATISTICS FROM 1970 CENSUS OF POPULATION AND HOUSING, SECOND PLUS FOURTH COUNT (ALL FILES)

STATISTIC DESCRIPTION	NUMBER, PERCENT, RATIO OR VALUE FOR THE TOTAL AREA	DENOMINATOR FOR PERCENTS AND RATIOS OR POPULATION FOR MEDIANS AND AVERAGES	COMPARISONS WITH OTHER AREAS: SMSA RATIO OR VALUE (A)	STATE RATIO OR VALUE	U.S. RATIO OR VALUE	PERCENT OF VALUES LESS THAN THE COUNTY (B)	MEDIAN VALUE OF ALL COUNTIES IN MARYLAND (B)
GENERAL POPULATION DATA							
(1) TOTAL POPULATION	660567	...	2861123	3922399	203211905	92	53764
(2) NUMBER OF MALES (IN HOUSEHOLDS)	316987	...	1331512	1851019	95456663	92	24804
(3) NUMBER OF FEMALES (IN HOUSEHOLDS)	327069	...	1443391	1966563	101943250	92	26225
(4) POPULATION IN GROUP QUARTERS	16511	...	86220	104817	5812013	88	1589
(5) POPULATION WHITE	561476	...	2124903	3194888	177748975	92	46423
(6) POPULATION NEGRO	91808	...	703745	699479	22580289	92	7424
SOCIOECONOMIC STATUS							
ECONOMIC STATUS							
(7) INCOME OF FAMILIES AND UNRELATED INDIVIDUALS: MEDIAN INCOME OF FAMILIES AND UNRELATED INDIVIDUALS	$10467	218761	$10079	$ 9130	$ 7,699	83	0.0
(8) FAMILIES IN POVERTY: PERCENT OF ALL FAMILIES BELOW POVERTY LEVEL	4.3	163400	6.1	7.7	10.7	13	11.9
SOCIAL STATUS							
(9) LOW OCCUPATIONAL STATUS, MALES: PERCENT OF EMPLOYED MALES 16 AND OVER WHO ARE OPERATIVES, SERVICE WORKERS, AND LABORERS INCLUDING FARM LABORERS	22.6	160697	23.5	30.3	36.0	8	37.2
(10) HIGH OCCUPATIONAL STATUS, MALES: PERCENT OF EMPLOYED MALES 16 AND OVER WHO ARE PROFESSIONALS, TECHNICAL AND KINDRED WORKERS, AND MANAGERS EXCEPT FARM	37.2	160697	42.2	31.4	25.4	88	20.6
EDUCATIONAL STATUS							
(11) SCHOOL YEARS COMPLETED: MEDIAN SCHOOL YEARS COMPLETED BY PERSONS 25 AND OVER	12.5	319843	12.6	12.1	12.1	92	11.0
ETHNIC COMPOSITION							
(12) NEGRO: PERCENT OF HOUSEHOLD POPULATION NEGRO	13.9	644056	24.6	17.7	11.1	46	15.9
(13) OTHER NONWHITE: PERCENT OF HOUSEHOLD POPULATION NONWHITE AND NON-NEGRO	1.1	644056	1.1	0.7	1.4	92	0.2
(14) FOREIGN STOCK: PERCENT OF POPULATION WHO ARE FOREIGN BORN OR NATIVE BORN OF FOREIGN OR MIXED PARENTAGE	12.4	660564	13.5	11.6	16.5	88	5.0
HOUSEHOLD COMPOSITION AND FAMILY STRUCTURE							
(15) HUSBAND-WIFE HOUSEHOLDS: PERCENT OF ALL HOUSEHOLDS WITH HUSBAND-WIFE FAMILIES	75.1	192962	64.8	70.9	69.4	54	74.3
(16) AGE OF HOUSEHOLD HEADS: MEDIAN AGE OF HOUSEHOLD HEADS	39.2	192962	45.0	45.7	48.2	4	48.7
(17) YOUTH DEPENDENCY RATIO: PERSONS UNDER 18 PER 100 PERSONS 18-64 IN HOUSEHOLD POPULATION	65.3	374295	60.1	63.4	63.4	63	64.1
(18) AGED DEPENDENCY RATIO: PERSONS 65 AND OVER PER 100 PERSONS 18-64 IN HOUSEHOLD POPULATION	6.8	374295	9.8	13.1	17.4	0	17.2
TYPE OF HOUSING (URBANIZATION)							
(19) SINGLE DWELLING UNITS: PERCENT OF ALL YEAR-ROUND HOUSING UNITS THAT ARE SINGLE DETACHED (EXCLUDING MOBILE HOMES AND TRAILERS)	49.8	200179	43.1	51.1	(NA)	4	75.4
(20) HIGH RISE APARTMENTS: PERCENT OF ALL YEAR-ROUND HOUSING UNITS THAT ARE IN STRUCTURES OF 7 OR MORE STORIES	2.7	200185	10.0	2.8	(NA)	88	0.0
CONDITION OF HOUSING							
(21) OVERCROWDING: PERCENT OF PERSONS IN HOUSEHOLDS IN HOUSING UNITS WITH 1.01 OR MORE PERSONS PER ROOM	12.5	644056	13.3	13.6	16.9	25	14.5
(22) STANDARD HOUSING: PERCENT OF OCCUPIED HOUSING UNITS WITH DIRECT ACCESS/COMPLETE PLUMBING AND KITCHEN FACILITIES FOR EXCLUSIVE USE	98.2	192962	97.8	95.4	93.4	92	88.9
COMMUNITY INSTABILITY							
(23) RECENT MOVERS: PERCENT OF POPULATION WHO MOVED INTO PRESENT RESIDENCE 1969-1970	29.4	660564	27.8	22.6	23.5	92	20.5

(A) Medians estimated from SMSA county medians and county populations.
(B) Based on 23 counties and Baltimore city.

APPENDIX TABLE 4
PRINCE GEORGE'S COUNTY, MARYLAND
ADDITIONAL STATISTICS FROM 1970 CENSUS OF POPULATION AND HOUSING, SECOND PLUS FOURTH COUNT (ALL FILES)

STATISTIC DESCRIPTION	NUMBER, PERCENT, RATIO OR VALUE FOR THE TOTAL AREA	DENOMINATOR FOR PERCENTS AND RATIOS OR POPULATION FOR MEDIANS AND AVERAGES	COMPARISONS WITH OTHER AREAS: SMSA RATIO OR VALUE (A)	STATE RATIO OR VALUE	U.S. RATIO OR VALUE	PERCENT OF VALUES LESS THAN THE COUNTY (B)	MEDIAN VALUE OF ALL COUNTIES IN MARYLAND (B)
GENERAL POPULATION DATA							
SOCIOECONOMIC STATUS							
ECONOMIC STATUS							
INCOME							
(1) INCOME OF FAMILIES, WHITE: MEDIAN INCOME OF WHITE FAMILIES	$12748	141258	$14841	$11635	$ 9,961	88	$ 9333.
(2) INCOME OF FAMILIES, NEGRO: MEDIAN INCOME OF NEGRO FAMILIES	$10624	20578	$ 8513	$ 7701	$ 6,067	96	$ 6431.
(3) INCOME OF UNRELATED INDIVIDUALS: MEDIAN INCOME OF UNRELATED INDIVIDUALS 14 AND OVER	$ 3809	55361	$ 4539	$ 3099	$ 2,489	83	$ 2207.
(4) INCOME OF UNRELATED INDIVIDUALS, WHITE: MEDIAN INCOME OF WHITE UNRELATED INDIVIDUALS 14 AND OVER	$ 3859	49813	$ 5274	$ 3349	$ 2,568	83	$ 2352.
(5) INCOME OF UNRELATED INDIVIDUALS, NEGRO: MEDIAN INCOME OF NEGRO UNRELATED INDIVIDUALS 14 AND OVER	$ 3436	4789	$ 3839	$ 2325	$ 1,936	96	$ 1834.
(6) FAMILIES IN POVERTY, WHITE: PERCENT OF WHITE FAMILIES BELOW POVERTY LEVEL	3.8	141258	3.5	5.3	8.6	13	8.9
(7) FAMILIES IN POVERTY, NEGRO: PERCENT OF NEGRO FAMILIES BELOW POVERTY LEVEL	8.0	20578	14.6	20.9	29.8	4	24.5
(8) POPULATION IN POVERTY: PERCENT OF POPULATION BELOW POVERTY LEVEL	5.8	644486	8.3	10.1	13.7	13	15.4
(9) POPULATION IN POVERTY, WHITE: PERCENT OF WHITE POPULATION BELOW POVERTY LEVEL	5.0	548636	5.0	6.9	10.9	13	11.3
(10) POPULATION IN POVERTY, NEGRO: PERCENT OF NEGRO POPULATION BELOW POVERTY LEVEL	10.4	89320	18.4	24.7	35.0	4	29.1
(11) HIGH INCOME FAMILIES: UPPER QUARTILE FAMILY INCOME	$18311	163400	$22103	$16674	$14176	88	$12805.
VALUE OF HOUSING							
(12) HOUSE VALUE, NON-NEGRO: MEDIAN VALUE OF NON-NEGRO OCCUPIED HOUSING UNITS	$24170	79799	$32394	$19637	(NA)	88	$16484.
(13) HOUSE VALUE, NEGRO: MEDIAN VALUE OF NEGRO OWNER-OCCUPIED HOUSING UNITS	$19783	10816	$18732	$11107	(NA)	88	$ 9560.
(14) RENT, NON-NEGRO: MEDIAN MONTHLY RENTAL OF NON-NEGRO RENTER-OCCUPIED HOUSING UNITS	$145	82107	$145	$121	(NA)	92	$ 76.
(15) RENT, NEGRO: MEDIAN MONTHLY RENTAL OF NEGRO RENTER-OCCUPIED HOUSING UNITS	$130	10740	$ 98	$ 85	(NA)	88	$ 49.
EMPLOYMENT LEVEL AND LABOR FORCE PARTICIPATION							
(16) UNEMPLOYMENT: PERCENT OF LABOR FORCE 16 AND OVER UNEMPLOYED	2.2	293776	2.5	3.1	4.4	8	2.9
(17) UNEMPLOYMENT, WHITE: PERCENT OF WHITE LABOR FORCE 16 AND OVER UNEMPLOYED	2.1	251506	2.0	2.6	4.1	21	2.5
(18) UNEMPLOYMENT, NEGRO: PERCENT OF NEGRO LABOR FORCE 16 AND OVER UNEMPLOYED	3.2	39351	3.9	5.5	7.0	21	4.7
(19) UNDER EMPLOYMENT, MALES 25-64: PERCENT OF MALE LABOR FORCE 25-64 WHO WORKED LESS THAN 40 WEEKS IN 1969	3.7	139022	3.9	4.1	(NA)	29	4.6
(20) UNDER EMPLOYMENT, WHITE MALES 25-64: PERCENT OF WHITE MALE LABOR FORCE 25-64 WHO WORKED LESS THAN 40 WEEKS IN 1969	3.3	120425	3.2	3.5	(NA)	33	4.0
(21) UNDER EMPLOYMENT, NEGRO MALES 25-64: PERCENT OF NEGRO MALE LABOR FORCE 25-64 WHO WORKED LESS THAN 40 WEEKS IN 1969	6.0	17139	6.1	7.2	(NA)	38	7.4
(22) LABOR FORCE PARTICIPATION, FEMALES: PERCENT OF FEMALES 16 AND OVER IN THE LABOR FORCE (INCLUDING ARMED FORCES)	50.2	224804	50.3	44.4	41.4	92	42.0
(23) LABOR FORCE PARTICIPATION, WHITE FEMALES: PERCENT OF WHITE FEMALES 16 AND OVER IN THE LABOR FORCE (INCLUDING ARMED FORCES)	48.4	194410	47.5	42.6	40.6	96	39.4
(24) LABOR FORCE PARTICIPATION, NEGRO FEMALES: PERCENT OF NEGRO FEMALES 16 AND OVER IN THE LABOR FORCE (INCLUDING ARMED FORCES)	63.0	28092	58.9	53.1	47.5	96	51.1
SOCIAL STATUS							
(25) LOW OCCUPATIONAL STATUS, WHITE MALES: PERCENT OF EMPLOYED WHITE MALES 16 AND OVER WHO ARE OPERATIVES, SERVICE WORKERS, AND LABORERS INCLUDING FARM LABORERS	19.3	139433	15.7	25.0	33.2	4	30.1
(26) HIGH OCCUPATIONAL STATUS, WHITE MALES: PERCENT OF EMPLOYED WHITE MALES 16 AND OVER WHO ARE PROFESSIONALS, TECHNICAL AND KINDRED WORKERS, AND MANAGERS EXCEPT FARM	39.4	139433	49.9	34.7	27.0	88	25.35
(27) LOW OCCUPATIONAL STATUS, NEGRO MALES: PERCENT OF EMPLOYED NEGRO MALES 16 AND OVER WHO ARE OPERATIVES, SERVICE WORKERS, AND LABORERS INCLUDING FARM LABORERS	45.7	19698	50.1	61.4	64.9	8	68.0
(28) HIGH OCCUPATIONAL STATUS, NEGRO MALES: PERCENT OF EMPLOYED NEGRO MALES 16 AND OVER WHO ARE PROFESSIONALS, TECHNICAL AND KINDRED WORKERS, AND MANAGERS EXCEPT FARM	19.9	19698	15.6	11.3	8.9	88	7.8
(29) LOW OCCUPATIONAL STATUS, FEMALES: PERCENT OF EMPLOYED FEMALES 16 AND OVER WHO ARE OPERATIVES, SERVICE WORKERS, AND LABORERS INCLUDING FARM LABORERS	15.9	108708	20.8	28.8	36.2	4	42.3

SEE FOOTNOTES AT END OF TABLE.

APPENDIX TABLE 4--CONTINUED
PRINCE GEORGE'S COUNTY, MARYLAND
ADDITIONAL STATISTICS FROM 1970 CENSUS OF POPULATION AND HOUSING, SECOND PLUS FOURTH COUNT (ALL FILES)

STATISTIC DESCRIPTION	NUMBER, PERCENT, RATIO OR VALUE FOR THE TOTAL AREA	DENOMINATOR FOR PERCENTS AND RATIOS OR POPULATION FOR MEDIANS AND AVERAGES	COMPARISONS WITH OTHER AREAS: SMSA RATIO OR VALUE (A)	STATE RATIO OR VALUE	U.S. RATIO OR VALUE	PERCENT OF VALUES LESS THAN THE COUNTY (B)	MEDIAN VALUE OF ALL COUNTIES IN MARYLAND (B)
SOCIAL STATUS--CONTINUED							
(30) LOW OCCUPATIONAL STATUS, WHITE FEMALES: PERCENT OF EMPLOYED WHITE FEMALES 16 AND OVER WHO ARE OPERATIVES, SERVICE WORKERS, AND LABORERS INCLUDING FARM LABORERS	13.1	90810	41.7	22.5	32.7	4	32.5
(31) MID OCCUPATIONAL STATUS, WHITE FEMALES: PERCENT OF EMPLOYED WHITE FEMALES 16 AND OVER WHO ARE SALES, CLERICAL AND KINDRED, WORKERS, AND CRAFTSMEN, FOREMEN AND KINDRED WORKERS	62.5	90810	57.7	54.5	46.8	96	48.5
(32) LOW OCCUPATIONAL STATUS, NEGRO FEMALES: PERCENT OF EMPLOYED NEGRO FEMALES 16 AND OVER WHO ARE OPERATIVES, SERVICE WORKERS, AND LABORERS INCLUDING FARM LABORERS	30.4	16830	40.1	55.1	62.4	4	74.6
(33) MID OCCUPATIONAL STATUS, NEGRO FEMALES: PERCENT OF EMPLOYED NEGRO FEMALES 16 AND OVER WHO ARE SALES, CLERICAL AND KINDRED WORKERS, AND CRAFTSMEN, FOREMEN AND KINDRED WORKERS	46.3	16830	43.2	29.7	24.7	96	17.3
EDUCATIONAL STATUS							
(34) SCHOOL YEARS COMPLETED, WHITE: MEDIAN SCHOOL YEARS COMPLETED BY WHITE PERSONS 25 AND OVER	12.5	276886	12.9	12.2	12.1	92	11.4
(35) SCHOOL YEARS COMPLETED, NEGRO: MEDIAN SCHOOL YEARS COMPLETED BY NEGRO PERSONS 25 AND OVER	12.2	39413	11.4	9.9	9.8	96	9.0
(36) LOW EDUCATIONAL STATUS: PERCENT OF PERSONS 25 AND OVER WITH 8 YEARS OR LESS EDUCATION	15.1	319843	15.9	27.4	28.3	4	32.6
(37) LOW EDUCATIONAL STATUS, WHITE: PERCENT OF WHITE PERSONS 25 AND OVER WITH 8 YEARS OR LESS EDUCATION	14.2	276886	11.9	24.7	26.6	4	31.3
(38) LOW EDUCATIONAL STATUS, NEGRO: PERCENT OF NEGRO PERSONS 25 AND OVER WITH 8 YEARS OR LESS EDUCATION	21.9	39413	29.3	42.1	43.8	4	49.1
(39) HIGH SCHOOL COMPLETION, WHITE: PERCENT OF WHITE PERSONS 18 AND OVER WHO HAVE COMPLETED AT LEAST 4 YEARS OF HIGH SCHOOL	69.5	359682	74.4	58.2	(NA)	92	48.5
(40) HIGH SCHOOL COMPLETION, NEGRO: PERCENT OF NEGRO PERSONS 18 AND OVER WHO HAVE COMPLETED AT LEAST 4 YEARS OF HIGH SCHOOL	57.9	50476	48.0	36.3	(NA)	96	29.4
(41) HIGH SCHOOL COMPLETION, YOUNG WHITE ADULTS: PERCENT OF WHITE PERSONS 18-24 WHO HAVE COMPLETED AT LEAST 4 YEARS OF HIGH SCHOOL	72.7	82796	69.3	68.8	75.0	88	68.7
(42) HIGH SCHOOL COMPLETION, YOUNG NEGRO ADULTS: PERCENT OF NEGRO PERSONS 18-24 WHO HAVE COMPLETED AT LEAST 4 YEARS OF HIGH SCHOOL	65.4	11063	59.4	52.6	55.2	83	50.5
(43) HIGH EDUCATIONAL STATUS: PERCENT OF POPULATION 25 AND OVER WITH 4 OR MORE YEARS OF COLLEGE	17.1	319843	23.4	13.9	10.7	88	7.0
ETHNIC COMPOSITION							
(44) SOUTHERN OR EASTERN EUROPEAN STOCK: PERCENT OF POPULATION WHO ARE FOREIGN STOCK OF SOUTHERN OR EASTERN EUROPEAN BACKGROUND (POLAND, CZECHOSLOVAKIA, AUSTRIA, HUNGARY, U.S.S.R., AND ITALY)	3.3	660567	3.4	4.1	6.6	83	1.1
(45) SPANISH AMERICANS: PERCENT OF POPULATION WHO ARE OF SPANISH HERITAGE	2.2	660567	2.5	1.4	4.6	92	0.6
HOUSEHOLD COMPOSITION AND FAMILY STRUCTURE							
GENERAL CHARACTERISTICS							
(46) MEDIAN HOUSEHOLD SIZE	3.1	192962	2.9	2.9	2.7	58	2.9
(47) SMALL HOUSEHOLDS: PERCENT OF HOUSEHOLDS WITH ONLY ONE PERSON	12.0	192962	19.4	14.9	17.6	29	13.4
(48) LARGE HOUSEHOLDS: PERCENT OF HOUSEHOLDS WITH 6 OR MORE PERSONS	11.1	192962	9.9	11.3	10.4	50	11.0
(49) CHILDREN LIVING WITH THEIR PARENTS: PERCENT OF PERSONS UNDER 18 LIVING WITH BOTH PARENTS	86.0	244446	81.2	81.0	82.7	75	83.8
(50) SEX RATIO: MALES PER 100 FEMALES IN HOUSEHOLD POPULATION	96.9	327069	92.3	94.1	94.8	58	95.5
(51) FERTILITY RATIO: CHILDREN UNDER 5 PER 1000 FEMALES 15-44 IN HOUSEHOLD POPULATION	443.4	155706	387.7	412.6	404.2	71	420.9
(52) HUSBAND-WIFE HOUSEHOLDS, WHITE: PERCENT OF WHITE HOUSEHOLDS WITH HUSBAND-WIFE FAMILIES	75.3	167889	68.9	73.8	71.3	42	75.6
(53) HUSBAND-WIFE HOUSEHOLDS, NEGRO: PERCENT OF NEGRO HOUSEHOLDS WITH HUSBAND-WIFE FAMILIES	73.3	23170	52.4	54.9	52.6	96	55.3
FAMILY LIFE CYCLE							
(54) AGE OF HOUSEHOLD HEADS, WHITE: MEDIAN AGE OF WHITE HOUSEHOLD HEADS	39.7	167889	42.9	46.2	48.5	4	49.0
(55) AGE OF HOUSEHOLD HEADS, NEGRO: MEDIAN AGE OF NEGRO HOUSEHOLD HEADS	36.2	23170	42.0	43.5	45.5	4	48.0
(56) YOUTH DEPENDENCY RATIO, WHITE: WHITE PERSONS UNDER 18 PER 100 WHITE PERSONS 18-64 IN HOUSEHOLD POPULATION	62.3	323170	57.5	59.8	60.8	50	61.1

SEE FOOTNOTES AT END OF TABLE.

APPENDIX TABLE 4--CONTINUED
PRINCE GEORGE'S COUNTY, MARYLAND
ADDITIONAL STATISTICS FROM 1970 CENSUS OF POPULATION AND HOUSING, SECOND PLUS FOURTH COUNT (ALL FILES)

STATISTIC DESCRIPTION	NUMBER, PERCENT, RATIO OR VALUE FOR THE TOTAL AREA	DENOMINATOR FOR PERCENTS AND RATIOS OR POPULATION FOR MEDIANS AND AVERAGES	COMPARISONS WITH OTHER AREAS: SMSA RATIO OR VALUE (A)	STATE RATIO OR VALUE	U.S. RATIO OR VALUE	PERCENT OF VALUES LESS THAN THE COUNTY (B)	MEDIAN VALUE OF ALL COUNTIES IN MARYLAND (B)
FAMILY LIFE CYCLE--CONTINUED							
(57) YOUTH DEPENDENCY RATIO, NEGRO: NEGRO PERSONS UNDER 18 PER 100 NEGRO PERSONS 18-64 IN HOUSEHOLD POPULATION	87.1	46737	68.7	82.3	85.5	63	83.3
(58) AGED DEPENDENCY RATIO, WHITE: WHITE PERSONS 65 AND OVER PER 100 WHITE PERSONS 18-64 IN HOUSEHOLD POPULATION	7.1	323170	10.2	13.6	17.9	0	18.1
(59) AGED DEPENDENCY RATIO, NEGRO: NEGRO PERSONS 65 AND OVER PER 100 NEGRO PERSONS 18-64 IN HOUSEHOLD POPULATION	4.8	46737	8.7	10.7	13.6	4	15.1
(60) FAMILIES WITH CHILDREN: PERCENT OF FAMILIES WITH OWN CHILDREN UNDER 18	63.5	163379	59.7	57.4	54.9	88	55.3
(61) CHILDBEARING ONLY FAMILIES: PERCENT OF FAMILIES WITH OWN CHILDREN UNDER 6, NO CHILDREN 6-17	18.4	163379	15.3	13.5	12.7	92	11.9
(62) CHILDBEARING AND CHILDREARING FAMILIES: PERCENT OF FAMILIES WITH OWN CHILDREN BOTH UNDER 6 AND 6-17	15.7	163379	14.1	14.1	13.4	71	13.6
(63) CHILDREARING ONLY FAMILIES: PERCENT OF FAMILIES WITH OWN CHILDREN 6-17, NO CHILDREN UNDER 6	29.3	163379	30.3	29.8	28.7	67	28.3
(64) CHILDREARING COMPLETED: PERCENT OF HUSBAND-WIFE FAMILIES WITH HUSBAND 45 OR OVER AND WITH NO CHILDREN PRESENT	20.7	144880	26.0	30.7	(NA)	4	35.5
PERSONS NOT IN FAMILIES							
(65) GROUP QUARTERS: PERCENT OF TOTAL POPULATION WHO LIVE IN GROUP QUARTERS	2.5	660567	3.0	2.7	2.9	54	2.3
(66) INMATES OF INSTITUTIONS: PERCENT OF POPULATION IN GROUP QUARTERS WHO ARE INMATES OF INSTITUTIONS	14.5	16511	23.2	37.9	36.5	13	42.0
(67) INMATES OF MENTAL HOSPITALS: PERCENT OF POPULATION IN GROUP QUARTERS WHO ARE INMATES OF MENTAL HOSPITALS	0.2	16286	6.3	8.6	7.3	46	0.3
(68) GROUP QUARTERS EXCLUDING INSTITUTIONS: PERCENT OF TOTAL POPULATION WHO LIVE IN ROOMING HOUSES AND OTHER NONINSTITUTIONAL GROUP QUARTERS	2.1	660567	2.1	1.7	1.8	63	1.1
(69) HOUSEHOLD HEADS PRIMARY INDIVIDUALS: PERCENT OF HOUSEHOLDS WITH HEADS WHO ARE PRIMARY INDIVIDUALS	15.3	192962	23.3	17.3	19.7	46	15.5
(70) NON-RELATIVES: PERCENT OF PERSONS IN HOUSEHOLDS WHO ARE NOT RELATED TO THE HOUSEHOLD HEAD	2.2	644056	2.9	2.0	1.6	79	1.7
POPULATION NOT CURRENTLY MARRIED							
(71) SINGLE MALES: PERCENT OF MALES 25 AND OVER WHO HAVE NEVER MARRIED	7.8	156744	11.3	9.1	(NA)	25	8.8
(72) SINGLE FEMALES: PERCENT OF FEMALES 25 AND OVER WHO HAVE NEVER MARRIED	5.6	163183	10.2	7.1	(NA)	21	6.7
(73) DIVORCED OR SEPARATED MALES: PERCENT OF MALES 14 AND OVER WHO ARE DIVORCED OR SEPARATED	4.0	224616	5.5	5.0	4.3	38	4.1
(74) DIVORCED OR SEPARATED FEMALES: PERCENT OF FEMALES 14 AND OVER WHO ARE DIVORCED OR SEPARATED	6.3	237020	8.2	7.0	6.2	79	5.2
(75) WIDOWED FEMALES: PERCENT OF FEMALES 14 AND OVER WHO ARE WIDOWS	7.5	237020	9.6	11.1	12.4	0	12.0
NON HUSBAND-WIFE HOUSEHOLDS							
(76) FEMALE HEADED HOUSEHOLDS: PERCENT OF HOUSEHOLDS WITH FEMALE HEAD	15.9	192962	23.4	20.0	21.0	33	17.1
(77) FEMALE HEADED HOUSEHOLDS WITH OWN CHILDREN: PERCENT OF HOUSEHOLDS WITH OWN CHILDREN UNDER 18 THAT ARE HEADED BY FEMALES	9.0	103756	12.4	11.2	10.7	50	8.7
CONDITION OF HOUSING							
HOUSING							
(78) VACANCY INDEX: PERCENT OF ALL YEAR-ROUND HOUSING UNITS THAT ARE VACANT	3.6	200211	4.2	4.8	6.2	13	6.2
(79) STANDARD HOUSING, NON-NEGRO: PERCENT OF NON-NEGRO HOUSING UNITS WITH DIRECT ACCESS AND WITH COMPLETE PLUMBING AND KITCHEN FACILITIES FOR EXCLUSIVE USE	98.8	169792	98.3	96.6	(NA)	92	92.4
(80) STANDARD HOUSING, NEGRO: PERCENT OF NEGRO HOUSING UNITS WITH DIRECT ACESS AND WITH COMPLETE PLUMBING AND KITCHEN FACILITIES FOR EXCLUSIVE USE	93.5	23170	85.8	89.2	(NA)	88	58.5
DENSITY							
(81) OVERCROWDED HOUSING: PERCENT OF OCCUPIED HOUSING UNITS WITH 1.01 OR MORE PERSONS PER ROOM	6.4	192962	6.9	6.6	8.2	33	7.0
(82) NON-NEGRO POPULATION IN OVERCROWDED HOUSING: PERCENT OF NON-NEGRO HOUSEHOLD POPULATION IN HOUSING UNITS WITH 1.01 OR MORE MORE PERSONS PER ROOM	10.0	554364	7.5	9.7	14.3	42	10.3

SEE FOOTNOTES AT END OF TABLE.

APPENDIX TABLE 4--CONTINUED
PRINCE GEORGE'S COUNTY, MARYLAND
ADDITIONAL STATISTICS FROM 1970 CENSUS OF POPULATION AND HOUSING, SECOND PLUS FOURTH COUNT (ALL FILES)

STATISTIC DESCRIPTION	NUMBER, PERCENT, RATIO OR VALUE FOR THE TOTAL AREA	DENOMINATOR FOR PERCENTS AND RATIOS OR POPULATION FOR MEDIANS AND AVERAGES	COMPARISONS WITH OTHER AREAS: SMSA RATIO OR VALUE (A)	STATE RATIO OR VALUE	U.S. RATIO OR VALUE	PERCENT OF VALUES LESS THAN THE COUNTY (B)	MEDIAN VALUE OF ALL COUNTIES IN MARYLAND (B)
DENSITY--CONTINUED							
(83) NEGRO POPULATION IN OVERCROWDED HOUSING: PERCENT OF NEGRO HOUSEHOLD POPULATION IN HOUSING WITH 1.51 OR MORE PERSONS PER ROOM	27.9	89692	31.1	31.3	38.1	17	35.6
(84) PERSONS IN HIGHLY OVERCROWDED HOUSING: PERCENT OF HOUSEHOLD POPULATION IN HOUSING UNITS WITH 1.51 OR MORE PERSONS PER ROOM	2.7	644056	3.9	3.0	5.0	25	3.8
TYPE OF HOUSING(URBANIZATION)							
(85) RENTER OCCUPANCY: PERCENT OF OCCUPIED HOUSING UNITS THAT ARE RENTER OCCUPIED	49.9	192962	54.0	41.2	37.1	92	32.0
(86) TRAILERS: PERCENT OF ALL YEAR-ROUND HOUSING UNITS THAT ARE MOBILE HOMES OR TRAILERS	0.7	200211	0.7	1.5	2.7	8	3.5
(87) LARGE APARTMENT STRUCTURES: PERCENT OF ALL YEAR ROUND HOUSING UNITS THAT ARE IN STRUCTURES WITH 20 OR MORE UNITS	8.9	200179	17.3	5.9	(NA)	92	0.8
(88) SINGLE DWELLING UNITS, NON-NEGRO: PERCENT OF NON-NEGRO HOUSING UNITS THAT ARE SINGLE DETACHED	51.9	169988	52.2	57.6	(NA)	4	81.6
(89) SINGLE DWELLING UNITS, NEGRO: PERCENT OF NEGRO HOUSING UNITS THAT ARE SINGLE DETACHED	47.9	22974	18.2	29.1	(NA)	13	75.9
(90) RURAL POPULATION: PERCENT OF POPULATION	7.8	660567	7.6	23.4	26.5	67	
COMMUNITY INSTABILITY							
(91) RECENT MOVERS, WHITE: PERCENT OF WHITE POPULATION WHO MOVED INTO PRESENT RESIDENCE 1969-1970	28.3	562773	28.1	21.8	(NA)	92	21.1
(92) RECENT MOVERS, NEGRO: PERCENT OF NEGRO POPULATION WHO MOVED INTO PRESENT RESIDENCE 1969-1970	35.4	91390	26.1	25.5	(NA)	83	22.0
(93) MOBILE PERSONS: PERCENT OF POPULATION 5 YEARS AND OVER LIVING IN A DIFFERENT HOUSE THAN IN 1965	61.6	591440	58.3	48.5	47.0	96	43.1
(94) MIGRANTS: PERCENT OF POPULATION 5 AND OVER RESIDING IN A DIFFERENT COUNTY THAN IN 1965	34.3	591440	31.9	22.1	19.5	92	16.2
OTHER POPLUATIONS WITH HIGH POTENTIAL NEED FOR HEALTH, WELFARE AND RELATED SERVICES							
(95) TEENAGERS NOT IN SCHOOL: PERCENT OF POPULATION 14-17 NOT ENROLLED IN SCHOOL	4.3	47021	4.8	6.8	7.3	8	8.7
(96) TEENAGERS NOT IN SCHOOL, NEGRO: PERCENT OF NEGRO POPULATION 14-17 NOT ENROLLED IN SCHOOL	7.8	7177	8.9	11.2	(NA)	8	14.5
(97) WORKING MOTHERS OF CHILDREN UNDER 18: PERCENT OF WOMEN 16 AND OVER WITH CHILDREN OF THEIR OWN UNDER 18 WHO ARE IN THE LABOR FORCE	44.8	105566	44.6	42.0	40.8	58	43.2
(98) WORKING MOTHERS OF PRESCHOOL CHILDREN: PERCENT OF WOMEN 16 AND OVER WITH CHILDREN OF THEIR OWN UNDER 6 WHO ARE IN THE LABOR FORCE	35.6	57474	35.6	32.2	30.8	63	33.6
(99) AGED PERSONS LIVING ALONE: PERCENT OF HOUSEHOLDS THAT ARE 1 PERSON HOUSEHOLDS WITH HEAD 65 OR OVER	2.4	192962	4.3	5.3	(NA)	0	6.5
(100) AGED PERSONS IN POVERTY: PERCENT OF PERSONS 65 AND OVER BELOW POVERTY LEVEL	13.8	25492	0.0	21.9	19.2	4	27.4
(101) EXTREMELY CROWDED HOUSING UNITS LACKING PLUMBING FACILITIES: PERCENT OF OCCUPIED HOUSING UNITS WITH 1,51 OR MORE PERSONS PER ROOM AND WITHOUT COMPLETE PLUMBING FACILITIES	0.2	192962	21.8	0.3	(NA)	17	0.6
(103) FEMALE HEADED HOUSEHOLDS WITH OWN CHILDREN, NEGRO: PERCENT OF NEGRO HOUSEHOLDS WITH OWN CHILDREN THAT ARE HEADED BY FEMALES	13.1	16930	48.5	19.6	30.6	38	13.6
(104) LARGE HOUSEHOLDS WITH LOW INCOME: PERCENT OF HOUSEHOLDS WITH 6 OR MORE PERSONS THAT HAVE AN ANNUAL INCOME OF LESS THAN $7,000	11.9	21510	22.1	2 .3	(NA)	13	30.9
(105) FEMALE HEADED FAMILIES WITH CHILDREN IN POVERTY: PERCENT OF FAMILIES WITH AT LEAST ONE RELATED CHILD UNDER 18 THAT ARE FEMALE HEADED AND BELOW POVERTY LEVEL	2.1	107725	6.8	4.7	5.1	13	4.8
(106) DISABLED POPULATION: PERCENT OF PERSONS 16-64 NOT INMATES OF INSTITUTIONS AND NOT ATTENDING SCHOOL WHO ARE DISABLED OR HANDICAPPED	8.2	351414	11.2	9.9	11.2	21	9.2
(107) DISABLED POPULATION UNABLE TO WORK: PERCENT OF PERSONS 16-64 NOT INMATES OF INSTITUTIONS AND NOT ATTENDING SCHOOL WHO ARE DISABLED OR HANDICAPPED AND WHO ARE UNABLE TO WORK	2.6	351414	42.0	3.6	4.3	13	3.5
(108) CHILDREN IN POVERTY: PERCENT OF RELATED CHILDREN UNDER 18 BELOW POVERTY LEVEL	6.0	243269	32.2	11.5	15.1	13	15.4

(A) Medians estimated from SMSA county medians and county populations. (B) Based on 23 counties and Baltimore city.

App. Fig. 1A **TOTAL POPULATION OF TRACT 8052.01, COUNTY 33, MARYLAND, COMPREHENSIVE CMHC I**
TOTAL POPULATION 4674 MALES 2209 FEMALES 2465

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
14	0.30	85+			0.26	12
15	0.32	80-4		F	0.47	22
25	0.53	75-9	M	FF	0.73	34
43	0.92	70-4	MM	FFFFF	1.71	80
71	1.52	65-9	MMMM	FFFFF	1.71	80
89	1.90	60-4	MMMMM	FFFFFFF	2.40	112
96	2.05	55-9	MMMMMM	FFFFFFFF	2.70	126
126	2.70	50-4	MMMMMMMM	FFFFFFFFF	3.06	143
113	2.42	45-9	MMMMMMM	FFFFFFFFF	3.19	149
117	2.50	40-4	MMMMMMM	FFFFFFFF	2.89	135
120	2.57	35-9	MMMMMMM	FFFFFFF	2.55	119
164	3.51	30-4	MMMMMMMMM	FFFFFFFFF	3.27	153
196	4.19	25-9	MMMMMMMMMMM	FFFFFFFFFFFFFFF	5.54	259
198	4.24	20-4	MMMMMMMMMMM	FFFFFFFFFFFFFFF	5.50	257
169	3.62	15-9	MMMMMMMMM	FFFFFFFFFFF	4.04	189
165	3.53	10-4	MMMMMMMMM	FFFFFFFFFF	3.94	184
217	4.64	5-9	MMMMMMMMMMM	FFFFFFFFFFF	4.09	191
271	5.80	0-4	MMMMMMMMMMMMMMM	FFFFFFFFFFFFF	4.71	220

App. Fig. 1B **WHITE POPULATION OF TRACT 8052.01, COUNTY 33, MARYLAND, COMPREHENSIVE CMHC I**
TOTAL POPULATION 2249 MALES 1056 FEMALES 1193

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
13	0.58	85+	M M	F	0.44	10
15	0.67	80-4	MM	FF	0.93	21
23	1.02	75-9	MMM	FFFF	1.33	30
40	1.78	70-4	MMMM	FFFFFFFF	3.07	69
68	3.02	65-9	MMMMMMMM	FFFFFFF	2.93	66
76	3.38	60-4	MMMMMMMMM	FFFFFFFFFFF	4.45	100
81	3.60	55-9	MMMMMMMMM	FFFFFFFFFFFFF	5.02	113
97	4.31	50-4	MMMMMMMMMM	FFFFFFFFFFFFF	5.11	115
71	3.16	45-9	MMMMMMMM	FFFFFFFFFFF	4.45	100
46	2.05	40-4	MMMMM	FFFFFF	2.45	55
39	1.73	35-9	MMMM	FFFF	1.78	40
45	2.00	30-4	MMMMM	FFFF	1.69	38
58	2.58	25-9	MMMMMM	FFFFFFF	2.98	67
86	3.82	20-4	MMMMMMMMM	FFFFFFFF	3.11	70
100	4.45	15-9	MMMMMMMMMM	FFFFFFFFFFF	4.71	106
80	3.56	10-4	MMMMMMMMM	FFFFFFFFF	3.56	80
62	2.76	5-9	MMMMMMM	FFFFFFF	2.85	64
56	2.49	0-4	MMMMMM	FFFFF	2.18	49

App. Fig. 1C **NEGRO POPULATION OF TRACT 8052.01, COUNTY 33, MARYLAND, COMPREHENSIVE CMHC I**
TOTAL POPULATION 2372 MALES 1124 FEMALES 1248

NUMBER	%	AGE	9 8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	%	NUMBER
1	0.04	85+			0.08	2
0	0.0	80-4			0.04	1
2	0.08	75-9			0.17	4
3	0.13	70-4		F	0.46	11
2	0.08	65-9		F	0.59	14
12	0.51	60-4	M	F	0.51	12
12	0.51	55-9	M	F	0.55	13
28	1.18	50-4	MMM	FFF	1.18	28
41	1.73	45-9	MMMM	FFFFF	1.94	46
70	2.95	40-4	MMMMMMM	FFFFFFFFF	3.29	78
80	3.37	35-9	MMMMMMMM	FFFFFFFFF	3.29	78
118	4.97	30-4	MMMMMMMMMMMMM	FFFFFFFFFFFFFF	4.81	114
132	5.56	25-9	MMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFFFFFF	7.93	188
112	4.72	20-4	MMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFFFFFF	7.80	185
69	2.91	15-9	MMMMMMMM	FFFFFFFFFF	3.41	81
84	3.54	10-4	MMMMMMMMM	FFFFFFFFFFF	4.30	102
148	6.24	5-9	MMMMMMMMMMMMMMMM	FFFFFFFFFFFFFF	5.31	126
210	8.85	0-4	MMMMMMMMMMMMMMMMMMMMM	FFFFFFFFFFFFFFFFFFFF	6.96	165

☆ U. S. GOVERNMENT PRINTING OFFICE : 1973 O - 505-780

Social and Economic Statistics Administration
Washington, D.C. 20233