POPULATION PROJECTION: A DEMOGRAPHIC PROCEDURAL MANUAL FOR PLANNING PRACTITIONERS

by

JOHN ERIK HEDEEN

B. A., Bethany College, 1966

A NON-THESIS PROJECT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF REGIONAL AND COMMUNITY PLANNING

Department of Regional and Community Planning

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1973

Approved by:

Ray Bradley Weisenburger
Major Professor

TABLE OF CONTENTS

APPENDIXES | Page

LIST OF TABLES

Chapter 1

INTRODUCTION

Purpose

How often has one read a planning textbook or a journal article and upon completion wondered what was said or why nothing was said? The planning profession has many persons conducting research and formulating theory. Likewise, the planning profession has practitioners in all types of agencies, public and private. Unfortunately, researchers and practitioners are not in communication much of the time.

This sweeping generalization forms the scenario for the subsequent manual, namely to provide the planning practitioner in Kansas with some of the researcher's knowledge to undertake planning population projection. The information in this manual is only one approach to population projection and should not be treated as an inclusive discussion of this broad topic. The client for this manual is the practitioner, not the researcher; the subject of the manual is population projection as it relates to planning, not demography.

Population projection as it relates to planning has many weaknesses. Two reasons are paramount for these weaknesses. First, research in small area projection is lacking. The following quotation states this fact:

> . . . many demographers will not involve themselves in forecasts for small areas. Traditionally they have worked with large areas--the nation, regions, and states--but have been reluctant to engage in work on small areas because of problems of meeting their own standards of accuracy. They acknowledge the needs of city planners

> and recognize the dilemma, holding out hope that demographic research will in time remedy this situation.[1]

Not to be totally nihilist, one must recognize that techniques do exist to perform planning population projections and can, with care, be applied to small areas.

The second reason for the weakness in planning population projection lies in the planning profession itself. Planning practitioners generally do not understand existing projection methods. This weakness forms the purpose of this manual--to acquaint planning practitioners with the projection methodology. The four categories which follow account for the majority of misuse of projection methodology in the planning, profession and form the basis for the remaining chapters of this manual: (1) improper application of a given technique to a planning problem, (2) use of inaccurate or improper data sources, (3) improper operation of the basic technique, and (4) improper interpretation of the outcome of the projection.

Use of the Manual

This manual is to serve as a guide for proper methods of small area population projection for planning practitioners in Kansas. By its nature, this manual will be eclectic. No attempt will be made to give a review of the literature on the subject; however, the manual in total will reflect one approach to the major projection methodologies presented in current literature.

---

[1]F. Stuart Chapin, Jr., Urban Land Use Planning (2d ed.; Urbana: University of Illinois Press, 1965), p. 198.

The user should consider this manual as a single treatise and should use it as one approach to population projection. It is not a lexicon of demographic methods. For the above reason, the three general techniques presented are not given full treatment in any one subdivision of the manual. Rather, the essential elements of each stage in a population projection are explained in the subdivisions. The user must refer to all chapters for understanding of any individual technique.

The manual is written assuming the user has an understanding of algebraic notation and a basic understanding of planners' use of population projection. It is assumed that additional reading will be required for more advanced projection situations than the ones presented in this manual. It is recommended that The Methods and Materials of Demography be included in each planning library for required additional reading. This publication, in two volumes, is a lucid and complete explanation of demographic methods. The annotated bibliography of this manual can be used for further reading on population projection.

The manual assumes that planners use the following procedure for conducting planning population projections. The subdivisions of this manual will be helpful during each stage of that procedure.

1. Identify all possible uses for the population projections. e.g. housing need, school enrollment, swimming pool need, etc.

2. Identify which method or combination of methods are consistent with the uses identified. (Chapter 2)

3. Gather the raw data required to perform the projection and analyze the data for consistency. (Chapter 3)

4. Perform the projection either by hand computation or by computer. (Chapter 4)

5. Analyze the results and present the assumptions and results in clear, honest form. (Chapter 5)

Chapter 2

DESCRIPTION OF METHODS OF POPULATION PROJECTION

Three terms will occur repeatedly through this manual; these terms are essential elements in the practitioners' understanding of population projection: estimate, projection, and forecast. The definitions of these terms are derived from the U. S. Department of Commerce, Bureau of Census publication The Methods and Materials of Demography, Volume 1 and 2, published October 1971. It is fitting to use these definitions in the planning profession. The major source of data used in planning population projection is derived from the U. S. Department of Commerce census series publications.

Population Estimate: Any nonsurvey or analytic technique to approximate population figures involving the use of vital statistics data, immigration data, and data symptomatic of population change.

Population Projection: A population estimate which relates to approximating data outside the bounds of available census information. A projection is usually to derive data in the future but may include projections into the past.

Population Forecast: A population projection which is designated by the data user or data originator to be the most likely population for a given date.

From these definitions one can see the inherent relationship of the terms. All forecasts are projections and estimates; but, not all projections and estimates are forecasts. The subject matter of this manual is therefore to describe existing methods of estimating future, small area populations based on analytic techniques using symptomatic data of population change.

A second area of implied concern for all practitioners relates to the probability of error in population projection. The probability of error in projections is simply the probability that a stated projection will come true for the study area. Generally speaking, the probability of error for small area projections is great due to the need for added demographic research.

> Even with techniques, technological forecasting today is much more an art than a science. Virtually all users of special techniques in technological forecasting stress an aspect of their experience which they invariably consider to be the most valuable.[2]

Any projection technique has implied or explicitly made assumptions. The probability of error of a given technique is governed solely by the correctness of the assumptions made in relation to the trends in the projected study area. The art of population projection is in making assumptions not in manipulating numbers.

Very little research has been done concerning which projection technique has the lowest probability of error given a set of conditions in a study area. It is generally accepted that the probability of error is directly proportional to the projection period and indirectly proportional to the size of the study area. Inspite of the fact that it would be extremely helpful for planners to have at their disposal a documented algorithm which states what technique to use under given city size, historic growth pattern, and stated purpose for projection, the theory does not presently exist to construct that guide. However, with

---

[2]Erich Jantsch, "Technological Forecasting: Techniques in Perspective," Decision-Making in Urban Planning, ed. Ira M. Robinson (Beveraly Hills: Sage Publications, 1972), p. 151.

greater understanding of the projection techniques, the planner can make reasonable judgments, based on his immediate need, regarding which technique will fulfill his needs.

The planner should keep in mind that his immediate need for accuracy is not the same as the needs demographers presuppose. Planning population projections are always used with a given standard to derive a future need. Since planning standards vary greatly and are designed with inherent safety factors, the immediate need for an absolute projection does not exist. As further evidence to this fact, the planner often uses different projections within one report to forecast the various future requirements for a study area. A high limit projection is often used to describe the need for a new water system, where a medium or low limit projection may be used to demonstrate the need for a swimming pool. A third factor, unique to the planning profession, is the very nature of the planning process which allows restudy and review to adjust continually for error.

The remainder of this chapter will describe three broad-category projection techniques. The purpose of these descriptions is to acquaint the planner with techniques, their inherent assumptions, and their strengths and weaknesses, so that required planning judgments can be founded in current projection theory.

Ratio Techniques

The ratio technique is a method which can be used to project population of a study area which is one part of some known population aggregate referred to as the parent area. The computations follow the algebraic statement that if three of four dependent values are known, the fourth value can be computed. Stated differently, if the present and future population of the parent area are known as well as the present population of the study area, then the future population of the study area can be computed. The general formula for this technique follows:

$$\text{Study Area}_{fut} = \frac{\text{Parent Area}_{fut} \times \text{Study Area}_{pres}}{\text{Parent Area}_{pres}}$$

The simplicity of this technique is offset by the complexity of the assumptions inherent in the technique.

1. A stable relationship between the parent population and the projected study area must be accepted. Stated differently, fertility, mortality, and net migration rates for the two areas must be identical or proportional and projected to remain equal or proportional over time.

2. It is assumed that any growth or change in population which occurs in the parent population will also occur in the same proportion to the study area.

3. The user must assume that the future parent population projection is an accurate forecast of that area's future population.

The principle advantage of using the ratio technique of projection is that it is fast and relatively easy to compute. The method does not require detailed demographic data. In many planning situations, data availability becomes the overriding factor in the choice of a projection technique. Secondly, the ratio technique, by the nature of its computations, recognizes the dependency of a study area on its surrounding region in relation to population change. This dependency is extremely important when dealing with small area projections especially under 10,000 persons. This error would be a result of the known differences in births, deaths, and migration patterns for the two races.

There are statistical methods to test if the assumptions inherent in the ratio technique are true. Multiple regression of the various demographic rates for the parent and study area will reveal to what extent the two areas are related. If a high correlation exists between the two populations, the probability of error for the ratio technique should be small; the converse of this statement is also true. Computation of these values by multiple regression negates the two principal advantages of the ratio technique, speed and the need of less detailed information.

A second limitation of this technique is that it produces only aggregate population projections. If age, sex, and race projections are required as a part of the study, one must disaggregate the parent and study area population. Such disaggregation required detailed demographic data for both area and required much more time to project.

Mathematical Methods

Mathematical techniques of projection fit a line to past population values and extend that line into the future. For example, if a given study area has a population of 500 for each ten year census from 1900 to 1970, one would anticipate that, given certain assumptions, the population in 1980 would be 500. In this example the past data was fitted to an equation of a line; the projection was made by solving that equation for the desired projection year.

As in the ratio technique, assumptions play the important factor in determining the probability of error. The basic assumption inherent in these techniques is that the cause of change in the past will be the cause of change in the future. For example, if no war or national disaster occurred in the study area in the past, none will occur in the future. Or, if economic expansion occurred in the past, that expansion will continue in the future.

Under the general category of mathematical techniques, many sub-techniques have been used. All these techniques essentially perform the same task of fitting data to a line. The techniques are named by the shape of the line they describe. The actual number of techniques under this category is limited only by the number of different geometric lines possible. Some of the most commonly used techniques are linear regression, parabolic regression, polynomial regression, exponential curve, and logistics curve. The choice of which curve to use is governed simply by which one best fits the past population trend.

Mathematical techniques are relatively simple to compute. Since they do not rely on parent population assumptions, the data required for computation is less detailed than for the ratio method. The computation of the individual projections lends itself to computerization which greatly reduces the planner's time expenditure.

These techniques do assume that past trends will continue indefinitely into the future. Therefore, a long range projection for an area which has had a high rate of growth will have extremely high population projection. Since extreme growth rates cannot continue indefinitely, the probability of error for long range mathematical projections is great. These techniques do not account for variation in any of the demographic or socioeconomic factors of population change. If the migration rate for an area changed greatly while other factors remained constant, mathematical methods would not reflect that change.

Finally, mathematical methods do not give projections by age, sex or race. It is possible to disaggregate the population to arrive at an age-sex population projection. However, such disaggregation requires more detailed data while increasing the computations and time required to accomplish the projection.

Disaggregative Methods

The disaggregative or component methods of population projection are the latest evolution of projection techniques and are the most rigorous to perform. These methods involve disaggregation of the study area's population into components of age, sex, and race. To make a population projection the planner must project the death, birth, and migration patterns for each population component in the study area. These patterns expressed as rates are substituted into a formula which represents population change. The general formula for population change follows:

$$P_{z+1} = P_z - D + B + IM - OM$$

In this formula P_{z+1} equals projected population, P_z equals the starting population, D equals deaths in the period z to z+1, B equals the births during that period, IM equals in-migration during period, and OM equals out-migration during the period.

There are many subtechniques within the disaggregative category. The cohort-survival and net migration and natural increase methods are only two of many variations in current use. All methods involve disaggregation of the study area to some degree and description of individual factors which affect population change for a given study area.

Generally speaking there are no inherent assumptions involved with these techniques. The assumptions made are all related to the methods used to project the individual factors of population change. One usually uses mathematical projection, line fitting, to make inferences about death, birth and migration patterns. However, in any given projection, one may use three different line patterns to project the three demographic factors, eg. linear regression for migration, static death rate, and logistics curve for birth rates.

The strength of disaggregative techniques lies in the fact that the planner is actually modeling the factors that effect population change in the study area. If an age-sex result is needed in the planning study for a school projection, disaggregative methods can give age-sex projections. The user is not limited to one assumption which covers all factors of demographic change. One can change assumptions to fit each factor of population change.

The disaggregative methods require complete demographic data to perform. This data is not available for cities of less than 2,500 persons in Kansas. The computation of a component projection, although not individually complicated, is time consuming to perform unless computerized. The probability of error for any component method may be as great as in using the ratio or mathematical methods, a fact not commonly known. The error depends on the various assumptions used and how clearly these assumptions model the actual study area.

There are some generalized estimating principles which can serve as a rough guide for the planner's choice of projection method in relation to the above explanations. These guidelines were derived from the publication The Methods and Materials of Demography, Volume 2.

1. More accurate estimates can generally be made from an entire country than for geographic subdivisions of the country.

2. More accurate estimates can generally be made for the total population of an area than for individual components of the area such as age and sex.

3. In general directly derived data which is complete, internally consistent, and recent is preferred to data which is not or which is derived indirectly by survey.

4. An estimate can be checked by comparing it with another estimate derived by a different method using different data and assumptions. If the two estimates are consistent, one should have greater confidence in both. If the two estimates differ, the planner should doubt both estimates.

5. The poorer the quality of data and the longer the estimating period, the less reliable are resulting population estimates.

These guidelines, as given, are tailored for estimating intercensal values; however, since projection is a form of estimating, these points are quite applicable to the planner's projection problems.

Inspite of the fact that further research is needed to develop an algorithm to state which projection technique to use in a given situation, some principles do exist to suggest basic distinctions for that algorithm. The practitioner must first describe the study area according to some broad criteria of classification. These criteria are city size, historical growth trends and reason for historical growth.

Classification according to city size is the easiest decision to make. The U. S. Census gives varying degrees of information depending on the population of the area. For this reason, the classification system used by the Census is appropriate to use in this manual since different projection techniques require varying degrees of input data. The Census classes are as follows: less than 1,000; 1,000 to 2,500; 2,500 to 10,000; 10,000 to 50,000; and greater than 50,000. The U. S. Census gives age-sex counts for areas with a population of 2,500 and above.

The second criterion of classification is the historic growth pattern of the study area. This classification is more difficult to make than the one regarding size. Three general categories - declining, stable, and increasing - should be used. The historic trend should be judged on the population values of the past twenty years. No clear guidelines can be given regarding the exact distinction between the three categories. Rather, they must be thought of on a continuum which has no absolute boundaries. Study areas which have grown or declined drastically in the past twenty years must be treated separately. However, as the trends approach a stable condition, the need for making an exact classification is lessened; techniques used for a stable area will accomodate slightly increasing or decreasing populations.

The third criterion of classification is the reason for past population change. If the population changes of the study area were a result of natural variances in births, deaths, and migration caused by the general economic conditions in the study area, no special treatment is required. However, if changes in population are a result of known, past events which have little chance of recurring the the future, special treatment of that study area is mandatory. There are many examples of these single-factor population changes in Kansas, eg. basic industry closure or opening, major study area annexation, or definitional changes in data. Salina had a basic industry closure when Shilling Air Force Base was closed. Dodge City annexed major amounts of territory which affected their population. Manhattan increased population drastically in 1950 due to a U. S. Census definition change regarding students.

Once the classification of the study area has been made regarding size, historic trends, and reasons for historic trends, the practitioner can choose the technique which fits his requirements. The guidelines suggested below match the classification made to the three broad category techniques outlines in this manual. These guidelines will not insure a lower probability of error. Rather, they match a study area to the technique which is theoretically best suited to handle the particular situation in the study area.

The following list states these general guidelines for matching a study area to a technique:

1. Disaggregative methods should not be used to project populations of areas of less than 2,500 persons unless the practitioner has extensive knowledge of the theory and use of these methods.

2. Disaggregative methods should <u>always</u> be used for study areas with a population over 50,000 persons. Other techniques can be performed as simple checks on the projections derived.

3. Disaggregative methods should be used to project areas which have demonstrated a single-factor population change in the last twenty years. If the study area is less than 10,000 persons and has shown a single-factor change, the practitioner should investigate other disaggregative methods which are beyond the scope of this manual.

4. Mathematical methods should <u>not</u> be used for projections over ten years unless it is to be used as a simple check on other derived projections.

5. Mathematical methods should <u>not</u> be used to project cities which have a single-factor change in the last twenty years. These methods can be used if the time series population values are adjusted for the effects of the single-factor change.

6. Ratio techniques are applicable to all city classes excluding only the single-factor change cities. The practitioner must bear in mind the extensive assumptions needed to apply these procedures to a study area.

Table 2-1 indicates these guidelines in relation to given techniques and study area classification. The order in which the techniques are listed is the preferred order. The techniques indicated relate to the broad category methods described earlier in this chapter rather than to any particular subtechnique described later in this manual. The practitioner must weigh the specific guidelines above to judge if additional research is needed to identify a particular subtechnique which will fulfill his requirements. The practitioner must further weigh his selection of technique with the uses he has anticipated for the derived projection.

TABLE 2-1

STUDY AREA CLASSIFICATION VERSUS PROJECTION TECHNIQUE

Size Class	Past Growth Trends			Single Factor Change
	Declining	Stable	Growing	
Less than 1,000	RM; MM	MM; RM	RM; MM	DM*
1,000 - 2,500	RM; MM	MM; RM	RM; MM	DM*
2,500 - 10,000	DM*; RM; MM	DM*; MM; RM	DM*; RM; MM	DM*
10,000 - 50,000	DM; RM; MM	DM; RM; MM	DM; RM; MM	DM
Greater than 50,000	DM; RM	DM; RM	DM; RM	

RM - Ratio Methods

MM - Mathematic Methods

DM - Disaggregative Methods

* - With Added Research Beyond Scope of This Manual

Chapter 3

DEMOGRAPHIC DATA SOURCES

All projection techniques require demographic data. The probability of error in a projection, although directly governed by assumptions made, has a relationship to the quality of demographic data used in the projection. This chapter will indicate types, sources, and evaluation of the demographic data available to planners in Kansas for use in population projections.

One universal principle exists when analyzing the use of demographic data - consistency. The classification of individuals and events into data is accomplished by set definitions. If these definitions differ from one census to another the data is useless. For example, the definition of a resident must be consistent from one census to the next. If a service man stationed in county A, but a legal resident of another state is included in county A's census in one year but not the next, the raw population counts of county A are inconsistent and relatively useless as demographic data.

A planner using census and vital statistics information must understand the collecting agencies exact definitions used to classify persons or events. Two terms are used by these demographic data collection agencies to classify persons or events - de facto and de jure. The following definition of these terms is generally accepted throughout demographic texts:

De facto: persons are counted where they are found on census day.

De jure: persons are counted at their usual place of residence.

The definitions actually used by any given agency do not follow this distinction exactly; however, the distinction clearly shows that data derived by a de facto method is not immediately consistent with data derived by de jure definitions. Recording of vital statistics also uses this distinction. A death can be recorded either where it occurs or where the person usually resides.

Census Data

Keeping these basic principles in mind, one can now analyze the demographic information available to planners in Kansas. The first, and most widely used, data source is the Department of Commerce, Bureau of the Census publications. The information derived by the Bureau of the Census is collected on April 1, every 10 years. The definition used to classify persons enumerated is based on de facto classifications. The only exception to this fact occurs if persons were traveling on census day and were out of their usual place of residence. These persons were allocated to their usual place of residence. The Federal Census has maintained excellent consistency in the use of definitions from census to census. However, one change has occurred relating to classification of college students. In 1950 the definition was changed to count students where they go to school rather than where their parents live.

The complete series of Federal census publications can be purchased from the Department of Commerce or a firm can be placed on automatic distribution for direct mailing as new publications are printed. Although there is considerable time lag in publication of census reports from the date of the census, first count reports are the only information needed to make population projections and are available at a reasonable time after the census date.

There are occasions when population figures are not available in published form which are needed in planning studies. In Kansas there are three Summary Tape Processing Centers which hold various census tape counts for various areas. These centers are not franchised, established or supported by the Bureau of the Census; they are developed through local initiative in response to individual needs. These centers can be contacted to get required census printouts for a price established by the local center. The following is a list of the centers in Kansas and an indication of what tapes they hold:

Institute for Social and Environmental Studies
University of Kansas
Lawrence, Kansas 66044
(group of states)

Center for Business and Economic Research
337 Clinton Hall
Wichita State University
Wichita, Kansas 67208
(Kansas)

Boeing Computer Services, Inc.
Sutton Place Building
Wichita, Kansas 67202
(Kansas)

The second source of census information in Kansas is from the Board of Agriculture yearly county census. This information is gathered by the individual county assessors on January 1, each year except during Federal census years. Kansas is one of only two states which undertakes a state-wide census yearly. The census, under state law, is the only officially recognized population data source in the state. Even the Federal census is not granted reciprocity during national census years.

The Agricultural census, generally speaking, classifies residents on a de jure basis. Therefore, military persons are not enumerated if their place of residence is out of the state. College students are enumerated in the county in which their parents live. The exact criterion used by the individual counties to classify residents is established by the Board of Agriculture and is distributed to the local counties on a yearly basis.

As a result of political implications, the definitions used in the Kansas census have varied from one year to the next. Also, by research funded by the Kansas Department of Economic Development, the definitions used throughout the counties of Kansas vary within any given year. The information collected in the Agricultural Census includes name, address, sex, and age. However, the only information published is the total population count, negating it as a source of current detailed information for planning studies.

In order to compare two sources of census information, the following criteria of a census were established by the United Nations. The essential features of a population census are "individual enumeration, universality within a defined territory, simultaneity, and defined periodicity."[1] As a fifth point the U.N. recommends that the census be conducted by the national government with active cooperation of state governments.

---

[1]United Nations, Principles and Recommendations for the 1970 Population Censuses, Statistical Papers, Series M., No. 44, 1967, p. 3.

Table 3-1 gives a comparison of the definitions used in the two census sources indicating hypothetical situations and the manner in which each census would count the individual described. The responses for Federal and State were supplied by the author by applying published guidelines for the data collection. The two counties' responses are a result of a mailed survey, questioning the assessors regarding how they classified the described individuals.

From the U. N. criteria and the survey of Kansas census practices, one can see that the Agricultural Census has only limited application as a demographic data source. There is evidence that the Kansas census is not universally applied through the state violating the one basic premise for demographic data - consistency. Even though common in planning reports, Agricultural Census data and U.S. Census data should not be combined in projection computations. Further analysis of the Agricultural Census is contained in the Kansas Department of Economic Development publication "Some Considerations for a Population Information System in Kansas" by Joseph E. Disanto, January 1970.

Vital Statistics

The third source of demographic data in Kansas is the Department of Health, Division of Vital Statistics in Topeka. This agency is a member of the national vital statistics recording agencies. This membership is important since the national office established the definitions and forms used in registration of events. For this reason, vital statistics are compatible in all states.

TABLE 3-1

COMPARISON OF CENSUS CONTACT CLASSIFICATION

Situation	Federal	Kansas	Dickenson County[1]	Labette County[1]
Minor child living with his divorced mother the childs father is a legal resident of other county	X	X	X	X
Single college student attending college in your community whose parents reside in other county	X			
Minor child living with his divorced mother in another county whose father lives in your county				
Single college student attending college in another county but parents reside in your county		X	X	X
Foreign student attending college in your county	X			X
Military dependent living off-post in your county but not legal resident of Kansas	X			X
Military person stationed in your county but legal resident elsewhere	X			
Military person who is legal resident of your county but stationed elsewhere		X	X	X
An adult in a federal, state, county or municiple institution in your county but who resided in another county before entering institution	X			
A person in a home for the aged who has a spouse living in another county	X			X
A widow (widower) in a home for the aged in your county whose prior address was elsewhere	X	X	X	X
A person in a Kansas penitentiary or reform school whose official residence was another county	X			
A widow (widower) in a home for the aged in another county whose prior address was your county			X	

X indicates classified as resident

[1]Counties chosen at random; responses are results of mailed survey in February, 1973.

The state office records the number of births, deaths, marriages, and divorces which occurred to residents of counties and cities throughout the state. The data is recorded on a de jure basis in that a death is recorded in the usual place of residence of the individual. This definition is the only reasonable one to use for vital statistics since most deaths and births occur in hospitals. If a different criterion were used, cities with hospitals would have an extremely high birth and death rate compared to cities without hospitals.

The department publishes a yearly summary of vital statistics for the state. Since this publication is only a summary, very little information is included about the various localities within the states. Appendix E indicates all the vital statistics collected and available through the Kansas office. If any of this information is needed, one can write the following address stating the exact description of the data desired:

Kansas State Department of Health
Division of Vital Statistics
Research and Analysis Section
535 Kansas Avenue
Topeka, Kansas 66603

The information which can be obtained through the recording of vital statistics in Kansas is extremely useful to planners in population projection problems and can be considered quite reliable. No other source of information is available to planners to derive age specific death rates or birth rates for given locations.

Data Types

The remainder of this chapter will discuss the exact data requirements for the projection methods described in this manual. This discussion will also indicate common misuses of demographic data in current planning practice.

Aggregate Population Counts. All projection techniques in this manual require this data type. Aggregate counts represent the total number of persons classified as residents of the study area in a specified census year. For the planner to understand the historical growth of his study area, he should gather data as far back as possible, which in Kansas is around 1860. Many areas in Kansas had their maximum population in the period 1900-1920 and have declined to the present time. The general rule of thumb is to gather enough historic data to demonstrate one of three alternatives: that population has shown continual increase; that the population has reached at least one inflection point; or that population has steadily declined.

After the discussion of data sources, it should be obvious that all data should be internally compatible. The compatibility of data can be affected by three points:

1. If data sources differ, eg. U. S. Census and Agricultural census, the data is incompatible.

2. If census definition of resident has changed from one census year to the next, the data is not compatible.

3. If the boundaries of the study area have changed from one census year to the next, eg. annexation, the data is not compatible.

Using these criteria of compatibility, there are many examples of misuse of aggregative population counts in planning practice. A population table which lists populations for 1968, 1969, 1970, and 1971 is in error due to incompatible data. The 1968-1969 and 1971 data was derived from the Agricultural Census and the 1970 from U. S. Census.

If the population table shows a sharp increase between 1940 and 1950, part of that increase may be a result of incompatible data. As stated earlier, the Federal census changed the resident classification of college students in 1950 to include them as part of the college town. This fact can have a dramatic effect on any population projection which would span the years of that definition change to derive a trend.

A third area of difficulty in using aggregate population counts occurs from the effects of city annexation. Any acquisition of land by a city changes the study area. The planner must be aware that a study area may have a decreasing population even though aggregate counts indicate growth. This fact is often hidden by the effects of annexation.

Disaggregative Population Counts. The population of the study area by age, sex and race is the second major data input for population projections. Even though this type of data is subject to problems of incompatibility, the planner is less likely to be confronted by it since the only source is the U. S. Census. These disaggregative counts are required only for component methods of projection.

Race counts are one differential which must be looked at closely. In the U. S., the proportion of the population which is non-white is twelve and one-half percent. If the study area has a greater percentage of non-white population than the U. S., race should be included as one component of the projection. As a point of information, Spanish surname populations are considered white under Federal census definitions. The agricultural census gathers no race information.

Vital Statistics. Birth data, available through the Division of Vital Statistics, is required by the component methods of projection. This data is collected by sex, age of mother, and residence of mother. As a point of information, this data is available for all cities with a population of 10,000 or more for all years since 1965, for counties, and for the state.

Death data, likewise available through the Division of Vital Statistics, is required by the component methods of projection. Deaths are recorded by sex and age of deceased and place of residence. This data is available for cities of 2,500 persons and above, for counties, and for the state.

Chapter 4

PROJECTION METHODOLOGIES

The selection of the method of projection to be used for a given study area should be based on the demographic data available and on the planner's expertise in matching required assumptions of the various techniques to the characteristics of the study area. Once a choice has been made, the planner should gather the data and proceed to compute the projections.

The methodologies outlined in this chapter are only prototype projections representing one proper application of the techniques. The variations which can be made on any one of the methodologies are subject only to the planner's imagination and ability to interpret the data derived. What should be kept in mind is that changes in the methods of calculation may also make basic changes in the assumptions of the projection. As has been emphasized earlier, since population projection is not a clear science, the sole important element of a projection is the assumption. A projection which is presented to the public without clearly stating the assumptions of the projection has no validity.

Each of the examples cited in this chapter are derived from acceptable demographic techniques. Region three, Riley County, and Manhattan, Kansas were chosen as the study areas since these areas have many problems common to projection of other areas in Kansas. Some of these problems include declining population, high military related population, high student population, and examples of high in and out migration. Each of the projections are made to 1975 to show the actual computations for a five year projection. It must be emphasized that these projections for 1975 do not represent any official value nor do they attempt to be normative for any other purpose than to illustrate computation. It should be further noted that some liberties have been taken within the computations for illustrative purposes.

Ratio Method

This projection technique is particularly suited for study areas which lack accurate or detailed demographic data. This technique can also be used to produce middle and long range projections (15 years plus) where the mathematical technique may give erroneous results for extended projections. The planner should bear in mind the inherent assumptions for the ratio techniques stated in Chapter 2. Inspite of the less detailed raw data needed and the technique's suitability for longer projection periods, the assumptions made are the major governing factor between an accurate and inaccurate projection.

It is proper to state at this time that the ratio technique often employed in planning reports which says that if Kansas grows eleven percent between 1970 and 1980, then city X will grow eleven percent is a questionable use of ratio methods. Close analysis of the assumptions inherent in this projection reveals that the true sympathetic relationship between the study area and Kansas has not been represented. Such a projection should only be made if the study area has shown a historic trend of conformance to the Kansas growth rate.

The example which follows is an application of ratio projection by percent of population. A second subtechnique, ratio projection by percent of growth, could also be used and computed by substituting percent of growth figures for percent of population figures. The method for computing the ratio projection is outlined below in procedural order.

Data Collection. This technique requires census counts for the study area and parent population for two past census years as well as a forecast for the parent population. The planner should use U. S. Census information for these aggregate counts and either a U. S. Census population projection or an independently derived projection of good quality for the parent population. The U. S. Bureau of the Census publication Series P-25, Number 477, dated March 1972, (or current edition) entitled "Population Projections and Estimates" gives projections for the state of Kansas 1975 to 1990 based on four sets of assumptions. The example given in this manual used this publication as the source of the parent area projection. The trend base in the example is 1960 to 1970. The planner could use a longer trend but must keep in mind that a longer trend may render the data insensitive to cumulative changes in birth, death, and net migration rates as well as sensitive to definitional changes in the census counts.

Construct Projection Matrix. The next step in the procedure is to construct the projection matrix as illustrated in Table 4-1. This matrix is for use only in computation and should not be included as a part of the final projection report unless included in an appendix. The matrix

TABLE 4-1

RATIO PROJECTION MATRIX

LOCATION	1960 Population (1)	1970 Population (2)	1960 Percent of State (3)	1970 Percent of State (4)	Average Yearly Rate Of Change (5)	PROJECTED RATE OF CHANGE 1970-71 (6)	1971-72 (7)	1972-73 (8)	1973-74 (9)	1974-75 (10)	1975 Prel. Percent of State (11)	1975 Adj. Percent of State (12)	1975 Projected Population (13)
KANSAS	2,178,611	2,249,071	---	---	---	---	---	---	---	---	---	---	2,287,000
CHASE CO.	3,921	3,409	.0018	.0015	-.0182	-.0178	-.0175	-.0171	-.0167	-.0164	.0014	.0014	3,202
CLAY CO.	10,675	9,890	.0049	.0044	-.0108	-.0106	-.0104	-.0102	-.0099	-.0097	.0042	.0042	9,605
DICKINSON CO.	21,572	19,993	.0099	.0089	-.0106	-.0104	-.0102	-.0100	-.0098	-.0095	.0085	.0084	19,211
GEARY CO.	28,779	28,111	.0132	.0125	-.0054	-.0053	-.0052	-.0051	-.0050	-.0049	.0122	.0121	27,673
LYON CO.	26,928	32,071	.0124	.0143	+.0142	+.0139	+.0136	+.0133	+.0131	+.0128	.0153	.0151	34,534
MARION CO.	15,143	13,935	.0070	.0062	-.0121	-.0119	-.0116	-.0114	-.0111	-.0109	.0059	.0058	13,265
MORRIS CO.	7,392	6,432	.0034	.0029	-.0159	-.0156	-.0153	-.0149	-.0146	-.0143	.0027	.0027	6,175
POTTAWATOMIE CO.	11,957	11,755	.0055	.0052	-.0056	-.0055	-.0054	-.0053	-.0052	-.0050	.0051	.0050	11,435
RILEY CO.	41,914	56,788	.0192	.0252	+.0770	+.0265	+.0259	+.0254	+.0248	+.0243	.0286	.0282	64,492
WABAUNSEE CO.	6,648	6,397	.0031	.0028	-.0102	-.0100	-.0098	-.0096	-.0094	-.0092	.0027	.0027	6,175
REGION 03	174,929	188.780	.0803	.0839	+.0044	+.0043	+.0042	+.0041	+.0040	+.0040	.0866	.0856	195,767
RILEY CO.	41,914	56,788	---	---	---	---	---	---	---	---	---	---	64,492
MANHATTAN	22,993	27,575	.5486	.4856	-.0122	-.0120	-.0117	-.0115	-.0112	-.0110	.4584	.4584	29,563

(1). (2) - Source: U. S Bureau of the Census should be carefully labeled by both column and row to insure that each piece of data is recorded correctly. A separate matrix should be used for each five year projection to be made.

Computation of Projection Matrix. Fill in the raw data of population for the parent and study areas in columns one and two. Insert the parent area forecast in column thirteen.

Compute what percentage the study area's population is of the parent area for the two base years and place figures in column three and four.

$$\frac{\text{Study area population 1960}}{\text{Parent area population 1960}} = \frac{174{,}929}{2{,}178{,}611} = .0803$$

The study area in this example is the entire region.

Compute the average annual change of percentage from 1960 to 1970 and insert in column five. The computation of this value is more complicated than it would initially seem. Assuming that the change is the same over each unit of time, the yearly compounding interest formula must be used to derive annual rate of change. Due to the mathematic rigor of solving that formula, this example uses an arithmatic approximation procedure to derive the value. This approximation method divides the average annual change in raw numbers by the average population during the period.

$$\text{Average Annual Change} = \frac{2(\text{Percent}_{70} - \text{Percent}_{60})}{\text{Number of Years} \times (\text{Percent}_{70} + \text{Percent}_{60})}$$

$$= \frac{2(.0839 - .0803)}{10(.0839 + .0803)}$$

$$= \frac{2(.0036)}{10(.1642)}$$

Average Annual Change = + .0044

If the exact rate of change is desired, one can use the following formula instead of the above procedure:

$$P_{1970} = P_{1960} (1 + R)^n$$

The value of R is the average annual rate of change.

Project the average annual rate of change for each year of the projection period and insert values in columns six through ten. At this stage of the computation the planner must make some assumptions about how this rate of change will vary over time. There are many assumptions that can be made regarding these rates of change; however, the extremes of logical possibilities are represented by the following:

1. That the rate of change will become zero at some future year indicating that the effect of births, deaths, and migration between the parent area and study area will become zero at that future date, eg. 50 years.

2. That the rate of change will remain constant at the 1960-1970 base year period. This assumption indicates that the differential effect of birth, deaths and net migration will remain the same as the 1960-1970 trend.

For purposes of this example the matrix will be computed on the assumption that the annual average rate of change will become zero in 50 years. Compute annual rate of change (ARC) in the following manner:

$$ARC_{70-71} = ARC_{60-70} \text{ X } 49/50 = .0044 \text{ X } 49/50 = .0043$$

$$ARC_{71-72} = ARC_{60-70} \text{ X } 48/50 = .0044 \text{ X } 48/50 = .0042$$

$$ARC_{72-73} = ARC_{60-70} \text{ X } 47/50 = .0044 \text{ X } 47/50 = .0041$$

$$ARC_{73-74} = ARC_{60-70} \text{ X } 46/50 = .0044 \text{ X } 46/50 = .0040$$

$$ARC_{74-75} = ARC_{60-70} \text{ X } 45/50 = .0044 \text{ X } 45/50 = .0040$$

These computations are continued until an annual rate of change is derived for each year of the projection period.

Calculate what percent the study area is of the parent area for the projection year. Calculate the percent by the following equation:

$$\text{Percent}_{75} = \text{Percent}_{70} \times (1 + \text{ARC}_{71}) \times (1 + \text{ARC}_{72}) \times (1 + \text{ARC}_{73}) \times (1 + \text{ARC}_{74}) \times (1 + \text{ARC}_{75})$$

$$\text{Percent}_{75} = .0839 \times 1.0043 \times 1.0042 \times 1.0041 \times 1.0040 \times 1.0040$$

$$\text{Percent}_{75} = .0856$$

Care must be taken in substituting the annual rate of change in the formula. If the rate of change is declining (minus value), then the formula is $(1 - \text{ARC}_x)$.

Compute the projected population. The final stage of the projection is to apply the percentage of parent area derived to the forecasted population of the parent area. Again the planner can make alternate assumptions to derive a high and low projection based on different forecasts for the parent area. The U. S. Census publication mentioned earlier derives four projections for the state based on the following assumptions:

1. High fertility rate and migration equal to the 1960-70 trend.
2. High fertility rate and no net migration.
3. Low fertility rate and migration equal to the 1960-1970 trend.
4. Low fertility rate and no net migration.

For purposes of this example the low fertility and migration equal to the 1960-1970 trend forecast was used. The projection is derived by multiplying the forecast by the projected percent of population.

Some general notes on the use of this ratio technique need to be made. At two points the planner must make explicit assumptions about the relationship of the study area to the parent population. The planner should carefully examine these assumptions and report these assumptions when presenting the projection. The accuracy of the projection is a sole result of the appropriateness of the assumption to the study area. Need for revision of the projection in subsequent years can be judged only in relationship to assumptions made and actual trends measured.

Numerous planning reports use ratio techniques which use national population projection, ratio them to the state population, then to the county and finally the study area. This use, although technically proper, should be questioned due to the great number of ratio links between parent areas. As a general rule, use the smallest parent area for which an acceptable forecast is available to minimize compounding errors from making multiple assumptions.

Mathematical Methods

Mathematical methods are also particularly suited for study areas which lack detailed demographic data. One drawback to these techniques is that they lose accuracy in middle and long range projections. The basic principle of these methods is to fit some form of a line expressed as a mathematical formula to time series data and extrapolate the formula to a future date. The assumption inherent in mathematical methods is that past trends exist and that past trends will continue into the future.

There are two basic methods of fitting a line to past population data--exact fit and approximate fit. The first method, exact fitting, is based on the general mathematical theorem which states that a polynomial equation of the nth degree will pass through n+1 given points. Therefore, given three points, the general form of the equation $Y=A+Bx+Cx^2$ will pass through those points. The drawback to these exact fit equations is that, even though they fit the given data exactly, one has no assurance that he has uncovered the underlying historical trend. The second method, approximate fitting, is based on using any number of given past data points and approximating a line to those points. At any given point, the line may or may not pass through that point exactly. However, the method minimizes the sum of the squares distance between the line and any given point.

In this manual, straight and parabolic lines are given as examples of exact and approximate line fitting techniques. Even though there are an infinite number of possibilities which could be used, these two are the most commonly used and widely applicable curves. Other specialized curves such as logistics and exponential are useful in some special planning situations; however, they are extremely complicated to compute and beyond the scope of this manual.

The first example given is that of the exact fit, line fitting technique. This method is fast and can have some value to the planner as a quick method to estimate short term projections. The examples given use Waring-Lagrange formulas to simplify computations. The procedure for computing exact fit projection is outlined on the following page.

Data Collection. Aggregate census counts for the study area for each past year should be gathered from the U. S. Census of Population. If Federal census information is not available, Kansas Agricultural census figures should be gathered. Under no circumstances should the Federal census data be mixed with Kansas Agricultural census data. The formulas for exact fit projection given in this manual are applicable only to 1960 and 1970 data for linear projection and 1950, 1960 and 1970 for parabolic projection. The formulas were limited to those dates because of the definitional change which occurred in the 1940-1950 Federal census.

Compute Linear Exact Fit Projection. By substituting the derived data into the following formulas, compute the population projections to the indicated years.

$$POP_{75} = 1.5(POP_{70}) - 0.5(POP_{60})$$

$$POP_{80} = 2.0(POP_{70}) - 1.0(POP_{60})$$

$$POP_{85} = 2.5(POP_{70}) - 1.5(POP_{60})$$

$$POP_{90} = 3.0(POP_{70}) - 2.0(POP_{60})$$

$$POP_{95} = 3.5(POP_{70}) - 2.5(POP_{60})$$

$$POP_{00} = 4.0(POP_{70}) - 3.0(POP_{60})$$

The following example of a projection to 1975 for Manhattan, Kansas, illustrates the use of these formulas:

$$POP_{75} = 1.5(POP_{70}) - 0.5(POP_{60})$$

$$= 1.5(27{,}575) - 0.5\ (22{,}993)$$

$$= 41{,}362.5 - 11{,}496.5$$

$$POP_{75} = 29{,}866$$

Compute Parabolic Exact Fit Projection. By substituting the derived data into the following formulas, compute the parabolic projection to the indicated years.

$$POP_{75} = 1.875(POP_{70}) - 1.250(POP_{60}) + 0.375(POP_{50})$$

$$POP_{80} = 3.000(POP_{70}) - 3.000(POP_{60}) + 1.000(POP_{50})$$

$$POP_{85} = 4.375(POP_{70}) - 5.250(POP_{60}) + 1.875(POP_{50})$$

$$POP_{90} = 6.000(POP_{70}) - 8.000(POP_{60}) + 3.000(POP_{50})$$

$$POP_{95} = 7.875(POP_{70}) - 11.250(POP_{60}) + 4.375(POP_{50})$$

$$POP_{00} = 10.000(POP_{70}) - 15.000(POP_{60}) + 6.000(POP_{50})$$

The following example of a projection to 1975 for Manhattan, Kansas illustrates the use of these parabolic equations:

$$POP_{75} = 1.875(POP_{70}) - 1.250(POP_{60}) + 0.375(POP_{50})$$

$$= 1.875(27{,}575) - 1.250(22{,}993) + 0.375(19{,}056)$$

$$= 51{,}703.125 - 28{,}741.25 + 7{,}146$$

$$POP_{75} = 30{,}108$$

The methods outlined above for exact fit projections are only short-cut techniques for approximating detailed projections. One must assume that the net effects of births, deaths, and migration will vary in the future in accordance with the 1950 or 1960 to 1970 trends and that these trends will continue indefinitely into the future. Such an assumption when carried to long range projections is very questionable.

The second method of fitting a curve to data approximates the fit to the past data points. The computation of these projections is more rigorous than the exact fit methods. However, the planner is more certain to uncover the historical trends of population growth in the study area by his added computations. The approximate curve fitting techniques presented in this manual are called linear and parabolic regression. They minimize the sum of squares difference between the given data and the estimated curve.

Data Collection. Raw population counts for the study area should be collected from the U. S. Census for all past recorded years. In Kansas, this data generally dates back to 1970. Construct a simple time versus population graph of the data to analyze whether a point of highest or lowest population has been reached in the study area. If an inflection point occurred prior to 1930, delete all data prior to the inflection point. If no inflection point occurred, use all data.

Construct Regression Matrix. Construct a regression matrix as illustrated in Table 4-2. This matrix should be labeled by both columns and rows so that data can be clearly understood. The population values are the "Y" values. The years are the "X" values. To simplify computation, the actual dates are reduced to proportional numbers which represent the spacing of the data points. Any values can be chosen as long as the entire series of substitute numbers is proportional to the entire series of actual dates.

Computation of Regression Matrix. The computation of the values in the regression matrix is relatively simple with an electronic calculator. When all rows are computed for each year, each column should be totaled to derive the sumation (Σ) of each value. The data included in Table 4-2 is for Manhattan, Kansas, 1870-1970.

TABLE 4-2

REGRESSION MATRIX

YEAR	X	X^2	X^3	X^4	Population Y (1)	XY	X^2Y
1870	1	1	1	1	1,173	1,173	1,173
1880	2	4	8	16	2,105	4,210	8,420
1890	3	9	27	81	3,004	9,012	27,036
1900	4	16	64	256	3,438	13,752	55,008
1910	5	25	125	625	5,722	28,610	143,050
1920	6	36	216	1,296	7,987	47,922	287,532
1930	7	49	343	2,401	10,136	70,952	496,664
1940	8	64	512	4,096	11,659	93,272	746,176
1950	9	81	729	6,561	19,056	171,504	1,543,536
1960	10	100	1,000	10,000	22,993	229,930	2,299,300
1970	11	121	1,331	14,641	27,575	303,325	3,336,575
N	ΣX	ΣX^2	ΣX^3	ΣX^4	ΣY	ΣXY	ΣX^2Y
11	56	506	4,356	39,974	114,848	973,662	8,944,470

(1) Source: U. S. Bureau of the Census

Compute Linear Regression Population Projection. The following procedure is a simplified, electronic calculator method for making the projection. Care must be taken not to confuse large "X" which is the raw population values with little "x" which is the difference between any given value of "X" and the mean of all "X's".

Compute the summation of x^2 by the following series of formulas. Some of the values required can be taken from the regression matrix and other require further computation.

$$\Sigma X = X_1 + X_2 + X_3 + \ldots X_n = 66$$

$$\overline{X} = \Sigma X / n = 6$$

$$\Sigma X^2 = X_1^2 + X_2^2 + X_3^2 + \ldots X_n^2 = 506$$

$$(\Sigma X)^2/n = (66)^2/11 = 396$$

$$\Sigma x^2 = \Sigma X^2 - (\Sigma X)^2/n = 110$$

Compute the mean population $\overline{Y}$ by the following formula:

$$\overline{Y} = \Sigma Y/n = 114{,}848/11 = 10{,}441$$

Compute the summation of little xy, Σxy, by the following formulas:

$$\Sigma XY = X_1Y_1 + X_2Y_2 + X_3Y_3 + \ldots X_nY_n = 973{,}662$$

$$(\Sigma X)(\Sigma Y)/n = (66)(114{,}848)/11 = 689{,}088$$

$$\Sigma xy = \Sigma XY - (\Sigma X)(\Sigma Y)/n = 284{,}574$$

Compute "b", the slope of the regression line by the following equations:

$$b = \Sigma xy/\Sigma x^2 = 284{,}574/110 = 2{,}587$$

This value can be either plus or minus; a minus indicates a decreasing population and a plus indicates an increasing population.

Compute the equation for the regression line by the following formula. $\hat{Y}$ is the population projection for the desired year X. The values $\bar{Y}$, b, and $\bar{X}$ have been derived above.

$$\hat{Y} = \bar{Y} + b(X - \bar{X})$$

$$\hat{Y} = 10{,}441 + 2{,}587(x - 6)$$

$$\hat{Y} = 2{,}587X = 5{,}081$$

Compute the Population Projection. Compute the population projection, $\hat{Y}$, for the desired years, X. Care must be used to ascertain the proper X value corresponding to the year desired if substitution was done for the actual years in the regression matrix. For example, if 1960 equals X(10) and 1970 equals X(11), then 1970 would equal an X value of 11.5 and 1980 would equal 12.0. The following example is a linear regression projection for Manhattan, Kansas in accordance with the above example computations:

$$\hat{Y}_{75} = 2{,}587\ (11.5) = 5{,}081$$

$$\hat{Y}_{75} = 24{,}680$$

As one can see this projection is below the actual 1970 population, and is an example of the erroneous conclusions which can be drawn from using a straight line to predict a past population trend which is not straight. If this method were applied to a study area which had linear past growth, the projection would be more reasonable.

The computation of parabolic regression projections use the regression matrix developed for linear regression. Referring to the values in Table 4-2, parabolic regression is computed by the following procedure.

Establish Simultaneous Equations. Establish three simultaneous equations by substituting values derived in the regression matrix in the proper locations of the following general equations:

$$\Sigma Y = An + B\Sigma X + C\Sigma X^2$$

$$\Sigma XY = A\Sigma X + B\Sigma X^2 + C\Sigma X^3$$

$$\Sigma X^2Y = A\Sigma X^2 + B\Sigma X^3 + C\Sigma X^4$$

By substitution of data from the regression matrix, these general formulas become the simultaneous equations for solution of parabolic regression for Manhattan, Kansas, 1870 to 1970:

1) $114{,}848 = 11A + 66B + 506C$

2) $973{,}662 = 66A + 506B + 4{,}356C$

3) $8{,}944{,}470 = 506A + 4{,}356B + 39{,}974C$

Solve the Simultaneous Equations. Solve the equations for the values of A, B, and C. The procedure for solving these equations for the required values is presented in an abbreviated form as follows:

1. Multiply equation one by minus six to make the coefficients of A in equation one and two equal and opposite in sign.

2. Add the resultant equation from step one to equation two which gives the following equation:

   $$284{,}574 = 110B + 1{,}320C$$

3. Multiply original equation two by minus 506 and original equation three by 66 to make coefficients of A in the two equations equal and opposite in sign.

4. Add the resultant equations from above step which gives the following equation:

   $$9{,}763{,}060 = 31{,}460B + 434{,}148C$$

5. Multiply the resultant equation of step two by minus 286 and add the resultant equation of step four to it. The resultant

equation is as follows:

$$16{,}275{,}468 = 56{,}628C$$

6. By solving the above equation for C and the other resultant equations for A and B, the following values were derived:

$$A = 2{,}391.27$$

$$B = -861.90$$

$$C = 287.41$$

Derive the Parabolic Regression Equation. Derive the equation which represents the parabolic trends of population growth by substituting the above values in the general parabolic equation.

$$\hat{Y}_x = A + BX + CX^2$$

$$\hat{Y}_x = 2{,}391.27 - 861.90X + 287.41X^2$$

Derive the Parabolic Regression Projection. The projected population (Y_x) is derived by substituting the X value for the year desired. Care must be taken as in linear regression to arrive at the proper X value which corresponds to the desired year. The following projection is for Manhattan, Kansas for the year 1975.

$$\hat{Y}_{75} = 2{,}391.27 - 861.9\ X + 287.41\ X^2$$

$$= 2{,}391.27 - 861.9\ (11.5) + 287.41\ (11.5)^2$$

$$\hat{Y}_{75} = 30{,}489$$

The results of the parabolic regression are very close to the projection derived by the ratio technique for 1975. By graphing the derived values for both the linear regression and parabolic regression, one can see that in the case of Manhattan, Kansas, the parabolic regression is the best curve fit of the two. For this reason, the projection derived by the parabolic regression is the more realistic of the two for 1975. One must keep in mind that for a different study area, one may find that the linear regression is the better fit.

Disaggregative Techniques

General Description. The disaggregative methods of population projection are the most commonly used methods of projecting national and regional populations. At present, little evidence exists to support whether these more analytical tools are more accurate than the simpler ratio and mathematic methods. However, as was true for mathematical and ratio techniques, the accuracy of the projection is governed by the assumptions made, not the method of computation.

There are numerous techniques of varying rigor which are classified as disaggregative projections. Some techniques use the components of births, deaths, and net migration applied to aggregate population figures to derive the projection. Others use censal survival rates expressing both deaths and migration as one function to derive the projection. Inspite of the actual technique employed, all these techniques use a formula for projection which models actual population change.

$$P_1 = P_0 + B - D + NM$$

In this formula, P_1 is the population at the end of the period, P_0 is the population at the beginning of the period, B is the number of births during the period, D is the deaths during the period, and NM is the net migration during the period.

This manual will present one technique of the many general disaggregative techniques known as the Cohort-Survival-Migration or Cohort-

Survival projection. The central important point to understand in this technique is the concept of cohorts. The general definition of a cohort is a group of individuals who experience the same demographic event during a specified time period and who may be identified as a group at later time periods due to the effects of that demographic event. In the technique presented in this manual, a cohort is a five year age group of males or females, for example, 5-9 years old who five years later will be 10-14 years old. The Cohort-Survival technique takes the number of persons who were living in the study area in a given age group and applies probabilities of death, probabilities of child birth, and probabilities of migration to derive the number of persons who will be living five years later.

The required data for the Cohort-Survival projection is first, an age-sex cohort population of the study area; second, probabilities of living expressed as a decimal number for each age-sex cohort; third, probabilities of births for females in the child bearing cohorts; and fourth, probabilities of migrating for each age-sex cohort.

The age-sex breakdown of population is available in the U. S. Census for the state, all counties, and cities of 2,500 population or greater. Since the Agricultural Census does not gather age specific data, it is not suitable for use as a data source. The accepted convention for these cohort populations is to disaggregate the population into five year cohorts starting 0-4 years and ending 85+ years old. If more detailed projections are needed, disaggregation may be by one year age groups. The census publications give five year cohorts while census tapes can give one year age groups.

The probabilities of living are known as survival rates and indicate the proportion of persons who are living at age group X who will live to age group X+5. There are three major sources of survival information, two of which must be computed. The first source is the U. S. Bureau of the Census publication Series P-25, Number 493, dated December 1972, entitled "Projection of Population of the U. S. by Age and Sex: 1972-2020". This publication lists five-year survival rates for each age group to the year 2005. These rates are derived from vital statistics data for the United States. These rates are tailored for the nation as a whole but can be applied to individual study areas without correction. However, if the study area has significantly higher or lower mortality than the U. S., a correction factor must be applied to each rate. These rates are used, uncorrected, in this manual.

The second source of survival rates is from survival data from the actual census counts. These rates are called censal survival rates and include the effects of both deaths and migration. Censal survival rates are derived by the following formula:

$$\text{Censal Survival Rate}_x = \frac{1970^P\ x+10}{1960^P\ x}$$

To calculate this value, the 1970 population of a cohort, for example, 20-24 age group is divided by the 1960 population of the 10-14 cohort. Survival rates computed in this manner account for combined probability of dying and migrating and are derived for ten year rates rather than five. One common usage of censal survival rates in projection is to use the national censal survival rates for a study area and also apply migration rates computed for the study area. This technique has value in that the net effects of census taking error for the period is canceled.

The third source of survival rates is an abridged life table. By using vital statistics for the study area, this method develops exact survival rates for the study area. However, the results represent only one time period and further computations must be performed to project these rates into the future.

The second data requirement is the probability of birth. There are two sources of fertility data--U. S. Census publications and vital statistics. The publications mentioned earlier under survival rates also lists age specific fertility rates by five year age groups of females. These rates represent the national trends in fertility. If the study area patterns do not differ greatly from the national average, these rates can be applied to the study area. These rates, uncorrected are used in the example in this manual. The second source of fertility data is vital statistics. Age-specific fertility rates can be derived for the study area by the following formula:

$$\text{Fertility rate} = \frac{\text{Births}}{\text{Population of female cohort}}$$

In this formula the births are the number of births occurring to a given age group of mothers divided by the number of females in that age group. These rates are relatively easy to compute; however, since the rates represent only one time period, one must make further assumptions and computations to project fertility into the future.

The final data required is migration data. Unlike the first three data requirements, migration patterns are not recorded in any comprehensive manner. Therefore, past migration trends must be estimated for the study area. In population studies in Kansas, migration is the source of greatest variance in population from year to year. For this reason the examples given in this manual go into detail regarding computation of past trends and explanation of various possible assumptions to project migration trends.

Migration data is derived by computing the amount of population change which is a result of deaths and then assigns the remainder of the variance to migration. The variance due to migration is expressed as a five year, age specific probability of migration.

The remainder of this chapter will deal with the exact procedure for gathering and computing the data requirements for the Cohort-Survival projection. The examples given are for Manhattan, Kansas.

Population Data. The age-sex population counts for the study area should be collected from the U. S. Census for 1960 and 1970. This data should be assembled by five year cohorts from ages 0-4 to 85+ for both male and female. In addition, if the study area has a large college student population in relation to the total population of the town, eg. Manhattan, the census data about the number of college students in each age group should be collected. The purpose of this student data is to enable the planner to adjust the total population counts for student enumeration.

Since the main purpose of component methods is to disaggregate the population into cohorts of persons with like demographic trends, it is necessary to adjust the total population counts for cohorts that display dissimilar demographic trends. College students have a lower fertility rate than the general population and have extremely predictable migration patterns which are unlike the general population. For these reasons the projection of college students should be done separately from the general population and added to the base figures on completion of the projection. The following sections give a procedure for adjusting total population for college students.

The school enrollment figures are given in the "Social Characteristics" volume of the U. S. Census under the general heading of education. Collect the required information and place it in the appropriate location in the College Student Adjustment Matrix, Table 4-3. Columns one and two are the total enumeration counts for the ages indicated. The census does not indicate the exact population of persons 21 years old and that age group should be estimated by the following formula:

$$P_{21} = \frac{P_{21-24}}{4}$$

The population of the 21 to 24 cohort must be derived by subtracting P_{20} from P_{20-24}. Column three is the proportion of persons in that age group that are classified students. Column four and five are computed by multiplying the proportion of college students (column 3) by the number of persons in the age group (column 1 and 2). Column six and seven are computed by subtracting the number of students (column 4 and 5) from the number of persons in the cohort (column 1 and 2). The adjusted

TABLE 4-3

COLLEGE STUDENT ADJUSTMENT MATRIX

AGE GROUP	1970 Male Population (1)	1970 Female Population (2)	Proportion College Students (3)	Male Students (4)	Female Students (5)	Adjusted Male Population (6)	Adjusted Female Population (7)
15-17	558	589	.000	0	0	558	589
18-19	1,076	877	.733	789	643	287	234
15-19	1,634	1,466	--	--	--	845	823
20-21	1,739	1,338	.666	1,158	891	581	447
22-24	2,260	1,585	.395	893	626	1,367	959
20-24	3,999	2,923	--	--	--	1,948	1,406
25-29	1,363	958	.221	301	212	1,062	746
30-34	650	586	.221	144	129	506	457

(1), (2), (3) - Source: U. S. Bureau of the Census cohorts are used throughout the remainder of the projection. The student population is projected separately later in this chapter.

As a note of caution, the planner should use the above method to compensate for students in lieu of using off-campus student population derived from the school itself. Schools do not use the same definitions as the census for classifying residency and student status. Therefore, using school figures would be using incompatible data to effect the needed correction.

Survival Data. The age specific death patterns for the study area are derived from the data given in the Bureau of the Census publication mentioned above. These figures represent national trends; however, since death rates are essentially equal throughout the U. S., they can be used in most circumstances for individual study areas. The convenience of using these tables with their projections more than offsets any inaccuracy incurred by assuming that the study area death patterns equals national trends.

These rates express the probability of living to the next five year age group; for this reason, they are called survival rates. The component projection technique presented in this manual uses two sets of survival rates--ten year rates and five year rates. The ten year rates are used to derive the migration patterns from 1960 to 1970. These rates are derived from the five year rates published by the census. The procedure is to multiply the five year rate for age X to X+5 by the five year rate for age X+5 to X+10. The resultant rate represents the probability of living from age X to age X+10 or a ten year survival rate.

The five year rates used in the examples in this manual are taken directly from the census publication. Ten year survival rates are given in Appendix C.

The planner should be cautious when applying a given rate to a population. The survival rates for the initial age should be applied to that age group. The result of the multiplication gives the population of the terminal age group. For convenience, the tables from the census publication are reproduced in Appendix A and B. These tables should be updated each year when the Bureau of the Census issues changes.

Birth Data. The age-specific birth rates used in this manual are derived from the Bureau of the Census publication mentioned earlier. These rates indicate the national average probability of having a child given the age group of the mother. As with death rates the convenience of having tabulated and projected data more than offsets any error incurred by applying national trends to the study area.

The rates presented in this manual represent yearly rates and the resultant births derived by these rates must be multiplied by five to get five year births. These births must then be divided into sex of the child and be multiplied by the survival rate for births which represents the infant mortality rate.

Due to the difficulty of projecting birth rate changes, the census publication lists four projected alternatives. These alternatives represent possible changes in the total level of fertility and also in the average age of mother at child birth. Since the factors that effect birth rates are extremely complex, it is suggested that at least two of these projection series be used to develop a planning projection. This multiple computation will bracket the probable trends for the study area in the future.

For convenience the tables from the census publication are reproduced in Appendix D. These tables should be updated each year as revised projections are made by the Bureau of the Census.

Migration Data. The Cohort-Survival projection uses age specific migration rates derived by the residual method from 1960 and 1970 census data. The planner should keep in mind that migration rates derived have the largest effect on probability of error for the derived projection. Further, even though past migration rates can be estimated with some accuracy, there is no guarantee that the past trends will extend to the future. Many elaborate models have been developed to predict migration tendencies. Some have been successful but none have been foolproof.

The method of computing migration rates presented in this manual simply uncovers the past trends of migration for the study area. The planner should use these rates as one alternative assumption in his projections. In addition, however, he should vary the derived rates for future contingencies such as increased economic expansion or decreased job opportunities. In this manner, the derived projections cover the realm of logical possibilities for the study area.

Secondly, the planner must be aware of the fact that some specific age groups are the most probable migrants over time. As a general rule, the 20-45 and 60-65 year age groups are most mobile and subject to greatest migration. For some study areas, it may be appropriate to vary the rates of the 20-45 year groups while leaving the other ages in conformance with past trends.

Third, the planner can use the Cohort-Survival projection to test alternatives of policy decisions for the study area. For example, what effect would a net in migration of twenty persons a year of age group 65 and over have on the population patterns of the study area? Such a migration pattern could be a result of construction of retirement homes in the area. A second question might be, what migration rates are required of what ages to reverse the past trend of declining population?

The remainder of this section will give the procedure for computing the 1960-1970 age specific migration rates for a study area. These rates are important in the population projection and are the starting point for detailed analysis of migration implications.

The first step is to set up the Migration Rate Matrix as illustrated in Table 4-4. One table will be required for males and one for females. Fill in columns one and six with the actual 1960 and 1970 census population for the study area. If adjustment has been made for students, use the adjusted figures for age groups 15-34. Be sure to locate the figures adjacent to proper age groups as there are three age shifts in the table.

Place the ten year survival rates in column three. These rates represent the probability of living for the next ten years and are used to derive the expected 1970 population assuming no one migrated from the study area. The expected population for 1970 (column 5) is derived by multiplying the 1960 population (column 2) by the ten year survival rates (column 3). Note that the probability of living ten more years

TABLE 4-4

MIGRATION RATE MATRIX

1960 Age (1)	1960 Pop. (2)	10 year Survival Rates (3)	1970 Age (4)	1970 Exp. Pop. (5)	1970 Act. Pop. (6)	Net Migration 1960-70 (7)	$(5)_x + (5)_{x+5} =$ (8)	$(6)_x + (6)_{x+5} =$ (9)	$\frac{(9)}{(8)} =$ (10)	Initial Age Migration Rate (11)	5 year Migration Rate (12)
			0-4	--	940	--					
			5-9	--	990	--				0-4	-.1485
0-4	1,282	.9955	10-14	1,276	897	-379	2,207	-487	-.2207	5-9	-.1104
5-9	934	.9964	15-19	931	823	-108	1,692	+536	+.3168	10-14	+.1584
10-14	765	.9946	20-24	761	1,406	+644	1,270	+881	+.6937	15-19	+.3469
15-19	513	.9929	25-29	509	746	+237	1,047	+157	+.1500	20-24	+.0750
20-24	543	.9911	30-34	538	458	- 80	1,139	-153	-.1343	25-29	-.0672
25-29	609	.9875	35-39	601	528	- 73	1,110	+ 34	+.0306	30-34	+.0153
30-34	519	.9813	40-44	509	616	+107	1,117	+ 72	+.0645	35-39	+.0323
35-39	625	.9720	45-49	608	573	- 35	1,114	- 5	-.0045	40-44	-.0023
40-44	528	.9582	50-54	506	536	+ 30	971	+ 69	+.0711	45-49	+.0356
45-49	495	.9388	55-59	465	504	+ 39	880	+ 55	+.0625	50-54	+.0313
50-54	455	.9114	60-64	415	431	+ 16	779	+ 45	+.0579	55-59	+.0289
55-59	419	.8682	65-69	364	393	+ 29	708	+ 60	+.0847	60-64	+.0424
60-64	430	.8003	70-74	344	375	+ 31	594	+ 63	+.1061	65-69	+.0531
65-69	356	.7035	75-79	250	282	+ 32	425	+ 55	+.1294	70-74	+.0647
70-74	308	.5669	80-84	175	198	+ 23	252	+ 97	+.3849	75-79	+.1925
75-79	250	.3075	85 +	77	151	+ 74	77	+ 74	+.9610	80-84	+.4805
80-84	158	.0000								85 +	.0000
85 +	84	.0000									

for age groups 80-84 and 85+ is zero. All computations in this migration matrix are carried in a straight line across the table. The three shifts in age patterns account for the proper age change in the computations. For example, the 1,282 persons 0-4 years old in 1960 became the 1,276 persons 10-14 years old in 1970 after applying the ten year survival rate.

Compute the net number of persons who migrated (column 7) by subtracting the expected 1970 population (column 5) from the actual 1970 population (column 6). It is important to retain the sign of this computation as a plus indicates in-migration while a minus indicates out-migration.

The next stage of the computation is designed to convert ten year migration patterns to five year migration patterns. Explanation of the theory behind the precise computations is beyond the scope of this manual. However, the idea behind the addition of two age groups and two migration groups is to reflect the fact that five year rates involve two cohorts over time instead of one cohort for ten year rates.

Column eight is derived by adding the value from column five and the value for the age group immediately below it. The formula for this operation follows:

$$\text{Column } 8_{\text{age } x} = \text{Column } 5_{\text{age } x} + \text{Column } 5_{\text{age } x+5}$$

One exception of this formula occurs for age group 85+ where column eight equals column five.

Column nine is derived in a similar manner by algebraically adding the value in column seven and the value for the age group immediately below it. The formula for this computation follows:

$$\text{Column } 9_{\text{age } x} = \text{Column } 7_{\text{age } x} + \text{Column } 7_{\text{age } x+5}$$

Again for age group 85+ column nine equals column seven.

Column ten is derived by dividing column nine by column eight expressed as a decimal number with the sign retained. This value represents a hypothetical ten year migration rate for the ten year period. However, due to the additions of adjacent age groups, the value does not represent a true ten year migration rate.

Finally, the five year age specific migration rates based on 1960-1970 trends (column 12) are derived by dividing the values in column ten by 2. The value for the 0-4 age group is computed by dividing the 10-14 value in column seven by that value in column five. This result is divided by two and placed opposite the 0-4 age group in column twelve. The rates expressed as plus or minus indicate the proportion of the expected population which migrated into or out of the study area during a five year period.

These computed age-specific migration rates approximate the past trends of migration for the study area but do <u>not</u> represent what will happen in the future. As was stated earlier, these rates represent a starting point from which the planner can project changes in migration for future years.

<u>Projection of Population</u>. After the planner has gathered and computed the required data as outlined above, the actual computation of the projection is relatively simple. The following formula indicates the process of projection:

TABLE 4-5

COHORT-SURVIVAL PROJECTION MATRIX

FEMALES, MANHATTAN, KANSAS 1970-75

1970 Age (1)	1970 Population (2)	Survival Rate (3)	1975 Exp. Population (4)	Migration Rate (5)	Net Migrants (6)	1975 Proj. Population (7)	1975 Age (8)
			1,011	--	--	1,071	0-4
0-4	940	.9970	937	-.1485	-139	798	5-9
5-9	990	.9985	989	-.1104	-109	880	10-14
10-14	897	.9979	895	+.1584	+142	1,037	15-19
15-19	823	.9968	820	+.3469	+284	1,104	20-24
20-24	1,405	.9962	1,400	+.0750	+105	1,505	25-29
25-29	746	.9951	742	-.0672	- 50	692	30-34
30-34	458	.9927	459	+.0153	+ 7	466	35-39
35-39	528	.9891	522	+.0323	+ 17	539	40-44
40-44	616	.9835	606	-.0023	- 1	605	45-49
45-49	573	.9754	559	+.0356	+ 20	579	50-54
50-54	536	.9643	517	+.0313	+ 16	533	55-59
55-59	504	.9474	477	+.0289	+ 14	491	60-64
60-64	431	.9195	396	+.0424	+ 17	413	65-69
65-69	393	.8747	344	+.0531	+ 18	362	70-74
70-74	375	.8081	303	+.0647	+ 20	323	75-79
75-79	282	.7018	198	+.1925	+ 38	236	80-84
80-84	198	.4448	88	+.4805	+ 42	130	85 +
85 +	151	.0000	0	.0000	0	0	
TOTAL	10,846	.9452	11,323	+.0390	+441	11,764	

$$P_{x+5} = P_x - D + B + NM$$

This formula is computed for each age group until the projection is completed for male and female to the projection year.

The first stage of the computation is to construct the Projection Matrix, Table 4-5. Place the actual 1970 population in column two of the matrix by the appropriate age groups. As in the migration computation, use the adjusted population figures if college students are present. Place the five year survival rates by the appropriate initial age group in column three. Place the derived migration rates by the appropriate initial age group in column five.

Compute the expected 1975 population (column 4) by multiplying the population 1970 (column 2) by the survival rate (column 3). Calculate the number of net migrants (column 6) by multiplying the expected population (column 4) by the migration rate (column 5). Retain the sign of the migration rate in this computation. Calculate the projected population (column 8) by algebraically adding the net migration (column 6) to the expected population 1975 (column 4).

The final stage of the computation is to calculate the number of births during the period to complete the 0-4 age group in the projection. The first step of this computation is to construct a Birth Matrix as illustrated in Table 4-6. Place the number of females in the child bearing ages by the appropriate age groups in column two. The population figures are for 1970, <u>not</u> the derived 1975 population projection. Place the age specific birth rates for the projection period in column three. The number of births which occurs to an age group of mothers in

TABLE 4-6

BIRTH PROJECTION MATRIX

Age Group (1)	1970 Female Population (2)	Age-Specific Birth Rate (3)	Births In One Year (4)
10-14	897	.0011	1
15-19	823	.0694	57
20-24	1,406	.1634	230
25-29	746	.1383	103
30-34	458	.0716	33
35-39	528	.0328	17
40-44	616	.0085	5
45-49	573	.0005	0
TOTAL	--	--	446
Five-Year Total	--	--	2,230

(3) - 1970 U. S. Census Fertility Rates

TABLE 4-7

SEX PARTITIONING MATRIX

Sex	Prop By Sex	Births By Sex	Survival Rate Births	0-4 Cohort
Male	.51	1,137	.9742	1,108
Female	.49	1,093	.9799	1,071

one year (column 4) is computed by multiplying the population (column 2) by the birth rate (column 3). The total births for the year is derived by summing the values in column four. This value is multiplied by five to achieve the total births for a five year period.

The final stage of the birth projection is to divide the births by sex and apply a survival rate to that result, Table 4-7. In general, 51% of all births are male and 49% are female. The total five year births are partitioned in that manner. Finally, the survival rate for births, derived from the census publication, is multiplied by the male and female births to derive the value which is placed in columns four and seven, age group 0-4 of the Projection Matrix.

The projection is continued for five year increments for both males and females using the results of one matrix as the starting base of the next matrix. Column six of the Projection Matrix can be added to get the number of net migrants during the projection period. Column four of that matrix can be added to get the projection of 1975 with no net migration. The overall net migration rate during the period can be computed by dividing total net migrants (column 6) by total expected population (column 4).

If students were adjusted from the base population, they must be projected separately by age and sex and added to the base projection. For purposes of this manual, the student population was projected by the ratio technique illustrated in this manual in Table 4-8. The assumptions made were that the student population of Manhattan would peak in 1980 at 6,000. The male-female ratio would vary according to the 1960-1970 trend and that the rate of change would become zero in 1980. For complete explanation of the ratio technique refer to the appropriate section of this manual.

TABLE 4-8

RATIO PROJECTION OF COLLEGE STUDENTS

MANHATTAN, KANSAS 1970-75

	1960 Pop.	1970 Pop.	1960 % of Total	1970 % of Total	Ave. Annual Rate Change	1970-71	1971-72	1972-73	1973-74	1974-75	Prelim. 1975 %	Adj. 1975 Percent	Projected 1975 Pop.
TOTAL STUDENTS	4,354	5,786	--	--	--	--	--	--	--	--	--	--	6,000
15-19 F	424	643	.0974	.1111	+.0131	+.0118	+.0105	+.0092	+.0079	+.0066	.1163	.1153	692
20-24 F	873	1,517	.2005	.2622	+.0267	+.0240	+.0214	+.0187	+.0160	+.0134	.2876	.2852	1,711
25-29 F	149	212	.0342	.0366	+.0068	+.0061	+.0054	+.0048	+.0041	+.0034	.0375	.0372	223
30-34 F	126	129	.0289	.0223	-.0258	-.0232	-.0206	-.0181	-.0155	-.0129	.0204	.0203	122
TOTAL F	1,572	2,501	.3610	.4323	+.0180	+.0162	+.0144	+.0126	+.0108	+.0090	.4602	.4580	2,748
15-19 M	933	789	.2143	.1364	-.0444	-.0400	-.0355	-.0311	-.0266	-.0222	.1165	.1147	688
20-24 M	1,496	2,051	.3436	.3545	+.0031	+.0028	+.0025	+.0022	+.0019	.0016	.3584	.3527	2,116
25-29 M	218	301	.0501	.0520	+.0037	+.0033	+.0030	+.0026	+.0022	+.0019	.0527	.0519	311
30-34 M	135	144	.0310	.0249	-.0218	-.0196	-.0174	-.0153	-.0131	-.0109	.0231	.0227	137
TOTAL M	2,782	3,285	.6390	.5677	-.0118	-.0106	-.0094	-.0083	-.0071	-.0059	.5446	.5420	3,252

Chapter 5

ANALYSIS OF PROJECTIONS

Once a series of projections are derived by any of the methods described earlier, the planner must analyze the results in view of the specific uses of his projections. The problem which faces the planner is that he must make normative statements about future needs in a community based on the results of his projections. The one statement which can be made about the series of projections derived is that the planner cannot assume that any one of the projections will come true. To reconcile the probabilities of error in any given projection with the need to make normative statements, the planner must completely understand the projections and the purposes for making the projections. He must analyze what probability of error is acceptable for each use.

At the risk of oversimplification, there are three general purposes for making planning population projections:

1. To project the need for facilities which have a fixed maximum capacity such as water or sewer systems.
2. To project the needs which involve little public expenditure such as private housing or future land use categories.
3. To project the needs of a community which are based on market demand such as swimming pools or values such as assessed valuation.

Given these three purposes for making planning population projections, the planner must analyze what standard of error is acceptable for each purpose.

The public consequences of making an error of underestimating future population would be extremely great if the purpose of the projection were to project the need for fixed capacity systems. The particular projection designated as a forecast to demonstrate maximum design capacity of a water system should be a high level projection. The assumptions made in developing a projection for fixed capacity facilities should be based on high fertility rates and migration rates which reflect the greatest possibility of growth for the study area. However, caution must be exercised to insure that the high level projection is a viable projection for the study area. The consequences of over-designing a fixed capacity system can also be great.

Projections for land use plans and private housing expansion have only minimal financial consequences in the public sector. The projection which is designated as a forecast to identify needs of this nature is generally based on mid-level projections for the study area. Often these projections involve making fertility and migration assumptions which equal the historical trend. Also needs identified of this nature are the most open to future restudy and revision.

The final generalized purpose for planning population projections involves identification of needs which are based on market demand, eg. swimming pool needs or projection of assessed valuation. In these areas the consequences of overestimating population growth can be great. Projections derived for these purposes would generally involve assumptions of low fertility and migration patterns which would minimize the potential growth in relation to past trends. As with the first general class of purposes of projection, the low level values must be viable for the study area to avoid grossly underestimating market demand.

To summarize the above statements, the planner should attempt to develop accurate projections for the study area. However, at the same time, the planner should manage the projection results so that he places the greatest probability of error below the projection, symmetrical to the projection line, or above the line depending upon the purposes for which he is making the projection.

The form of analysis outlined above should occur upon completion of any series of projections. However, due to the nature of the planning process, the analysis of projections must be a recurring element of the planners' work program. In the present scenario of 701 Comprehensive Planning, the element of recurring analysis is often not included. The typical agency designates time early in the planning process to develop projections of population. Based on these projections, the other elements of the comprehensive plan are developed, needs projected, and budgets developed. The problem of analysis occurs at this point. Many times the consequences of changes in city policy are not taken as recurring feedback into the analysis of population changes. "Such feedback in the planning process is consistent with the fact that population does change in response to environmental changes."[1]

Therefore, a second level of projection analysis is needed which tests the population growth implications of policy decisions. This level of analysis was mentioned in the discussion of migration trends in the Cohort-Survival technique in Chapter 4. The planner should model the

---

[1]International City Managers Association, Principles and Practice of Urban Planning, ed. William I. Goodman (4th ed.; U. S.: ICMA, 1968), p. 75.

implication of an expansion of retirement homes as an example. In a different form, the planner should identify exact migration goals for a community if the policy concern is reversing a decreasing population trend. Too often planners imply reversal of the decreasing trends by making the presumptious statement that growth will occur if planning proposals are implemented.

The third stage of projection analysis and revision should occur as more detailed or more current base year data becomes available. Unfortunately, the base data is updated only every ten years by the U. S. Census. Planners have a vested interest in developing demographic information systems which serve the needs of accurately updating population data. The advent of Revenue Sharing apportioned on population data is expanding the constituency of such a demographic system for the nation. Only through accurate and current data can the planner judge the errors of his past assumptions and thereby the error of his projections.

The final stage of the population projection is to represent the results in a manner which is clear to all future users of the data. This display can be accomplished by the use of time series graphs and comparative tables. However, the most important element of the representation is to list in detail the assumptions used to derive the projection. A great many reports list the general assumptions which apply to all projections, eg. no depression will be experienced during the period or no war or national disaster will occur. What is important is to describe the exact assumptions made about the levels of fertility, death, and migration. Secondly, the purpose for which the projection was made should be indicated.

As a matter of professional honesty, the results of any series of projections should not be portrayed as figures which will come true in the future; no planner has that level of insight. Rather, by relating assumptions made to the results, the user of the projection can relate to actual trends in the study area to judge the accuracy of the projection. The assumptions stated should not use jargon from the demographic or planning profession, eg. total fertility rate, age specific migration rate, or survival rate. These terms are meaningless to the average person. The assumptions should be stated in simple terms such as, "The town will gain new residents at the net rate of 20 per year."

Finally, lengthy tables and discussion of detailed techniques are often more confusing than enlightening. The presentation of such information, if required, should be included only in appendices to the basic report.

Conclusions

This manual is meant to give an explanation of some of the current population projection techniques as they apply to planners' needs. As is the case in any discussion of this sort, many topics have been touched briefly, but none in detail. The basic premise used to justify this approach was that planners should use more analytic tools to investigate population change in their practice. This manual presents demographic techniques which are simplified in some respects yet correct. This manual can serve as a basis to upgrade the analytic techniques available to the planning practitioner. It is not an end but rather a step towards the analytic level required by modern planning practice.

BIBLIOGRAPHY

Chapin, F. Stuart Jr. Urban Land Use Planning. 2n ed. Urbana: University of Illinois Press, 1965.

A traditional approach to land use planning with a chapter discussing general methods of population analysis and projection.

DiSanto, Joseph E. Some Considerations for a Population Information System in Kansas. Topeka: Kansas Department of Economic Development, 1970.

A critical analysis of the Kansas Agricultural Census with concrete suggestions for establishing a demographic information base in Kansas.

Goodman, W. I. and Eric C. Freund (eds.). Principles and Practice of Urban Planning. 4th ed. Washington: International City Managers' Association, 1965.

An indepth discussion of the elements of urban planning with a chapter discussing general population analysis and reasonably detailed discussion of population projection.

Robinson, Ira M. Decision-Making in Urban Planning: An Introduction to New Methodologies. Beverly Hills: Sage Publications, 1972.

A discussion of new planning methodologies by various authors with a chapter on technological forecasting.

Rogers, Andrei. Matrix Methods in Urban and Regional Analysis. San Francisco: Holden-Day, 1971.

A rigorous discussion of the application of matrix methods to planning analysis. A matrix model of interregional migration analysis is included with actual examples of computations.

Snedecor, G. W., and W. G. Cochran. Statistical Methods. 6th ed. Ames: The Iowa State University Press, 1967.

A general text book on statistical methods with complete presentation of regression methods and application.

U. S. Bureau of the Census. Census of Population: 1970 General Population Characteristics, Final Report PC(1)-B18 Kansas. Washington: U. S. Government Printing Office, 1971.

Lists population by age, sex and race for the State of Kansas, counties, and cities.

________, Census of Population: 1970 General Social and Economic Characteristics, Final Report PC(1)-C18 Kansas. Washington: U. S. Government Printing Office, 1972.

Lists the college population of enumerated areas in Kansas.

______, Current Population Reports, Series P-25, No. 470, "Projections of the Population of the United States by Age and Sex: 1970 to 2020," Washington: U. S. Government Printing Office, 1971.

Lists population projections for the United States and birth and survival rates for use in Cohort-Survival projection techniques.

______, Current Population Reports, Series P-25, No. 477, "Preliminary Projections of the Population of States: 1975 to 1990". Washington: U. S. Government Printing Office, 1972.

Lists population projections for the state of Kansas based on four projections of the level of fertility and migration.

______, Current Population Reports, Series P-25, No. 493, "Projections of the Population of the U. S. by Age and Sex: 1972 to 2020". Washington: U. S. Government Printing Office, 1972.

Lists population projections for the United States as well as projections of birth and death rates for use in the Cohort-Survival projection technique. The terminal age of death rates is 75+.

______, The Methods and Materials of Demography, by Henry Shryock, Jacob S. Siegel, and Associates. Washington: U. S. Government Printing Office, 1971.

An indepth and well documented description of all facets of demography. Exact procedures are given along with actual examples of their use. Suggested readings at the end of each chapter are complete and extremely useful.

______, U. S. Census of Population: 1960, Vol. I, Characteristics of the Population, Part 18, Kansas. Washington: U. S. Government Printing Office, 1963.

Lists population by age, sex and race for the State of Kansas, counties, and cities as well as the social and economic characteristics of those areas.

APPENDIX A

FIVE YEAR U. S. MALE SURVIVAL RATES 1970-2000

INITIAL	PROJECTION PERIOD					
AGE	1970-75	1975-80	1980-85	1985-90	1990-95	1995-2000
BIRTHS	.9742	.9789	.9799	.9808	.9819	.9829
<5	.9962	.9967	.9968	.9970	.9972	.9974
5 - 9	.9978	.9978	.9979	.9979	.9980	.9980
10 - 14	.9954	.9953	.9956	.9958	.9960	.9962
15 - 19	.9915	.9910	.9915	.9919	.9924	.9929
20 - 24	.9905	.9901	.9906	.9910	.9915	.9920
25 - 29	.9904	.9903	.9906	.9910	.9914	.9918
30 - 34	.9876	.9878	.9883	.9887	.9892	.9896
35 - 39	.9815	.9819	.9826	.9832	.9838	.9843
40 - 44	.9713	.9717	.9727	.9734	.9743	.9750
45 - 49	.9541	.9559	.9567	.9577	.9586	.9595
50 - 54	.9289	.9326	.9338	.9349	.9364	.9376
55 - 59	.8938	.8971	.8992	.9017	.9042	.9072
60 - 64	.8461	.8486	.8525	.8556	.8593	.8629
65 - 69	.7812	.7867	.7912	.7965	.8006	.8055
70 - 74	.7031	.7081	.7132	.7181	.7232	.7279
75 - 79	.6055	.6079	.6095	.6111	.6123	.6138
80 - 84	.4032	.3992	.3978	.3984	.3975	.3965
85 +	.0000	.0000	.0000	.0000	.0000	.0000

Source: U. S. Bureau of the Census, Current Population Reports, Series P-25, No. 470, 1971 and No. 477, 1972.

APPENDIX B

FIVE YEAR U. S. FEMALE SURVIVAL RATES 1970-2000

INITIAL	PROJECTION PERIOD					
AGE	1970-75	1975-80	1980-85	1985-90	1990-95	1995-2000
BIRTHS	.9799	.9833	.9841	.9850	.9858	.9867
<5	.9970	.9974	.9975	.9977	.9978	.9979
5 - 9	.9985	.9986	.9986	.9986	.9986	.9987
10 - 14	.9979	.9979	.9980	.9981	.9982	.9983
15 - 19	.9968	.9968	.9969	.9971	.9972	.9974
20 - 24	.9962	.9962	.9963	.9964	.9966	.9967
25 - 29	.9951	.9950	.9951	.9953	.9954	.9956
30 - 34	.9927	.9929	.9931	.9933	.9935	.9937
35 - 39	.9891	.9893	.9897	.9900	.9903	.9906
40 - 44	.9835	.9839	.9844	.9849	.9854	.9858
45 - 49	.9754	.9760	.9766	.9774	.9780	.9787
50 - 54	.9643	.9661	.9667	.9673	.9681	.9687
55 - 59	.9474	.9498	.9502	.9509	.9516	.9525
60 - 64	.9195	.9225	.9237	.9243	.9252	.9261
65 - 69	.8747	.8807	.8815	.8829	.8835	.8845
70 - 74	.8081	.8093	.8105	.8113	.8125	.8130
75 - 79	.7018	.7007	.6998	.6987	.6971	.6959
80 - 84	.4448	.4404	.4366	.4361	.4350	.4335
85 +	.0000	.0000	.0000	.0000	.0000	.0000

Source: U. S. Bureau of the Census, Current Population Reports, Series P-25, No. 470, 1971 and NO. 477, 1972.

APPENDIX C

TEN YEAR U. S. MALE-FEMALE SURVIVAL RATES 1960-1970
(Table Computed From 1967, Five Year Rates)

INITIAL	MALE	FEMALE
0 - 4	.9939	.9955
5 - 9	.9930	.9964
10 - 14	.9864	.9946
15 - 19	.9814	.9929
20 - 24	.9803	.9911
25 - 29	.9772	.9875
30 - 34	.9683	.9813
35 - 39	.9521	.9720
40 - 44	.9249	.9582
45 - 49	.8833	.9388
50 - 54	.8255	.9114
55 - 59	.7498	.8682
60 - 64	.6526	.8003
65 - 69	.5406	.7035
70 - 74	.4217	.5669
75 - 79	.2403	.3075
80 - 84	.0000	.0000
85 +	.0000	.0000

Source: U. S. Bureau of the Census, Current Population Reports, Series P-25, No. 470, 1971.

APPENDIX D

SINGLE YEAR CENTRAL BIRTH RATES FOR U. S. 1970-1995

SERIES C (HIGH)

YEAR	AGE OF MOTHER							
	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49
1970	.0011	.0694	.1634	.1383	.0716	.0328	.0085	.0005
1975	.0007	.0679	.1691	.1375	.0678	.0285	.0070	.0004
1980	.0007	.0719	.1942	.1502	.0719	.0285	.0068	.0004
1985	.0007	.0717	.2042	.1670	.0802	.0287	.0065	.0003
1990	.0007	.0717	.2042	.1670	.0802	.0287	.0065	.0003
1995	.0007	.0717	.2042	.1670	.0813	.0288	.0063	.0003

SERIES D (MID-HIGH)

YEAR	AGE OF MOTHER							
	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49
1970	.0011	.0694	.1634	.1383	.0716	.0328	.0085	.0005
1975	.0007	.0622	.1583	.1308	.0649	.0275	.0068	.0004
1980	.0007	.0640	.1754	.1386	.0673	.0271	.0065	.0004
1985	.0007	.0640	.1823	.1465	.0697	.0267	.0062	.0003
1990	.0007	.0640	.1823	.1491	.0720	.0263	.0060	.0003
1995	.0007	.0640	.1823	.1491	.0726	.0259	.0058	.0003

SERIES E (MID-LOW)

YEAR	AGE OF MOTHER							
	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49
1970	.0011	.0694	.1634	.1383	.0716	.0328	.0085	.0005
1975	.0006	.0522	.1352	.1133	.0567	.0242	.0060	.0004
1980	.0006	.0538	.1481	.1184	.0581	.0236	.0057	.0003
1985	.0006	.0538	.1531	.1236	.0594	.0230	.0054	.0003
1990	.0006	.0538	.1531	.1253	.0607	.0224	.0052	.0003
1995	.0006	.0538	.1531	.1253	.0610	.0218	.0049	.0002

SERIES F (LOW)

YEAR	AGE OF MOTHER							
	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49
1970	.0011	.0694	.1634	.1383	.0716	.0328	.0085	.0004
1975	.0005	.0467	.1245	.1063	.0538	.0232	.0057	.0004
1980	.0005	.0461	.1293	.1068	.0534	.0221	.0054	.0003
1985	.0005	.0461	.1312	.1072	.0530	.0210	.0051	.0003
1990	.0005	.0461	.1312	.1074	.0526	.0200	.0047	.0002
1995	.0005	.0461	.1312	.1074	.0522	.0189	.0043	.0002

Source: U. S. Bureau of the Census, Current Population Reports, Series P-25, No. 477, 1972.

APPENDIX E

LIST OF VITAL STATISTICS GATHERED BY KANSAS STATE BOARD OF HEALTH

BIRTHS

1. Births by County of Residence, by Month of Birth, Attendant, and Sex
2. Births by City of Residence, by Month of Birth, Attendant, and Sex
3. Births by County of Residence, by Race, Length of Gestation, Inst., & Plural Births
4. Births by City of Residence, by Race, Length of Gestation, Inst., & Plural Births
5. Births by County of Residence, by Birth Weight in Grams
6. Births by City of Residence, by Birth Weight in Grams
7. Births by County of Birth, by County of Residence
8. Births by City of Birth, by County of Residence
9. Births by County of Residence, by Age Group of Mother, and Legitimacy
10. Births by City of Residence, by Age Group of Mother, and Legitimacy
11. Births by County of Residence, by Total Birth Order
12. Births by City of Residence, by Total Birth Order
13. Births by Month & Sex, by Birth Weight, and Length of Gestation
14. Births by Age Group of Mother, by Birth Order of Child, and Birth Weight in Grams
15. Births by Age Group of Mother, by Weeks of Gestation, and Legitimacy
16. Births by Age Group of Father, Sex of Child, by Race and Age Group of Mother
17. Births by Age Group of Father, by Age Group of Mother
18. Births by Race of Father, by Race of Mother
19. Births by Sex, by Race of Mother
20. Births by Education of Mother, by Education of Father

ILLEGITIMATE BIRTHS

21. Illegitimate Births by County of Residence, by Age Group, and Race of Mother
22. Illegitimate Births by City of Residence, by Age Group, and Race of Mother
23. Illegitimate Births by Age Group and Race of Father, by Age Group and Race of Mother
24. Illegitimate Births by Length of Gestation and Sex, by Birth Weight in Grams
25. Illegitimate Births by Education of Mother by Education of Father

STILLBIRTHS

26. Stillbirths by Month of Birth, by Sex

27. Stillbirths by County of Res., by Sex, Inst., Plural Births, Attendant, & Legitimacy
28. Stillbirths by City of Res., by Sex, Inst., Plural Births, Attendant, & Legitimacy
29. Stillbirths by Age Group of Mother, by Race of Mother
30. Stillbirths by Cause of Death, Weeks of Gestation, and Sex
31. Stillbirths by Cause of Death, by Birth Weight
32. Stillbirths by Cause of Death, by Birth Order, Age Group of Mother, & Legitimacy

DEATHS

33. Deaths by County of Residence, by Month of Death
34. Deaths by City of Residence, by Month of Death
35. Deaths by County of Residence, by Sex, Race, and Marital Status
36. Deaths by City of Residence, by Sex, Race, and Marital Status
37 Deaths by County of Residence, by Institution, Attendant, and Performance of Autopsy
38. Deaths by City of Residence, by Institution, Attendant, and Performance of Autopsy
39. Deaths by County of Residence and Sex, by Age Group
40. Deaths by County of Residence, by County of Death
41. Deaths by City of Residence, by County of Death
42. Deaths by City of Occurrence
43. Deaths by Age Group and Sex, by Race
44. Deaths by Cause and Sex, by Age Group (4 digits)
45. Deaths by Cause and Sex, by Age Group (3 digits)
46. Deaths From Selected Causes and Sex, by Age Group
47. Deaths by Cause, by County of Residence
48. Deaths by Cause, by City of Residence
49. Deaths by Cause, by Month of Death
50. Deaths by Cause, by Institution
51. Deaths by Cause and Sex, by Race, and Marital Status
52. Deaths by Cause, by Age Group (105 Counties)

MATERNAL DEATHS

53. Maternal Deaths by County of Occurrence
54. Maternal Deaths by City of Occurrence

INFANT DEATHS

55. Infant Deaths by County of Residence, by Month of Death
56. Infant Deaths by City of Residence, by Month of Death
57. Infant Deaths by County of Residence, by Age Group, by Race, and Sex
58. Infant Deaths by City of Residence, by Age Group, Race, and Sex

NEONATAL DEATHS

MARRIAGES

DIVORCES