AN EXAMINATION OF GROWTH POTENTIAL IN SMALL TOWNS

by

J. RONALD NEWMAN

B. S., University of Central Arkansas, 1971

-

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

1975

Approved by:

Major Professor

CONTENTS

CHAPTER I

PURPOSE OF THE STUDY

Predicting the Future

The purpose of this study is to determine if growth potential can be predicted in a small town at a given time by examining variables of growth indication that exist in that town at that given time.

Predicting the future is of utmost concern to city and regional planners for several reasons. The basic one is the premise that in a dynamic situation, one cannot adequately plan for the future if an abstract of the future is not foreseen. The planner attempts to be aware of looming "dangers" such as overpopulation, pollution, depletion of natural resources and uncontrolled sprawl. It is, therefore, necessary to have a clear picture of the future, or more exactly, of possible futures.[1]

There are various useful methods available to planners with which one might predict growth in a given area. The term "growth", however, is an ambigious term. The dictionary defines the word as "the process of gradual increase". In terms of an area inhabited by humans, however, growth could take the form of population increase, economic base increase, social and cultural enlightenment, increase in land area, or a combination of these and other functions. It is common practice by the layman, however, to equate "growth" with population increase. Population is, in fact, a basic indicator of growth in that an increase in population usually accompanies other types of urban and regional growth. Moreover, a decline in population is usually due to a decrease in those urban functions.

For the purposes of this study, therefore, "growth" is defined as "an increase in population". This is not to say that population increase is desirable in all areas or that it is the ultimate goal of planning endeavors. Planners attempt to predict population increase or decrease as a basic indicator of the future of an area due to the fact that the human being is the element for which we plan.

There are two basic techniques for determining future population in an area: the estimate and the projection.[2] Estimating techniques basically use facts about the population of the area at earlier dates to reach an estimation of population in that area for future points in time. The simplest of these methods would be the "extrapolation" or "straight-line" technique. This method employs a single group of parameters, past population, to determine future population. It is performed by plotting the past population on graph paper, year by year, and fitting a straightline to the median of the plotted points. This line, extended to future years on the graph, designates population estimates for those years. This method may be satisfactory for very general, long-range estimates but should not be used to forecast population to a specific year. Assuming that a steady rate of growth or decline will occur is an invalid assumption to make.

The extrapolation method may be mathematically performed by utilizing a linear regression analysis. This technique attempts to explain how past population would explain the variance of future population in a given area. The result of the analysis is a formula which defines a line of least squares for population, over time. It is basically a more precise "straight-line" projection. A population estimate should never be used to determine population for a precise point in time, but should be used to ascertain possible future trends in population increase or decrease.[3]

A second technique to determine future population is the projection. This technique is used to reach conclusions referring to a particular point in time which lies in the future. In general, projections assume that "if A, B, and C, are true in the future, then the population will be D for a particular year".

Perhaps the most widely used and the most sophisticated method of population projection is the cohort survival model of population dynamics. This method utilizes age groups by race and sex, survival rates, birth rates, and migration rates to project population. The "cohort" is used to project changes in those factors that affect the size and composition of an area's population over time; namely births, deaths, and migration. The results are projections by age groups, race, and sex. The cohort survival model requires a substantial amount of computation. Due to the time, effort, and human error involved, computer programs have been developed to run the model. Due to the fact that the majority of planning firms have no computer hardware, this approach to population projection is neglected by many planners.[4]

There are various other methods for both estimating and projecting future population and future growth in an area. The preceeding discussion of technique presented two widely used methods and does not represent a comprehensive view of population estimation or projection. It was meant to qualify type in each category.

Definitions:

The following definitions are meant to clarify the terminology used in this report.

Growth is increase in population. It is represented in this study in terms of absolute increase, positive or negative over time.

Growth potential is the propensity for a town to increase in population. This growth, or increase in human habitation, will not be defined further as controlled or contained growth versus uncontrolled or sprawling growth. The methodology does not take into account the possible ramifications of the future predicted growth potential.

Small town is defined in this research as a town containing a population of 2,500 to 10,400. This size range was selected due to the availability of data and will be discussed in a later chapter.

Variables of growth indication are selected variables that were tested to determine the extent to which they might predict growth potential. These variables were examined collectively for each town at a given year. Then a comparison was made between the extent of growth prediction by these variables and the actual growth that took place after the base year.

Hypothesis:

It is the hypothesis of this study that existing variables of growth indication may be used as a partial predictor of growth potential in a small town. The study was conducted in the state of Kansas.

A selected sample (n=36) of towns in Kansas was used, restricting those towns to a population size of 2,500 to 10,400.

CHAPTER II

MODELING THEORY

Purpose of Modeling

In his article " A Short Course in Model Design,"[5] Lowry discusses modeling urban systems and the ramifications of those models toward planning decisions. He states that the computer is not wiser than the human mind but rather, it is capable of performing repetitive tasks at high speed with absolute accuracy. This implies that models may be built to emulate the urban environment but they are, in fact, a product of the human mind and not of the computer's binomial brain.

Britton Harris defines a model as "a simplified abstraction from the real world."[6] In another vein, a model may be called a mathematical description of the real world. At any rate, regardless of the method of modeling, it is an attempt to recreate the real world or a part of the real world, systematically, by means of mathematical or statistical methods.

Types of Models

Models may be classified into three categories: descriptive models, predictive models, and planning models.[7]

Descriptive models attempt to replicate the features of an already observed process. This type of model attempts to reduce the complexity of the observed world to a relationship of mathematical language. It may offer a means to determine values for hard to measure variables from input data of easy to measure variables. This type of model emulates the existing world but does not satisfy the planner's need for information about the future.

Predictive models, as the name implies, attempt to predict the future. This type endeavors to determine the cause-effect relationship between two or more variables. "Effect" might be predicted if the present and future value of "cause" is known (e.g., a one unit change in the value of X will cause the value of Y to change by 5 units).

The third model category is the planning model. The planning model incorporates prediction with evaluation of that prediction in terms of planning goals. The planning model might approach a problem by suggesting alternate means of solution. Then the consequence of each course of action is forecast into the future in an attempt to predict the results. The model might assign a "score" to each result and the highest score would designate the optimum solution to the problem. Planning models are still in their infancy, although much academic research has been devoted to their development.

Of these three types of models, the predictive type would be the one this study is concerned with. In an attempt to prove the hypothesis of this study, a predictive model will be calibrated, and tested.

Construction of a Predictive Model

Steger categorizes predictive models two ways: controlled variation of independent variables and simulation forecasting models.[8]

The first category of predictive models projects relationships between relevant independent and dependent variables. Then the independent variables are changed and an alternative result is determined.

Simulation forecasting models attempt to simulate urban or regional development by means of mathematical procedure. The types of procedure range from regression equations to equations based on probability of occurance.

Steger includes mathematical optimizing models in his categories of predictive models, although this type assumes development has already been predicted.

This type of predictive model that seems to best fit the needs of the hypothesis in this study is the simulation forecasting model which will utilize multi-variate component factor analysis as the mathematical testing procedure.

The model will utimately explain to what degree the selected variables predict growth potential in the selected small towns.

CHAPTER III

SCOPE OF THE STUDY

Approach

As was stated in Chapter I, it is the hypothesis of this study that existing variables of growth indication may be used as a partial predictor of growth potential in a small town.

The study was approached by asking the question, "Is it possible to examine a number of variables that exist during a given year in a small town and by collectively analyzing those variables, to determine if that town has the potential to grow or more specifically, to increase in population?"

At this point, it was necessary to select the variables which might be indicators of growth potential. In an attempt to build a workable model for prediction of growth potential, variables were selected which could be readily obtained by the user of the proposed model. Therefore, the test variables were limited to those included in the U.S. Census and similar publications such as city-county data books.

The next step in approaching the problem was to select a number of "small towns". As was previously stated, observed towns were limited to those with a population of 2,500 to 10,400 persons. The lower limit is due to the non-availability of detailed census information for towns smaller than 2,500. The upper limit was chosen in that towns much larger than 10,000 begin to have characteristics and services that tend to suggest "largeness" rather than "smallness". (This observation was strictly biased on the part of the author and can be debated, as everyone's concept of "small" or "large" is a personal bias.)

Towns which fluctuated between 2,500 and 10,400 from 1950 to 1970 in the State of Kansas were observed to test the hypothesis. A total of 36 towns qualified for the testing procedure. These towns were listed in alphabetical order and were assigned a code number of T_1 to T_{36}.

A total of 11 independent variables were tested as growth indicators. These variables were assigned codes of V_2 to V_{12}.

The year 1950 was selected as the first base year for the test. Each set of variables (V_n) were tested by multivariate analysis for the observed towns to determine the extent to which they predicted the growth (or decline) which took place after 1950. The same test was performed for 1960 to determine the prediction of the growth after 1960.

The following chapter will further discuss the methodology used in the study.

CHAPTER IV

METHODOLOGY

Definition of the Variables

The following variables were built into the growth potential model: population increase; median value of one-dwelling unit structures; assessed valuation per capita; percent of rental units to total dwelling units; percent of employment to labor force; percent of employment in construction industries to labor force; percent of employment in wholesale and retail trade to labor force; median family income; median school years completed; distance to the nearest small city; distance to the nearest large city; distance to the nearest interstate highway. These tweleve variables were chosen for inclusion in the model because of their relative availability and because they are possibly characteristic of the viability of the sample town. These variables are listed in Appendix II.

The dependent component of the model is Variable 1, population increase. The analysis will indicate whether the variance of population increase can be explained by the independent variables.

Variable 1: Population Increase

Source: U.S. Census[9]

Population increase is defined as the difference in a city's population in a base year and the population of that city in a future year. The model was tested using two 10 year growth periods and one 20 year growth period. Measurements of population increase for the ten year periods were from base year 1950 to 1960 and from base year 1960 to 1970. Measurements of population increase for the ten year periods were from base year 1950 to 1960 and from base year 1960 to 1970.

Measurements of population increase for the 20 year period were from 1950 to 1970. If a tested city showed a decrease in population over a growth period it is termed negative growth and is represented as a negative number.

There are eleven independent components of the model. The independent variables of the base year (1950 for instance) were tested to see if they could explain any of the variance of the dependent variable (or 1950 to 1960 population increase in this case).

Variable 2: Median Value of One-Dwelling Unit Structures

Source: U.S. Census[10]

Median value of one-dwelling unit structures was selected as a possible indicator of standard of living, quality of life, or affluence of a sample town.

Variable 3: Assessed Valuation Per Capita

Source: Kansas Government Journal[11]

Assessed valuation is a possible indication of the monetary worth of personal goods and an indicator of quality of life or affluence.

Variable 4: Percent of Rental Units to Total Dwelling Units

Source: U.S. Census[12]

This percentage was tested as a possible indicator of available housing in a town and was calculated by dividing total rental units by total dwelling units.

Variable 5: Percent of Employment to Labor Force

Source: U.S. Census[13]

This variable was tested as an indicator of employment-unemployment in a town. It is a measure of possible job availability and economic conditions and was calculated by dividing total employment by total labor force of a town.

Variable 6: Percent of Employment in Construction Industries To Labor Force

Source: U.S. Census[14]

Activity in construction industries was tested as an indicator of "progressiveness" or physical growth. This phenomenon might account for housing starts, industrial construction or private construction, which might indicate available loans and a healthy economy. It was calculated by dividing total employment in construction industries by labor force.

Variable 7: Percent of Employment in Wholesale and Retail Trades To Labor Force

Source: U.S. Census[15]

This percentage was tested as a possible measure of economic activity of a town and was calculated by dividing employment in wholesale and retail trades by labor force.

Variable 8: Median Family Income

Source: U.S. Census[16]

Median family income is a measure of wage earnings and a possible indicator of affluence and standard of living in a town.

Variable 9: Median School Years Completed

Source: U.S. Census[17]

This is a measure of educational level of a town's citizens and was tested as a possible affluence or quality of life factor.

Variable 10: Distance to the Nearest Small City

Source: Kansas Highway Department[18]

This characteristic is a highway measurement from the sample town to the nearest city of over 15,000 people (1970 population). It was tested as a positive or negative attractor of the sample small towns.

Variable 11: <u>Distance to the Nearest Large City</u>

Source: Kansas Highway Department[19]

This characteristic is a highway measurement from the sample town to the nearest city of over 30,000 (1970 population).

Variable 12: <u>Distance to Nearest Interstate Highway</u>

Source: Kansas Highway Department[20]

This characteristic was selected as a possible "friction factor" or propensity to live in a town due to its proximity to an interstate highway or a planned interstate highway. It is represented in miles; 1 representing the closest distance. It was assumed here, that in 1950, if the highway was not yet built, plans for that Interstate might have been an indicator of growth potential.

Construction of the Model

The growth potential model was constructed as follows:

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4 + b_5 X_5 + b_6 X_6 + b_7 X_7 + b_8 X_8 + b_9 X_9 + b_{10} X_{10} + b_{11} X_{11}$$

X= independent variables 2 through 12

Y= dependent variable 1

a= intercept constant

$b_1 \ldots b_{11}$ = regression coefficients

This multivariate formula was used to analyze the differences in ten year and 20 year growth for a sample of 36 Kansas towns, ranging in population from 2,500 to 10,400. These are the smallest size towns in which census data is available in detailed listings. No towns were tested that exist close to a metropolitan area. It was felt that, in most cases, those towns are not primarily self-sufficient in terms of economic activities and services. They are largely dependent on the close metropolitan areas.

Testing the Model

The above model was tested using the multiple regression and correlation analysis. This statistical routine investigates the association between the dependent variable (Y) and the independent variable (X). An equation is derived in which the known value of one variable is used to estimate the unknown value of the other. Correlation analysis expresses the degree in which tow variables are related and estimates the efficiency achieved by using one variable to predict the value of another. The formulated equation is stated as a regression of Y on X.

Using a computer program this model was tested with relative ease and mathematical accuracy. The program tested each Y on X using stepwise deletion of X variables that were determined (by t-test) not to be significant in contributing to the regression. In other words, if an X did contribute to the estimation of Y but not significantly, it was deleted. In each "run", independent variables were deleted until only those which were significant remained.

The growth potential model was tested and calibrated in three separate procedures as follows:

Model 1. 1950 Growth Potential

In the first test, the pretense was made that the current year is 1950 and the objective is to determine if the 36 towns have a potential to grow during the next ten years, or to 1960. Independent variables (X), or growth indicators 2 through 12, represent 1950 or "current" conditions. Therefore 1950 median value of one-dwelling unit structures, 1950 assessed valuation per capita, et cetera, are utilized as X variables. In actuality, the Y variable, or population increase from 1950 to 1960 (whether positive or negative), is known and may be tested against the independent variables. Therefore, the equation is calibrated. It will express the degree by which the 1950 independent variables (X) predict actual growth from 1950 to 1960 in the 36 sample towns.

Model 2. <u>1960 Growth Potential</u>

This procedure is exactly the same as the previous one except that here, the model attempts to predict 1960 to 1970 growth with known 1960 variables. The dependent variable is 1960 to 1970 ten year population increase and the independent variables are variables 2 through 12 for 1960.

Model 3. <u>1950 Twenty Year Growth Potential</u>

This procedure was used to test the predicting power of the model for a twenty year period instead of ten years, as in steps one and two. Here, the model attempts to examine 1950 data or X variables and to determine how this relates to the dependent variable, 1950 to 1970 population increase for each sample town.

Stratification

The sample of 36 towns in Kansas (N = 36) were used in testing the regression of Y on X. In each of the three previous models, the towns were divided or stratified in the following manner:

Step A. <u>Aggregated</u> (N = 36)

In this step, each growth potential model was tested using all 36 towns, ranging in population from 2,500 to 10,400 in a single regression formula.

Step B. Upper Stratification (N = 15)

Here, the 15 towns with the largest populations were tested in one formula. The upper 15 towns were not necessarily the same in each model -- 1950, 1960, or 1950 20 year model.

Step C. Lower Stratification (N = 21)

This step alludes to the testing of the 21 towns with the lower populations.

(NOTE: The computer program would accept a minimum sample size of 12.)

As a result of the preceding breakdown, nine computer "runs" were made to test the nine components of the model. They will be referred to in the following way:

Model 1A. 1950 Growth Potential, Aggregated

Model 1B. 1950 Growth Potential, Upper Stratification

Model 1C. 1950 Growth Potential, Lower Stratification

Model 2A. 1960 Growth Potential, Aggregated

Model 2B. 1960 Growth Potential, Upper Stratification

Model 2C. 1960 Growth Potential, Lower Stratification

Model 3A. 1950 Twenty Year Growth Potential, Aggregated

Model 3B. 1950 Twenty Year Growth Potential, Upper Stratification

Model 3C. 1950 Twenty Year Growth Potential, Lower Stratification

CHAPTER V

FINDINGS

Overview of Methodology

Each of the nine growth potential models were tested using multiple regression with stepwise deletion. The computer program was designed such that it calculated each equation in two major steps. After the preliminary calculations, analysis of regression for Variable 1, the dependent variable, was performed as step 1. This step yielded an analysis of variance table, F-test, and coeffecient of determination (R^2), or the percentage of the variance of Y that can be attributed to X; and F ratio, the effect of the respective independent variables on each other. Step two analyzed the relative importance of each X in estimating Y (partial standard B's) and t-test values indicated the significance of each X in the regression equation. If an independent variable proved little importance in predicting Y, it was deleted. The deletion process was designed to maximize the F ratio so that the significance of F would be as low as possible. As variables were deleted, the significance of F (the probability that the F-ratio occured by chance) was minimized due to the deletion of variables. It was assumed in this research, however, that 0.01 is a valid significance level of F for the regression equations used. This is stating that the probability of F occuring by chance is 1% or less. Therefore, the deletion process was stopped when and if the significance of F reached 0.01.

Analysis

Each of the nine equations were subjected to the above process and each resulted in a number of independent variables being deleted. An analysis of each model follows. The complete results of each analysis are shown in Appendices 1 through 9.

Model 1A. 1950 Growth Potential Aggregated

Table 1 illustrates the multiple regression of Model 1A.

Table 1

Multiple Regression of Model 1A
1950 Growth Potential, Aggregated

$Y = a + b_1 X_1$

$R^2 = 0.15$

$F = 6.0844$

Significance of $F = 0.0188$

Dependent Variable: 1 - Population Increase

Independent Variable: 2 - Median Value of One-Dwelling Unit Structures

This model deleted all independent variables except Variable 2, Median Value of One Dwelling Unit Structures. The coefficient of determination, $R^2 = .15$ reveals that only 15% of the variance of Y can be accounted for by X. This suggests that no combination of the eleven independent variables can accurately be used to predict 1950 to 1960 population increase using all 36 sample towns. An R^2 value of .15 is generally considered to be too low for prediction purposes. Even so, the F ratio of 6.08 indicates that this relationship ($R^2 = .15$) is significant at the .019 level.

As a result of the multiple regression equation, it is assumed that Model 1A would not be an accurate predictor of growth potential. The stepwise deletion procedure reduced the model to a simple regression formula with a low coefficient of determination. This model resulted in the least significant analysis and shall be discounted as a possible predicting model.

Model 2A. 1960 Growth Potential Aggregated

Table 2 illustrates the multiple regression analysis of Model 2A.

Table 2

Multiple Regression of Model 2A
1960 Growth Potential, Aggregated

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 + b_9X_9 + b_{10}X_{10} + b_{11}X_{11}$$

$R^2 = 0.69$

F = 4.9692

Significance of F = 0.0005

Dependent Variable: 1 - Population Increase

Independent Variables: 2 - Median Value of One-Dwelling Unit Structures

3 - Assessed Valuation Per Capita

4 - Percent of Rental Units to Total Dwelling Units

5 - Percent of Employment to Labor Force

6 - Percent of Employment in Wholesale and Retail Trade to Labor Force

7 - Percent of Employment in Construction to Labor Force

8 - Median Family Income

9 - Median School Years Completed

10 - Distance to Nearest Small City

11 - Distance to Nearest Large City

12 - Distance to Nearest Interstate Highway

This model emerged as the only one in which no variables were deleted to reach a 1% significance of F. All independent variables are included in the model yielding a coefficient of determination of $R^2 = .69$. This indicates that 69% of the variance of 1960 to 1970 population increase for all 36 sample towns is associated with the variance of Variables 2,3,4,5,6,7,8,9,10,11, and 12. The F ratio of 4.9692 indicates that this relationship is significant at the 0.0005 level, a high level for prediction purposes.

As a predictive model, 2A is possibly significant. In contributing to the regression, variables 4,7,11, and 12 rank high in predicting power. This may indicate that housing, employment and spatial variables are significant in predicting growth potential.

Model 3A. <u>1950 Twenty Year Growth Potential, Aggregated</u>

Table 3 illustrates the multiple regression of Model 3A.

Table 3

Multiple Regression of Model 3A
1950 Twenty Year Growth Potential
Aggregated

$Y = a + b_1X_1 + b_5X_5 + b_{10}X_{10} + b_{11}X_{11}$

$R^2 = 0.36$

$F = 4.3791$

Significance of F = 0.0064

Dependent Variable: 1 - Population Increase

Independent Variables: 2 - Median Value of One-Dwelling Unit Structures

6 - Percent of Employment in Wholesale and Retail Trade to Labor Force

11 - Distance to the Nearest Large City

12 - Distance to the Nearest Interstate Highway

This model deleted all independent variables except Variables 2, Median Value of One-Dwelling Unit Structures; 6, Percent of Retail and Wholesale Trade to Labor Force; 11, Distance to the Nearest Large Town; and 12, Distance to the Nearest Interstate Highway. The coefficient of determination is $R^2 = .36$, indicating that 36% of the variance of 20 year growth in 36 sample towns, can be explained by the variance of Variables 2,7,11, and 12. The F-test ratio of 4.3791 indicates that this relationship is significant at the 0.0064 level or there is greater than 99% probability that this did not occur by chance.

The independent variables in Model 3A can explain less than one half of the variance of Y. In that 50% is arbitrarily being used as a conficence level in this procedure, 36% falls below that criteria and 64% of the variance is unexplained. By partial standard regression coefficients, the variables rank 6,12,11, and 2 in importance of singular predictors. The same rank applies to the t-test value or the predicting power in combination. None of the three can be singled out as substantially higher in importance.

Of the three aggregated models, only 2A seems to be useful in predicting growth potential with R^2 = .69. It seems that testing all 36 sample towns in one regression equation may not be the best way to analyze the data. It does relate that 1960 to 1970 may be the most accurate testing model. Variable r, 6, 11, and 12 seem to be the most significant in causing Y to vary. Variable 4 showed up very strongly in Model 2A and 11 and 12 were predictive in two models, 2A and 3A.

The aggregations suggest that in testing a large population range, spatial or distance variables are the most important. These variables relate to travel which will prove to be a very important variable in this study.

Model 1B. 1950 Growth Potential, Upper Stratification

Table 4 illustrates the multiple regression analysis of Model 1B.

Table 4

Multiple Regression of Model 1B
1950 Growth Potential, Upper Stratification

$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_{10} + X_{10}$

$R^2 = 0.70$

$F = 5.7257$

Significance of F = 0.0116

Dependent Variable: 1 - Population Increase

Independent Variables: 2 - Median Value of One-Dwelling Unit Structures

3 - Assessed Valuation Per Capita

4 - Percent of Rental Units to Total Dwelling Units

11 - Distance to the Nearest Large City

This model deleted all independent variables but Variables 2,Median Value of One Dwelling Unit Structures; 3,Assessed Valuation Per Capita; 4,Percent Rental Units to Total Units; and 11,Distance to the Nearest Large City. The coefficient of determination, $R^2 = .70$ suggests that these variables explain 70% of the variation of Y, 1950 to 1960 growth of the upper 15 cities in the sample. The F-test ratio of 5.73 indicates that this is significant at the 0.012 level, meaning that there is a .91% probability that it did not occur by chance.

The coefficient of determination in this case is substantial enough to warrant further examination. The relative importance of each independent variable, both alone and combined, rank them 2,3,4, and 11. Variables 2 and 3 both allude to housing and property valuation. This may indicate that housing prices or tax base might influence growth. Variables 4 and 11 again (as in Model 2A) appear as significant, possibly again influencing each other.

Model 2B. 1960 Growth Potential, Upper Stratification

Table 5 illustrates the multiple regression analysis of Model 2B.

Table 5

Multiple Regression of Model 2B
1960 Growth Potential, Upper Stratification

$$Y = a + b_3X_3 + b_4X_4 + b_6X_6 + b_7X_7 + b_9X_9 + b_{11}X_{11}$$

$R^2 = 0.85$

$F = 7.7984$

Significance of F = 0.0053

Dependent Variable: 1 - Population Increase

Independent Variables: 4 - Percent of Rental Units to Total Dwelling Units

5 — Percent of Employment to Labor Force

6 - Percent of Employment in Construction to Labor Force

8 - Median Family Income

10 - Distance to the Nearest Small City

12 - Distance to the Nearest Interstate Highway

This model deleted all independent variables except 4,Percent of Rental Units to Total Dwelling Units; 5,Percent of Employment to Labor Force; 6,Percent of Employment in Construction to Labor Force; 8,Median Family Income; 10,Distance to the Nearest Interstate Highway. With a coefficient of determination of R^2 = 0.85, this model was very significant in that it indicates that the above six variables explain 85% of the variance of Y.

The relative importance of the independent variables in estimating Y is 12,8,10,6,4,and 5. Distance variables 12 and 10 again contribute heavily to the variation of Y.

Model 3B. 1950 Twenty Year Growth Potential, Upper Stratification

Table 6 illustrates the multiple regression analysis of Model 3B.

Table 6

Multiple Regression of Model 3B
1950 Twenty Year Growth Potential, Upper Stratification

$$Y = a + b_4X_4 + b_5X_5 + b_6X_6 + b_8X_8 + b_9X_9 + b_{10}X_{10} + b_{11}X_{11}$$

$R^2 = 0.90$

$F = 9.3420$

Significance of F = 0.0043

Dependent Variable : 1 - Population Increase

Independent Variables: 5 - Percent of Employment to Labor Force

6 - Percent of Employment in Wholesale and Retail Trade to Labor Force

7 - Percent of Employment in Construction to Labor Force

9 - Median School Years Completed

10 - Distance to the Nearest Small City

11 - Distance to the Nearest Large City

12 - Distance to the Nearest Interstate Highway

Of all the regression models, 3B emerged as the one with the greatest predicting power. The model deleted all independent variables except 5, Percent of Employment to Labor Force; 7, Percent of Employment in Wholesale and Retail Trade to Labor Force; 6, Percent of Employment in Construction to Labor Force; 9, Median School Years Completed; 10, Distance to the Nearest Small City; 11, Distance to the Nearest Large City; and 12, Distance to the Nearest Interstate Highway. The coefficient of determination, $R^2 = 0.90$, indicates that these seven variables account for 90% of the variation of 20 year growth in the upper 15 towns. This is a very high R^2 for prediction purposes.

The variables rank 11,10,12,7,9,5,6, in importance in predicting Y. This would indicate that the spatial variables were the most important followed by the economic variables.

Model 1C. 1950 Growth Potential, Lower Stratification

Table 7 illustrates the multiple regression analysis of Model 1C.

Table 7

Multiple Regression of Model 1C
1950 Growth Potential, Lower Stratification

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_9X_9 + b_{10}X_{10} + b_{11}X_{11}$$

$R^2 = 0.79$

F = 4.6362

Significance of F = 0.0100

Dependent Variable: 1 - Population Increase

Independent Variables: 2 - Median Value of One-Dwelling Unit Structures

3 - Assessed Valuation Per Capita

4 - Percent of Rental Units to Total Dwelling Units

5 - Percent of Employment to Labor Force

6 - Percent of Employment in Construction to Labor Force

7 - Percent of Employment in Wholesale and Retail Trade to Labor Force

10 - Distance to the Nearest Small City

11 - Distance to the Nearest Large City

12 - Distance to the Nearest Interstate Highway

This model only deleted variables 8,Median Family Income and 9,Median School Years Completed. The coefficient of determination $R^2 = 0.79$, indicates that variables 2,3,4,5,6,7,10,11, and 12 explain 79% of the variation of Y. In relative importance these rank 4,10,11,5,2,3,12,6, and 7. Of these, variable 4, Percent of Rental Units to Total Dwelling Units, seems to be the most heavily important in terms of partial standard regression coefficients. Spatial variables 10 and 11 also rank high.

Model 2C. 1960 Growth Potential, Lower Stratification

Table 8 illustrates the multiple regression analysis of Model 2C.

Table 8

Multiple Regression of Model 2C
1960 Growth Potential, Lower Stratification

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_7X_7 + b_8X_8 + b_9X_9 + b_{11}X_{11}$$

$R^2 = 0.73$

$F = 5.0256$

Significance of F = 0.0060

Dependent Variable: 1 - Population Increase

Independent Variables: 2 - Median Value of One-Dwelling Unit Structures

3 - Assessed Valuation Per Capita

4 - Percent of Rental Units to Total Dwelling Units

8 - Median Family Income

9 - Median School Years Completed

10 - Distance to the Nearest Small City

12 - Distance to the Nearest Interstate Highway

This model deleted all independent variables except variables 2, Median Value of One-Dwelling Unit Structures; 3, Assessed Valuation Per Capita; 4, Percent of Rental Units to Total Dwelling Units; 8,Median Family Income; 9, Median School Years Completed; 10, Distance to the Nearest Small City; and 12, Distance to the Nearest Interstate Highway. The coefficient of determination, $R^2 = .73$, indicates that these variables are related to 73% of the variance of Y.

These variables rank in importance as 4,2,10,8,9,12,and 3. Here again, variable 4 is revealed as an important variable related to housing along with variable 2.

Model 3C. 1950 Twenty Year Growth Potential, Lower Stratification

Table 9 illustrates the multiple regression analysis of Model 3C.

Table 9

Multiple Regression of Model 3C
1950 Twenty Year Growth Potential, Lower Stratification

$Y = a + b_9X_9$

$R^2 = 0.29$

$F = 7.6319$

Significance of $F = 0.0124$

Dependent Variable: 1 - Population Increase

Independent Variable: 10 - Distance to the Nearest Small City

This analysis deleted all independent variables but Variable 10, Distance to the Nearest Small City. The R value of .29 suggests that this is a poor estimating model. Less than 50% of the variance of Y can be explained by X. Furthermore, this model is reduced to a simple regression formula, which decreases the value further as a predictive equation. The remaining variable, however, is a spatial variable which supports the trend to this point.

It can be seen that this battery of equations might possibly lend support to the hypothesis of this study. The following chapter will summarize the findings and offer a conclusion to this research.

CHAPTER VI

SUMMARY AND CONCLUSIONS

Summary of the Findings

The hypothesis of this study, as stated previously, is that existing variables of growth indication may be used as a partial predictor of growth potential in a small town. The variables of growth indication are represented by variables 2 through 12 and growth potential is represented in the testing procedure by actual growth.

Model 3B, 1950 Twenty Year Grwoth Potential, Upper Stratification, emerged as the most reliable predictive model and the most supportive of the hypothesis. In this model, 90% of the variance of 1950 to 1970 growth can be related to the regression of variables 5,6,7,9,10,11, and 12. The population range for the upper stratification of the sample towns (N = 15) was 4,355 to 8,967, 1960 population.

As the previous chapter on findings indicates, the coefficient of determination of the nine growth models range from the high of R^2 = .90 to a low of R^2 = .15. However, six of the nine models resulted in a reliable predictive equation. The main concern, however, is that the X values or the variables of growth indication (2 through 12) which remained after the stepwise deletion procedure, differed in many models.

Table 10 illustrates the frequency that Variables 2 through 12 were significant in the growth potential models. Each marked box indicates that the variable on the left was found to be a significant independent variable in the model along the upper margin. The right margin depicts the total number of times each variable remained "non-deleted" and the lower margin contains the coefficients of determination for each growth potential model. As can be noted, Variables 2, 10,

Table 10

Frequency of Occurance of Variables

	VARIABLE NUMBER	MODELS									
		1950 AGGREGATED	1950 UPPER STRAT.	1950 LOWER STRAT.	1960 AGGREGATED	1960 UPPER STRAT.	1960 LOWER STRAT.	1950 (20Y) AGGREGATED	1950 (20Y) UPPER	1950 (20Y) LOWER	TOTAL FREQUENCY OF OCCURANCE
		1A	1B	1C	2A	2B	2C	3A	3B	3C	T
(X_1)	2	1	1	1	1		1	1			6
(X_2)	3		1	1	1		1				4
(X_3)	4		1	1	1	1	1				5
(X_4)	5			1	1	1			1		4
(X_5)	6			1	1			1	1		4
(X_6)	7			1	1	1			1		4
(X_7)	8				1	1	1				3
(X_8)	9				1		1		1		3
(X_9)	10			1	1	1	1		1	1	6
(X_{10})	11		1	1	1			1	1		5
(X_{11})	12			1	1	1	1	1	1		6
	R^2	.15	.70	.79	.69	.85	.73	.36	.90	.29	

and 12 each appeared in six different models. This is not to say, however, that they appeared in the same six models. Variable 2, Median Value of One-Dwelling Unit Structures is a measure of housing quality. Variables 10, and 12 are spatial variables that represent distance from a small town, and an interstate highway respectively. Variables 4 and 11 each appeared in five different models. Variable 4, Percent of Rental Units to Total Dwelling Units is a measure of rental house availability and variable 11, Distance to the Nearest Large City is a spatial variable.

Variables 2 through 12 in varying widely in significance among the models, might be better analyzed in a collectice vein. If the eleven independent variables are clustered into six major groups they might be arranged as in Table 11.

Table 11

VARIABLE GROUP	VARIABLE
Property Valuation	2 - Median Value of One-Dwelling Unit Structures 3 - Assessed Valuation Per Capita
Housing Characteristics	2 - Median Value of One-Dwelling Unit Structures 4 - Percent Rental to Total Units
Dollar Characteristics	2 - Median Value of One-Dwelling Unit Structures 3 - Assessed Valuation Per Capita 8 - Median Family Income
Employment	5 - Percent Employment to Labor Force 6 - Percent Employment in Construction to Labor Force 7 - Percent Employment in Wholesale & Retail Trade to Labor Force
Income Level	8 - Median Family Income 9 - Median School Years Completed
Spatial Characteristics	10 - Distance to Nearest Small City 11 - Distance to Nearest Large City 12 - Distance to Nearest Interstate Highway

This table lists the variable groups and the variables that would fall under each group. For instance, the Property Valuation group would contain variables 2 and 3 — Median Value of One-Dwelling Unit Structures and Assessed Valuation Per Capita. Table 12 illustrates the number of models (out of nine) in which at least one variable of a variable group appeared.

Table 12

Frequency of Occruance of Variable Groups

VARIABLE GROUP	MODELS									
	1950 AGGREGATED	1950 UPPER STRAT.	1950 LOWER STRAT.	1960 AGGREGATED	1960 UPPER STRAT.	1960 LOWER STRAT.	1950 (20Y) AGGREGATED	1950 (20Y) UPPER	1950 (20Y) LOWER	TOTAL FREQUENCY OF OCCURANCE
	1A	1B	1C	2A	2B	2C	3A	3B	3C	T
Property Valuation	1	1	1	1		1	1			6
Housing Characteristics	1	1	1	1	1	1	1			7
Dollar Characteristics	1	1	1	1	1	1	1			7
Employment			1	1	1		1	1		5
Income Level				1	1	1		1		4
Spatial Characteristics		1	1	1	1	1	1	1	1	8
R^2	.15	.70	.79	.69	.85	.73	.36	.90	.29	

This table suggests that the variable group of Spatial Characteristics may be the most significant in determining growth potential due to the fact that a spatial variable was significant in defining the variance of Y in eight of the nine models. Only two of these eight had a coefficient of determination of less than $R^2 = .69$. In six of nine models, a spatial variable helped to account for at least 69% of the variance of Y. This may be explained by the substantial coverage of interstate highways in the sample area of Kansas. This would influence the propensity to travel.

The two variable groups, housing characteristics and dollar characteristics were most significant groups. A housing variable was significant in defining the variance of Y in seven of the nine models. Only two of these seven models had $R^2 = .69$. Therefore in five models, a housing variable and a dollar variable helped account for at least 69% of the variance of Y. This was expected by the researcher as housing is a basic function of an urban system and the price, condition and provision of housing is commonly felt to be a determination of growth. In addition, dollar characteristics are significant in any economy.

The remaining groups had four or less models of the nine in which a group variable helped to explain at least 69% of the variation of Y.

Application of the Model

At this point in the research, it is necessary to approach the question, "Has a predictive model been developed?" This question shall be addressed at this time.

The nine regression models resulted in various significance levels and different predicting variables for each size town and time period tested. Therefore, it seems necessary to examine the "best" modesl and to evaluate their predicting power. This evaluation will be done in terms of town size and "best" will be defined by coefficient of determination (R^2).

The model that emerged as the most accurate predictor of growth potential in towns within the upper stratification is Model 3B, 1950 Twenty Year Growth Potential, Upper Stratification. This model displayed a coefficient of determination of R^2 = .90, the highest of any model. Model 3B was applied to the city of Fort Scott, Kansas to determine the extent to which it might explain the growth that occured in that city after 1950. The model is constructed as follows:

$$Y = a + b_4X_4 + b_5X_5 + b_6X_6 + b_8X_8 + b_9X_9 + b_{10}X_{10} + b_{11}X_{11}$$

$\hat{Y} = -1092$

Y: 1950 to 1960 = -925

Y: 1950 to 1960 = -1368

The prediction power of this model is most accurate at the ten year level. The model came within 82% of predicting the actual growth of Fort Scott from 1950 to 1960. It came within 80% of predicting the growth from 1950 to 1970.

The model that appears to be the best predictor of growth potential in the lower stratification of towns is Model 1C, 1950 Growth Potential, Lower Stratification. This model resulted in a coefficient of determination of R^2 = .79. It was applied to Kingman, Kansas to determine how close the model might explain the growth that actually occured in that town after 1950. The model is constructed as follows:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_9X_9 + b_{10}X_{10} + b_{11}X_{11}$$

$\hat{Y} = 344$

Y: 1950 to 1960 = 382

Y: 1950 to 1970 = 422

The predictive power of this model is the most accurate at the ten year level. It came within 90% of predicting the actual growth of Kingman from 1950 to 1960 as opposed to 82% for 1950 to 1970.

This research has been directed toward predicting growth potential, or the propensity for a town to grow or decline. Due to this, models 3B and 1C appear to be the best predictive models.

Model 3B seems to be the most accurate in predicting growth potential in towns of the upper stratification. This model utilizes variables 5,6,7, (employment variables); 9 (income variable); and 10,11, and 12 (spatial variables).

For towns of the lower stratification, Model 1C appears to be the most accurate. This model applies variables 2,3,4,5,6,7,10,11, and 12. Variables 2,3, and 4 are property and housing characteristics; 5,6, and 7 are employment variables; 10,11, and 12 are spatial variables.

It is interesting to note that in the smaller towns, housing is a more important consideration for predicting growth potential. This is due, perhaps, to the fact that small towns lack quality housing but larger towns, overall, have this characteristic.

For a futher comparison, a simple linear regression technique was utilized to project the population of the above cities from 1950 using only past population figures. This formula is constructed as $Y = a + b X$, where Y is projected population and X is past population. This is commonly known as a "past trends" technique. In other words, the population for 1920, 1930, 1940, and 1950 was used to project population to 1960 and 1970 for Fort Scott and Kingman. This will be compared to the actual population increase and to the results of the growth potential model.

The simple regression projection of Fort Scott resulted as follows:

$Y = a + b X$

$\hat{Y} = 9,787$ for 1960, a decrease of 68 from 1950

$\hat{Y} = 9,467$ for 1970, a decrease of 196 from 1950

Y for 1960 = 9,410, a decrease of 925 from 1950
Y for 1970 = 8,967, a decrease of 1368 from 1950

This procedure came within only 7% of predicting actual growth (or decline) of Fort Scott from 1950 to 1960 and 14% of predicting 1950 to 1970 growth. As compared to the growth potential model's high of 82%, the simple regression formula is far less accurate.

Applying the same simple regression formula to Kingman resulted as follows:

$Y = a + b X$

$\hat{Y} = 3{,}603$ for 1960, an increase of 403 from 1950

$\hat{Y} = 3{,}887$ for 1970, an increase of 687 from 1950

Y for 1960 = 3,582 an increase of 382 from 1950

Y for 1970 = 3,622 an increase of 422 from 1950

This procedure came within 95% of predicting 1950 to 1960 growth in Kingman but within only 37% of predicting 1950 to 1970 growth. The growth potential model came within 90% and 82% respectively.

This comparison indicates that the overall predicting power of the growth potential model is more accurate than the simple linear regression model. This indicates that the present year variables are superior to past population as predicting parameters.

Concluding Remarks

As a result of the analysis, credibility is lent to the hypothesis that existing variables of growth indication may be used as a partial predictor of growth potential in a small town. This is due to the discovery that Spatial Variables, Housing Variables, and Dollar Chacteristics are significant in explaining the variance of the dependent variable, growth. All variables in these groups are economic variables as opposed to measures of social values. Perhaps social characteristics would be significant in this research if they were not so difficult to obtain for small towns.

It should be noted that this research leaves way for further study of growth potential prediction. This exercise resulted in a predictive model, with an indication that one model possibly can be developed that would predict growth potential in small towns by utilizing existing data rather than "past trend" data. Following are suggestions for future research:

1. Select and test additional spatial, housing, and monetary variables as well as more "quality of life" indicators. Explore predictive data other than that easily obtained through federal publications, such as social variables.

2. Select a larger sample size with different levels of stratification. If a larger sample were tested, cities might be stratified by 2,500's or 1,000's of people which might uncover new correlations.

3. Select sample from different areas of the United States. Perhaps the Federal Regional Commission delineation could be utilized in attempting to select sample small towns with varying characteristics.

In conclusion, a model such as the one under study, would assist a regional planner in predicting growth potential in small towns and in defining "hot spots" of development in the area of concern. This type of model would ignore past trends which are most often used today as predicting parameters. It would employ only current base year data to make the analysis. If growth potential was discovered, it would then become necessary to perform more indepth analyses of the town. However, this model would be a time saving device to locate growth potential areas and to define them as such, before actual plan preparation.

FOOTNOTES

[1]F. Stuart Chapin, Urban Land Use Planning, 1965, pp. 181 - 183.

[2]William I. Goodman and Eric C. Freund, Principles and Practice of Urban Planning, 1968, p. 54.

[3]Ibid.

[4]Goodman and Freund, op. cit., p. 61.

[5]Ira S. Lowry, "A Short Course in Model Design", Journal of the American Institute of Planners, Vol. 31, No. 2, May 1965, p. 158.

[6]Britton Harris, "New Tools for Planning", Journal of the American Institute of Planners, Vol. 31, No. 2, May 1965, p. 90.

[7]Lowry, op. cit., pp. 158 - 165.

[8]Wilbur A. Steger, "Review of Analytical Techniques for the C.R.P.". Journal of the American Institute of Planners, Vol. 31, No. 2, May 1965, pp. 167 - 169.

[9]U.S. Bureau of the Census, U.S. Census Population: 1970, 1960, 1950.

[10]Ibid.

[11]League of Kansas Municipalities, Kansas Government Journal, 1970, 1961, 1951.

[12]U.S. Bureau of the Census, op. cit.

[13]Ibid.

[14]Ibid.

[15]Ibid.

[16]Ibid.

[17]Ibid.

[18]Standard Highway Mileage Guide, 1966.

[19]Ibid.

[20]Ibid.

BIBLIOGRAPHY

Chao, Lincoln L. Statistics: Methods and Analysis. McGraw-Hill Book Company, 1969.

Chapin, F. Stuart, Jr. "A Model for Simulating Residential Development". Journal of the American Institute of Planners, Vol. 31, No. 2, May 1965.

Chapin, F. Stuart. Urban Land Use Planning. University of Illinois Press, 1965.

Ezekiel, Mordecai and Karl A. Fox. Methods of Correlation and Regression Analysis, John Wiley and Sons, Incorporated, 1959.

Forrester, Jay W. Urban Dynamics. MIT Press, 1969.

Goodman, William I. and Eric C. Freund. Principles and Practice Of Urban Planning, International City Managers Association, 1968.

Harris, Britton, "New Tools for Planning". Journal of the American Institute of Planners, Vol. 31, No. 2, May 1965.

Hill, Donald M. "A Growth Allocation Model for the Boston Region". Journal of the American Institute of Planners, Vol. 31, No. 2, May 1965.

Kerlinger, Fred N. Foundation of Behavioral Research. Holt, Rinehart and Winston, Incorporated, 1964.

League of Kansas Municipalities. Kansas Government Journal, ed. E.A. Mosher, Vol. 37, No. 1, January 1951.

League of Kansas Municipalities. Kansas Government Journal, ed. E.A. Mosher, Vol. 47, No. 1, January 1961.

League of Kansas Municipalities. Kansas Government Journal, ed. E.A. Mosher, Vol. 56, No. 1, January 1970.

Lowry, Ira S. "A Short Course in Model Design". Journal of the American Institute of Planners, Vol. 31, No. 2, May 1965.

Standard Highway Mileage Guide. Rand McNally and Company, 1966.

Steger, Wilbur A. "Review of Analytic Techniques for the C.R.P." Journal of the American Institute of Planners, Vol. 31, No. 2. May 1965.

U.S. Bureau of the Census. U.S. Census of Population: 1950. Vol. 2, Characteristics of the Population, Part 16, Kansas U.S. Government Printing Office, 1962.

U.S. Bureau of the Census. U.S. Census of Population: 1960. General Social and Economic Characteristics, Kansas, Final Report PC(1) - 18C, U.S. Government Printing Office. 1961.

U.S. Bureau of the Census. U.S. Census of Population: 1960. Vol. 1, Characteristics of the Population. Part 18, Kansas, U.S. Government Printing Office. 1963.

U.S. Bureau of the Census. U.S. Census of Population: 1970. Vol. 1, General Social and Economic Characteristics, Kansas Final Report PC(1) - 18C, U.S. Government Printing Office, 1972.

APPENDICES

APPENDIX 1

MULTIPLE REGRESSION AND ANALYSIS OF VARIANCE
MODEL 1A
1950 GROWTH POTENTIAL, AGGREGATED

Source	D.F.	Sum of Squares	Mean Square	F	Signif-F
Regression	1	1300329.1709	1300329.1709	6.0844	0.0188
Residual	34	7266337.3847	213715.8054		
Total	35	8566666.5556			

$R^2 = 0.15179$
Intercept is - 586.17411

	Variable Number	Partial B's	Partial STD. B's	Partial SSQ	T-Test Values
(X_1)	2	0.1434	0.3896	1300329.1709	2.4667

MULTIPLE REGRESSION AND ANALYSIS OF VARIANCE
MODEL 2A
1960 GROWTH POTENTIAL, AGGREGATED

Source	D.F.	Sum of Squares	Mean Square	F	Signif-F
Regression	11	4496827.1119	408802.4647	4.9692	0.0005
Residual	24	1974408.4436	82267.0185		
Total	35	6471235.5556			

R^2 = 0.69489
Intercept is - 560.39001

	Variable Number	Parital B's	Partial STD. B's	Partial SSQ	T-Test Values
(X_1)	2	0.0577	0.2357	51623.5328	-0.7922
(X_2)	3	0.2838	0.1425	68692.1244	0.9138
(X_3)	4	-31.1669	-0.3225	325886.1486	-1.9903
(X_4)	5	7.2974	0.0553	13350.2536	0.4028
(X_5)	6	4.7537	0.0482	6737.1259	0.2862
(X_6)	7	56.9614	0.3396	416628.6046	2.2504
(X_7)	8	-0.1556	-0.2546	133258.7657	-1.2727
(X_8)	9	79.9197	0.1727	57104.8669	0.8332
(X_9)	10	1.0793	0.0620	7498.6265	0.3019
(X_{10})	11	4.3442	0.5028	457391.6545	2.3579
(X_{11})	12	4.8443	-0.3158	390145.9715	-2.1777

MULTIPLE REGRESSION AND ANALYSIS OF VARIANCE
MODEL 3A
1950 TWENTY YEAR GROWTH POTENTIAL, AGGREGATED

Source	D.F.	Sum of Squares	Mean Square	F	Signif-F
Regression	4	4423251.3829	1105812.8457	4.3791	0.0064
Residual	31	7828143.8393	252520.7690		
Total	35	12251395.2222			

R^2 = 0.36104
Intercept is - 1244.61390

Variance Number	Partial B's	Partial STD. B's	Partial SSQ	T-Test Values
(X_1) 2	0.0554	0.1259	156051.8741	0.7861
(X_5) 6	39.5264	0.3524	1360724.7167	2.3213
(X_{10})11	3.5148	0.2961	846480.9083	1.8309
(X_{11})12	-7.2031	-o.3413	1203754.5599	-2.1833

APPENDIX 4

MULTIPLE REGRESSION AND ANALYSIS OF VARIANCE
MODEL 1B
1950 GROWTH POTENTIAL, UPPER STRATIFICATION

Source	D.F.	Sum of Squares	Mean Square	F	Signif-F
Regression	4	4972423.0689	1243105.7672	5.7257	0.0116
Residual	10	2171084.6645	217108.4664		
Total	14	7143507.7333			

$R^2 = 0.69608$

Intercept is 2176.14301

Variable Number	Partial B's	Partial STD. B's	Partial SSQ	T-Test Values
(X_1) 2	0.6911	1.0994	3894348.8002	4.2352
(X_2) 3	-2.9577	-0.8910	2711891.4843	-3.5343
(X_3) 4	-77.8610	-0.5390	1515768.3902	-2.6423
(X_{10}) 11	-6.4062	-0.4281	1253858.3978	-2.4032

MULTIPLE REGRESSION AND ANALYSIS OF VARIANCE
MODEL 2B
1960 GROWTH POTENTIAL, UPPER STRATIFICATION

Source	D.F.	Sum of Squares	Mean Square	F	Signif-F
Regression	6	4087235.9690	681205.9948	7.7984	0.0053
Residual	8	698813.3643	87351.6705		
Total	14	478604.3333			

$R^2 = 0.85399$

Intercept is 4810.22996

Variable Number	Partial B's	Partial STD. B's	Partial SSQ	T-Test Values
(X_3) 4	-16.5134	-0.1229	50163.4190	-0.7578
(X_4) 5	-30.8707	-0.0929	19009.9634	-0.4665
(X_6) 7	56.0195	0.2227	140696.0962	1.2691
(X_7) 8	- 0.4309	-0.5190	624201.7759	-2.6732
(X_9) 10	9.2046	0.5128	907646.6249	3.2235
(X_{11}) 12	-14.1760	-0.6040	842458.7155	-3.1056

APPENDIX 6

MULTIPLE REGRESSION AND ANALYSIS OF VARIANCE
MODEL 3B
1950 TWENTY YEAR GROWTH POTENTIAL, UPPER STRATIFICATION

Source	D.F.	Sum of Squares	Mean Square	F	Signif-F
Regression	7	8209408.8159	1172772.6880	9.3420	0.0043
Residual	7	878766.7841	125538.1120		
Total	14	9088175.6000			

R^2 = 0.90331

Intercept is -10460.03570

Variable Number	Partial B's	Partial STD. B's	Partial SSQ	T-Test Values
(X_4) 5	129.2425	0.2202	246363.8615	1.4009
(X_5) 6	57.2715	0.3883	968918.1483	2.7781
(X_6) 7	28.8582	0.1124	83425.1521	0.8152
(X_8) 9	-298.0278	-0.2888	575522.8511	-2.1411
(X_9) 10	20.2690	0.8194	499790.3261	1.9953
(X_{10}) 11	- 12.0051	-0.8830	536326.9655	-2.0669
(X_{11}) 12	- 26.6522	-0.8241	4324200.9323	-5.8690

MULTIPLE REGRESSION AND ANALYSIS OF VARIANCE
MODEL 1C
1950 GROWTH POTENTIAL, LOWER STRATIFICATION

Source	D.F.	Sum of Squares	Mean Square	F	Signif-F
Regression	9	1112339.5956	123593.2884	4.6362	0.0100
Residual	11	293242.6901	26658.4264		
Total	20	1405582.2857			

$R^2 = 0.79137$

Intercept is 1854.24591

Variable Number	Partial B's	Partial STD. B's	Partial SSQ	T-Test Values
(X_1) 2	0.0445	0.2440	24809.3410	0.9647
(X_2) 3	0.0921	0.0905	3264.7399	0.3500
(X_3) 4	37.3208	0.8013	453493.4657	4.1245
(X_4) 5	-33.2120	-0.3091	35030.8585	-1.1463
(X_5) 6	- 3.8628	-0.0750	3501.2663	-0.3624
(X_6) 7	- 3.3022	-0.0713	4831.0096	-0.4257
(X_9) 10	4.5453	0.3983	50319.2459	1.3739
(X_{10}) 11	- 1.8531	-0.3726	63789.8808	-1.5469
(X_{11}) 12	- 0.3258	-0.0769	4486.2628	-0.4102

APPENDIX 8

MULTIPLE REGRESSION AND ANALYSIS OF VARIANCE
MODEL 2C
1950 GROWTH POTENTIAL, AGGREGATED

Source	D.F.	Sum of Squares	Mean Square	F	Signif-F
Regression	7	1163885.5083	166269.3583	5.0256	0.0060
Residual	13	430095.6345	33084.2796		
Total	20	1593981.1429			

R^2 = 0.73018

Intercept is 273.75168

Variable Number	Partial B's	Partial STD. B's	Partial SSQ	T-Test Values
(X_1) 2	0.1064	0.6524	192295.0664	2.4109
(X_2) 3	-0.1373	-0.1011	5236.7097	-0.3978
(X_3) 4	-52.6945	-0.7480	468650.7188	-3.7637
(X_7) 8	-0.1477	-0.3738	90256.2404	-1.6517
(X_8) 9	77.1208	0.2681	33668.5833	1.0088
(X_9) 10	7.0157	0.4405	198234.3152	2.4478
(X_{11}) 12	1.1628	0.1142	16050.9990	0.6965

APPENDIX 9

MULTIPLE REGRESSION AND ANALYSIS OF VARIANCE
MODEL 3C
1950 TWENTY YEAR GROWTH POTENTIAL, LOWER STRATIFICATION

Source	D.F.	Sum of Squares	Mean Square	F	Signif-F
Regression	1	893287.6789	893287.6789	7.6319	0.0124
Residual	19	2223873.2735	117045.9618		
Total	20	3117160.9524			

R^2 = 0.28657

Intercept is -440.08618

Variable Number	Partial B's	Partial STD. B's	Partial SSQ	T-Test Values
(X_9) 10	11.9217	0.5353	893287.6789	2.7626

APPENDIX 10

SAMPLE AND TOWN SIZE

N	Town Name	1950 Population	1960 Population	1970 Population
1	Abiline	5,775	6,746	6,661
2	Anthony	2,792	2,744	2,653
3	Augusta	4,483	6,434	5,977
4	Baxter Spgs.	4,647	4,498	4,489
5	Belleville	2,858	2,940	3,063
6	Beloit	4,085	3,837	4,121
7	Cherryvale	2,952	2,783	2,609
8	Clay Center	4,528	4,613	4,963
9	Colby	,859	4,210	4,658
10	Columbus	3,490	3,395	3,356
11	Concordia	7,175	7,022	7,221
12	Eureka	3,958	4,055	3,576
13	Fort Scott	10,335	9,410	8,967
14	Fredonia	3,257	3,233	3,080
15	Galena	4,029	3,827	3,712
16	Garnett	2,693	3,034	3,169
17	Goodland	4,690	4,459	5,510
18	Herington	3,775	3,702	3,165
19	Hiawatha	3,294	3,391	3,365
20	Hoisington	4,012	4,248	3,710
21	Holton	2,705	3,028	3,063
22	Hugoton	2,781	2,912	2,739
23	Iola	7,094	6,885	6,493
24	Kingman	3,200	3,582	3,622
25	Larned	4,447	5,001	4,567
26	Lyons	4,545	4,592	4,355
27	Marysville	3,866	4,143	3,588
28	Neodesha	3,723	3,594	3,295
29	Norton	3,060	3,345	3,627
30	Osawatomie	4,347	4,622	4,294
31	Paola	3,972	4,784	4,622
32	Phillipsburg	2,589	3,233	3,241
33	Pratt	7,523	8,156	6,736
34	Russell	6,483	6,113	5,371
35	Scott City	3,204	3,555	4,001
36	Wellington	7,747	8,809	8,072

APPENDIX 11

SELECTED VARIABLES

Variable #	Variable Name
1	Population Increase (Dependent)
2	Median Value of One-Dwelling Unit Structures
3	Assessed Valuation Per Capita
4	Per Cent of Rental Units to Total Dwelling Units
5	Per Cent of Employment to Labor Force
6	Per Cent of Employment in Wholesale and Retail Trades to Labor Force
7	Per Cent of Employment in Construction Industries to Labor Force
8	Median Family Income
9	Median School Years Completed
10	Distance to the Nearest Small City
11	Distance to the Nearest Large City
12	Distance to the Nearest Interstate Highway

APPENDIX 12

STRATIFICATION

Model	Years Tested	Strat.	Population Range
1A	1950-1960	Aggregated	2,589-10,335
2A	1960-1970	Aggregated	2,744- 9,410
3A	1950-1970	Aggregated	2,609- 8,967
1B	1950-1960	Upper	4,085-10,335
2B	1960-1970	Upper	4,459- 9,410
3B	1950-1970	Upper	4,355- 8,976
1C	1950-1960	Lower	2,589- 4,029
2C	1960-1970	Lower	2,744- 4,248
3C	1950-1970	Lower	2,609- 4,194