CHARACTERIZING THE EVOLUTION OF URBAN PATTERNS AND TRAFFIC NETWORK PERFORMANCE

by

Hani S. Mahmassani, Robert Herman, C. Michael Walton
M. Hadi Baaj, Shekhar Govind, Chee-Chung Tong
and James C. Williams

Research Report GM-1986-F

Conducted for the Operating Systems Research Department
General Motors Research Laboratories

by the

CENTER FOR TRANSPORTATION RESEARCH
BUREAU OF ENGINEERING RESEARCH
THE UNIVERSITY OF TEXAS AT AUTIN

November 1986

PREFACE

This report summarizes the results of an eight-month study, funded by the General Motors Research Laboratory, "The Future of Cities: The Role of the Automobile", aimed at exploring ways of characterizing the evolution of urban development patterns and their implications for personal transportation and mobility.

The authors are appreciative of the support, encouragement and technical advice of Mr. William Spreitzer and Dr. Richard Rothery, of the GM Research Laboratory's Operating Systems Research Department, throughout the duration of the study.

Several students and other administrative and technical staff members have participated in the conduct of this research and preparation of this report. In particular, the contribution of Oddvar Steinsholt, undergraduate research assistant, in digitizing the census maps and helping develop the database for the study of spatial density patterns is acknowledged. The help of Donna Ogle in final report preparation is also appreciated.

Naturally, the authors remain solely responsible for the contents of this report, which do not necessarily reflect the views or opinions of the project sponsors.

EXECUTIVE SUMMARY

This report summarizes the results of a study aimed at exploring approaches for the characterization of urban patterns, their evolution and interaction with and effect on personal mobility and transportation. This objective was pursued at three levels: 1) long term (since the turn of the century) evolution of various infrastructural attributes of cities, with a focus on fundamental mechanisms and underlying constants; 2) urban structure, in terms of the spatial patterns of population density and other variables, and its evolution; and 3) urban traffic network performance, in terms of quality of traffic service, and its dependence on the physical and operational features of the network and the associated land use. The three levels address important components in an integrated framework linking the long-term evolution of urban characteristics, the spatial distribution of these characteristics (location of residences and business activities) and the physical and operational features of the transport infrastructure to the flow patterns and resulting quality of service.

In the first activity, the evolution of several infrastructural attributes of two cities (Austin and San Antonio), were tracked over about 85 years, in an attempt to understand the mechanisms underlying the overall development of these areas. A subset of these variables was also tracked over a shorter time period for eleven other cities (Atlanta, Boston, Chicago, Cincinatti, Denver, Los Angeles, Miami, New York, Phoenix, Seattle and St. Louis), providing a means for comparing these cities. The results suggest several hypotheses about future development, and begin to point at useful ways of characterizing urban development patterns. This approach would benefit from further investigation using a larger sample of cities and different variables.

In the second level of the research, the spatial patterns of population density, household automobile ownership and other socio-demographic variables, as a function of distance from the central city core, were found to provide a useful characterization of urban structure, and of its evolution when taken at different time intervals. The principal substantive conclusions from the four case cities (Austin, Atlanta, Dallas, Phoenix) considered for 1960, 1970 and 1980, are: continuing overall dispersion away from the traditional central core, accompanied by the densification of formerly low-density suburbs, with implications for high congestion levels in the densifying "suburban" communities, comparable to those typically associated with the CBD. In addition, the analysis has captured the continuing growth of average household automobile ownership, in all parts of the urban area, and revealed a distinct spatial pattern that seems to be robust across the case areas considered, as well as within radial corridors in the one case that was so analyzed (Austin). This component of the research should, in future effort, consider the spatial patterns of the location of firms and businesses of various types, which could then be considered, along with the spatial distribution of residents (and related variables at the home end), in describing the resulting flow patterns. In addition, it would benefit from the consideration of a richer collection of urban areas. Other suggestions for useful extensions of the approach are discussed in chapter 3 of the present report.

Suburban networks generally exhibit somewhat different characteristics from CBD systems, in terms of physical features, traffic control and adjoining land use and its generation of traffic-interfering activities. The implications of these differences on the performance of the traffic system, in terms of congestion and the associated quality of traffic service, can be addressed using the results generated by the third component of this work. Two principal objectives were accomplished: 1) the development of a system of relations that comprehensively describe the joint behavior of traffic variables such as the average speed, flow, concentration, average fraction of vehicles stopped in the network, and the two-fluid stopped and running time variables; and 2) the investigation of the sensitivity of these relations and their parameters to the network's characteristics, particularly its topological features, prevailing control strategies and degree of interference from the adjoining land use. Methodologically, both objectives were supported by microscopic simulation experiments, which provided useful macroscopic insights into network traffic behavior. The most important conclusion is that it is possible to characterize traffic flow in urban street networks with relatively simple relations among the principal networkwide traffic variables, and that these relations closely parallel those that have been established at the individual facility level, despite the complex interactions that take place in a network.

The analysis presented in chapter 5 suggested that these relations and their parameters appear to follow systematic trends in response to the network features. On the basis of the limited simulation experiments conducted for this study, it appears that the network's physical features have a greater effect on its performance than traffic control through conventional signal timing. Nevertheless, meaningful improvements could be achieved in the concentration ranges under which most actual networks operate.Interesting insights were also obtained regarding the effect of traffic interfering activities (short term events) typically present in an urban setting.

A better understanding of these mechanisms would be useful in trying to understand and deal with the scenarios of deteriorating traffic conditions in densifying suburban communities. The private automobile has been an essential element in the development of these suburban development patterns, which initially seemed to offer congestion-free circulation. Rapid degradation of traffic conditions in these communities could have important consequences on the outcomes of the personal location, mobility and travel decisions of area residents, particularly with regard to automobile ownership and utilization.

Particularly promising opportunities for substantive and methodological contributions through additional research on the various approaches explored in the present study are discussed in the concluding chapter of this report.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES, Continued

LIST OF FIGURES, Continued

LIST OF FIGURES, Continued

LIST OF FIGURES, Continued

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The future development of urban travel and associated car use patterns are intrinsically related to developments in city structure and the location of workplaces, activities and residences. Changes of this nature generally take place only over relatively long periods of time, as considerable inertia appears to be associated with the vast infrastructure already in place and with the lifestyle patterns acquired by residents over many years. At any given time, there are several forces and trends, associated with technological, economic, political and social phenomena, that shape such development. Recently, various indications seem to have generated concern among researchers and planners regarding the future development of urban systems and the potentially significant implications for personal transportation and mobility (Brotchie et al., 1985; Webster et al., 1985).

While the rate of change of urban patterns might seem "slow", it is probably greater than that required by the responsible agencies to plan and implement decisions aimed at the needs associated with the emerging patterns. Similarly, the lead times necessary for the conception, design, testing and production of transportation products and services adapted to the emerging urban patterns are also long compared to the time frames over which changes in these patterns might be taking place. It is thus desirable for the auto industry and other actors involved in urban transportation to characterize, understand and, to the extent possible, anticipate such development in order to provide valuable input to related decisions and actions.

While predicting the future undoubtedly involves a fair dose of speculation, exemplified by the recent collection of articles in Brotchie et al. (1985) or Lakshmanan and Chattergee (1986), understanding of the fundamental processes that have shaped past development provides an essential starting point and sound basis. Several studies have investigated the underlying structural mechanisms and phenomena shaping the evolution of urban systems, but these efforts have been fragmented, often along disciplinary lines, and rarely guided by a specific operational purpose. This is particularly true for those aspects that affect travel behavior and transportation system performance, and their interaction with the broader urban system. Many efforts, insightful in their own right, have been too microscopic in scope to usefully contribute to the macro-level characterization and understanding of complex urban systems.

This report presents the results of an eight-month effort directed at these complex questions. Naturally, it cannot claim to give all the answers, but rather indicates useful and promising directions for exploring, characterizing and understanding urban development patterns as they affect and interact with transport in cities. This work is pursued along three levels. The first, and broadest, level is an exploration of the long term (since the turn of the century) evolution of several infrastructural attributes of two cities (Austin and San Antonio), with a focus on the mechanisms underlying the overall development of these areas. A subset of these variables is also tracked over a shorter time period for eleven other cities (Atlanta, Boston, Chicago, Cincinatti, Denver, Los Angeles, Miami, New York, Phoenix, Seattle and St. Louis), providing a means for comparing these cities.

The second level is concerned with the combined effect of network shape and of the relative locations of activities, residences and workplaces on the characteristics of urban travel, particularly on average trip distances. Thus it incorporates a spatial dimension which is essential to the analysis of urban structure and travel interactions. The evolution of spatial density patterns over the past three decades is examined in four urban areas (Atlanta, Austin, Dallas and Phoenix). Comparisons over time and across cities are performed through the calibration of models of the spatial density as a function of distance from the traditional city centers, allowing the examination of hypotheses regarding suburbanization and spatial dispersion/concentration trends. In this report, only the spatial patterns of residential location and other residence-based variables are analyzed, primarily due to data availability and time limitations, though similar comparative analysis for the spatial density patterns of employment and business location would be particularly valuable to the description of the resulting intra-metropolitan flows.

At the third level, we focus on the characterization of traffic flow quality in urban networks, through the development of macroscopic network-level models relating the principal traffic flow variables in a simple and operationally useful manner. In particular, we examine, using extensive simulation experiments, the variation of the principal model parameters with the features of the network (network topology as well as traffic control), the associated trip patterns and a measure of urban activity. This work points to answers to questions such as: 1) What are the determinants of the quality of traffic service and congestion in an urban street network? 2) How can this quality of service can be influenced, and to what extent? 3) What is the "ultimate" service level that can be obtained in a particular network and urban environment?

The report is organized as follows. The next chapter presents the definition of and analysis of the historical data for about twenty variables for the cities of Austin and San Antonio (selected primarily on the basis of data accessibility to the reseach team), followed by the examination of several of these variables for a sample of eleven cities in the U.S. Chapter three presents the analysis of the spatial density patterns for the four initially selected urban areas: Austin, Dallas, Phoenix and Atlanta. Chapters four and five are devoted to the characterization of traffic service in urban street networks. Chapter four first presents the derivation, calibration, and comparative assessment of three interrelated sets of macroscopic network-level traffic models, while chapter five summarizes the results pertaining to the effect of various network features and urban characteristics on the model parameters (and consequently on traffic quality). The study's overall conclusions are summarized in Chapter six, along with suggestions for particularly promising further research.

CHAPTER 2

CHARACTERIZING THE EVOLUTION OF CITIES

2.1. INTRODUCTION

In a paper presented by John S. Adams (1986) at a workshop on The Future of Infrastructures sponsored by the National Academy of Engineering, the growth of the metropolitan framework is categorized into four eras or "epochs". The key feature that distinguishes between each "epoch" is the mode of transportation used ("Sail-Wagon", "Iron Horse", "Steel-Rail" and "Auto-Air..."). Each "epoch" sees unique metropolitan growth patterns being molded by the transportation available. Adams also speculates about the "epoch" yet to come :

> "Our current romance with computers and telecommunications prompts us to wonder about a fifth epoch, but their impact on the geography of metropolitan areas is still unclear."

One thing is clear. The automobile, like most new technology, brought about a virtual revolution in certain areas. In particular, it caused the spatial reorganization of human activities by redefining distances people could travel for their daily chores. By increasing personal mobility it also fostered the environment for the growth of the service sector - the popularity of the 'drive-in window service' is an excellent example of such a phenomenon.

For a peep into the crystal ball to examine the future of the automobile, we can hypothesize separately on two different time scales. Predictions can be made on a small time scale (roughly, the time to the implementation of the next new major substitute technology), or a much larger time scale (beyond the implementation of the next major substitute technology). Any attempts to predict the role of the automobile in society should be limited to the smaller of the two time scales. Perhaps, some day the automobile will go the way of the sailboat - from being a prime mover to being a pleasure craft. However, such conjectures about the distant future are akin to dabbling in science fiction.

Linked as it is to the spatial organization of the urban centers, the automobile's security in the short run depends primarily on the future of the metropolises and how they evolve. A number of pertinent questions can be asked in this regard. Will the metropolises continue to develop 'bedroom communities' necessitating a daily home-work-home trip of a considerable distance or will the communities reorganize and become self-sufficient in terms of providing adequate work opportunity (and space) at the local level? How will the home-market-home trip change over time? Will there be a conglomeration of a few large stores in key shopping areas removed from residential communities (such as seen in shopping malls and plazas) or will there be concentration of a number of small stores located much closer to the homes of the consumer (such as often seen in mixed development areas)? There are also some very interesting speculations about what telecommunications will do to the market for transportation and the extent to which "transportation" as we understand it today will be replaced by "communications".

For an urban center, any of these scenarios discussed in the previous paragraph could be equally likely. What eventually comes about will be due to the effects of a multitude of factors including economic, geographic, political, social and the infrastructural setup of the urban area. The automobile usage by the community would be different from one another as each area goes through diverse evolutionary patterns.

Consider the recent demographic shift towards cities. The bulk of the market for automobiles has been, is, and will continue to be, in the cities and urban areas. A study on the future of the automobile would have to be based on predictions about the future of the urban centers. One way to look at the future is to study the past in the hope that certain evolutionary trends may be established. Therefore, a look at the evolution of the metropolis during the "automobile...epoch" could not only provide valuable insights about how the automobile has shaped the growth of these cities, but also lead to predictions (for the time period left in this "epoch") about the future (usefulness) of the car in the urban context.

This study looks at historical data for variables that represent a few infrastructural attributes of two cities (Austin and San Antonio) and follow their evolutionary track from the turn of the century to the present. Similar variables are evaluated for eleven other cities (Atlanta, Boston, Chicago, Cincinnati, Denver, Los Angeles, Miami, New York, Phoenix, Seattle and St. Louis) over a shorter period of time. The data (on Austin and San Antonio) is initially viewed in the spirit that pictorial representations of data derived from a complex situation may provide insights into the mechanisms of the evolution of the system. An attempt is then made to characterize the evolution of these two cities by the use of "snow-flake diagrams" (a classification scheme introduced by Herman and Montroll (1972) for characterizing countries and representing them pictorially in the form of multi-axis phase diagrams). For the group of the eleven cities, similar diagrams are constructed for comparisons both across the cities and across time. Such an endeavor would allow us to examine how cities have evolved, and provide a basis for predicting the future of these cities.

2.2. DISCUSSION OF DATA

Acquisition of historical data is never an easy task. Even when data is available, there is virtually no means of checking its validity. It is not surprising therefore, that a large number of different publications had to be used to compile the data base for this study. As such, these sources can be divided into two distinctly different groups. The distinction between the types of sources depends on the fact that the first type of source provides disaggregate data where actual counts have to be made of the variables values while the second type of source has data already compiled in a format which can be used readily for analysis.

The first type of source consisted of the City Directories of Austin (1959 to 1985) and San Antonio (1900 to 1985). These directories are published each year and are available through the Public Library or the Chambers of Commerce of the two cities. These directories have a wealth of information for the cities in question and provide a fascinating glimpse into the antecedents of the modern metropolis, in terms of providing valuable (raw) statistical information and giving the researcher a panoramic view (across time) of changes in the city.

An actual count was made of all establishments and professions listed in the directories (e.g., the number of restaurants, number of doctors etc.). This data was complimented by using a number of sources of the second type. Data for the consumption for water and power were obtained from various publications of the respective city utility departments. Utility hookup data are listed in basic data sheets published by the city; for most years, these are also listed in the "City Directory". Street mileage was also obtained from the basic data sheets. For some years, when this data was not available, mileage of urban streets (on a county basis) was obtained from the Highway Department and reasonable adjustments were made to get data on the city level. A similar process was followed for years where data for the number of vehicles was available only on a county basis. Population and area estimates were found in census publications.

A major data source of the second type came from the Bureau of the Census in their publication "County and City Data Book" (1949-83). These volumes contain statistics on such varied items as vital statistics, housing, trade, manufacturing, crime, city government finances and even bond ratings. The drawbacks of using this source are that certain variables are not represented in each volume and that there are currently only seven published volumes (circa 1949, 1956, 1962, 1967, 1972, 1977 and 1983). Further complicating the issue is the fact that not all variables presented in any one volume are for a single common year. Despite these shortcomings, it remains the best single data source for a study of this nature. These seven volumes were utilized for building the data base for the group of the eleven cities mentioned earlier.

As can be seen from the preceding discussion, the sources that have been utilized in building the data base for this study are quite diverse. Not only does this raise the question of homogeneity of the data sources, but there is also the added problem of missing data for key variables for certain years. These features could reduce the confidence limits of the inferences drawn. However, in the view of the authors, the procedures being used in this study and the conclusions inferred, are well within the bounds of the limitations imposed by the use of this data set.

2.3. A TALE OF TWO CITIES

As noted in a previous section, we shall look at pictorial representations of the evolutionary path taken by some attributes of the city of Austin. Later in this section, similar aspects of the nature and growth of the city of San Antonio shall be presented. For a more detailed analysis of these two cities, the reader is referred to a publication by Ardekani, Dona, Govind and Herman (1986) for discussions on Austin and a paper by Govind, Herman and Walton (1986) for discussions pertaining to San Antonio. This section has been based in part on the aforementioned papers.

Comparisons, across both time and space, requires the use of data either normalized on some scale or reduced to a dimensionless quantity. It is a common practice to use population as a normalizing variable and represent attributes on a per capita (or its reciprocal) basis. Density functions in 2 dimensions (involving area) have also been used previously for normalizing. It would be worthwhile to investigate density functions formed in 1 dimension when normalization is done with respect to the total length of streets in a city. The following is a look at a partial history of Austin using the strategies outlined above.

Figures 2.1, 2.2 and 2.3 illustrate how the three basic variables population, area and street-miles, have changed over time. (It would be interesting to observe the changes in lane-miles rather than street-miles; unfortunately, such data are very difficult or impossible to obtain).

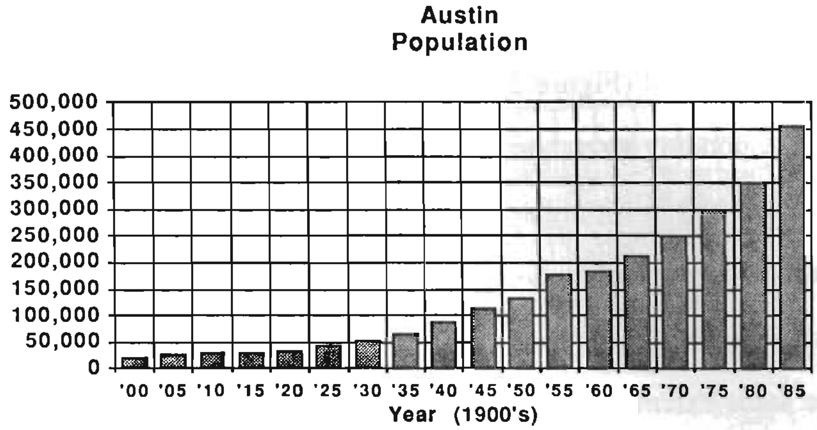


Figure 2.1

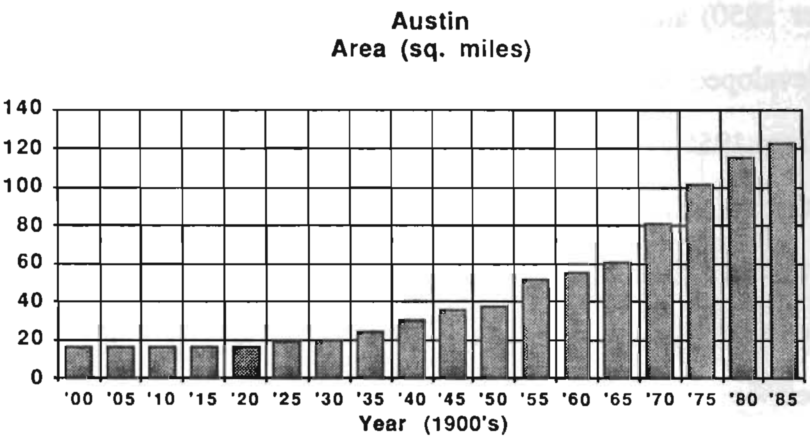


Figure 2.2

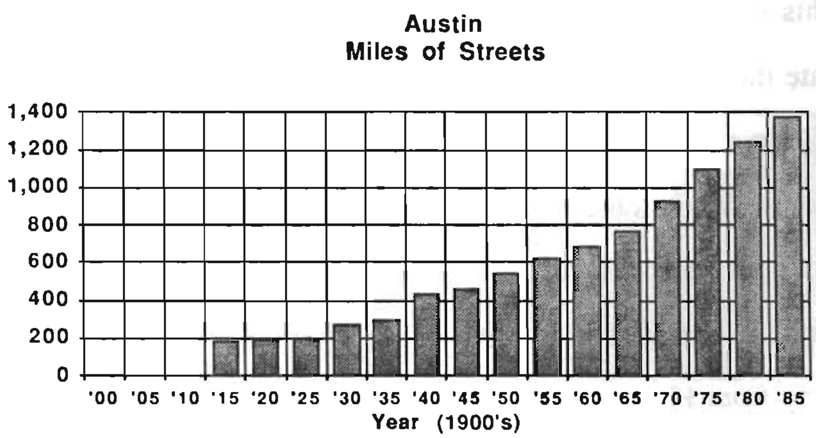


Figure 2.3

The next series of charts (Figures 2.4, 2.5 & 2.6), show how the basic variables change with respect to one another. The density of population shows an approximate linear increase from 1900 to 1950 (Figure 2.4). From 1950 to 1980, the growth in area was faster than the corresponding growth in population. The peak at 1985 represents the spurt of growth the city has had in the eighty's. The density of street-miles may give a general indication of the "efficiency" of land use (Figure 2.5). The decreasing trend from 1950 onwards suggests that even though the city did acquire more land, these have not been opened up to the same extent as land acquired previous to 1950. Another interpretation could be that the density of streets in the inner city (areas included in the city limits till 1950) may not be greater than the density of streets in the suburban areas adjoining the city (developed after 1950) and the decline comes simply from the inclusion of tracts of completely undeveloped land. The density of population on a street-mile basis remains fairly constant from 1955 to 1980 (Figure 2.6). This supports, though inconclusively, the latter argument that the inclusion of undeveloped land causes the decrease. One immediate conclusion drawn from Figure 2.6 is that the growth of streets in the city has lagged behind the growth in population over the last 70 years. This reflects a simple measure of the increasing number of people sharing the use of one unit of the transportation infrastructure.

An important variable to study with regard to this report would be the number of vehicles in a city. In the context of this study, the number of vehicles includes passenger vehicles, farm trucks, commercial trucks and motorcycles. The next four Figures 2.7 - 2.10 examine this variable. The total number of registered vehicles has grown at an "exponential" rate that is much faster than the growth in population (Figure 2.7). The number of vehicles per capita demonstrates this fact (Figure 2.8). As would be expected from the preceding discussions, both the number of vehicles per street-mile and the number of vehicles per square mile, have increased "exponentially" since 1950 (Figures 2.9 & 2.10). What is worthy of note is the fact that all the three charts have a similar characteristic hump from 1920 to 1945 and thereafter, from 1950, show an "exponential"

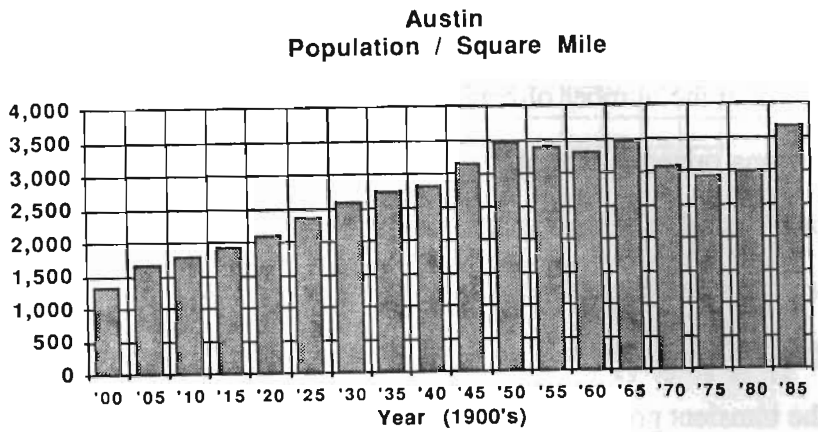


Figure 2.4

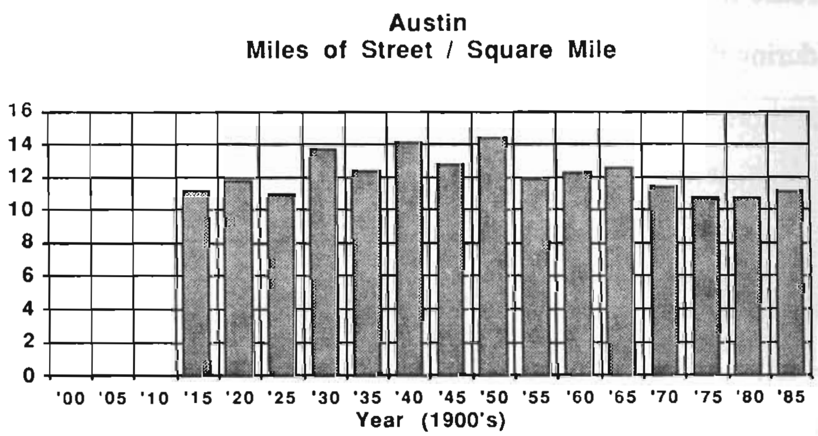


Figure 2.5

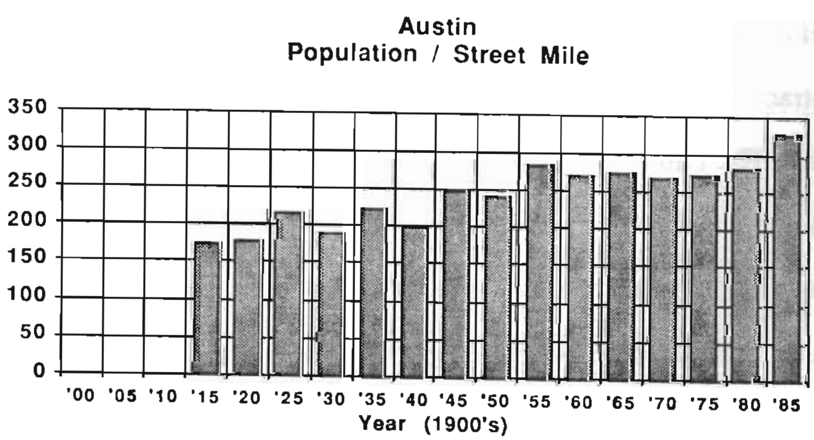


Figure 2.6

growth (Figures 2.8, 2.9, & 2.10). This phenomenon indicates that strong correlations may exist between the dependent variables.

When looking at the number of banks and hospitals (Figures 2.11 & 2.12), one of the pertinent questions raised is that of size. In the case of hospitals, data regarding the number of beds or the number of admissions may prove to be more revealing; similarly, for banks the number of customers or the net assets and liabilities may be more meaningful.

It is suggested that normalized variables derived from hotels/motels might give a fair measure of the transient population of a city (Figure 2.13). However, as in the case of the two previous charts, the question of size confounds the issue. A look at hotel revenues and the number of beds available may provide more definite answers. In terms of Figure 2.13, a sharp increase in the number of motels was noted after World War II while the number of hotels during the same period actually declined.

One of the more intriguing results to come out of this study involves different functions of the number of restaurants (Figures 2.14, 2.15 & 2.16). But for localized fluctuations, the number of people per restaurant has been remarkably steady from 1900 to 1985 (mean = 678, std. dev. = 159). Charts showing the number of restaurants normalized by area (mean = 4.2, std. dev. = 1.3) and street-miles (mean = 0.4, std. dev. = 0.1), also exhibit a similar consistency over the years (Figures 2.15 and 2.16). It would be interesting to compare these values for different cities assuming that other cities display such characteristic numbers as well.

The level of sharing of certain services - auto dealerships (Figure 2.17), doctors, lawyers and contractors (Figure 2.18) - by the population may represent key factors in the growth of the city. The number of doctors, which includes both physicians and dentists, indicates the level of health care available to the city. A lower population number per contractor will lead to the inference of a higher construction activity. Further, the number for law firms could indicate some essential aspects of the nature of the city and the type of activity it thrives on.

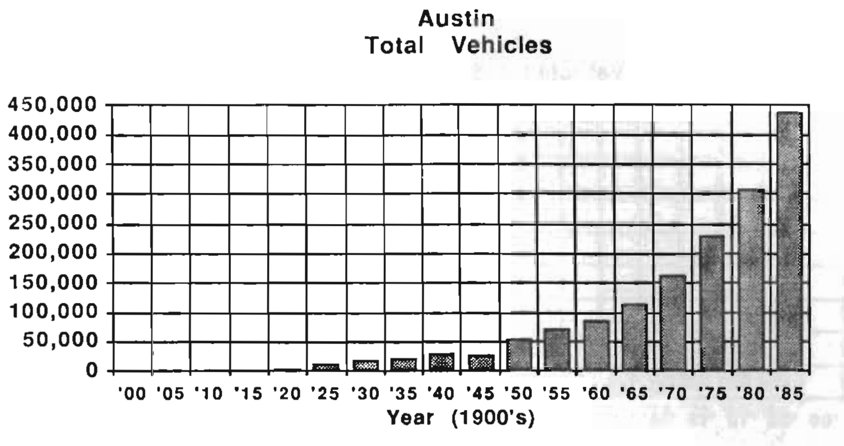


Figure 2.7

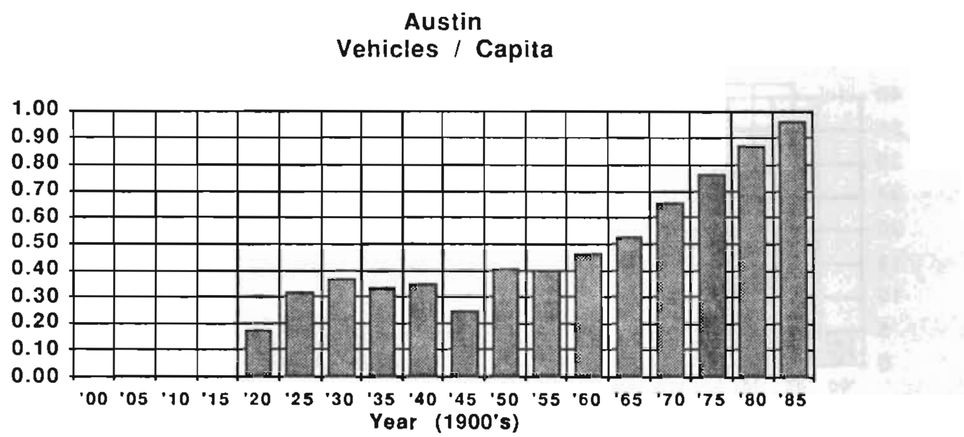


Figure 2.8

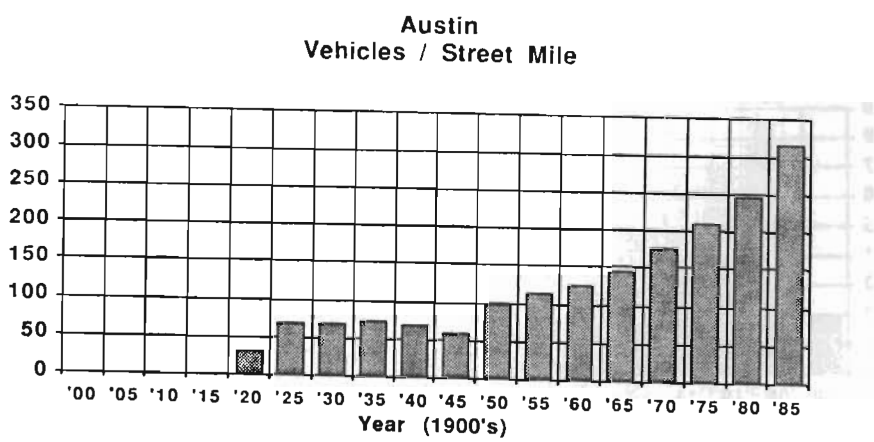


Figure 2.9

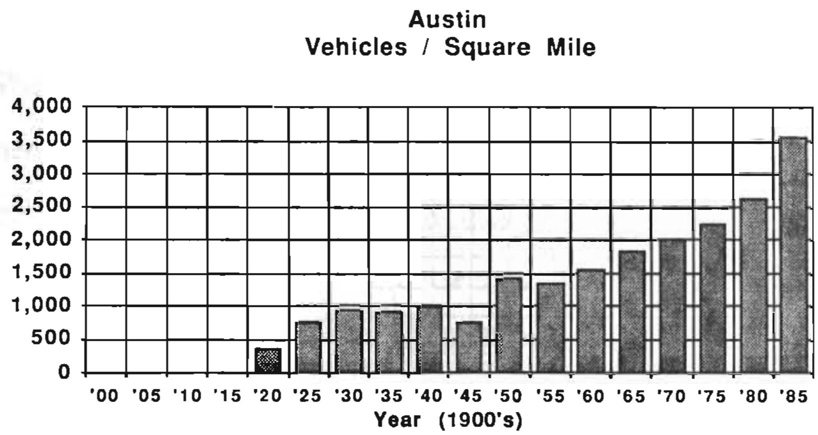


Figure 2.10

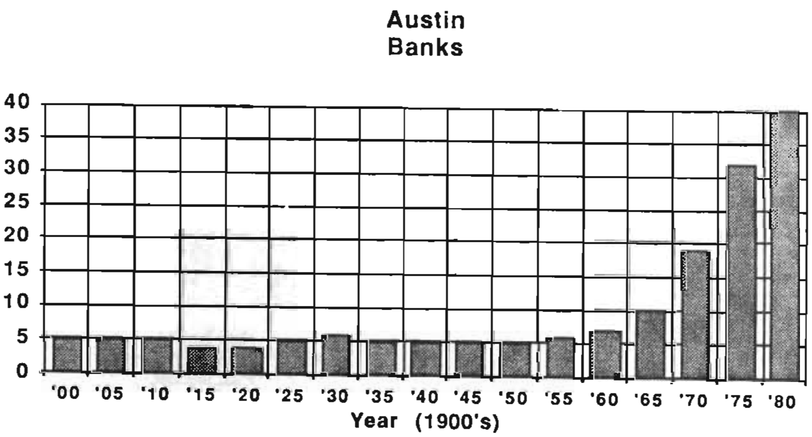


Figure 2.11

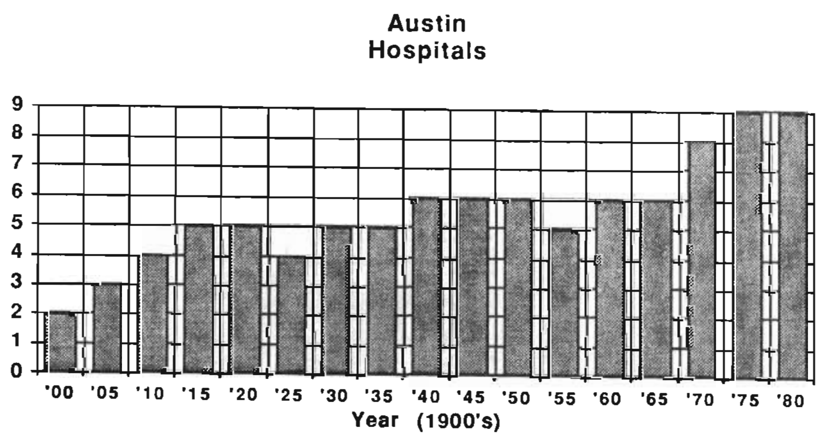


Figure 2.12

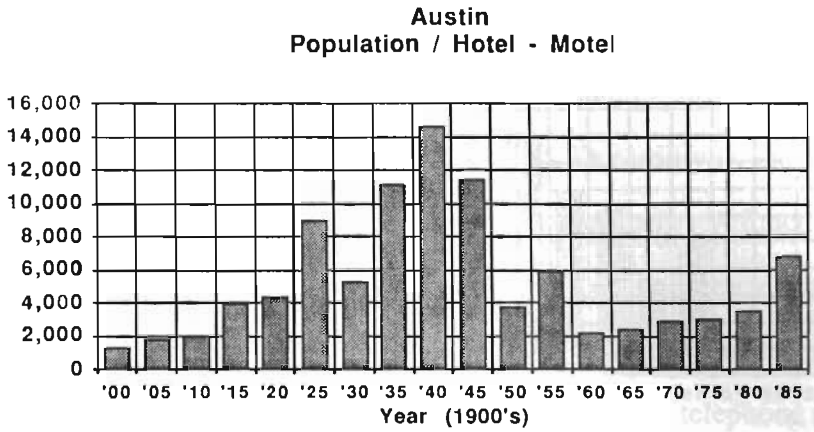


Figure 2.13

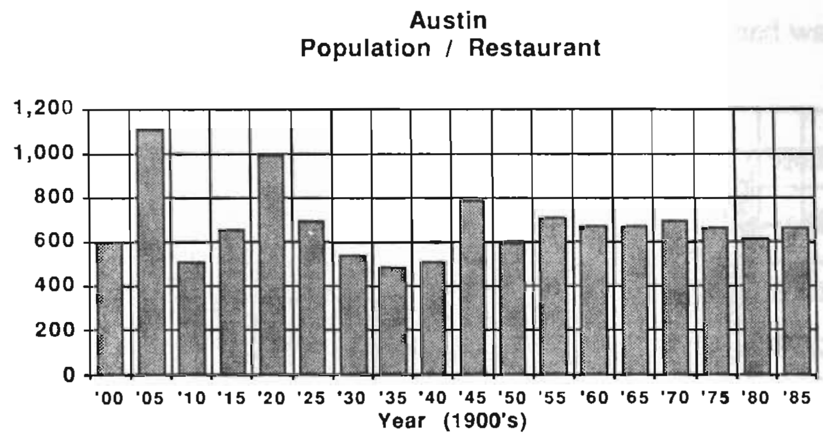


Figure 2.14

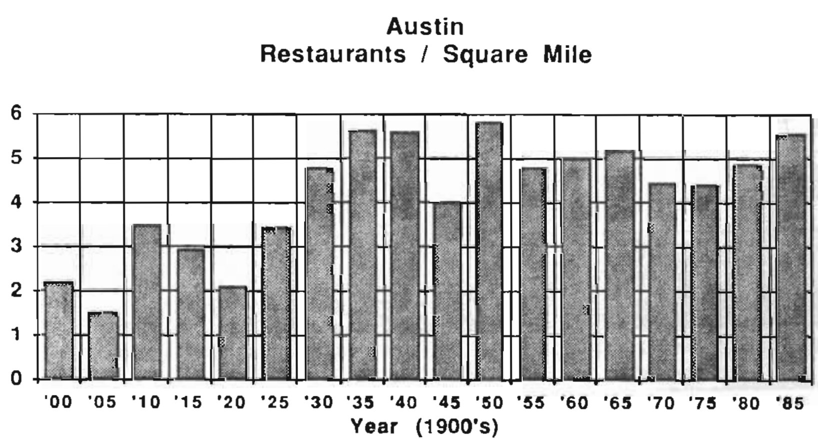


Figure 2.15

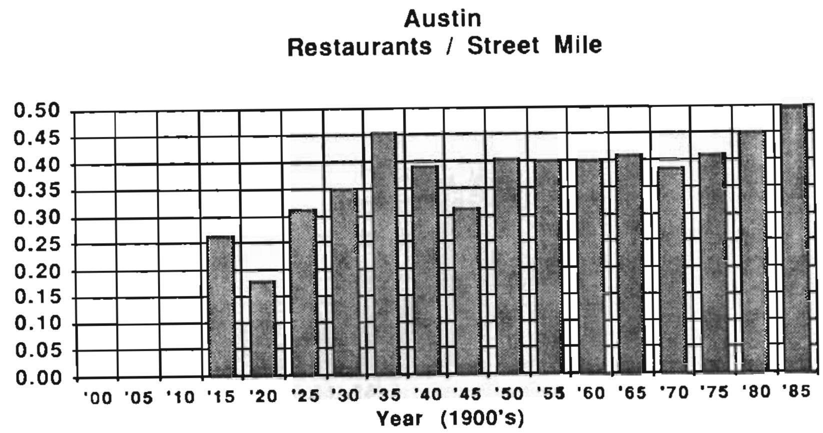


Figure 2.16

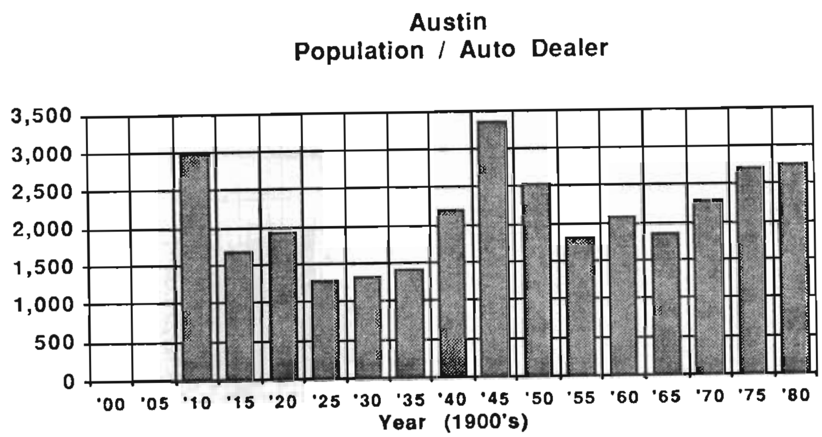


Figure 2.17

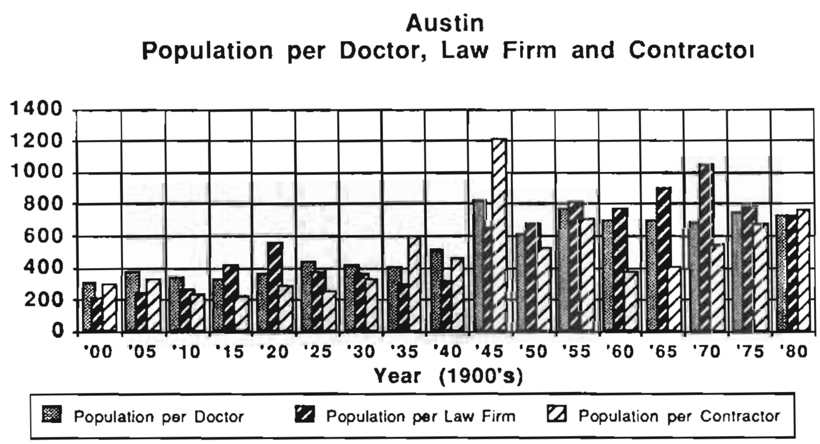


Figure 2.18

A number of conclusions may be drawn by looking at the utility hookups in a city (Figure 2.19 & 2.20). It can be reasoned that every dwelling unit would, most likely, have its own electric meter. However, in the case of almost all apartment complexes, the water meters are not uniquely assigned to each apartment. This utility is, mostly, charged for by a fixed rate (which may be included in the rent of the apartment). A function of the number of electric meters and water meters could, therefore, be used to characterize the composition of the housing market (Figure 2.19). The number of telephone numbers in use may be a reflection of not only the average family size (assuming one telephone number per family) but also the business activity of the city (Figure 2.20). Historical data on the number of business telephone lines would be interesting to look at. A spatial history of a city could also be captured by plotting miles of telephone lines, sewer lines and water pipe lines over time.

The annual per capita water consumption appears to have been kept stable at the 70,000 gallons mark since 1970 and shows no signs of increasing (Figure 2.21). This variable could be used as an indicator of geographical and climatological differences between cities. Another excellent scale for the comparison of cities would be the per capita annual power consumption since it measures an infrastructural attribute which is often of prime concern (Figure 2.22). Cities that use electrical energy for transportation (from the urban/suburban rapid transit systems to quite simply the escalator and elevators in a city full of high-rise buildings) may show higher level of consumption than the other cities. Weather and industry, (among others) affect this variable as well. The 'exponential' rise in postal receipts (Figure 2.23) may not be so steep if this chart was further normalized by the corresponding year's postage rates and indexed for inflation. Current levels of postal revenues could be a strong indicator of the nature of the businesses of the city.

Since normalization can be done by a number of different variables (e.g. area, population, street miles etc.), it can be seen that an almost endless number of such functions can be derived and plotted over time to represent a city's evolution. For the sake

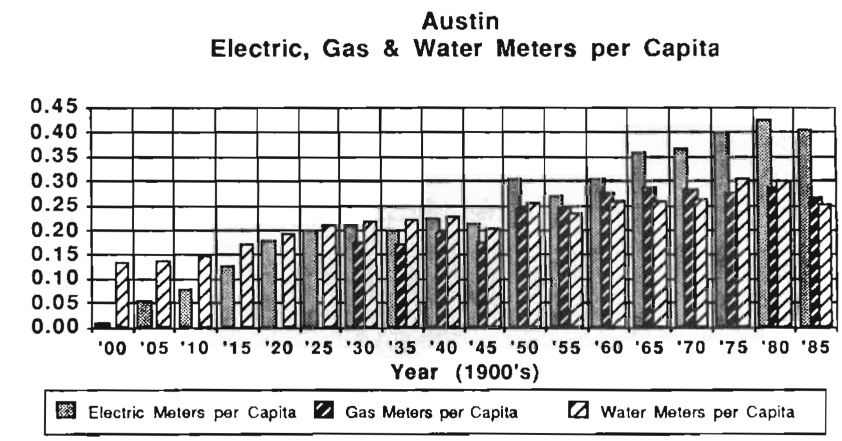


Figure 2.19

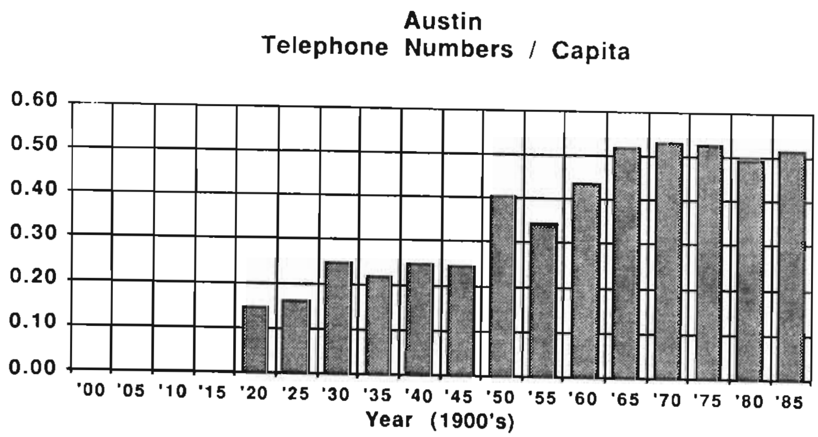


Figure 2.20

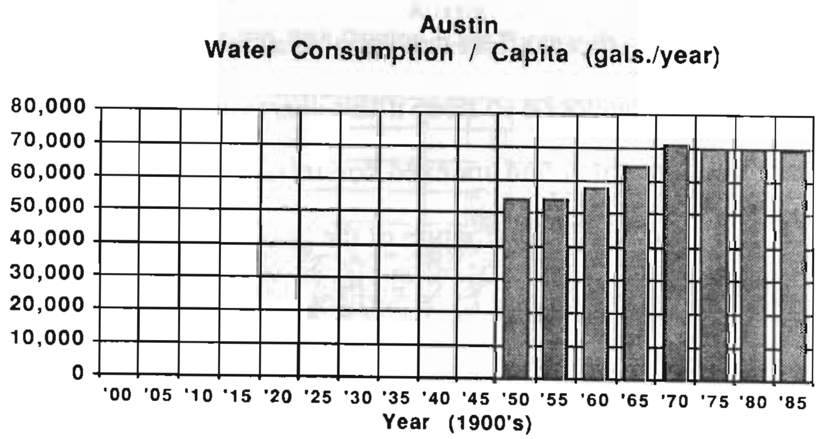


Figure 2.21

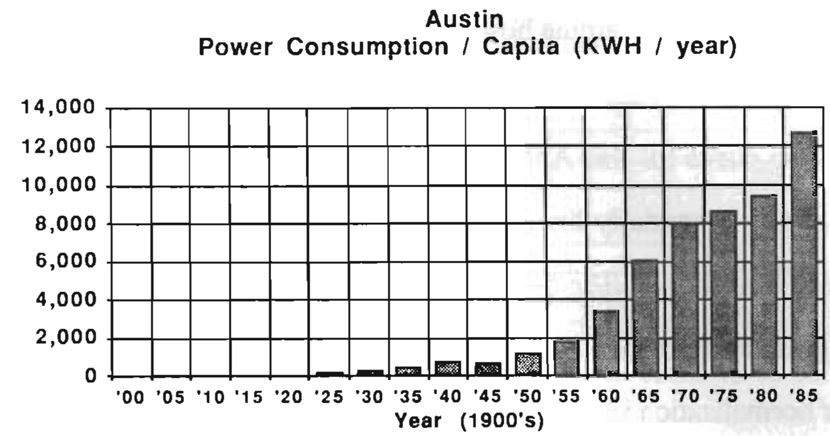


Figure 2.22

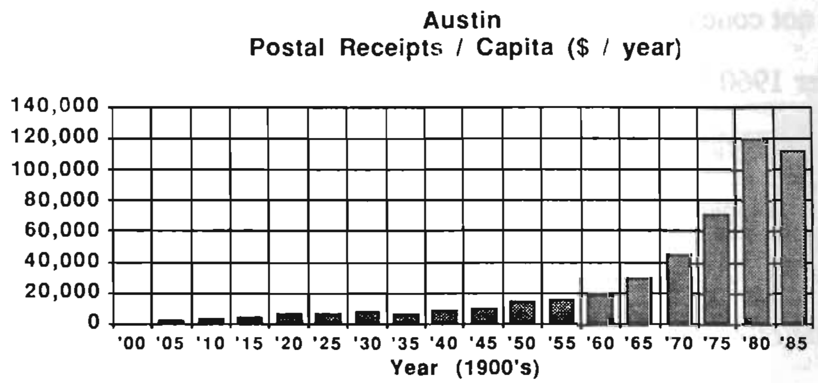


Figure 2.23

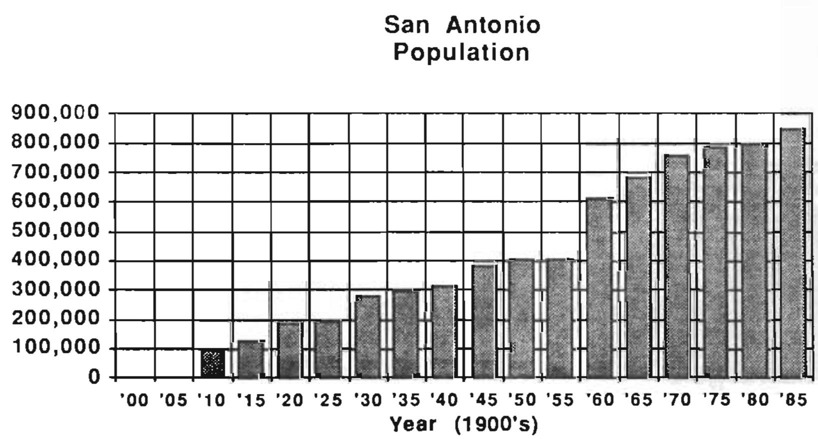


Figure 2.24

of brevity and conformity, we shall present the evolutionary track of only a few variables for the city of San Antonio, and examine how they appear when normalized by population alone.

The population curve for San Antonio (Figure 2.24) shows a distinct growth over time. The growth seems essentially linear in nature (unlike an "exponential" growth as seen for Austin - see Figure 2.1). The 50 % increase shown between the years 1955 and 1960 is a fairly high rate of growth and the effect of its presence in the charts where this variable is used for normalization is quite distinct.

Figure 2.25 is a plot of the number of vehicles in the city. The total number of registered vehicles has grown at an "exponential" rate. The number of vehicles per capita (Figure 2.26) does not conclusively establish a pattern. However, it may be said that the decline noticed after 1960 is simply due to the phenomenal growth in population noted earlier. The jump in vehicles per capita from 1975 to 1980 almost suggests an economic boom. It is interesting to speculate on the nature of the curve above a value of 1.

As in the case of Austin, one of the unresolved issues in Figure 2.27 (depicting hotels/motels) is that of size. It is suggested that World War II may be the cause of the fluctuation at the year 1940. The jump at 1960 is also observed here (see Figure 2.24). As in the case of Austin, a sharp increase in the number of motels was noted after the 1950's while the number of hotels started to decline. This effect could be attributed to a rise in the personal mobility brought about by the automobile. It would be worthwhile to plot gasoline sales revenue or airline activity versus the number of motels to see if any correlation exists among them.

The number of restaurants shows a period of stability sandwiched between two periods of growth (Figure 2.29). However, when normalized, the chart exhibits a tendency to be stable over fairly long periods (Figure 2.30). The year 1960 again provides a local anomaly. (Compare with Figure 2.14.)

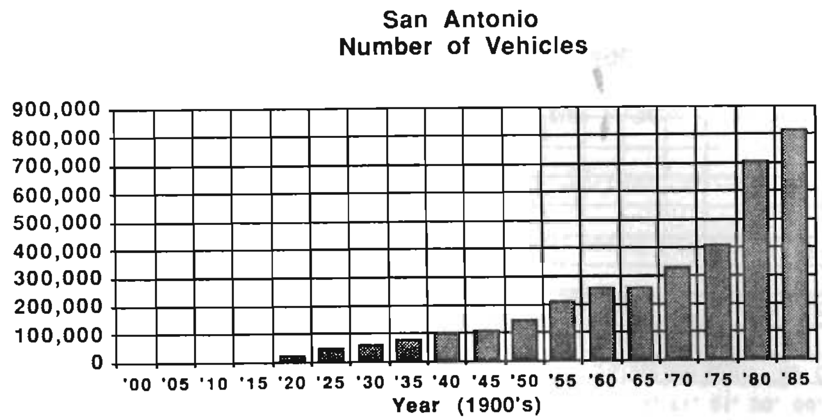


Figure 2.25

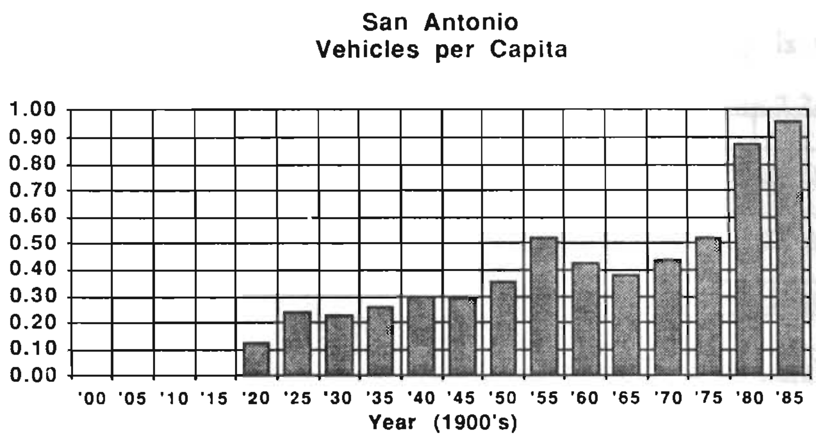


Figure 2.26

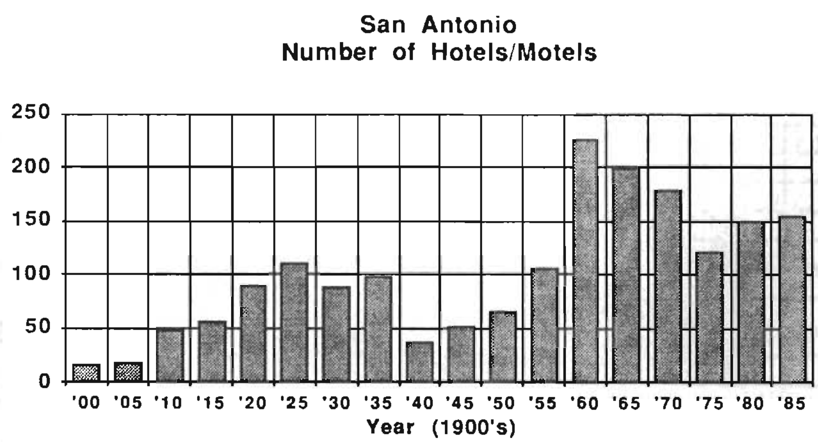


Figure 2.27

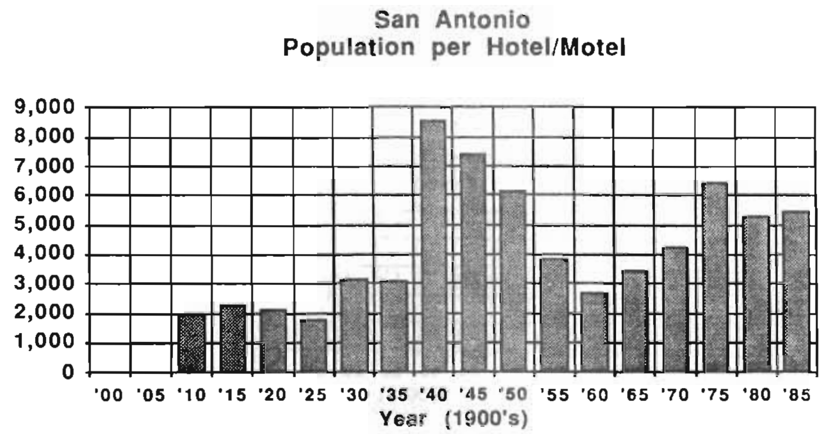


Figure 2.28

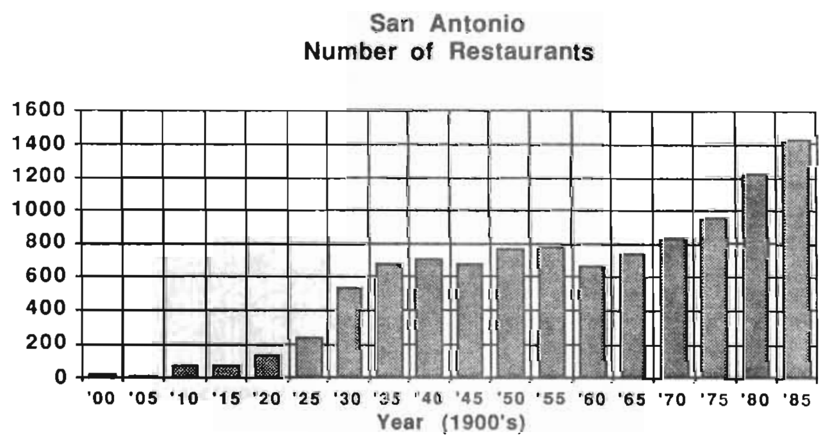


Figure 2.29

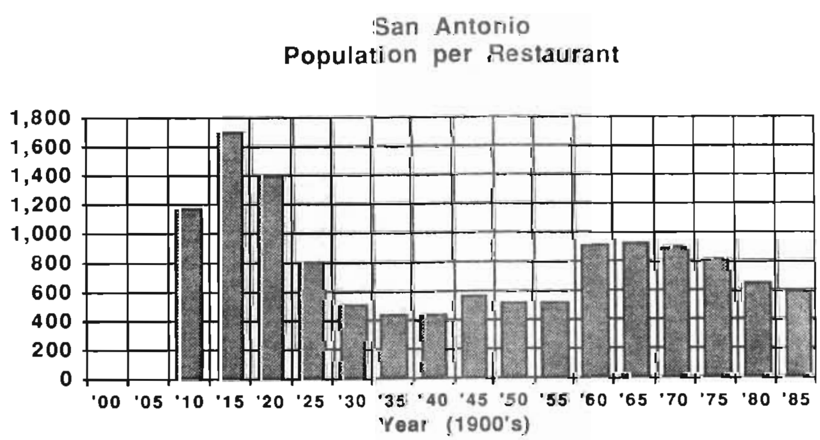


Figure 2.30

The history of auto dealerships (Figures 2.31 & 2.32) shows changes that can be reasoned in the following manner. With the advent of the automobile and its rise in popularity, the number of dealerships grew until the 1930's. The years of World War II saw a corresponding drop in their numbers followed by an increase in the post war years. A decline after 1950 can not be attributed to any one factor conclusively, though this could be resolved if the revenues from new cars sold was studied. A five fold increase in the number of auto dealerships in 1975 is most likely due to foreign car manufacturers competing successfully in a market dominated solely by domestic manufacturers until the 60's.

The number of doctors, lawyers and contractors seems to rise and fall (like a "wave") almost in unison (Figure 2.33). What is surprising is the fact that this phenomenon also evident when the variables are normalized (Figure 2.34).

For telephone numbers, an "exponential" growth is evident until 1980, both in Figure 2.35 and in Figure 2.36. This suggests that the increase in population is not the only factor responsible for the increase in the number of telephones. Historical data on the number of telephone lines for business establishments should provide more useful information.

As mentioned earlier, a function of the number of electric meters and water meters can be used to characterize the composition of the housing market (Figure 2.37). If Figure 2.38 were plotted only for residential connections, one could similarly derive data for the average occupancy of dwelling units.

Power consumption data shows an increase of usage not only in absolute values (Figure 2.39), but also when normalized by population (Figure 2.40). It would seem reasonable to expect both these data to correlate well with variables denoting the sales of electric goods. Water pumped annually has grown steadily over the years as can be seen in Figure 2.41. The commissioning of new sources of water to be pumped is suggested in or before the year 1980. When normalized by population in Figure 2.42, it holds a steady

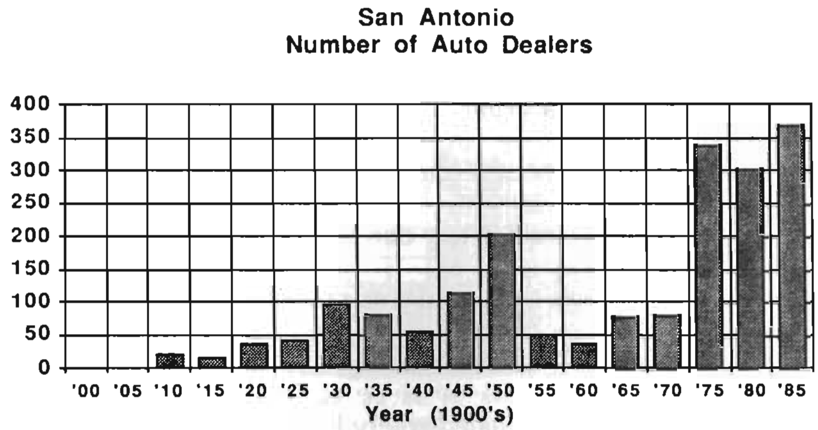


Figure 2.31

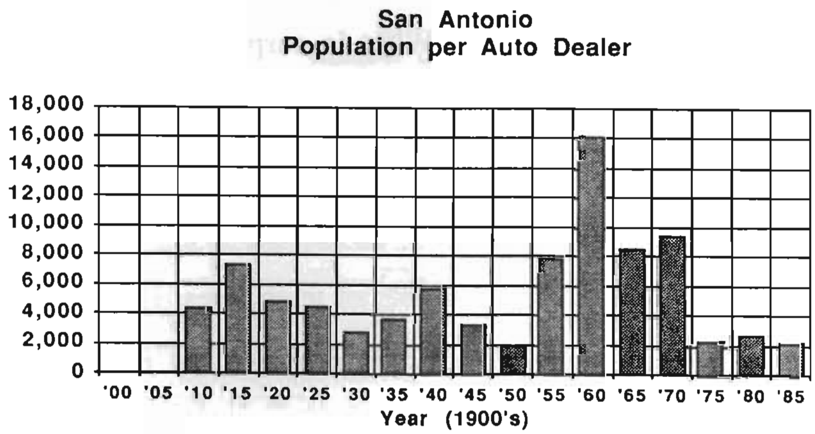


Figure 2.32

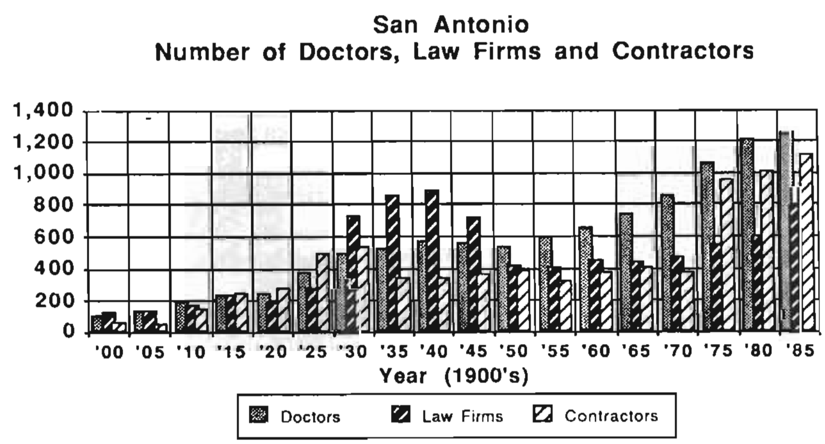


Figure 2.33

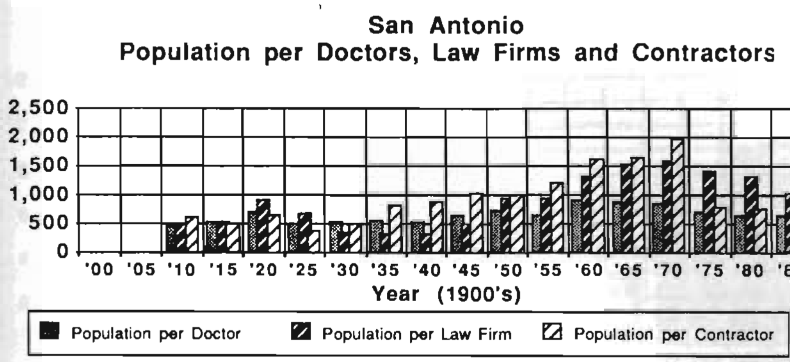


Figure 2.34

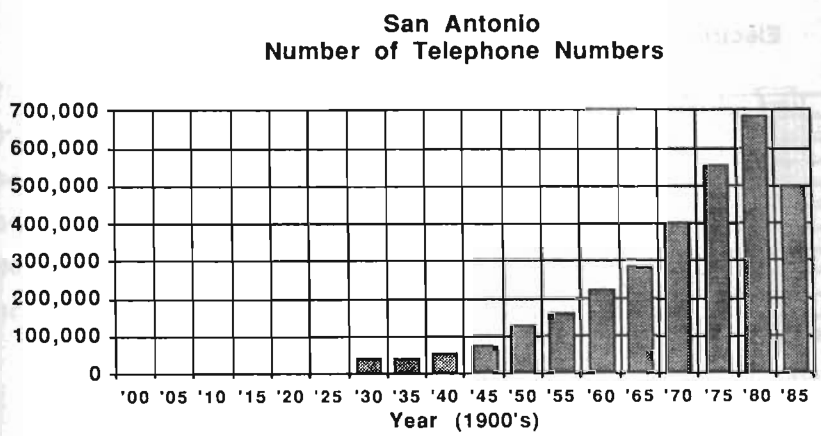


Figure 2.35

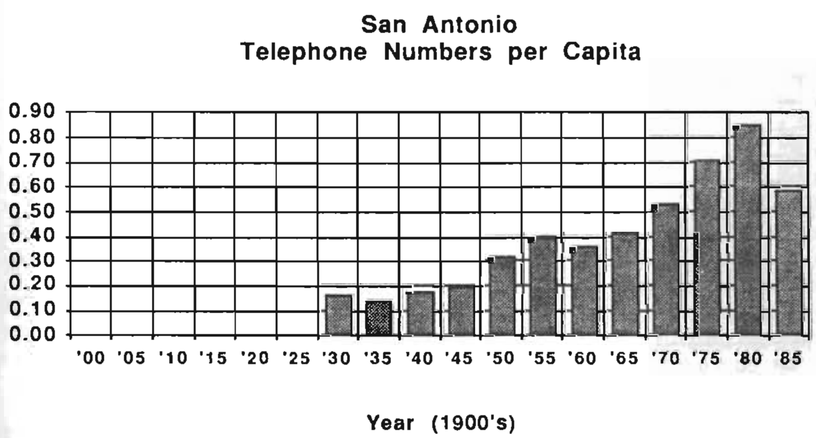


Figure 2.36

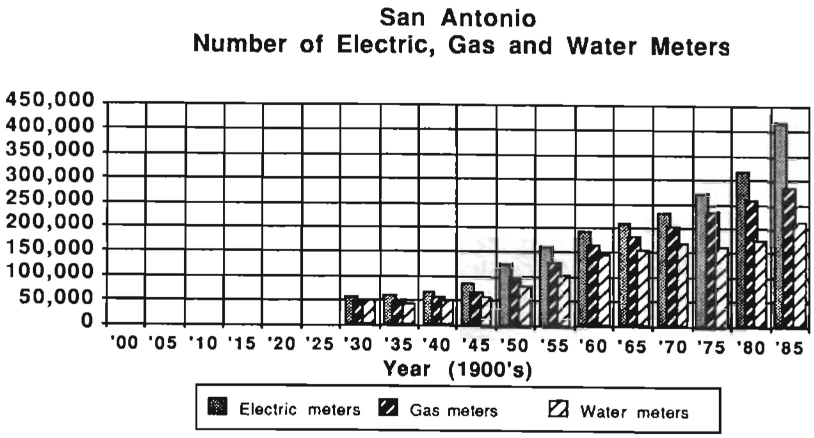


Figure 2.37

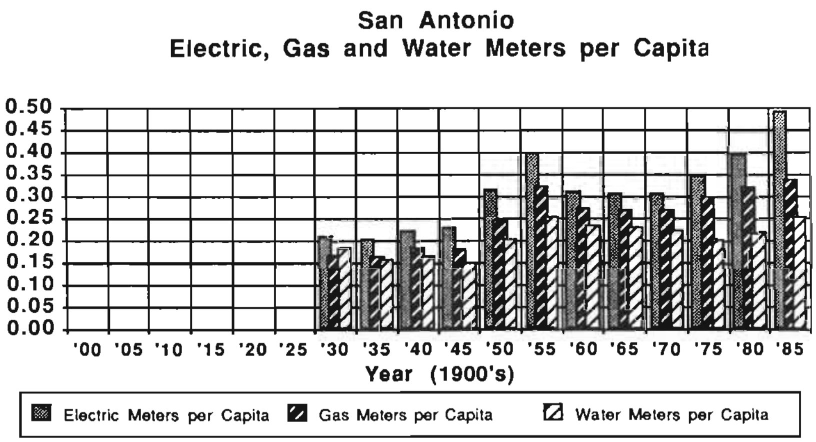


Figure 2.38

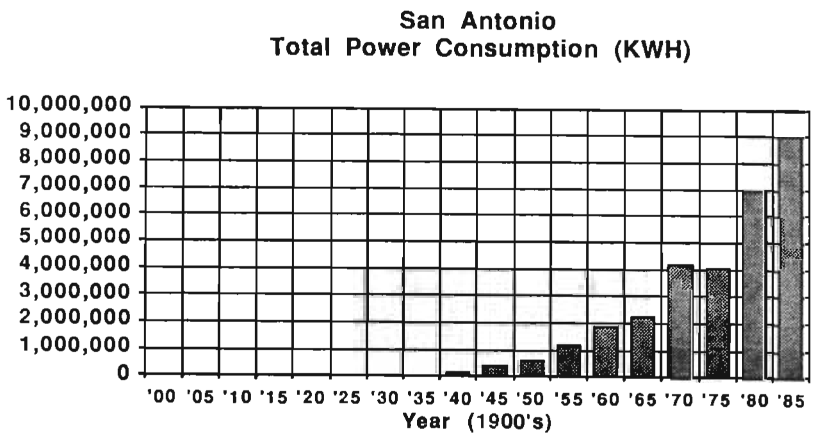


Figure 2.39

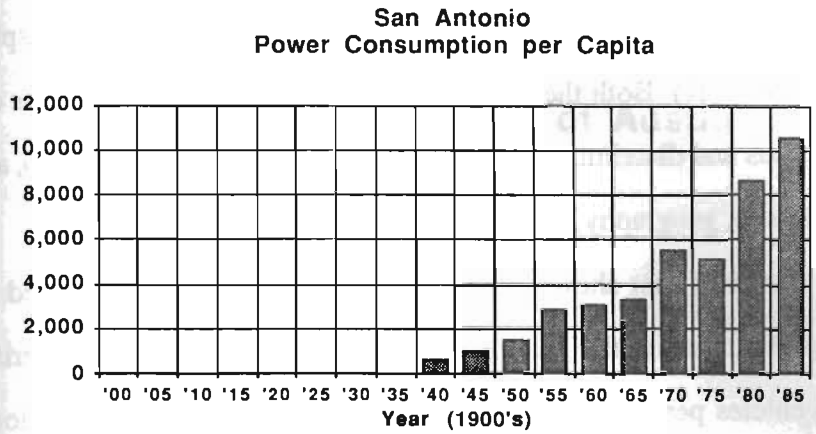


Figure 2.40

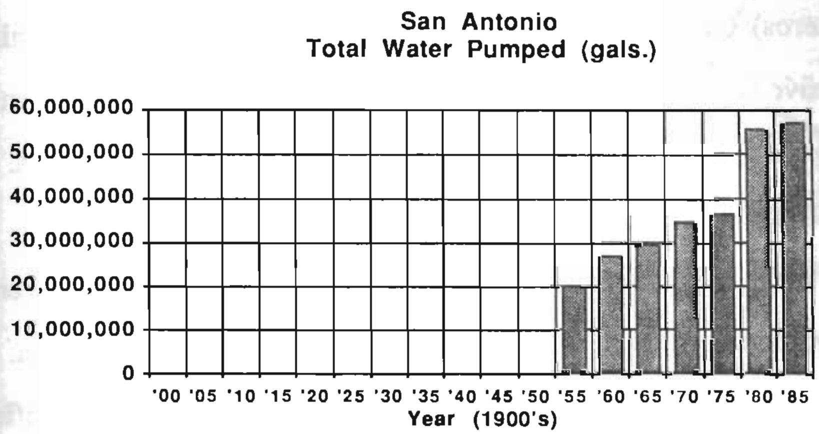


Figure 2.41

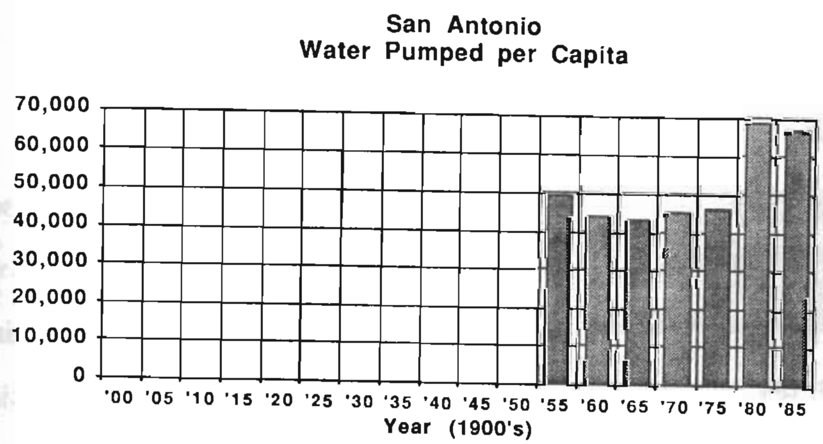


Figure 2.42

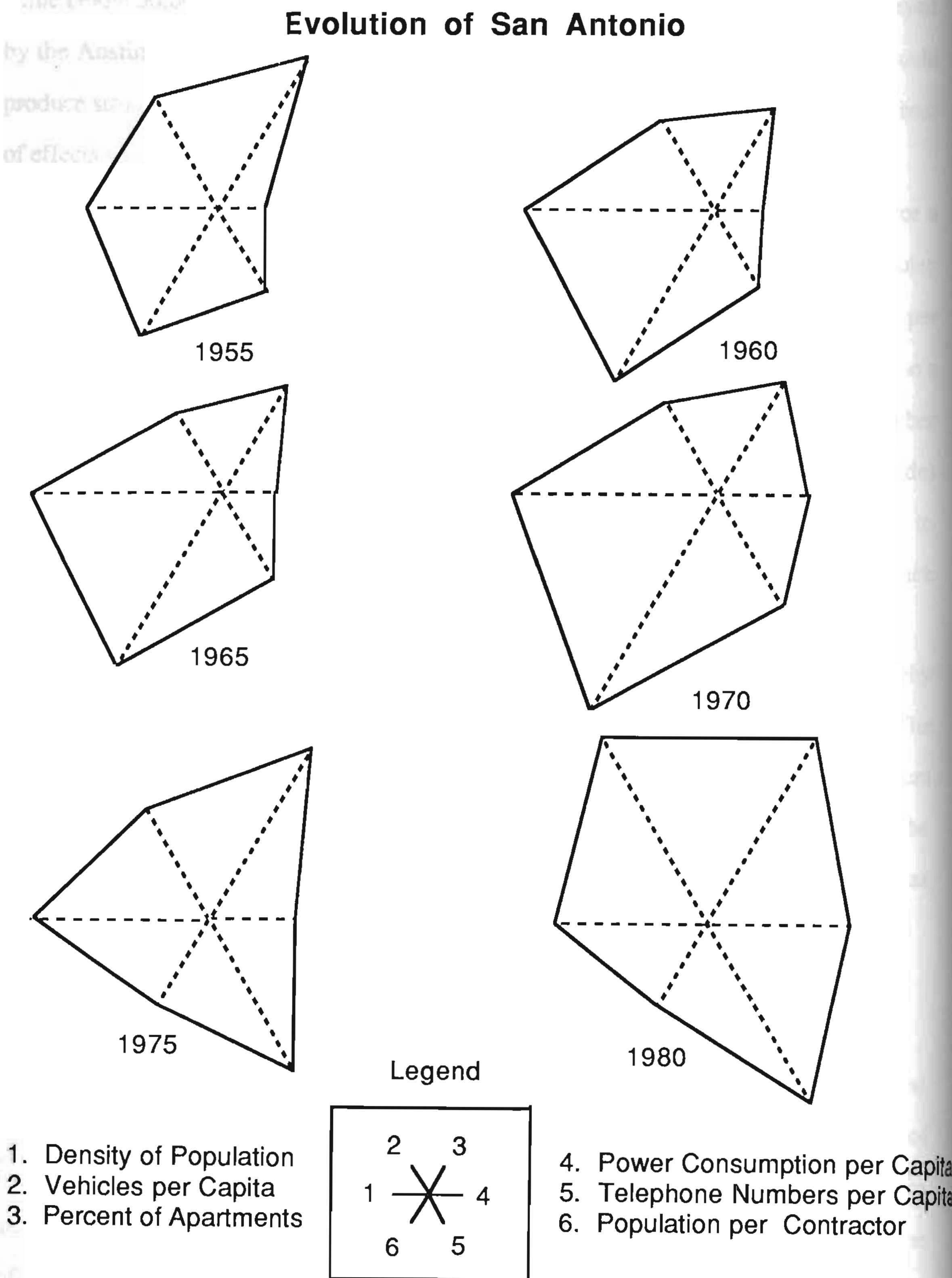

Evolution of San Antonio
1955
1960
1965
1970
1975
1980
Legend
1. Density of Population
2. Vehicles per Capita
3. Percent of Apartments
2 3
1 4
6 5
4. Power Consumption per Capita
5. Telephone Numbers per Capita
6. Population per Contractor

Figure 2.44

area bounded by the diagram. The aim, however, is not just to maximize information, but also to present it in such a manner that it can be easily assimilated by the reader. Situations where a large number variables are plotted on one diagram perhaps should be avoided. It is felt that four to six axes may be optimal for such representations in this study

The diagrams discussed in this section are scaled by a procedure similar to the one discussed in the previous section the values of the variable (for each of the eleven cities) on any one axis is divided by a round number marginally greater than the highest value among all cities attained by that variable for all the years under consideration. This insures that comparisons can be made across time for the cities as well as between cities (across space). Three sets of four-axis snow-flake diagrams are presented for discussion here.

Figure 2.45 shows a map of the U.S. with the locations of the cities being studied. Atlanta, Boston, Chicago, Cincinnati, Los Angeles, Miami, New York and Seattle - all are cities that fall in subdivisions created by Yeates (1976) for categorizing major urban areas on the continent. Denver, Phoenix and St. Louis are the only cities being studied here that are not located geographically in any of his regions. These three are included as being representative cities outside of the major urban areas.

The first set of diagrams for the evolution of the eleven cities is presented in Figures 2.46 through 2.50. The variables plotted on the four axes are - area, density of population, city budget per capita, and the city's highway department budget as a fraction of the total city budget. In addition to the information conveyed by the length of any one axis, it can be established that the space bounded by the axes area and density of population is proportional to the population of the city. Similarly, the space bounded by the axes density of population and city budget per capita would represent the city's budget per unit area and the space between the axes city budget per capita and the city's highway department budget as a fraction of the total city budget would indicate the amount spent by the highway department on a per capita basis.

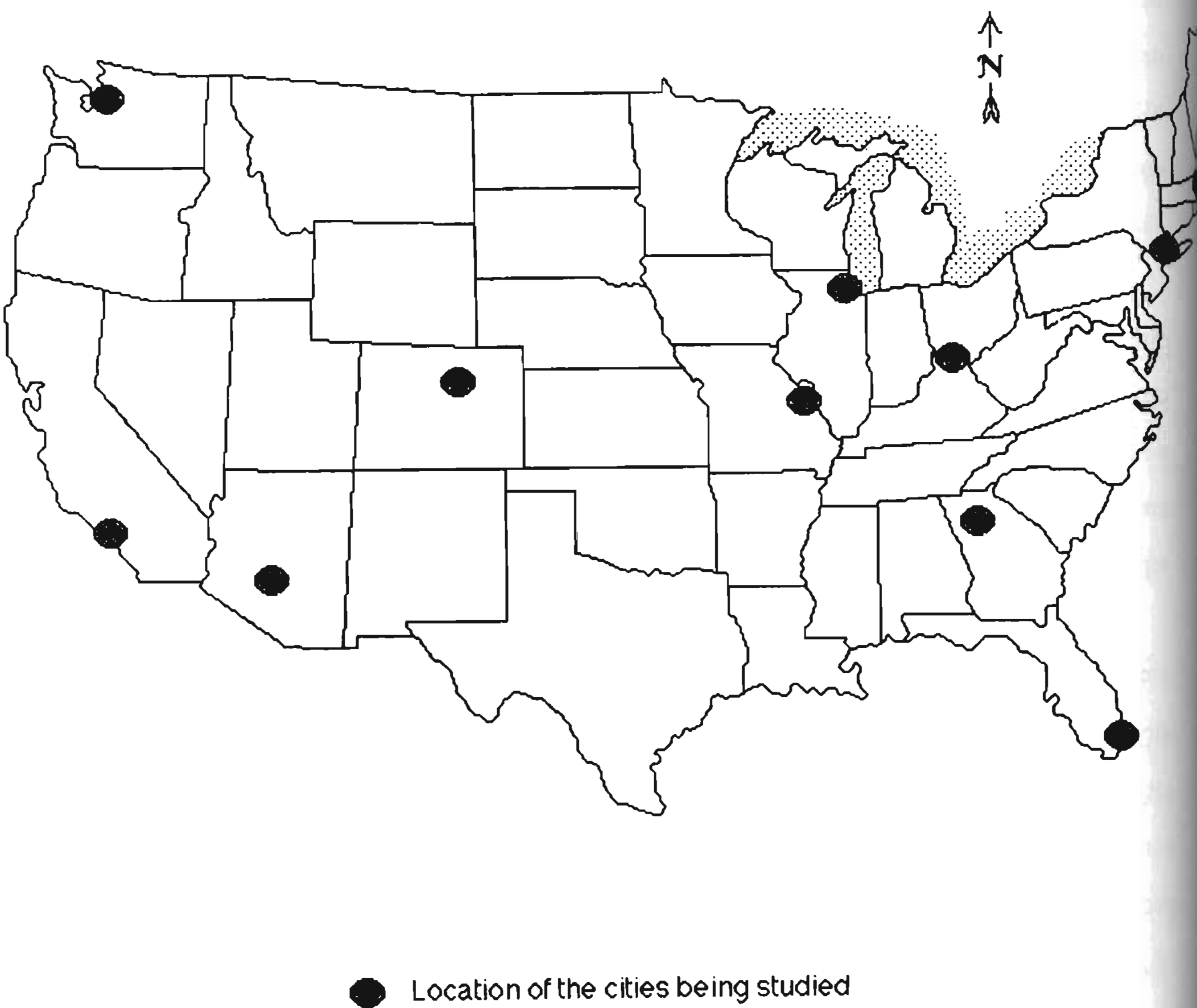


Figure 2.45

Figure 2.46 shows the situation existing in the year 1950. Cincinnati and Seattle are similar in all respects in this particular representation while Boston and St. Louis show similarities but for the fact that Boston's city budget per capita is almost three times that of St. Louis. New York, Chicago and Los Angeles stand out as being totally different from the rest of the cities. Among the three of them, while Los Angeles has the largest area, it has the smallest population. In fact, this is so disproportionate, that Los Angeles can be seen to have the lowest density of population among all the cities being studied. This is precisely what would be expected of a city of the "auto-air...epoch". This is brought home even more forcefully when Los Angeles is compared with older cities, like New York and Boston, both of which were thriving ports in the "sail-wagon epoch". Comparison with St. Louis and Chicago (both products of the "iron-horse epoch") also reinforces the differences. Even with no prior knowledge of the history and geography of these cities, the reader would still be able to deduce from the snow-flakes that the population was more dependant on the automobile for their mobility in Los Angeles than, say, in New York or Boston.

Phoenix, which we know to be another "auto-air....epoch" city, has undergone a drastic change in ten year from 1950 to 1960 (Figure 2.47). The area increases almost ten fold during this period, and the density of population in 1960 drops almost to a third of the 1950 value. Both the city budget and the fraction appropriated for the highway department increase. At this stage, Phoenix looks like a smaller version of Los Angeles. Miami has lowered the fraction spent on the highway department. However, an increase in its city budget per capita, with more or less the 1950 levels of area and population, now has this city spending a lot more on its highways per capita. New York, Chicago St. Louis and Seattle have increased not only their city budget per capita, but have also increased the fraction spent on the highway department. Atlanta too, has increased this fraction in a big way. At the same time, it also seems to have almost quadrupled its area without a corresponding rise in population. Boston, Cincinnati and Denver appear to have retained their character since 1950, only increasing their city budgets.

The 1970 diagrams are shown in Figure 2.48. New York, Chicago, Atlanta and St. Louis have actually decreased the fraction of the city's budget appropriated for the highway department. However, an increase in their city budget per capita (while the population and area remain a constant) offsets this decrease in terms of highway department dollars spent per capita. Boston and Cincinnati have increased both the city budget per capita and the fraction thereof spent on the highway department. Los Angeles and Phoenix seem to be travelling on the same evolutionary path, the current difference between these two being one of scale. Miami, Seattle and Denver have managed to retain the basic shape of their snow-flake diagram from 1960, while showing increases in their city budget per capita.

Figure 2.49, which is for the year 1975, has been presented here to show the state of the cities at the mid-point of the decade 1970 - 1980 in which all cities show a remarkable amount of change in terms of this snow-flake representation. All eleven cities show large increases in their city budgets per capita, while the other variables are more or less at the 1970 level. While some cities (e.g. Los Angeles, Seattle and Cincinnati) do show a decrease in their highway department budget as a fraction of the city budget, the highway department budget per unit area remains the same.

Before an an actual discussion of Figure 2.50 is presented, it wold be worthwhile to go back to Figure 2.46 and see what the cities 'looked like' at the beginning of the study period and what they 'look like' in Figure 2.50 - a time span of thirty years. Now consider the changes that are evident from 1975 to 1980. All cities except for Los Angeles, New York and Cincinnati, show significant increases in their city budget per capita. Atlanta, Los Angeles and Chicago show a decline in the highway department budget as a fraction of the city budget. Phoenix and Seattle are the only cities that show a considerable growth in area. Similarities between Los Angeles and Phoenix are no longer seen. In fact, Denver and

Cincinnati are the only two cities that have consistently seemed to have a basic similarity of shape over the study period.

The next set of variables discussed are number of manufacturing units, number of wholesale establishments, number of retail establishments and revenues from retail sales. These are presented in Figures 2.51 through 2.54 for the years 1948, 1958, 1967 and 1977 respectively. Due to the disproportionate sizes of the diagrams of some cities, the cities of Chicago, New York and Los Angeles have been reduced by a factor of 4 of their actual size. This means, for example, that when Chicago and St. Louis appear to be of the same size and shape in Figure 2.51, it actually implies that the values of the variables for Chicago are four times larger than those for St. Louis.

The change in these variables over time could indicate how the cities have evolved in terms of their manufacturing and trade characteristic. When taken in conjunction with the previous set of snow-flake diagrams (especially the variables that deal with area and population), trends could be established about the movement towards (or away from) large centralized establishments replacing smaller establishments. This could indicate an increase (or decrease) in the dependance on the automobile by the residents of the city.

The state of the cities in 1948 is shown in Figure 2.51. New York dominates in the size of the snow-flake, having more manufacturing units, more wholesale establishments and more retail establishments than any of other cities. In Figure 2.46, we have seen that the area of New York is actually smaller than that of Los Angeles. This implies that New York has a very high density of manufacturing units, wholesale establishments and retail establishments per unit area when compared with Los Angeles. St. Louis and Chicago are similar to each other in form (though in scale, Chicago is four times the size of St. Louis) in Figure 2.51. (Such a similarity, with difference in scale, is also evident in Figure 2.46). Similarities also exist between Denver and Seattle and between Miami and Phoenix in both the figures.

In the next ten years (ending 1958), the cities that show significant growth for these variables are Atlanta, Los Angeles, Miami and Phoenix (Figure 2.52). Further, with the exception of Los Angeles, they also show a corresponding growth in area and population (Figure 2.46 & 2.47). These cities can be identified as those of the "auto-air...epoch". St. Louis shows a decline in the number of manufacturing units. All cities show increases in retail revenues. Boston, Seattle, Cincinnati and Chicago remain stable for this decade. New York shows fewer manufacturing, wholesale and retail establishments with an increase in retail revenues. This may be representative of a trend moving away from small local units towards larger centralized establishments.

This trend is evident for New York in Figure 2.53 as well. By 1967, there is a marked decline in manufacturing, wholesale and retail establishments for Chicago, Boston and New York, while there is a definite increase in the retail revenues of these cities. Cincinnati and St. Louis too, **have** fewer retail stores in 1967 than in 1958. St. Louis, however, has an increased number of manufacturing units. Phoenix and Atlanta are the only cities that show increases in all four variables over this period.

Looking at Figure 2.54, one is stuck by the apparent similarity in the form of the snow-flakes for the various cities. What is it that causes these eleven diverse cities (that did not 'look alike' in 1948) to be so 'similar' to one another in 1977 ? There seem to be correlations that exist between the four variables (across the eleven cities) that gives each city a similar (kite-like) shape. Looking back at the snow-flakes for 1948 (Figure 2.51), we can see that Boston, Cincinnati, Chicago and St. Louis have 'shrunk' in terms of the number of manufacturing, wholesale and retail establishments. Over the same period (1950 to 1980), these cities do not show appreciable changes either in area or in population (Figures 2.46 & 2.50). Miami, and Phoenix show the most growth from 1948 to 1977 (Figures 2.51 & 2.54). Atlanta and Denver have evolved form a fairly similar state in 1948 to a point where they still resemble each other in 1977.

Figures 2.55 through 2.59 depict a new set of four variables number of hospital beds, revenues from new auto sales, revenues from gas stations, and revenues from restaurant and bar sales. While the grouping of these variables may seem unrelated at first glance, there is a common thread of information inherent in these four variables. One would expect that the slope of the line joining the auto sales and gas station sales would be a constant for all cities for a given year (if there are correlations between these two variable). If the slopes differ, it could indicate different auto usage patterns. The restaurant/bar sales and the number of hospital beds are indicators of differences between cities. When studied in conjunction with other variables, they can provide insights about the level of services available, which might lead to certain conclusions about the mobility of the population of these cities. As with the previous set of snow-flake diagrams, Chicago, New York and Los Angeles have been scaled down by a factor of 4 due to disproportionate sizes of the diagrams for the various cities.

Boston, Cincinnati and St. Louis show similarities both in the size and the shape of these diagrams in Figure 2.55. Seattle and Denver are also similar to each other for the year 1954. When looking at the other variables in the same general time span (1950 to 1960), we can see that Seattle and Denver have the same area and population (Figures 2.46 & 2.47). Boston and St. Louis also resembled each other in terms of area and population during the 1950's. The suggestion is almost as though the cities that developed during one era, did so with a remarkable similarity to each other in their ground plans.

Over the next four years, little seems to have changed for individual cities (Figure 2.56 for the year 1958). This may be due to the fact that only four years separate the two sets of data. Also, the variables under study here are expected to be fairly constant; after all it would be unreasonable to expect the number of hospital beds or the restaurant/bar sales to change drastically over such a short period of time. Despite these arguments, Atlanta, Phoenix, Los Angeles and Miami show distinct increases in each of these four variables.

As noted earlier, these are the cities of the "auto-air...epoch" and their rapid growth is hardly surprising.

Comparing the cities in 1963 (Figure 2.57) with their corresponding state in 1954 (Figure 2.55), we see that Boston, Chicago, Cincinnati, New York and St. Louis are more or less at their original (1954) levels. Atlanta, Los Angeles and Phoenix have all grown at a very rapid rate, while Denver, Miami and Seattle have grown at a slower pace. Figure 2.58 presents the snow-flakes for the year 1967. While changes are apparent in a few cities, no discussions will be presented here because of the small time span between this and the previous set of diagrams.

Looking at the changes that came about over twenty years is very revealing (Figures 2.59 & 2.55). Atlanta, Miami and Phoenix, the cities with the three smallest snow-flakes in 1954, are the three with the largest snow-flakes (excluding the scaled down heavy weights Chicago, Los Angeles and New York) in 1972. In terms of auto usage, we can see that even though the population of Los Angeles is about one third of the population of New York in 1970 (Figure 2.48), the revenues from auto sales and gas sales in these two cities are about equal in 1972. Among the cities drawn to its scale, Atlanta has the highest level of auto and gas station sales in 1972; in terms of population it is the third lowest ahead only of Miami and Cincinnati. An exact opposite is the case of Boston which has the lowest level of auto and gas station sales, yet has the largest population (excluding Chicago, Los Angeles and New York, the cities not in its scale).

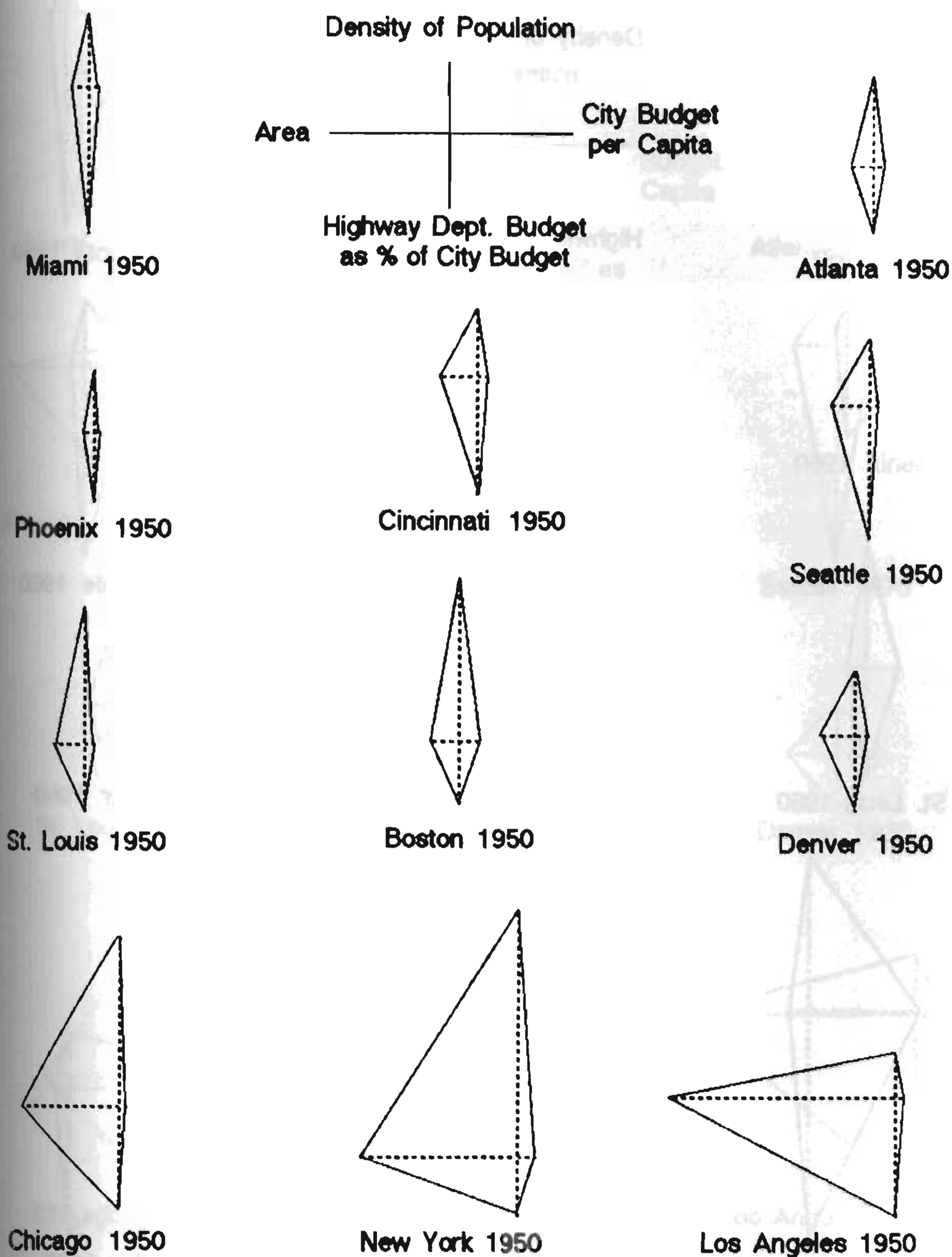

Density of Population
Area
City Budget per Capita
Highway Dept. Budget as % of City Budget
Miami 1950
Atlanta 1950
Phoenix 1950
Cincinnati 1950
Seattle 1950
St. Louis 1950
Boston 1950
Denver 1950
Chicago 1950
New York 1950
Los Angeles 1950

Figure 2.46

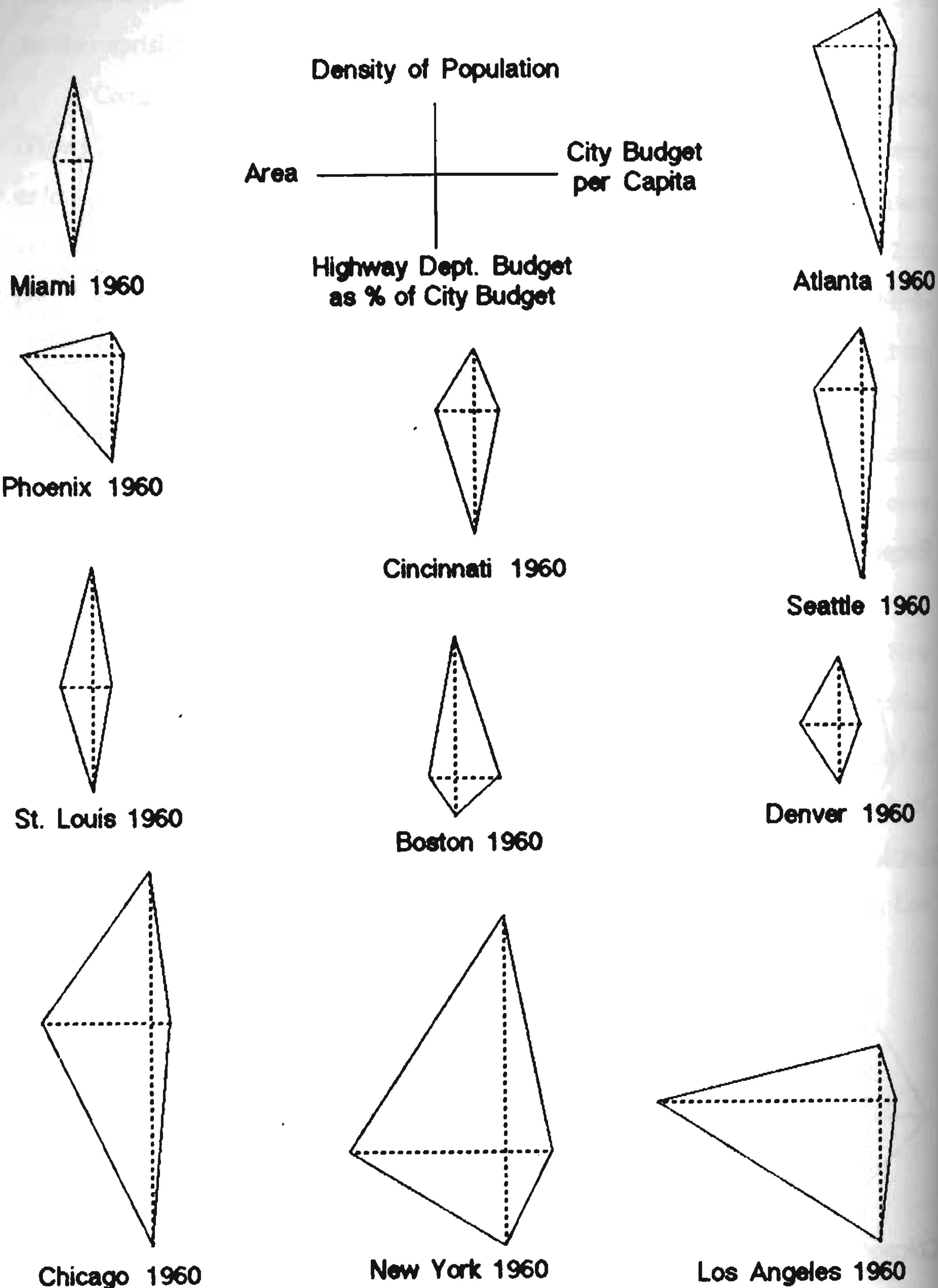

Density of Population
Area
City Budget per Capita
Highway Dept. Budget as % of City Budget
Miami 1960
Atlanta 1960
Phoenix 1960
Cincinnati 1960
Seattle 1960
St. Louis 1960
Boston 1960
Denver 1960
Chicago 1960
New York 1960
Los Angeles 1960

Figure 2.47

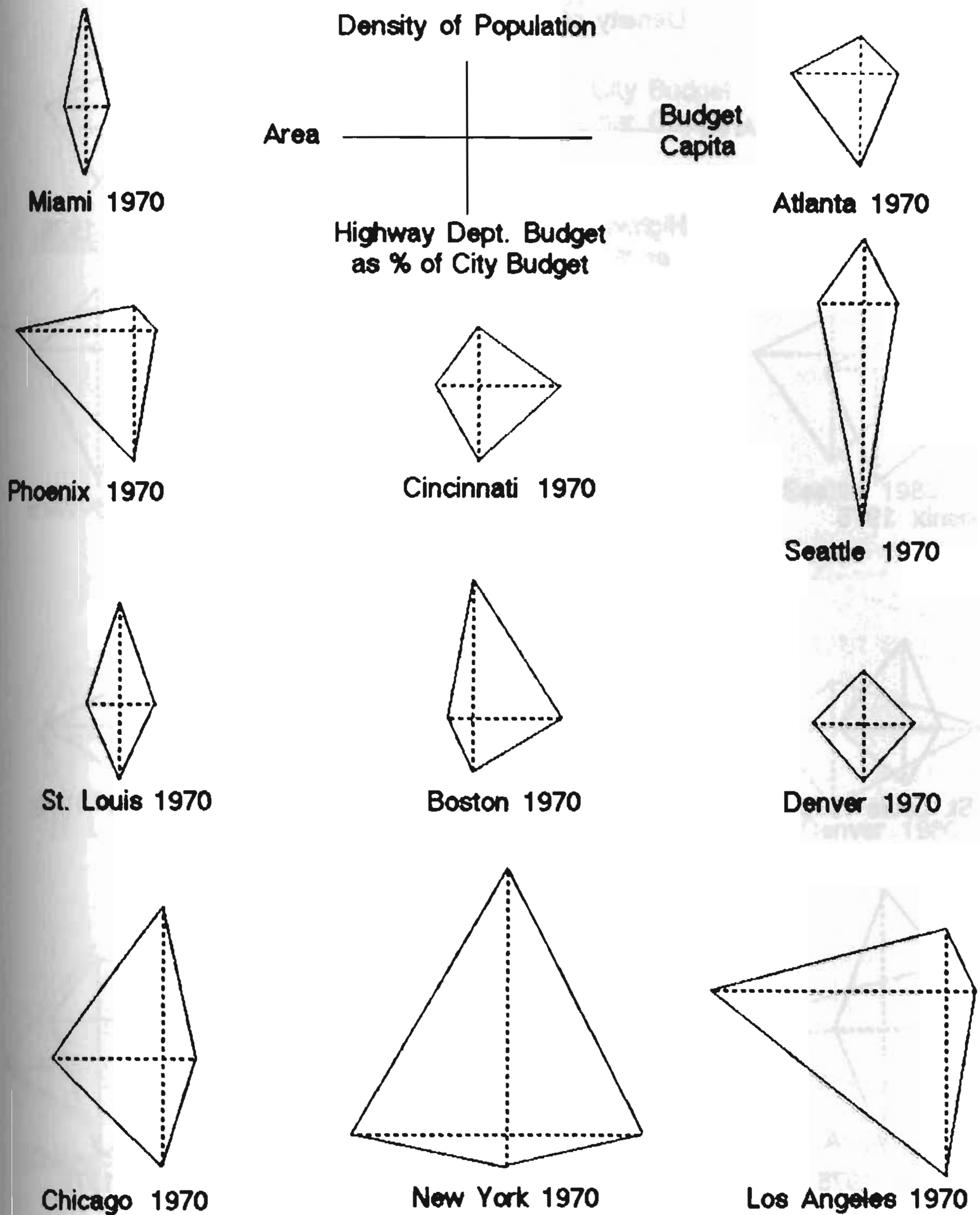

Density of Population
Area
Budget
Capita
Highway Dept. Budget
as % of City Budget
Miami 1970
Atlanta 1970
Phoenix 1970
Cincinnati 1970
Seattle 1970
St. Louis 1970
Boston 1970
Denver 1970
Chicago 1970
New York 1970
Los Angeles 1970

Figure 2.48

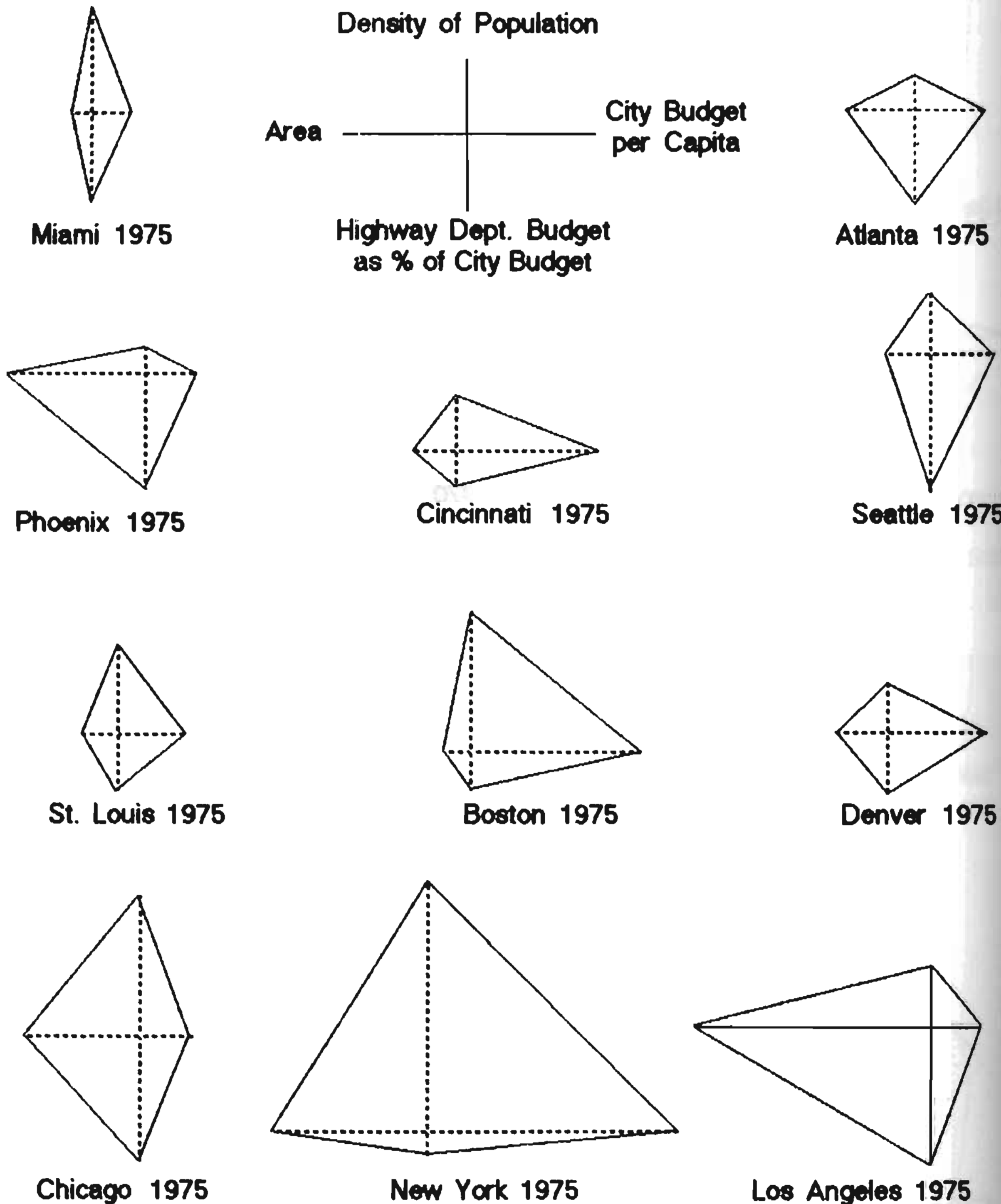

Density of Population
Area
City Budget per Capita
Highway Dept. Budget as % of City Budget
Miami 1975
Atlanta 1975
Phoenix 1975
Cincinnati 1975
Seattle 1975
St. Louis 1975
Boston 1975
Denver 1975
Chicago 1975
New York 1975
Los Angeles 1975

Figure 2.49

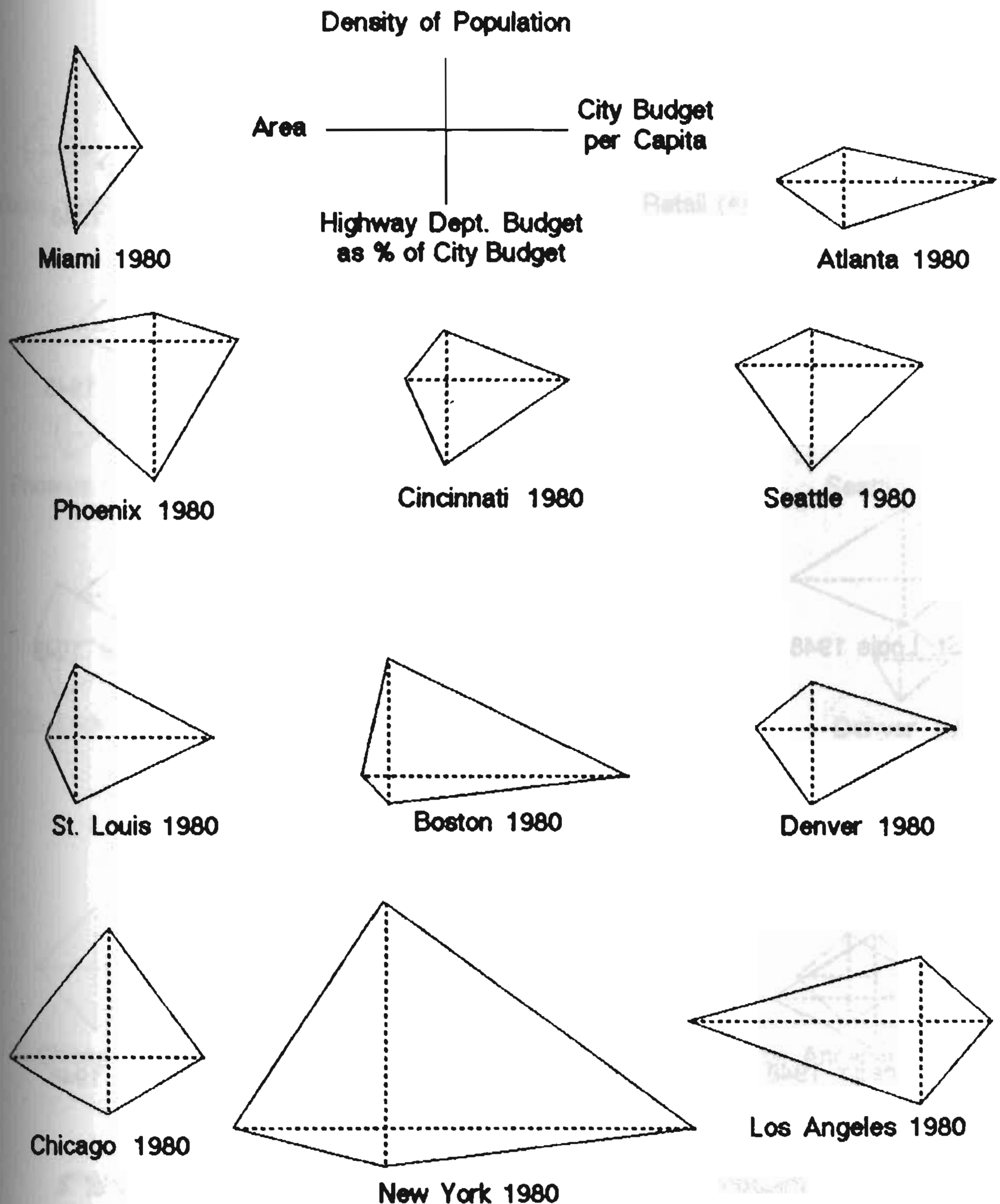
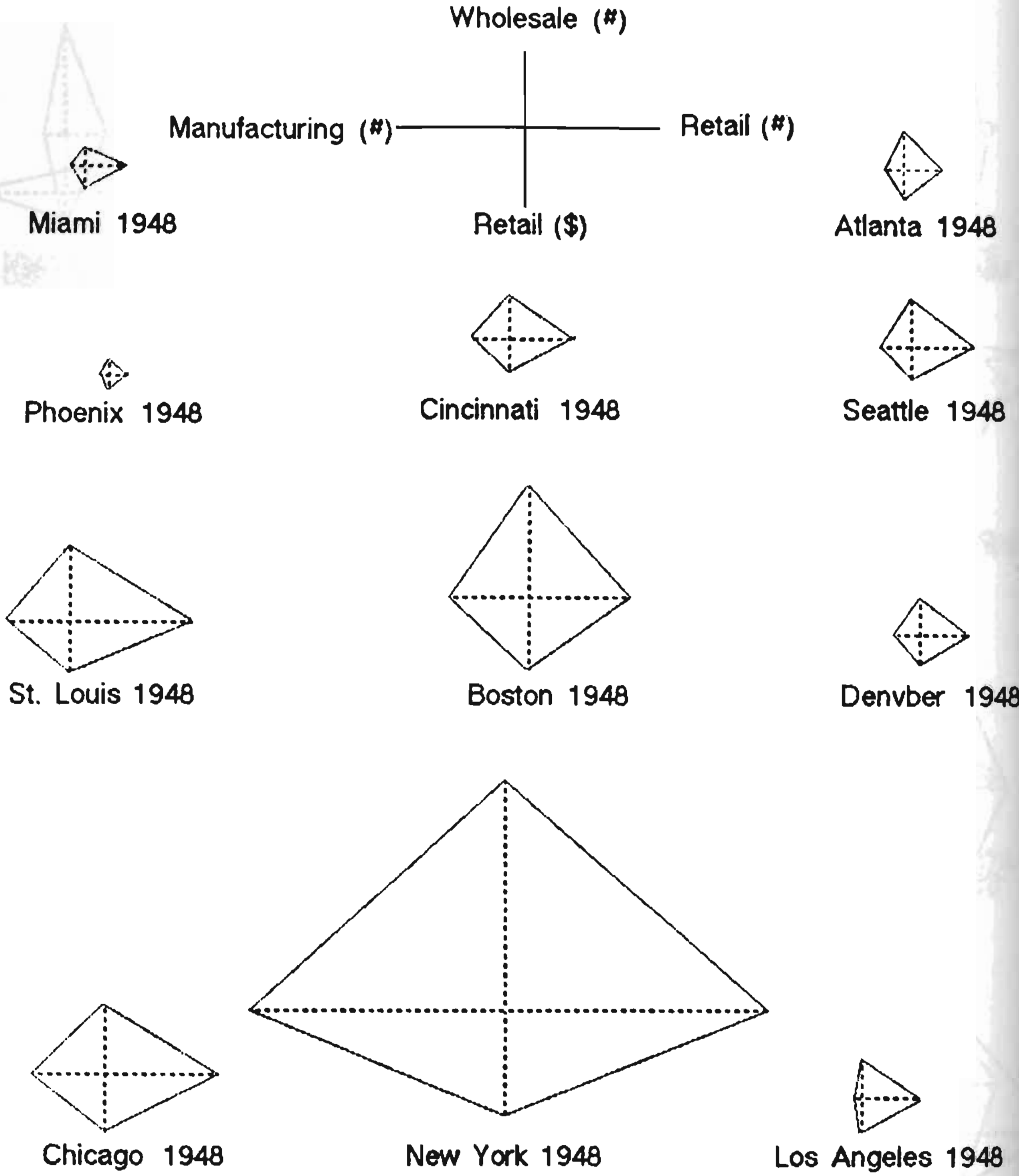
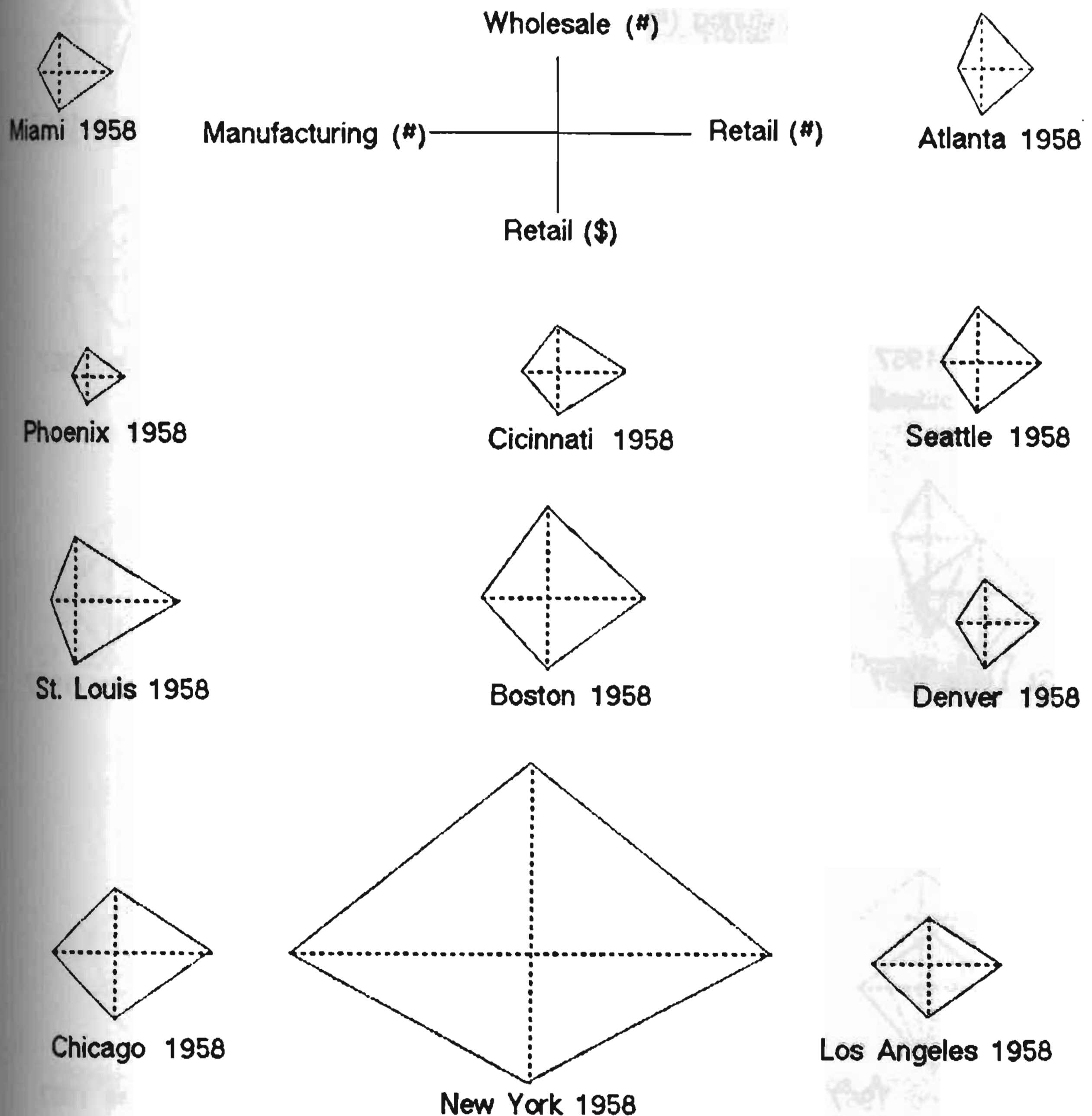

Density of Population
Area
City Budget
per Capita
Highway Dept. Budget
as % of City Budget
Miami 1980
Atlanta 1980
Phoenix 1980
Cincinnati 1980
Seattle 1980
St. Louis 1980
Boston 1980
Denver 1980
Chicago 1980
New York 1980
Los Angeles 1980

Figure 2.50



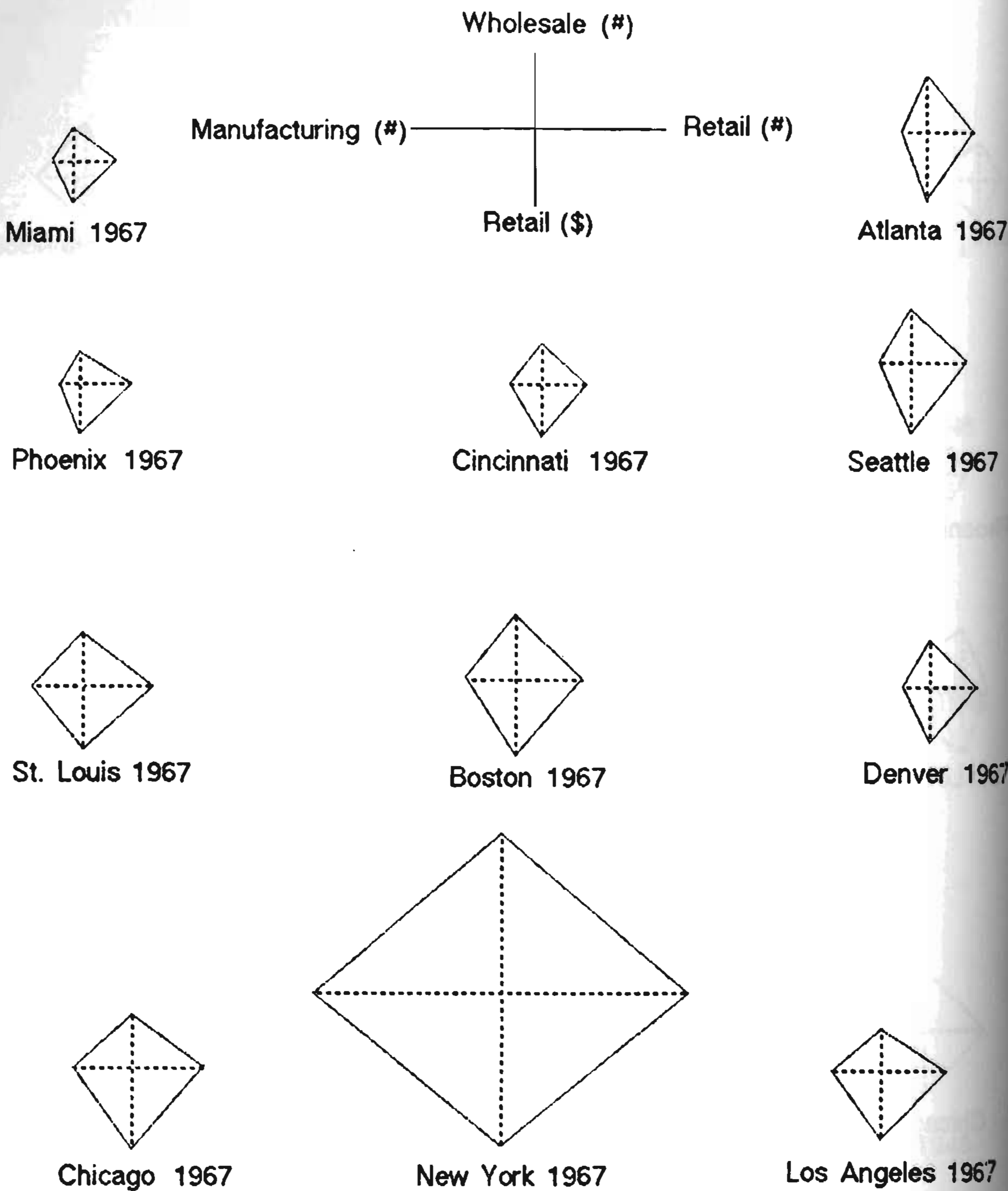
** Note : Chicago, New York and Los Angeles are reduced by a factor of 4.

Figure 2.51



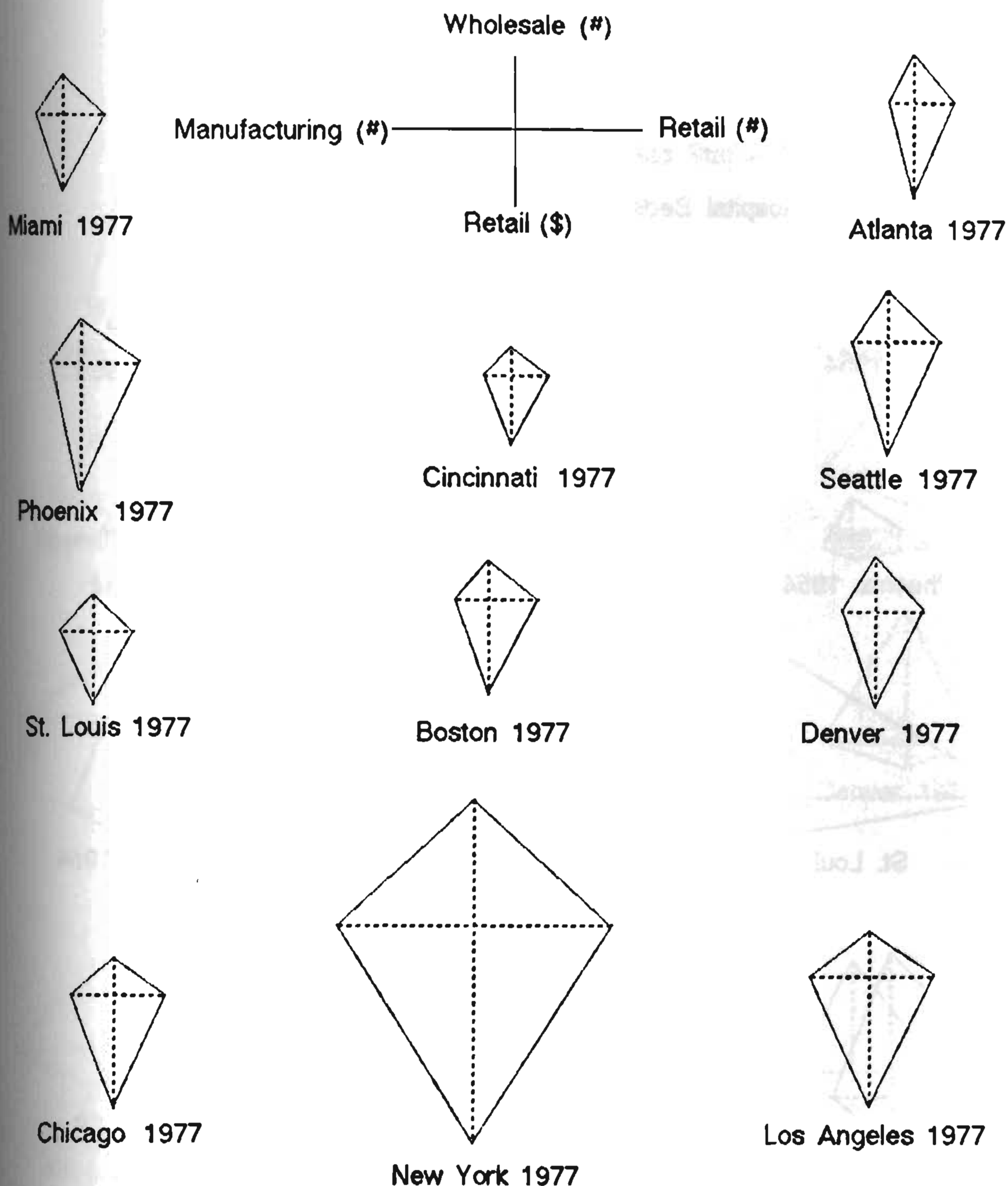
** Note : Chicago, New York and Los Angeles are reduced by a factor of 4.

Figure 2.52



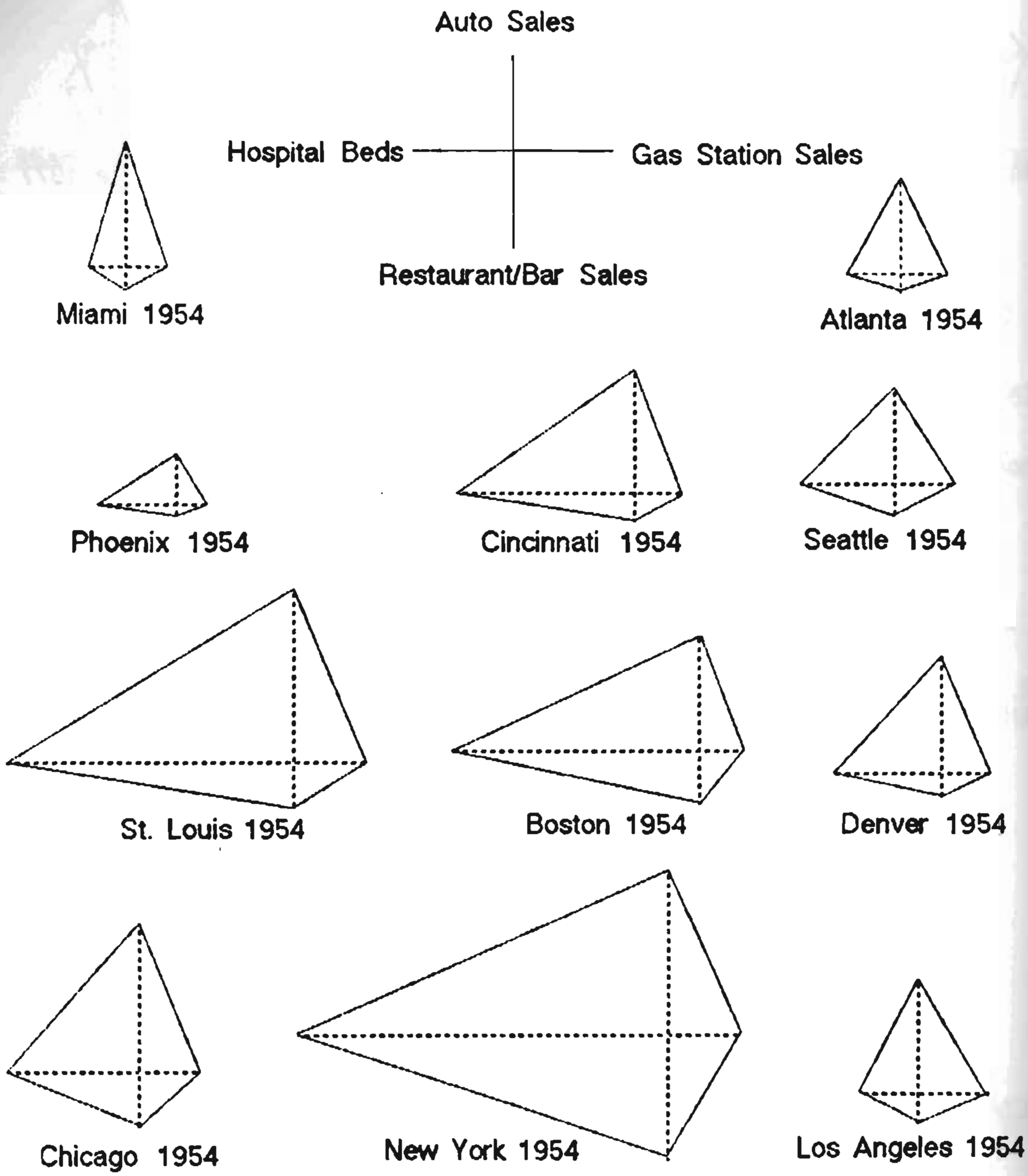
** Note : Chicago, New York and Los Angeles are reduced by a factor of 4.

Figure 2.53



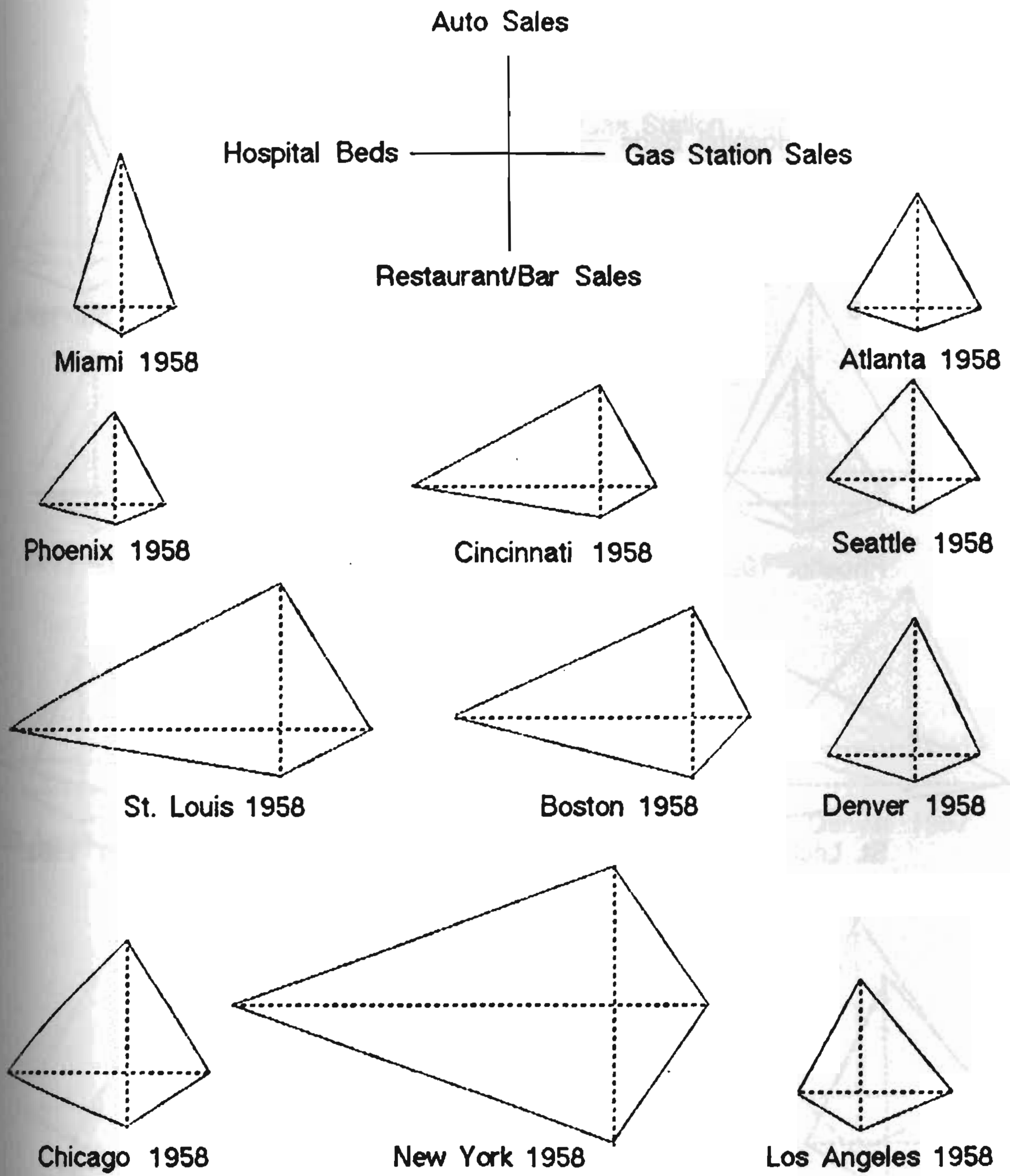
** Note : Chicago, New York and Los Angeles are reduced by a factor of 4.

Figure 2.54



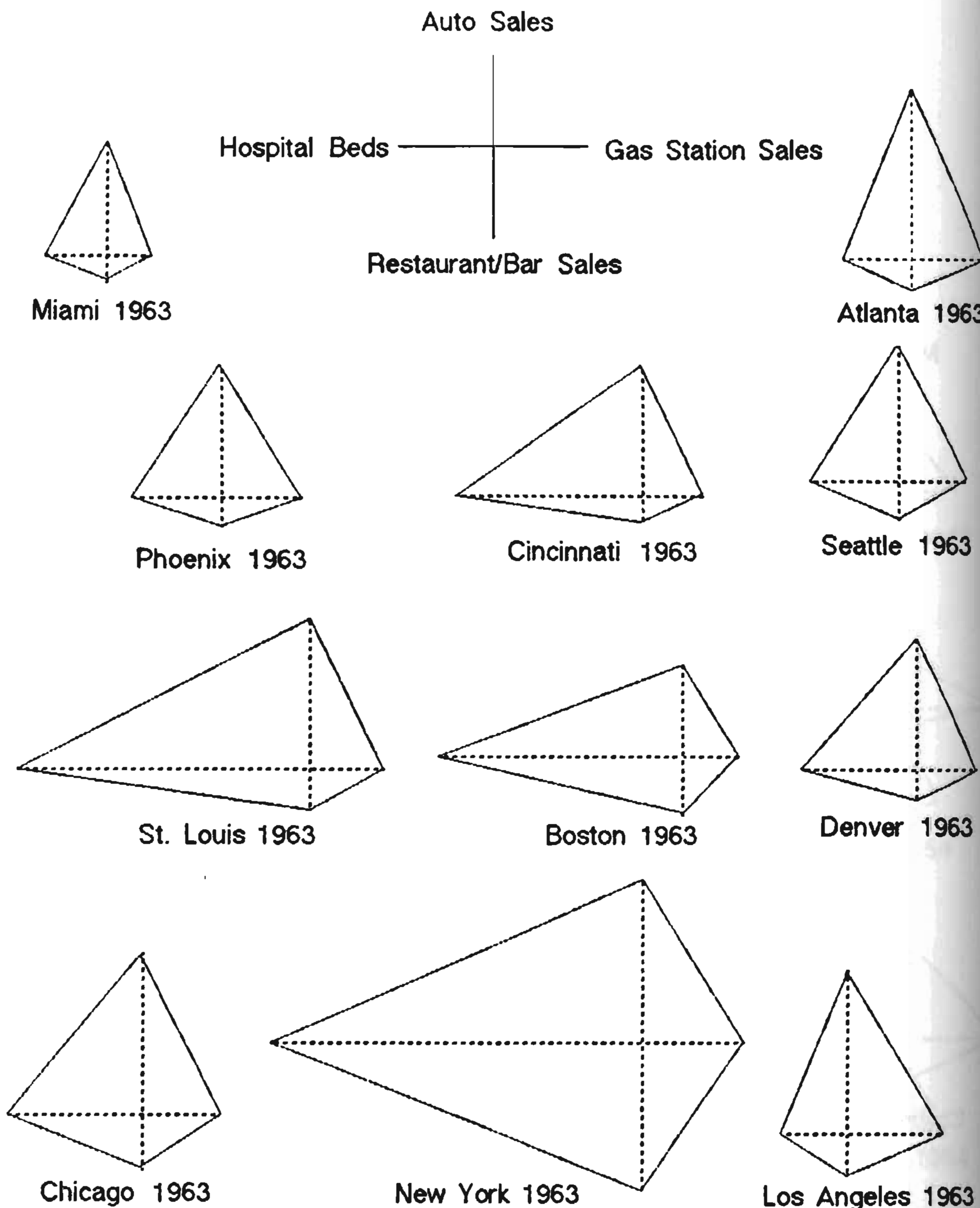
** Note : Chicago, New York and Los Angeles are reduced by a factor of 4.

Figure 2.55



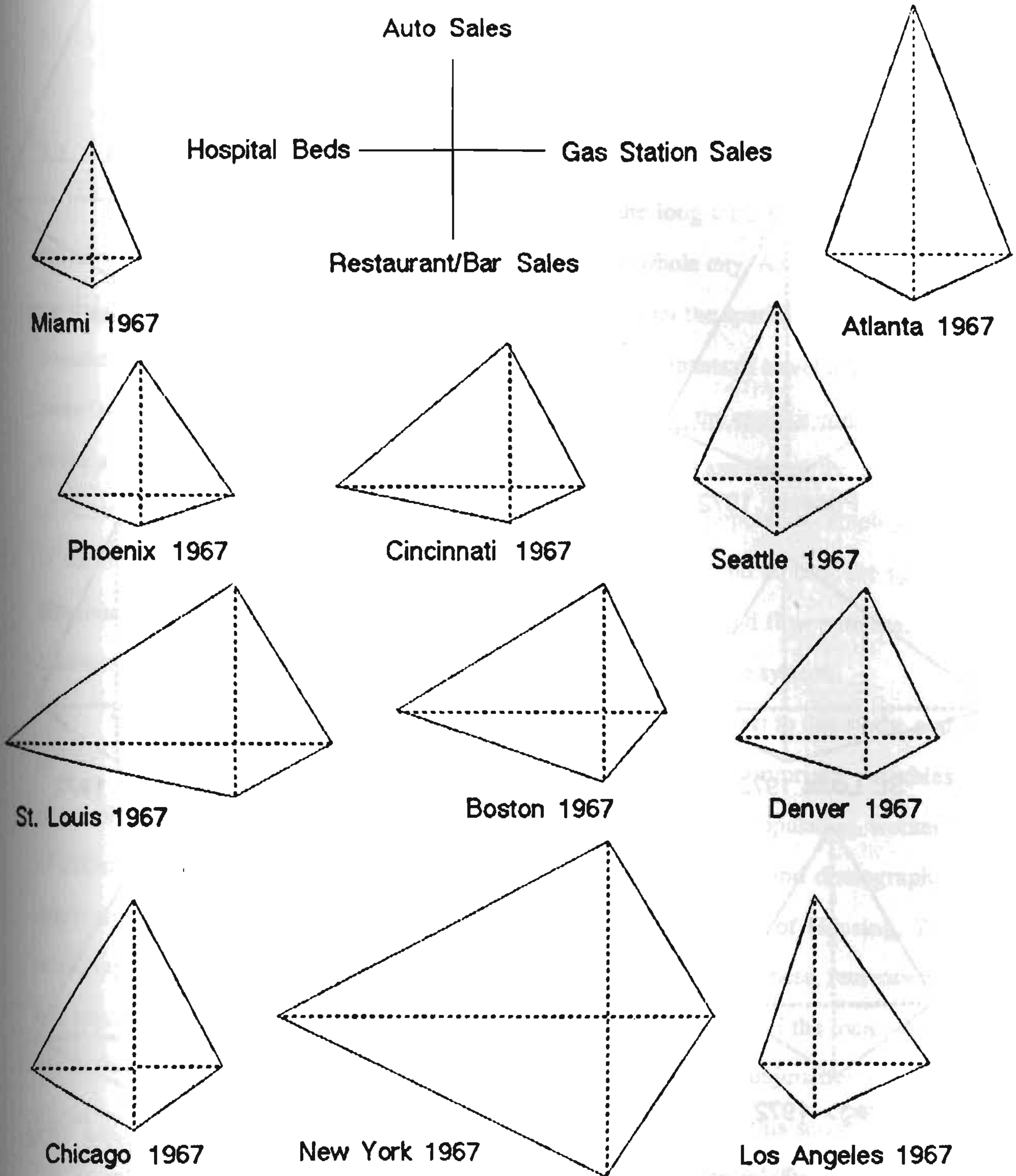
** Note : Chicago, New York and Los Angeles are reduced by a factor of 4.

Figure 2.56



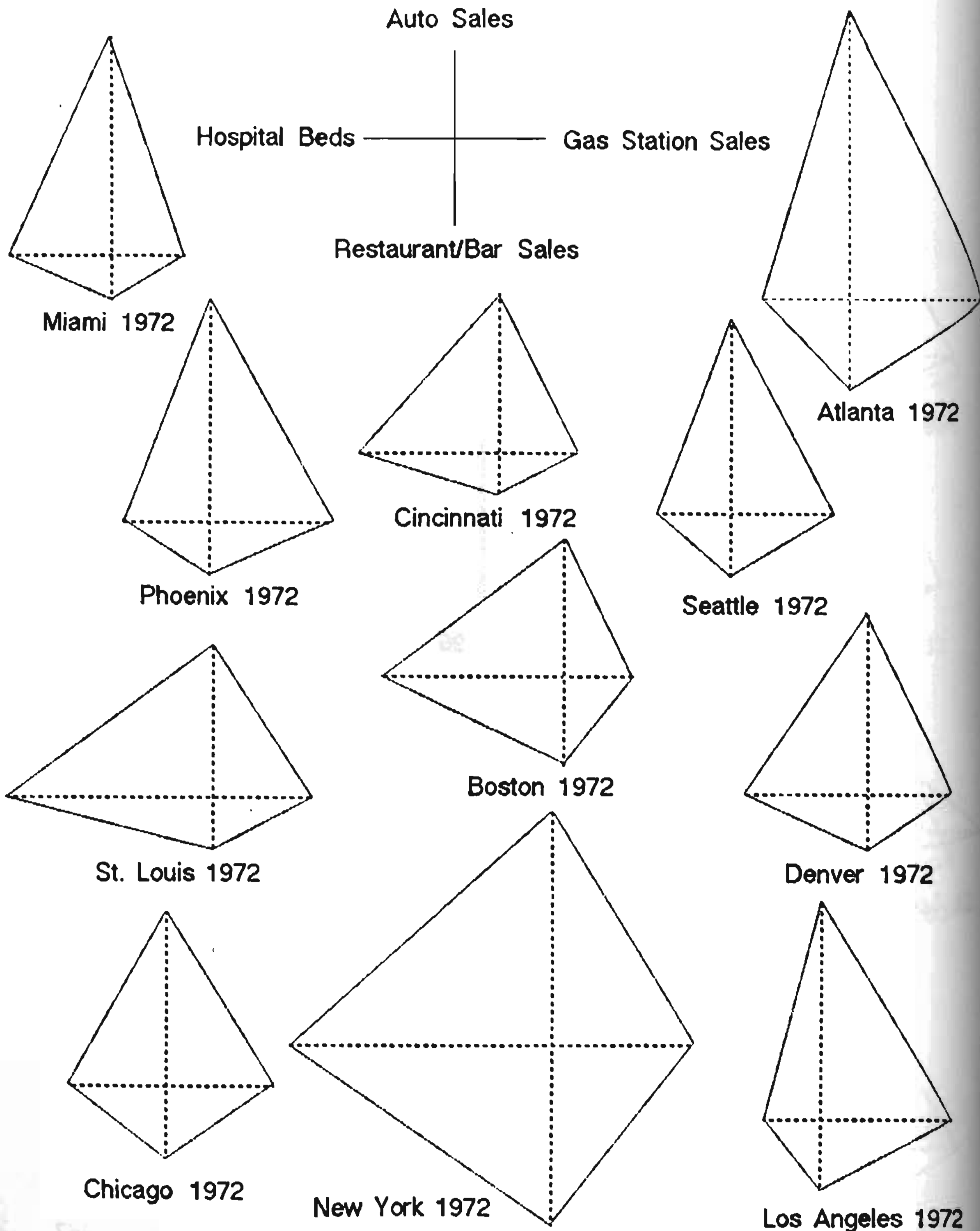
** Note : Chicago, New York and Los Angeles are reduced by a factor of 4.

Figure 2.57



** Note : Chicago, New York and Los Angeles are reduced by a factor of 4.

Figure 2.58



** Note : Chicago, New York and Los Angeles are reduced by a factor of 4.

Figure 2.59

CHAPTER 3

SPATIAL DENSITY PATTERNS

3.1. INTRODUCTION

In the previous chapter, the characterization of the long term evolution of cities relied on variables defined as aggregate quantities over the whole city. As such, no explicit spatial dimension was included in the analysis. It is clear that the spatial distributions of population, employment and other activities are critical determinants of travel intensity and associated flow patterns in the urban system. In our framework, the explicit consideration of the urban area's spatial characteristics provides an essential linkage among the evolution of variables such as those considered in the previous chapter (e.g. population, employment in various sectors, quantity of infrastructure of different types, and so on), the spatial distribution of these variables, the resulting travel characteristics and flow patterns, and ultimately the associated service levels experienced in the transportation system.

The spatial distributions of two types of variables are of interest to this study, and to any analysis of transport in an urban area. The first type comprises variables geographically identified by the dwellers' residential location, such as population, workers of different types, auto ownership, and many other socio-economic and demographic characteristics of households, typically provided by the U.S. Census of Housing. The second type consists of variables describing various employment, business, recreational and other activities in which residents engage; the spatial distribution of the locations at which these activities take place is the desired end result, for use in conjunction with the spatial patterns of the first type of variables. Unfortunately, data on this second type of variables at the desired spatial detail is not readily available, and requires the combination of information from several sources, such as various directories for each city. In this study, we focus on the first type of variables, primarily because the needed data gathering effort was more compatible with the study's time frame. This serves as an illustration of a general approach to examine and characterize the spatial patterns of interest, which could subsequently be extended to include other variables, and provides an initial indication of the changes in urban structure taking place over the past three decades.

The changing structure of cities and their transportation systems can be well illustrated by density functions. Population density functions, which capture the population density as a function of distance from the CBD, provide a good summary of the urban structure, and, when taken at different time periods, reveal the intervening changes in the urban landscape. In this chapter, the population density patterns of four urban areas (Atlanta, Austin, Dallas and Phoenix) are examined using the 1960, 1970 and 1980 U.S. Census data. Density functions are calibrated for these cities, using a cubic spline regression approach (recently suggested by Anderson [1985] for this type of application), providing a quantitative tool for the comparisons over time and across cities. The spatial patterns of the densities of other residence-based variables particularly relevant to passenger transport are also examined. For instance, we consider the density patterns of commuters by different transport modes, as well as the number of autos per square mile and per household, thereby capturing the variability of auto ownership with time and over space for a given city. In addition to overall density functions, specific directional density functions (i.e. within specific sectors or corridors) are important in revealing effects that may be unique to particular subareas within the city.

In the next section, we present the methodological aspects of the work presented in this chapter, including the background pertinent to urban density functions and the use of cubic spline regression models for this purpose, as well as the data sources and reduction procedures. This is followed by a discussion of the estimation results and the general trends in the evolution of the structure of each of the study areas. The examination of areawide density functions is followed in Section 3.5 by directional density functions for one of the study areas (Austin). Concluding comments are presented in section 3.6.

3.2. METHODOLOGICAL BACKGROUND

3.2.1. Density Functions and Their Estimation Using Cubic Splines

Population density functions are widely used to describe the changing structure of cities. Clark's (1951) pioneering work on the subject led to the early use of the negative exponential density function to represent urban structure, as follows:

$$Y_t = \alpha \exp[-\gamma X_t] \qquad (1)$$

where Y_t is the gross population density in census tract t, X_t is the distance in miles from the center of tract t to the CBD, α is a parameter representing the central business district (CBD) density, and γ a parameter known as the "density gradient", which reflects the degree of suburbanization.

The above form provides the classical mononucleated model of urban structure, with a smoothly decaying density from an urban area's CBD. However, as urban areas evolve, and suburbs continue their transformation from low-density bedroom communities for central city workers to active centers of retail, professional services and increasingly industrial activities, the negative exponential model's realism becomes questionable. Considerable empirical work has addressed the above functional form, such as Muth's (1969) investigation of its appropriateness for 46 cities. McDonald and Bowman (1976) tested 10 different functional forms, including the negative exponential, quadratic, and higher order polynomials, and evaluated their performance on several criteria, concluding that other forms outperformed the negative exponential in describing the urban density patterns in several cities. Recently, Anderson (1985) illustrated the use of cubic splines to the study of urban population density functions by applying the method to data from the Detroit metropolitan area. The presentation of the approach hereafter closely follows Anderson's paper (1985).

Cubic spline estimation is a procedure for estimating a relationship in several pieces (third-degree polynomials) in such a way that the pieces fit together smoothly (Poirier, 1973). This method involves dividing the x-axis (distance axis) into several segments, with latter type of information, i.e. to incorporate some feature of the area's transportation system into the spatial density function. The areas of the census tracts were obtained for this study from maps using a CAD automated planimeter system composed of an HP 1000 central processing unit, a CALCOMP 9000 digitizer (input device) and an HP2623A Graphics display (output display) unit.

In what follows, parameter estimates for the model of Eq. (4) are given only for the population density function. Graphical representations are given for the remaining variables; these graphs were generated using a cubic spline routine in the SAS statistical package.

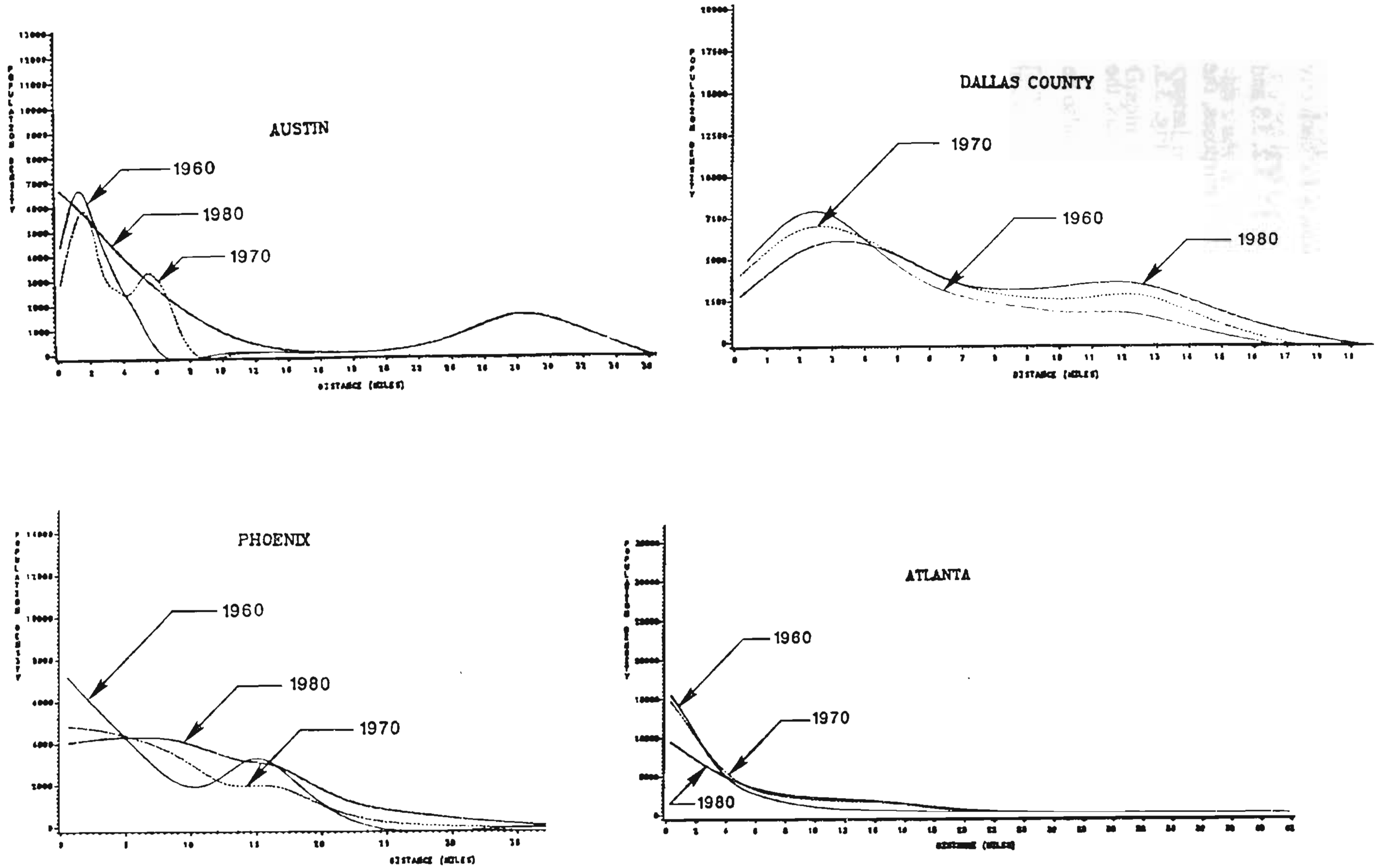
3.3. POPULATION DENSITY FUNCTIONS: ESTIMATION RESULTS

Figure 3.1 presents the fitted population density patterns in each of the four cities for 1960, 1970, and 1980, providing a graphical illustration of the evolution of urban structure in these cities. In all cases, there is a marked decrease over time in residential density in the core area, within 5 miles from the CBD, though not so evident in Austin, as discussed later. This is accompanied by an overall densification and "filling in" in the suburbs, whereby previous "troughs" between the highly peaked CBD and the much smaller outer peak in 1960 are being replaced a more uniform pattern of higher density. Also notable is the progressive "fattening" of the tail, away from the CBD, indicating the continuing move beyond the suburbs into the exurbs. Given the consistency of the patterns observed over the past three decades, these trends can be expected to continue, with formerly low-density suburbs reaching density and activity levels comparable to those previously associated with CBD's, and undoubtedly resulting in traffic congestion levels previously thought to be the exclusive domain of downtown streets.

However, while these general trends seem present in all four case areas, there are marked differences among these areas in terms of their evolution pattern and their current state along that process. These differences are more clearly visible in the parameter estimates for Eq. (4) than in the graphs. Estimation was performed separately for each city for each of the three time periods. Estimation results are shown in Tables 3.1, 3.2, 3.3 and 3.4 for Atlanta, Dallas, Austin and Phoenix, respectively. For comparison purposes, the results from a similar analysis for Detroit (Anderson, 1985) are reported in Fig. 3.5. Except for Austin, which has seen a dramatic growth in population in the past 10 years, the intercept a_1, interpreted as the density in the CBD, has decreased over time. It can also be noted that the magnitudes of the estimated coefficients corresponding to the most distant points (d_3-d_2) are smaller than those corresponding to closer points ([d_2 - d_1] or d_1).

Perhaps the most interesting indication in these estimates is the type of evolutionary pattern followed by each city, and, to some degree, its maturity. In particular, the results for Atlanta and Dallas exhibit clear trends which are consistent over time. For instance, all the signs of any given parameter behave consistently over time, and so do their magnitudes. This may be interpreted that the growth of these cities follows more stable patterns than those of either Phoenix or Austin. The results for the latter reveal the major structural change taking place in Austin and its transformation over the past decade. Similarly, while Phoenix shows consistent signs for the parameters of the lower degree terms (i.e. b_1 and c_1), the parameters of the higher degree terms (i.e. d_1, d_2-d_1 and d_3-d_2) exhibit fluctuating signs, unlike those of Atlanta, Dallas, or Detroit. The increased densification of the closer suburbs are evident through the decreasing magnitude of the parameters b_1 and c_1 over time, particularly apparent for Atlanta and Dallas, indicating the slower rate at which density decays away from the downtown. It would be useful to examine parameter estimates for a wider cross-section of cities, particularly in order to identify evolutionary paths that might provide a robust categorization of cities.

Another interesting set of quantities associated with the cubic splines consists of the location of the knots separating the various intervals, and the outward drift (away from the CBD) of these knots over time, reflecting the increased spatial dispersion and overall growth of the SMSA's over time.



POPULATION DENSITY PATTERN

Table 3.1 Estimated Coefficients for Austin Density Functions*

Year	$\hat{a}_1$	$\hat{b}_1$	$\hat{c}_1$	$\hat{d}_1$	d_2^d_1	d_3^d_2	Knots	R^2
1960**	3363.23 (3150.64)	6223.10 (5604.23)	-4691.05 (3147.75)	895.07 (574.51)	-904.73 (575.51)	- -	(0.16, 2.00, 15.47)	0.748
1970	2006.85 (3552.10)	9834.14 (5683.03)	-6340.18 (2861.44)	1031.61 (444.19)	-1129.61 (479.19)	97.00 (53.30)	(0.16, 2.50, 5.50, 12.75)	0.651
1980***	6066.40 (597.08)	-1154.78 (169.09)	72.38 (13.70)	-1.48 (0.33)	1.57 (0.47)	- -	(0.14, 17.00, 36.13)	0.583

* Estimated standard errors of coefficients are given in parentheses

** Due to insufficient data, only three knots were selected; hence, there is no estimate of (d_3 - d_2).

*** For 1980 data, three knots provided the best fit to the data.

Table 3.2 Estimated Coefficients for Phoenix Density Functions*

Year	$\hat{a}_1$	$\hat{b}_1$	$\hat{c}_1$	$\hat{d}_1$	d_2^d_1	d_3^d_2	Knots	R^2
1960	6264.55 (1215.64)	-1097.83 (658.72)	58.03 (104.66)	-0.69 (5.06)	-1.09 (5.54)	1.79 (0.61)	(0.43, 8.00, 16.00, 52.50)	0.505
1970	5172.48 (980.15)	-625.35 (476.83)	10.49 (68.66)	0.79 (3.00)	-2.02 (3.36)	1.22 (0.43)	(0.56, 9.00, 17.00, 55.68)	0.460
1980	4955.99 (902.50)	-1219.74 (334.68)	132.54 (37.36)	-4.62 (1.27)	6.18 (1.63)	-1.66 (0.43)	(0.56, 12.00, 19.00, 55.68)	0.378

* Estimated standard errors of coefficients are given in parentheses

Table 3.3 Estimated Coefficients for Atlanta Density Functions*

Year	$\hat{a}_1$	$\hat{b}_1$	$\hat{c}_1$	$\hat{d}_1$	$d_2{}^\wedge d_1$	$d_3{}^\wedge d_2$	Knots	R^2
1960	18867.75 (1599.27)	-7164.68 (1255.23)	1077.10 (289.30)	-59.95 (20.38)	51.35 (22.70)	8.42 (2.94)	(0.27, 5.50, 10.50, 32.85)	0.747
1970	11275.65 (776.52)	-2591.88 (348.74)	236.61 (46.16)	-7.65 (1.83)	9.43 (3.13)	-2.22 (1.60)	(0.27, 11.00, 15.50, 33.47)	0.773
1980	7786.83 (697.82)	-1517.37 (274.52)	130.54 (31.96)	-4.04 (1.12)	5.42 (1.77)	-1.60 (0.77)	(0.27, 12.00, 17.00, 41.84)	0.696

* Estimated standard errors of coefficients are given in parentheses

Table 3.4 Estimated Coefficients for Dallas Density Functions*

Year	$\hat{a}_1$	$\hat{b}_1$	$\hat{c}_1$	$\hat{d}_1$	$d_2{}^\wedge d_1$	$d_3{}^\wedge d_2$	Knots	R^2
1960	2805.15 (1611.52)	4672.63 (1617.84)	-1760.11 (483.84)	165.31 (43.86)	-185.93 (49.08)	30.10 (2.53)	(0.38, 4.50, 10.00, 16.68)	0.647
1970	2577.32 (1927.67)	3941.30 (1658.05)	-1285.71 (428.00)	105.66 (33.59)	-126.82 (39.47)	28.57 (11.00)	(0.15, 5.00, 10.00, 17.68)	0.611
1980	2375.23 (1489.30)	2699.41 (1223.98)	-833.53 (302.64)	65.61 (22.67)	-81.20 (26.85)	25.29 (8.68)	(0.16, 5.50, 11.00, 19.34)	0.602

* Estimated standard errors of coefficients are given in parentheses

TABLE 3.5 Estimated Coefficients for Detroit Density Functions*

Year	$\hat{a}_1$	$\hat{b}_1$	$\hat{c}_1$	$\hat{d}_1$	$\widehat{(d_2 - d_1)}$	$\widehat{(d_3 - d_2)}$	Knots	n	SE
1960	8401.05 (2896.32)	9776.91 (2888.60)	−2913.15 (844.44)	228.95 (73.86)	−226.01 (96.88)	−77.46 (85.76)	[0.10, 4.68, 9.26, 13.84]	484	3521.36
1970	3671.13 (2846.68)	7780.76 (2833.13)	−1790.37 (830.22)	120.76 (72.94)	−91.07 (96.36)	−102.47 (88.90)	[0.05, 4.62, 9.20, 13.77]	420	3122.82
1980	4572.69 (1702.09)	3148.62 (1860.03)	−615.98 (595.36)	40.79 (55.95)	−41.33 (77.54)	27.25 (85.73)	[0.38, 4.79, 9.21, 13.63]	343	2538.31

*Estimated standard errors of coefficients are given in parentheses.

Source: Anderson (1985)

Finally, the R^2 values for the regressions are quite reasonable in all cases, with the possible exception of Phoenix, where R^2 values ranged from 0.51 to 0.38. This underscores an important conceptual difficulty associated with this model's assumption of isotropic urban development, that can be explained exclusively by distance from the CBD. In actuality, asymmetries exist, with certain sectors or corridors experiencing greater densities than others, for a given distance from the center. For this reason, Anderson (1985) suggested the examination of "directional" density functions, taken only along one dimension of development away from the CBD. This type of directional analysis can reveal more about an urban area's spatial development patterns than areawide undifferentiated functions of the kind discussed in this section. For instance, the significance of an emerging relatively high-density non-downtown activity center would be more readily visible at the corridor or sector level than an areawide analysis, in which this center might be masked by the majority of less dense tracts located at the same distance from the CBD. A directional analysis of several corridors in the Austin area further illustrates this point in section 3.5. However, we first examine the spatial density patterns of the remaining residence-based variables, other than population, that have particular relevance to transport.

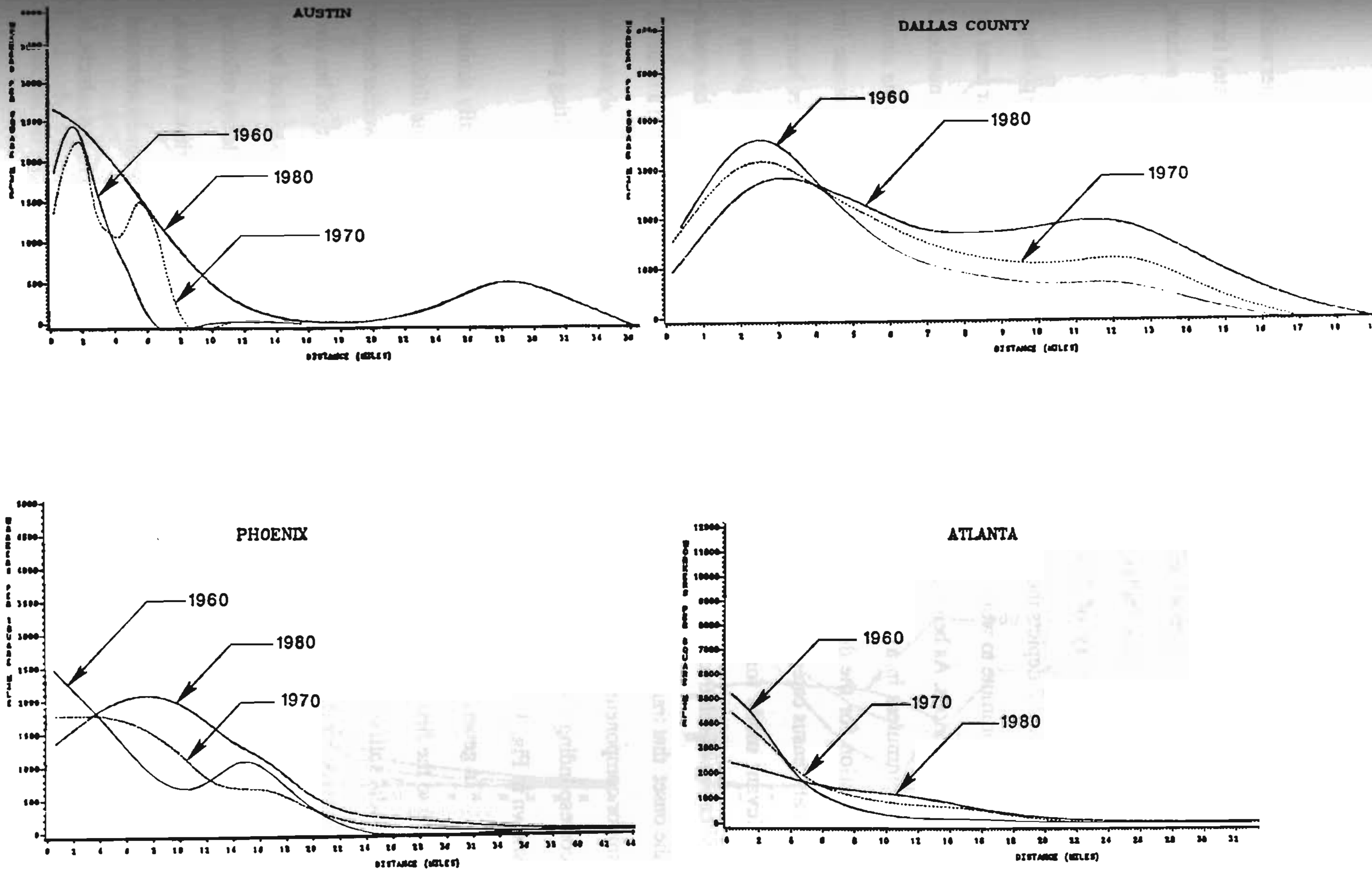
3.4. EVOLUTION OF URBAN STRUCTURE: OTHER VARIABLES

In this section, we continue the examination of the evolution of the spatial distribution of several demographic and transport related variables, geographically identified by residential location, for the four case areas. In particular, we consider the spatial density of workers, the density of private auto and carpooling commuters, the density of public transit users, the average number of autos per household, and the density of autos (measured as the number of autos in a given tract divided by the tract's area), all as a function of the distance from the CBD. The results of this analysis are in general fairly predictable given the above population density patterns, because most of these variables (with the exception of autos per household) tend to be highly correlated with population density.

3.4.1. Worker Density Patterns

Figure 3.2 depicts the results of the cubic spline regression fit to the census-tract level observations of the spatial density of workers, as a function of distance from CBD, for the four urban areas, in 1960, 1970 and 1980. The dependent variable was defined analogously to the population density case, by dividing the number of workers living in each tract by the area of that tract. In all four case areas, the patterns depicted in Fig. 3.2 very closely mirror those exhibited by the population density. Generally, worker densities tend to be around 40 to 50% of the corresponding population density. Discrepancies between the spatial patterns of population and worker densities thus reflect variations in the ratio of workers per person, or in the fraction of workers per household across the urban space. Such variation may be due to larger households with many young children, higher participation rates of women in the work force, the predominance of retirees in particular areas, or higher unemployment rates. Likely illustrations of each of the above can be detected by carefully comparing Fig. 3.1 to Fig. 3.2. Several observations can be made in this regard.

In Atlanta, the worker density near the origin, relative to the corresponding population density, is less than the relative worker densities as one moves away towards the more affluent suburbs, especially for the 1980 data. The lower workers per person ratio in all likelihood reflects higher unemployment rates in the center city, in which poorer racial minorities have tended to cluster over time. In Dallas, in the range of 7 to 13 miles from the CBD, it appears that the difference in worker density between 1970 and 1980 exceeds the difference in this quantity between 1960 and 1970 by a factor that is greater than the corresponding changes in population density. This indicates a marked increase in the workers per person ratio in the early 1980's, reflecting two simultaneous and continuing trends: more two-worker households due to higher labor force participation by women,



DENSITY PATTERN OF WORKERS

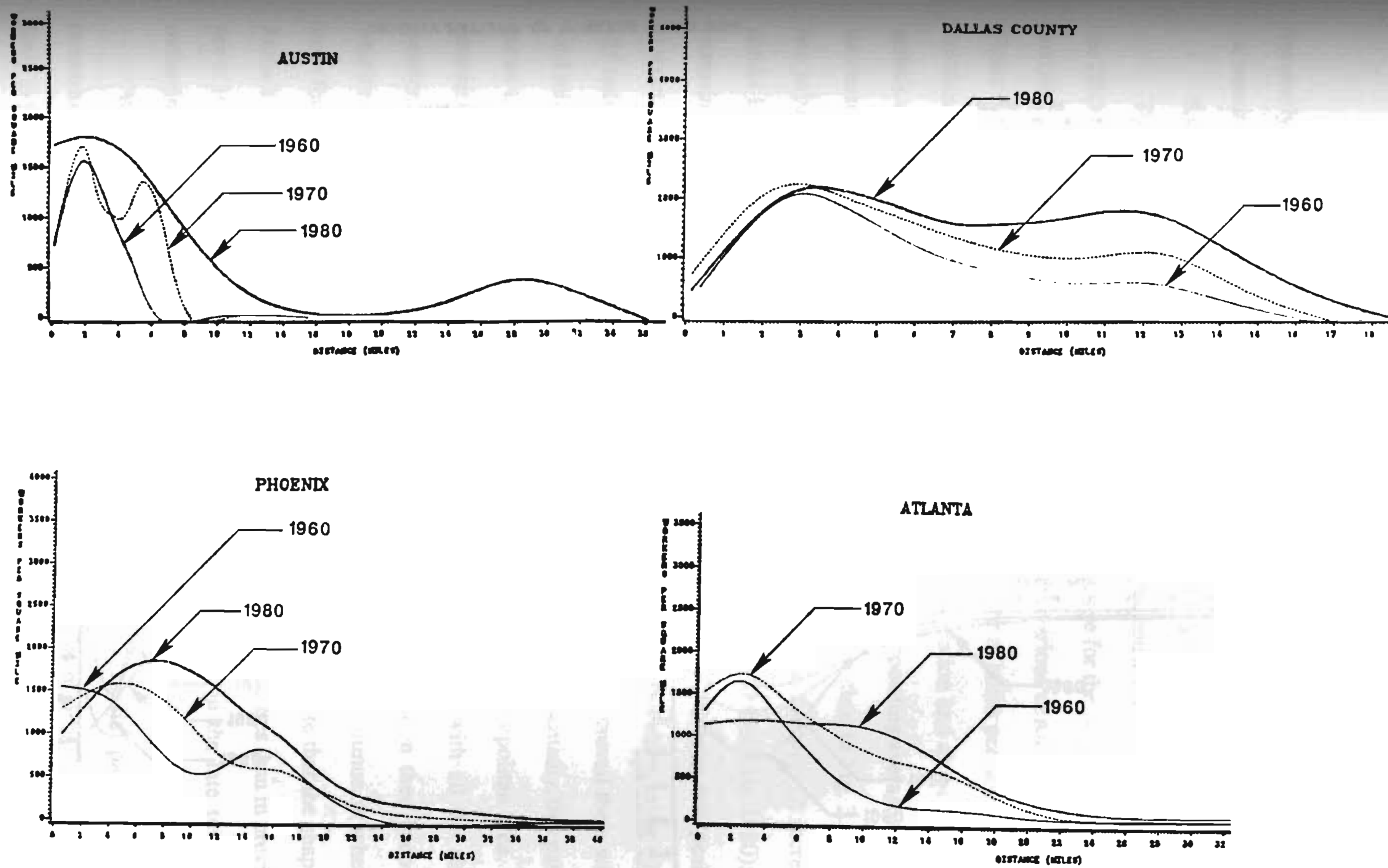
Figure 3.2

accompanied by smaller family sizes, among young couples in the middle to upper middle class that tend to locate in the suburban ring under consideration. Another related factor may be the larger number of working singles who live in the densifying suburbs, particularly in a city like Dallas whose growth was peaking in the early 1980's.

3.4.2. Density of Commuters by Transport Mode

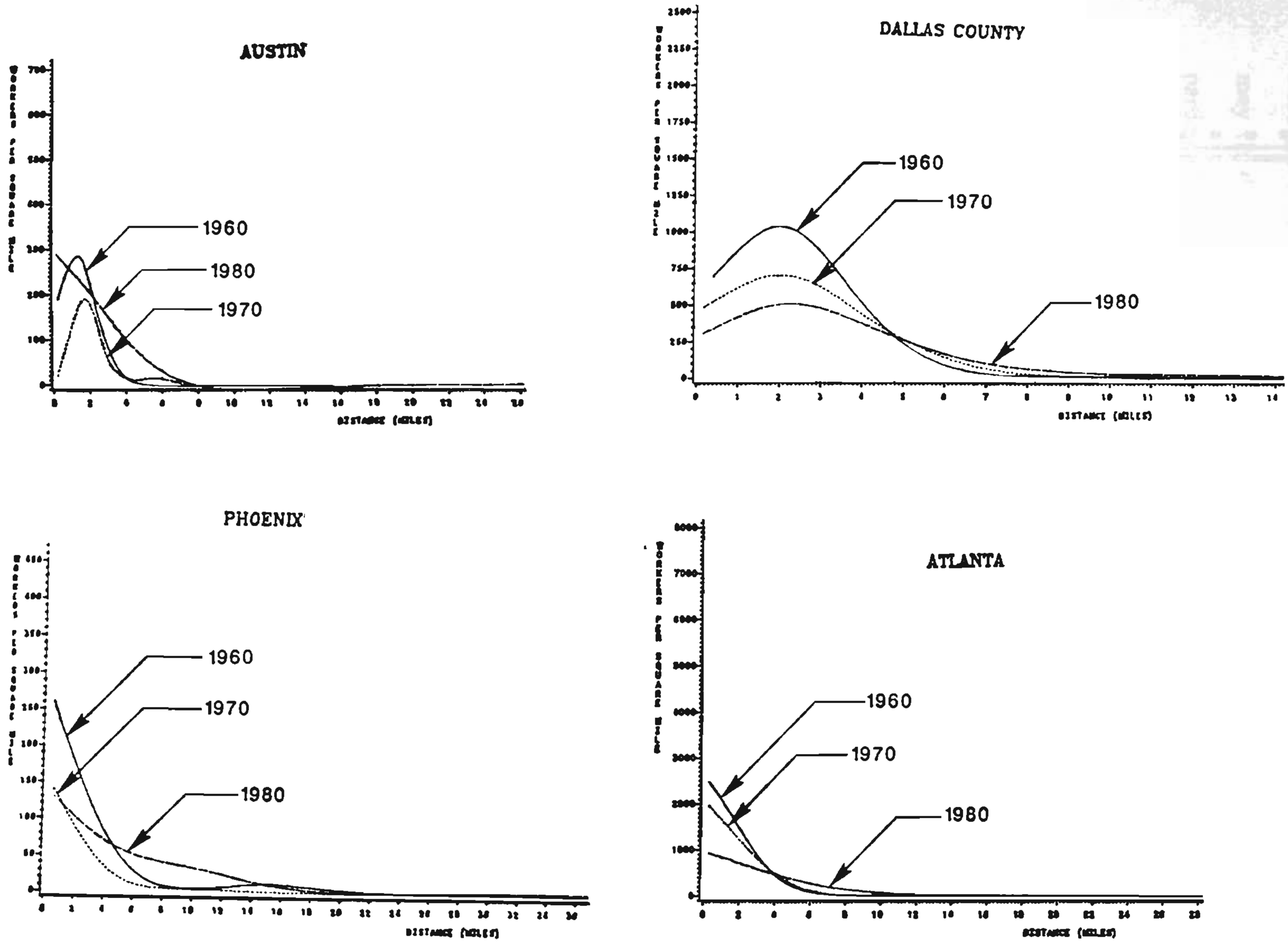
Figure 3.3 depicts the density of workers driving their automobile or carpooling in their daily commute to work, as a function of distance from the CBD, for all four areas and three time periods. As before, the dependent variable is obtained by dividing the number of such commuters in a given tract by the area of that tract. Figure 3.4 presents similar information for the density of workers using bus transit for their work commute (the official Census category consists workers using bus transit or streetcar, the latter being an irrelevant mode for the cities under consideration in this study), thereby giving an indication of relative mode usage as a function of distance from CBD. It must be noted at the onset that transit usage in the four case cities has been over the study period a relatively minor component of the total suburban travel picture, and that the statistical analysis of the corresponding observations might not be very meaningful.However, the resulting patterns shown in Fig. 3.4 appear to be meaningful and plausible.

In general the spatial density functions for both categories are generally similar in shape to the first two curves (worker density and population density). Possible differences in modal split over space and time did not perceptibly distort the underlying worker density patterns. Some differences are nevertheless apparent. For example, the density of bus users in Austin in 1970 does not exhibit the smaller though distinct peak exhibited by the population and worker densities at ~6 miles away from the CBD, thereby clearly reflecting the negligible transit usage in Austin among the suburban residents at that time. In Atlanta, relative transit usage among central city residents (up to ~4 miles) has remained substantial over time, at about 40%. This level rapidly drops as one moves away into the suburbs. In Phoenix, relative transit usage in the core area experienced a drop through the sixties, but



DENSITY PATTERN OF WORKERS USING PRIVATE AUTO OR CARPOOL

Figure 3.3



DENSITY PATTERN OF WORKERS USING BUS

Figure 3.4

apparently reversed some of it in the seventies, as seen by comparing the worker density to the transit users density patterns for the corresponding years.

3.4.3. Spatial Patterns of Automobile Ownership

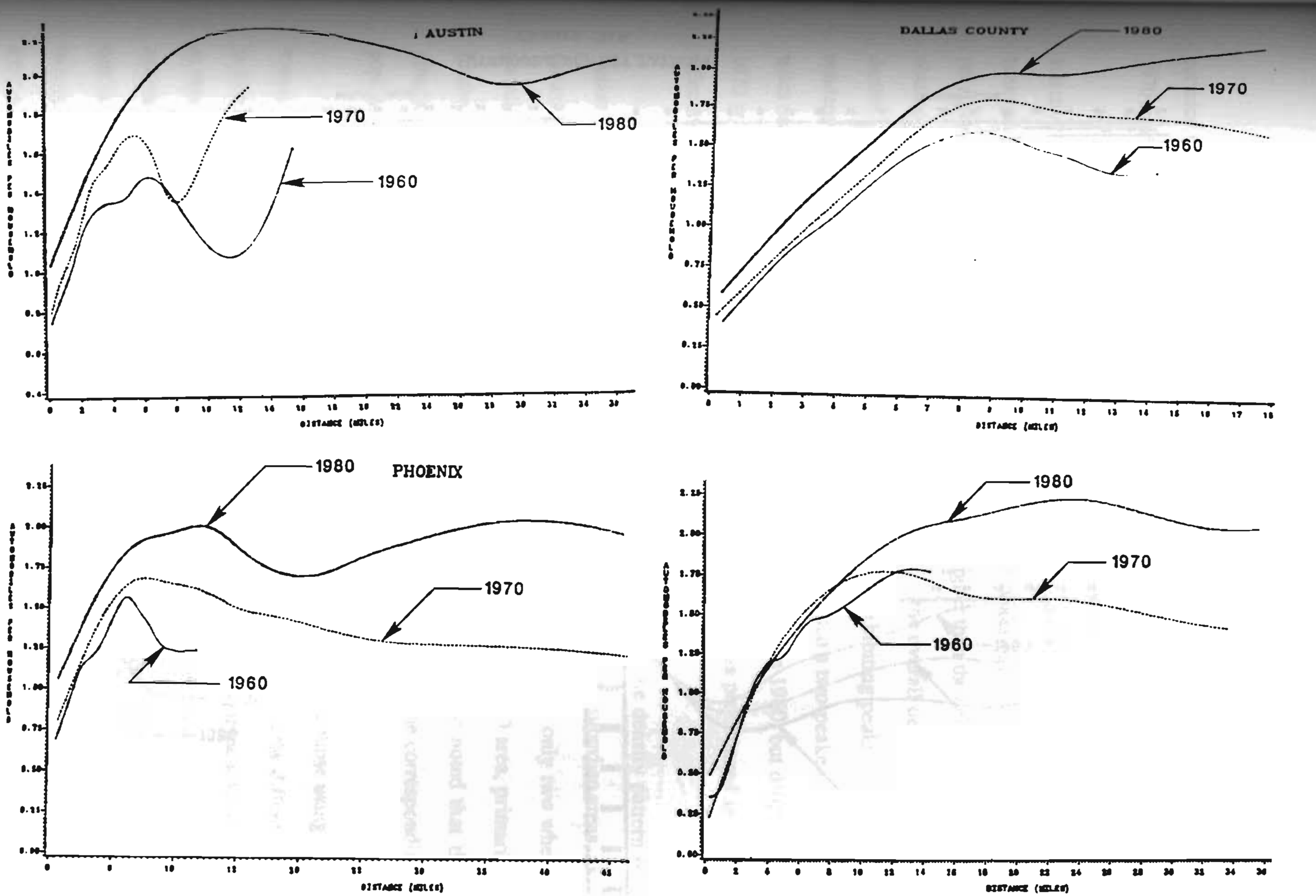
The variation of average automobile ownership per household with distance from the CBD provides a useful characterization of urban structure for the purpose of transport analysis, and one that is not as highly correlated with the various densities seen earlier. Figure 3.5 depicts the cubic spline fit to the average number of autos per household as a function of distance, for the four urban areas under consideration and at the three time periods considered so far. The dependent variables are the corresponding averages taken at the census tract level. Several important and consistent trends are clear in Fig. 3.5, for all four urban areas. Car ownership has increased steadily over time. Spatially, it increases almost linearly with distance from the CBD for about the first 10 miles (in 1980); the increasing trend with distance continues beyond that point, but at a decreasing rate, peaking at about 2 to 2.2 vehicles per household, at a distance that varies across cities, in the ~15 to ~25 miles range (for the 1980 data). A decreasing pattern appears to prevail then, with a local minimum somewhere in the ~25 to ~35 miles range (in 1980). Essentially, this pattern seems to mirror the spatial pattern of household income in the metropolitan area. The temporal evolution of this pattern also appears to be consistent, with all the above mentioned breakpoints progressively shifting outwards. The question then becomes whether the pattern observed over the past twenty or thirty years will continue, or whether some form of saturation has been reached. It is also interesting to note that the shape of Austin's 1980 pattern more closely resembles that of the other three cities than in previous years, reflecting the rapid growth and transformation of Austin in the late seventies, continuing onto the eighties.

While our sample is limited to four areas, the above spatial pattern of auto ownership and its evolution appear to be robust characteristics of urban areas, and are likely to hold in other mature cities as well. Additional data from a selected cross-section of areas could yield a useful tool for exploring the implications of evolving urban structure on automobile ownership.

The spatial pattern of automobile ownership density (autos per square mile, calculated by dividing the number of automobiles owned by residents of each given tract by the area of the tract) is shown in Fig. 3.6, as a function of distance from the CBD, for each of the areas and time periods of interest. These functions reflect 1) the underlying distribution of population (and/or workers) and 2) the household auto ownership patterns just discussed. The result is that while the general shape of the auto density functions is similar to that of population or workers, the differences in worker densities over space and time are magnified between the corresponding auto densities. Atlanta provides a good example of this phenomenon, where the differences in automobile ownership densities over time are relatively more substantial than the corresponding differences in population or worker densities.

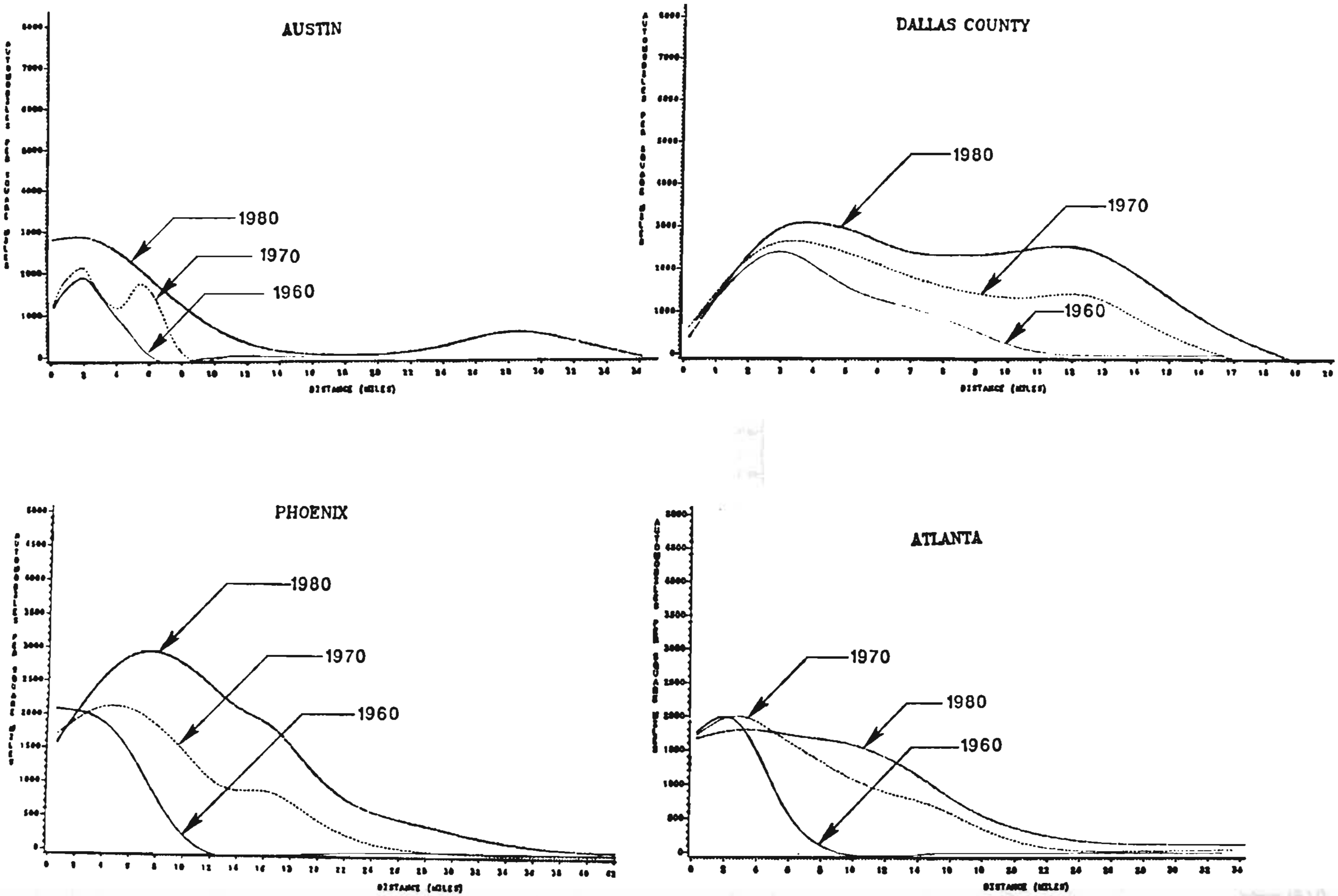
3.5. DIRECTIONAL DENSITY PATTERNS FOR AUSTIN

In order to further identify certain characteristics of urban structure that might have been masked when considering the overall study area, it is useful to examine some directional spatial density patterns, i.e. within specific travel corridors or sectors of the urban area. Emerging suburban centers unique to certain corridors would then be more likely to be uncovered by the analysis. To illustrate this approach, corridor-level analyses of the same quantities considered in the previous section were conducted for one of the case areas. Austin was selected for convenience, particularly since corridors had recently been defined for transit planning purposes by the appropriate agency. Six different corridors were considered in the Austin metropolitan area, with each corridor being radially oriented with one end being the CBD. These corridors are referred to as the North Central, Northwest, South Central, Southwest, Northeast and Northwest corridors, respectively. The cubic spline regression procedure was used here as well, and applied within each given



PATTERN OF AUTOMOBILES PER HOUSEHOLD

Figure 3.5



AUTOMOBILE DENSITY PATTERN

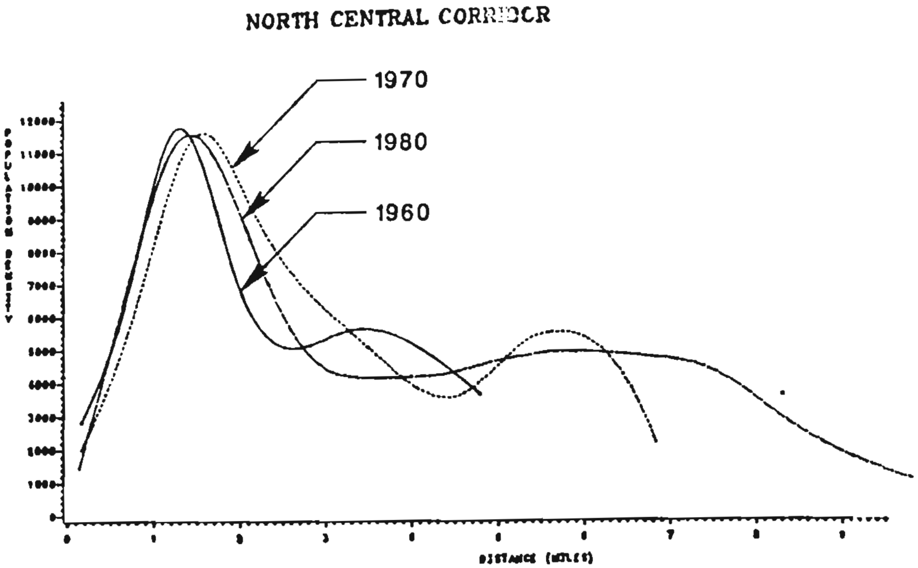
Figure 3.6

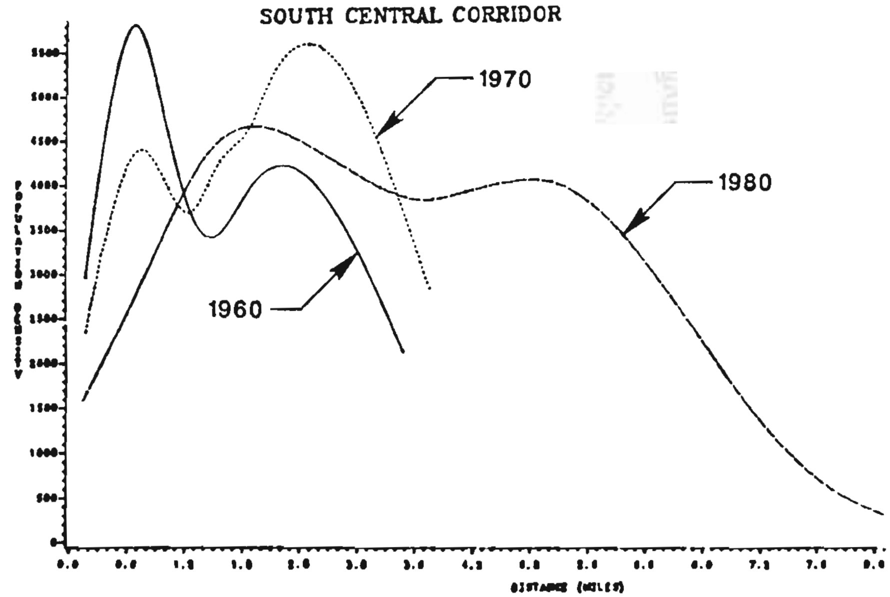
corridor (using only those census tracts that lie mostly within the corridor), for 1960, 1970, and 1980.

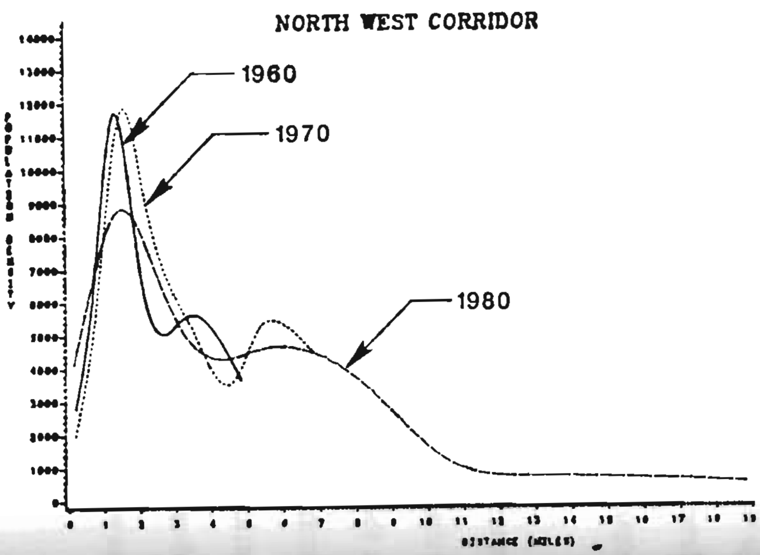
Figure 3.7 presents the population density patterns for the six corridors, indicating that the North Central and Northwest corridors are the most dense, both in terms of higher peaks and sustained high levels over long distances, with peaks of about 12,000 residents per square mile near the CBD area. This value is quite higher than the corresponding peak obtained (7,000 per square mile) for the cubic spline fit of the overall density pattern for Austin. The South Central and Southwest corridors, while exhibiting peaks below 6,000 residents per square mile, were interesting because they revealed a two-peaked city at 0.5 miles and 3 miles from the CBD. The first peak was dominant in 1960, but dropped by 1970 while the second peak had risen. The two peaks became less pronounced in 1980, reflecting the greater dispersion of residents within these corridors.

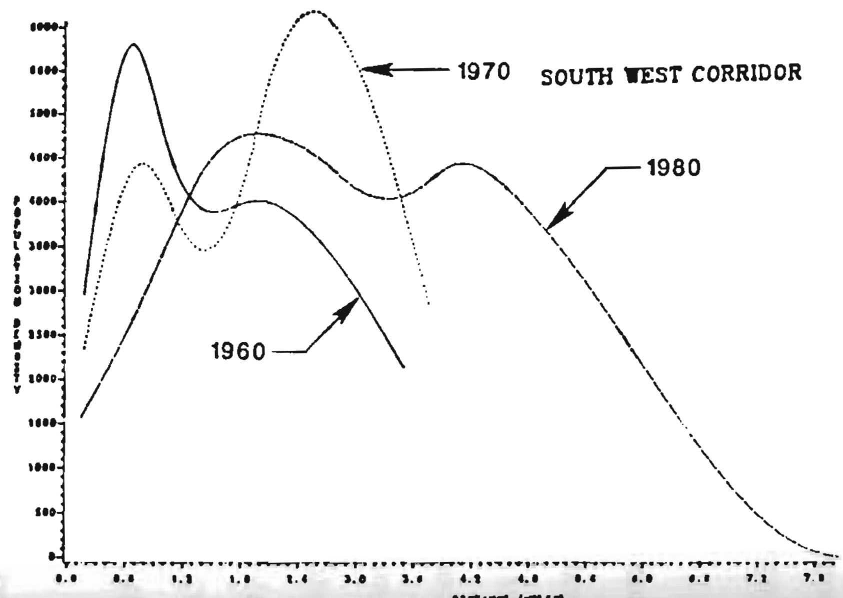
In all corridors, except the North Central and Northwest, the density pattern of workers exhibited a drop near the CBD, coupled with an increase in the suburban areas, as shown in Fig. 3.8. The North Central and Northwest corridors are the only two where densities of workers remained stable or increased slightly near the CBD area, primarily because they include the University (of Texas) area. It can again be noted that the directional worker density patterns maintained a high correlation with the corresponding population density patterns.

The density patterns for workers using private auto or carpool, and those using bus transit, are generally similar to those observed for the area as a whole, with the differences in distance that are better reflected in the directional densities. They are not included herein as their benefit is of a primarily local nature. Figure 3.9 depicts the spatial density of automobile ownership as a function of distance from the CBD, within each of the six corridors. The spatial pattern of household automobile ownership is shown in Fig. 3.10, indicating that the general trends discussed in the previous section in conjunction with the areawide data appear to hold at the corridor level as well. The increase in auto ownership



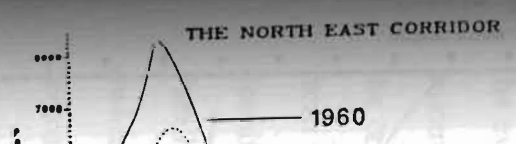


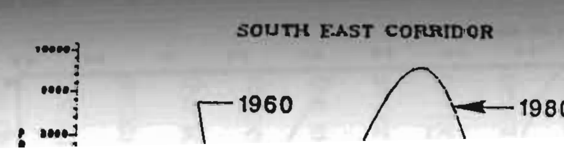




POPULATION DENSITY PATTERN

Figure 3.7





POPULATION DENSITY PATTERN

Figure 3.7

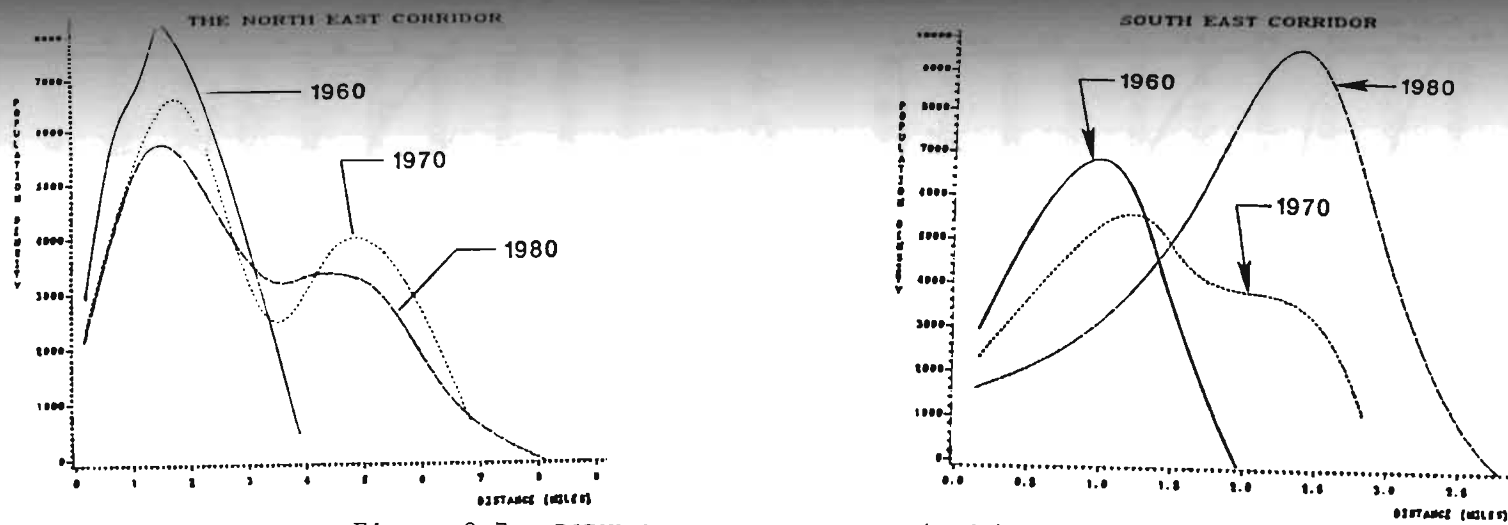
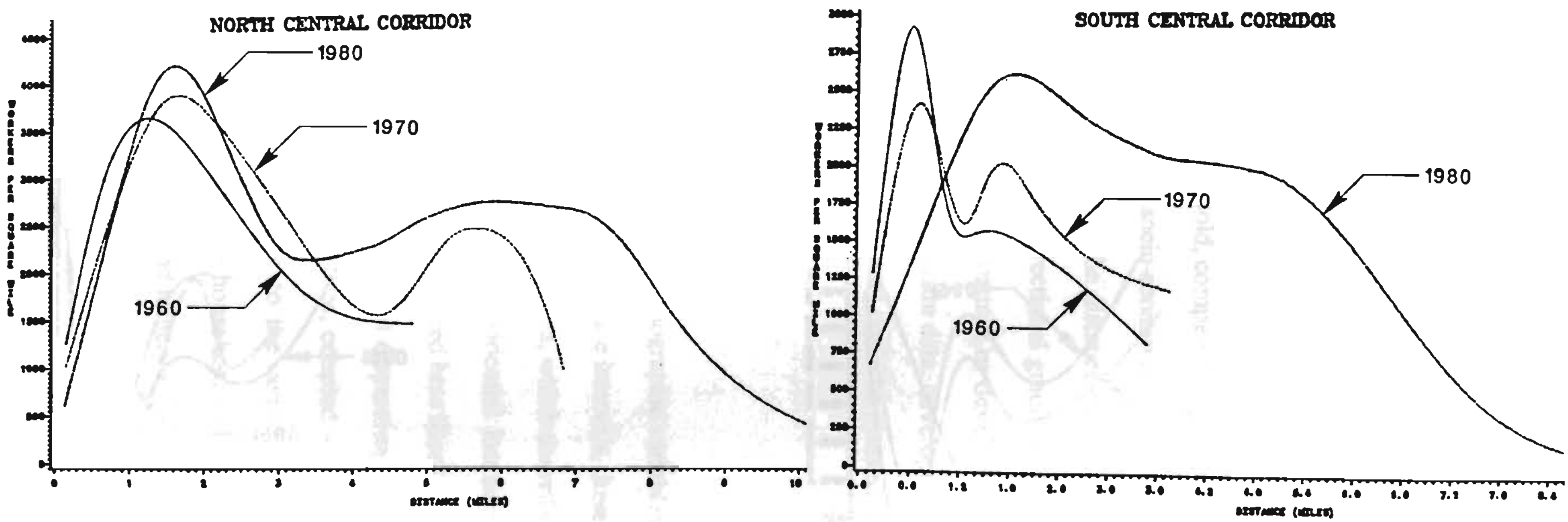


Figure 3.7 **POPULATION DENSITY PATTERN** (ctd.)



DENSITY PATTERN OF WORKERS

Figure 3.8

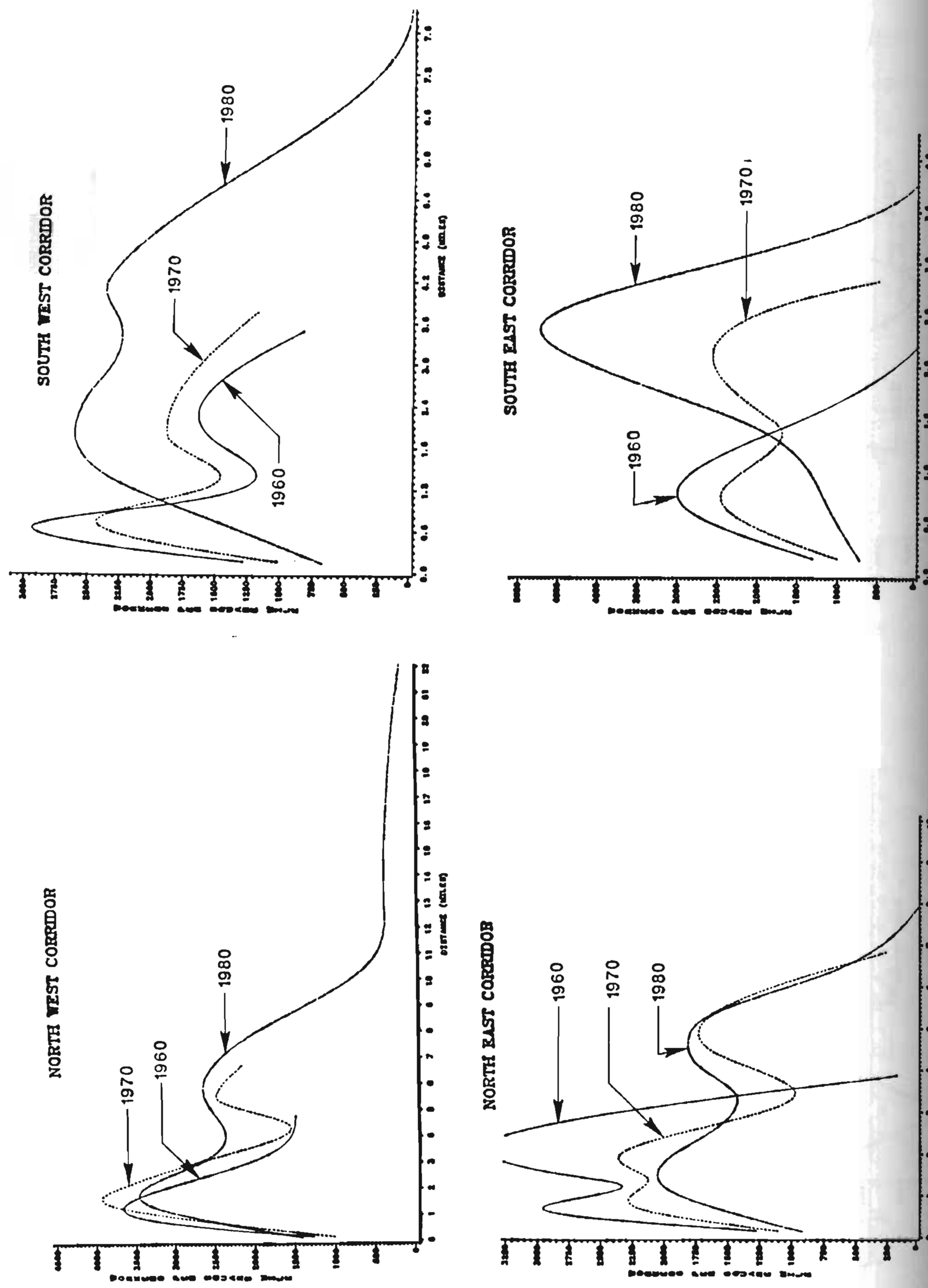


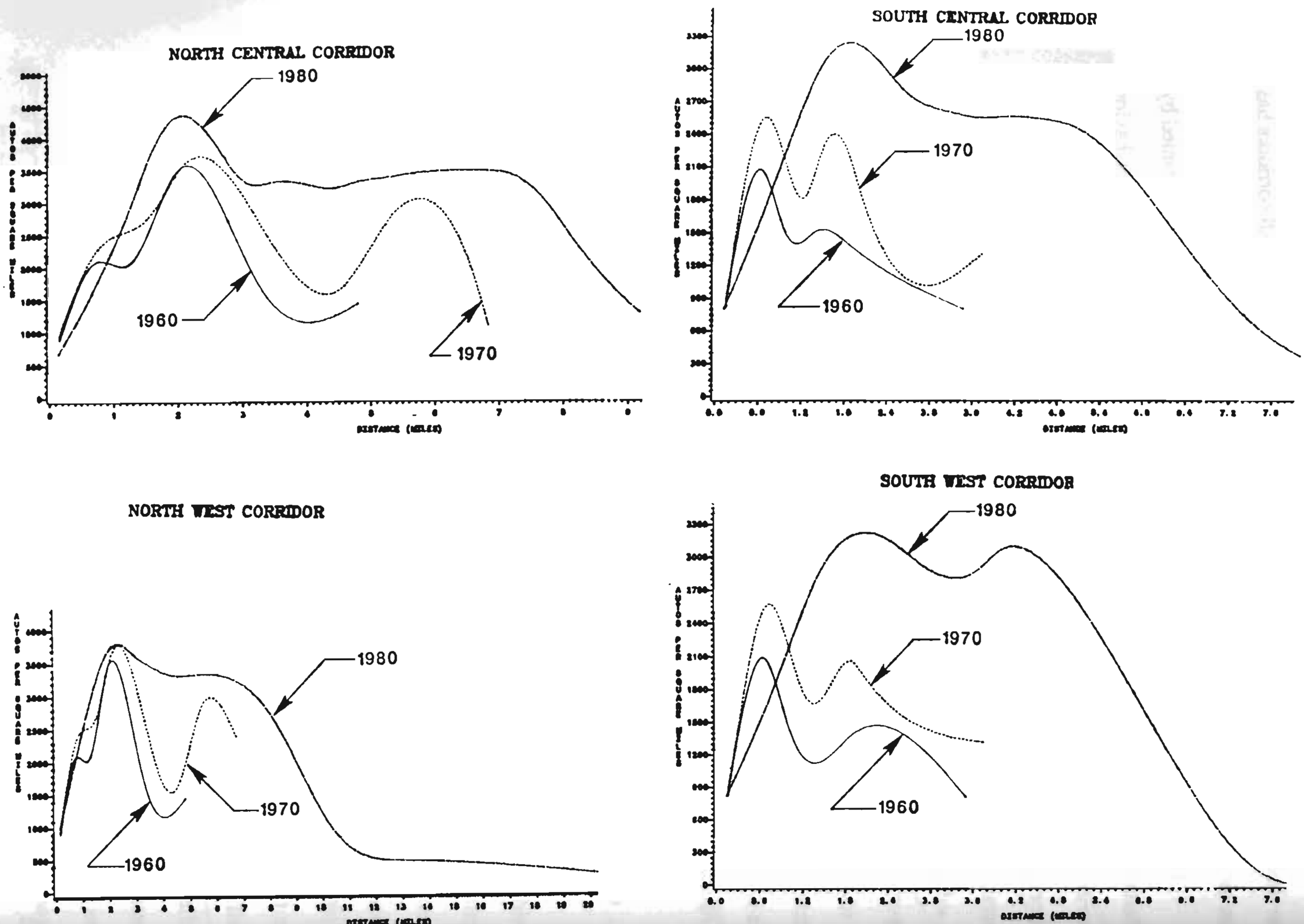
Figure 3.8 DENSITY PATTERN OF WORKERS (cont.)

over time is present in all corridors. It can also be noted that the increase in all corridors has been more rapid in the outer areas than near the center.

However, the plots also indicate that those general similarities are accompanied by differences in terms of the peak values and of the distance ranges over which the behavior described earlier applies. For instance , it can be seen that the peak average household auto ownership in the Southeast corridor is ~1.7 autos per household, compared to over 2.2 in several other corridors. These differences undoubtedly reflect socio-economic conditions in the subareas of interest. Spatial differences are also revealed in the distance at which the respective peaks occur, and the corresponding gradients. Thus directional gradients allow the identification of possibly substantial differences in urban structure or development patterns relative to the overall city pattern. As illustrated by the Austin data, development does not occur in a symmetric manner, and emerging suburban nuclei could be masked in the data for the overall area.

3.6. CONCLUDING COMMENTS

The spatial density functions of population and other socio-demographic variables provide a useful characterization of urban structure. Taken at different time intervals, these functions provide snapshots of the development pattern of an urban area, which describe the evolution of urban structure in terms of the spatial patterns of residential location, automobile ownership and other variables. The four case studies considered here illustrated the continuing trends, observed over the past three decades, of spatial dispersion away from the traditional central urban cores into the suburbs and exurbs, coupled with the densification of existing suburbs to levels almost comparable to the central cores. Parallelling this pattern has been a continuing growth of average household automobile ownership, in all parts of the urban area, with a distinct spatial and temporal pattern that seems to be robust across the case areas considered as well as within radial corridors in the one case that was so analyzed. While the resulting travel patterns cannot be determined in



PATTERN OF AUTOS PER SQUARE MILE

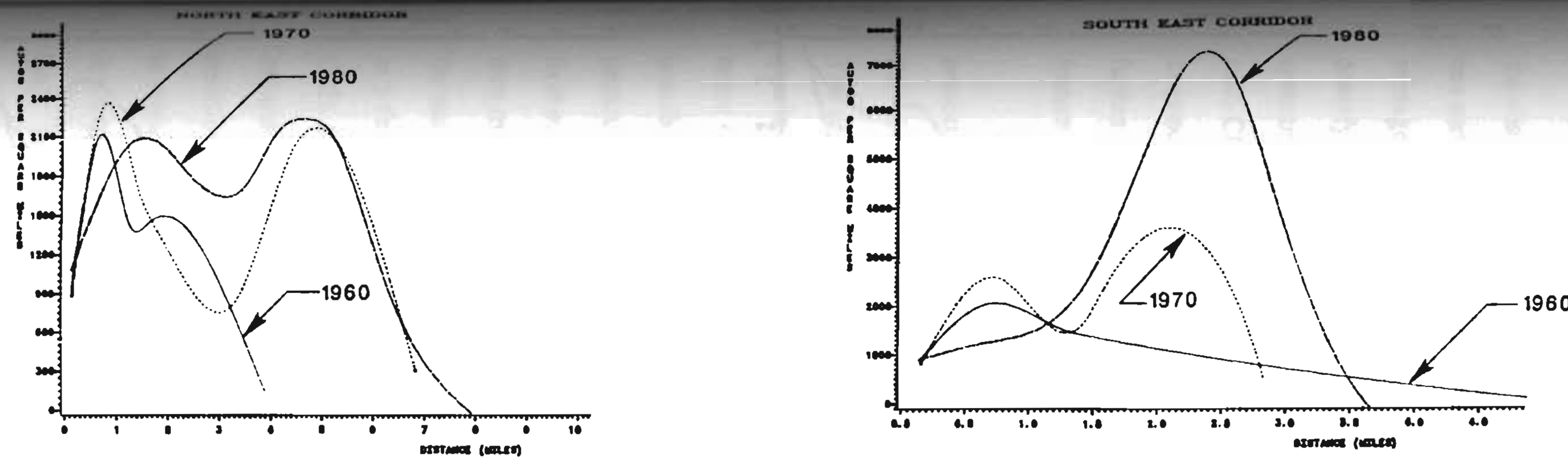
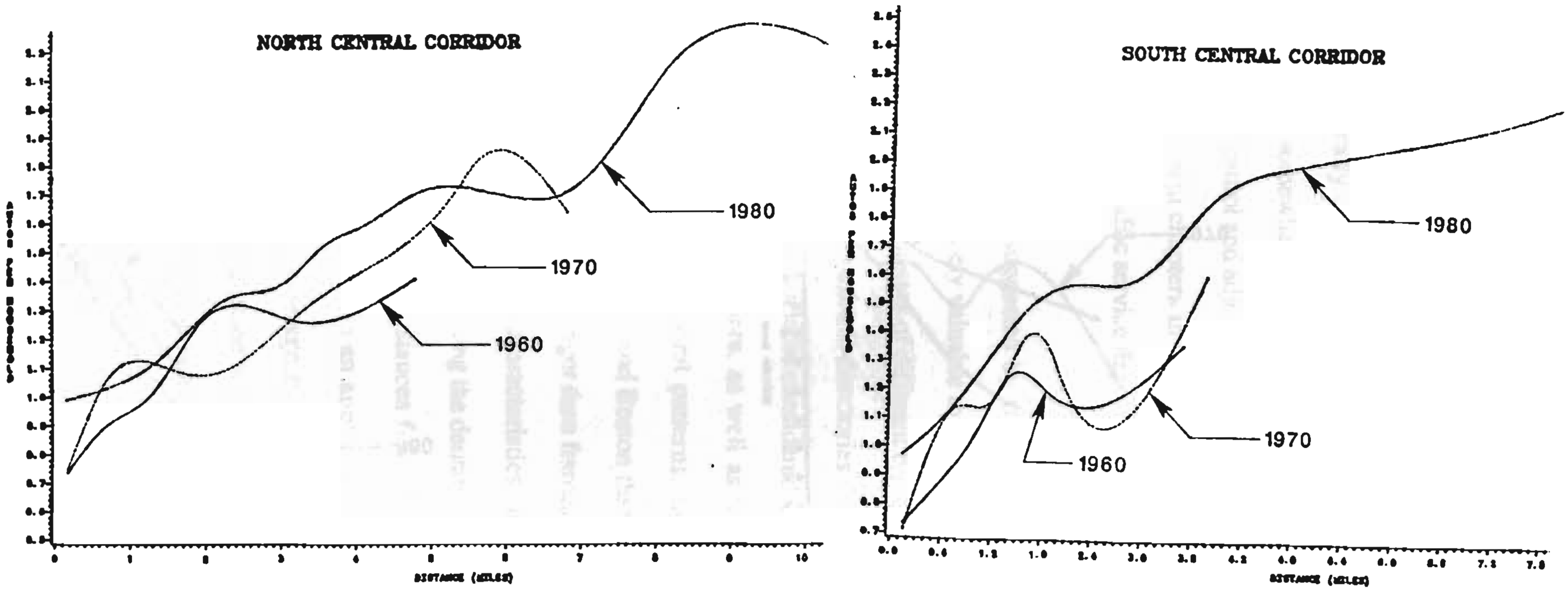


Figure 3.9 **PATTERN OF AUTOS PER SQUARE MILE** (ctd.)



PATTERN OF AUTOS PER HOUSEHOLD

Figure 3.10

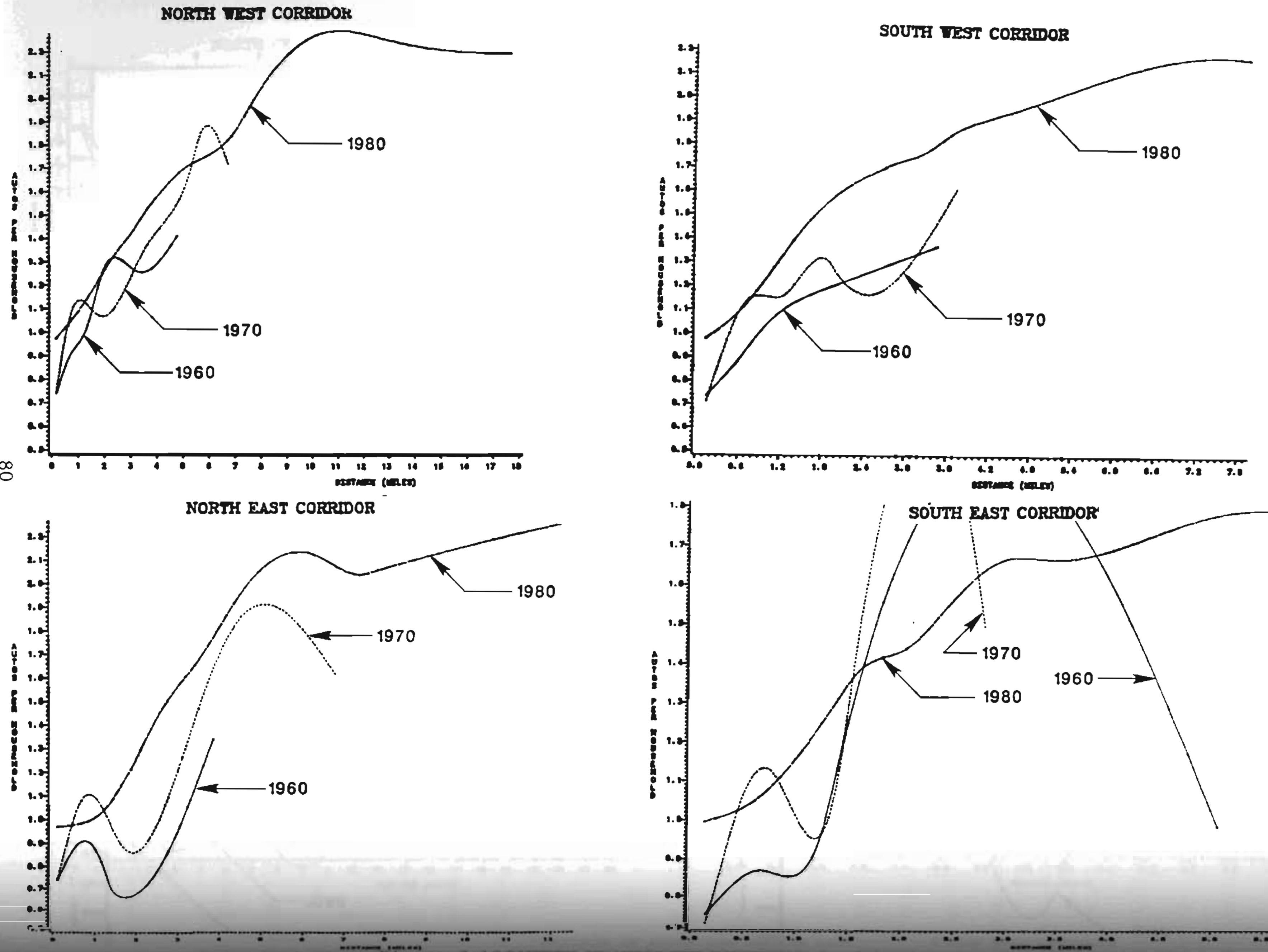


Figure 3.10 PATTERN OF AUTOS PER HOUSEHOLD (cont.)

the absence of the spatial patterns of the location of employment and other activities, one can surmise from looking at the patterns of residence-based variables examined herein that average trip distances are likely to continue to increase, and that congestion levels in the densifying "suburban" subareas will reach levels typically associated with CBD traffic systems. However, suburban networks generally exhibit somewhat different characteristics from CBD systems, in terms of physical features, traffic control and adjoining land use and its generation of traffic-interfering activities. The subsequent chapters in this report shed further light on how congestion and the associated quality of traffic service depend on such features of the traffic networks.

Several directions for meaningful additional work are suggested by the results presented in this chapter. First, as noted earlier, it would be extremely valuable to consider the corresponding spatial patterns of the location of firms and businesses of different types. This can be accomplished using data sources such as industrial and business directories that give information at the firm level. Second, a larger set of case areas should be considered in order to further assess the robustness of some of the patterns seen here, as well as to possibly classify areas in terms of their respective spatial and temporal patterns. In particular, it would be interesting to examine areas such as San Francisco and Boston that have relatively more elaborate spatial patterns that have developed over longer time frames than many other U.S. cities. Third, it would be interesting to capture the characteristics of the transportation network serving the areas under consideration by examining the densities as a function of network distance or travel time instead of Euclidian distances from the CBD. Finally, it is desirable to integrate the above spatial patterns into an approach to estimate flow patterns and aggregate travel descriptors (such as average trip lengths) resulting from the interaction between the spatial patterns of residence-based variables seen in this chapter, and those of the location of firms and businesses. This would provide an important linkage to the study of the transport implications of evolving urban structures. Such an approach could further incorporate the characteristics of the transport infrastructure. Preliminary analytic work along these lines, conducted in conjunction with this study, has been promising; it involves the derivation of intraurban flow patterns and average trip lengths under assumed spatial distributions of origins and destinations, for a given generic type of network.

CHAPTER 4

PERFORMANCE OF TRAFFIC NETWORKS: FUNDAMENTAL RELATIONS

4.1. INTRODUCTION

The previous chapters have examined transport from the demand side, in terms of those activities that generate the flow patterns in a given metropolitan area. This chapter is concerned more directly with the actual flows in the network, and particularly with describing the performance of an urban traffic network, in terms of the quality of traffic service that is or that could be offered by that network under different demand levels. In addition to developing relations among key network-level performance descriptors, the intent is to gain an understanding, and eventually a quantitative description, of how this performance is related to the various features of the network and of the urban area that it serves. This would then provide an operational tool for examining the implications on the quality of traffic service, and thus on the degree of mobility afforded by the traffic system, of the various predicted or proposed scenarios regarding the evolution of urban structure and associated travel patterns. The focus of the present work is limited to urban street networks, and does not include freeways; extensions in that direction constitute an important future step for this research.

Significant progress has been achieved over the past three decades in terms of characterizing and modelling fundamental traffic phenomena in arterials and intersections. Considerably less effort has been directed towards the development of a network-level traffic theory to characterize the performance of urban street networks. Macroscopic models of traffic flow in urban networks describe the behavior of and interrelation among traffic variables defined at the network level. Such characterization of traffic has important practical implications in terms of measuring the quality of traffic service in a network, and performing related diagnosis, monitoring (over time and space) and evaluation activities

(Ardekani and Herman, 1982; Mahmassani, Williams and Herman, 1984), particularly as these pertain to urban growth and development.

The complex interactions in a traffic network effectively preclude the analytic derivation of macroscopic network-level relationships from the basic principles governing microscopic traffic behavior or from link-level macroscopic models. In addition to questions regarding the very definition of the networkwide traffic descriptors, the existence of meaningful relations among these quantities cannot, a priori, be taken for granted. Some of these questions have been addressed in earlier work by Mahmassani *et al.* (1984) and Williams, Mahmassani and Herman (1985), providing definitions for average speed, concentration and flow, all at the network level, and indicating, on the basis of simulation experiments, that these quantities appear to be related in a manner not unlike their counterparts at the individual road or arterial level .

In the past five years, a network-level modelling approach has developed around Herman and Prigogine's two-fluid theory of town traffic (Herman and Prigogine, 1979; Herman and Ardekani, 1984), which views traffic in a street network (excluding freeway-type facilities) as consisting of two "fluids": moving (or running) vehicles versus stopped vehicles (though still in the traffic stream, such as at traffic signals, as opposed to parked vehicles). The theory postulates that $V_r = V_m (f_r)^n$, where V_r is the average speed of the moving vehicles, f_r the fraction of running vehicles in the network, and V_m and n are two parameters which describe the quality of traffic service in the system. The theory leads to a relation between the average running time per unit distance (i.e. the conditional expected trip time per unit distance taken only over moving vehicles) and the average total trip time per unit distance in the network. Field validation has been conducted, using data gathered by the chase-car technique, and, to a very limited extent, aerial photography, under operating conditions found in actual city networks (Ardekani and Herman, 1982; Herman and Ardekani, 1984; Ardekani, Torres-Verdin and Herman, 1985). The sensitivity of the model's parameters to various physical and operational network characteristics has also been explored by Williams *et al.* (1985) using (microscopic) simulation experiments.

The latter approach actually circumvents one of the major obstacles hindering the development of network-level models, namely the cost and difficulty of obtaining reliable data at the network level. The feasibility of using microscopic simulation as a tool to investigate network flow relations has been established in previous work by Mahmassani *et al.* (1984) and Williams *et al.* (1985), on the basis of the NETSIM simulation package. In addition to the obvious cost and resource considerations, this approach allows one a degree of experimental control that is not practical in actual traffic systems, as well as the ability to explore a wider range of situations than can be observed in field work.

Recently, Ardekani and Herman (1987) extended the two-fluid modelling framework to include a set of relations between the principal network traffic variables. These relations are derived from a postulated functional relation between the average fraction of vehicles stopped in a network and the prevailing concentration, slightly modifying an earlier form suggested in Herman and Prigogine's original work (1979). A speed-concentration model can then be derived from the postulated relation, given that the two-fluid assumptions hold. An exploratory calibration of the parameters of the postulated model was also reported, though only very limited aerial photographic data was available for this purpose.

Two principal objectives are adressed in this effort. First, we seek a system of relations that comprehensively describe the joint behavior of the average speed, flow and concentration, as well as the average fraction of vehicles stopped in the network, and the two-fluid stopped and running time variables. Three alternative sets of interrelated models, each (set) having a different starting postulate, are presented and tested against a series of microscopic simulation runs corresponding to different network concentration levels. Second, we analyze the sensitivity of these relations and their parameters to the network's characteristics, particularly its topological features and prevailing control strategies. This second objective is a modest step towards the more ambitious goal, stated earlier, of relating the quality of traffic service and associated mobility level, to both network features and spatial characteristics of urban activities. Both objectives are supported by microscopic simulation experiments, which provide useful macroscopic insights into network traffic behavior.

In the next section, we present the theoretical fundamentals underlying this work, including the definitions of the principal variables, and the derivation of the flow models from a given postulated relation. In section 4.3, we review the methodological approach followed in the supporting simulation experiments. The three sets of models are then presented, in turn, in section 4.4, and their parameters are calibrated using the results of the simulations, allowing the evaluation of their ability to provide a macroscopic characterization of network traffic flow phenomena. The second objective, namely the sensitivity of these relations and their parameters to network features, is addressed in the next chapter.

4.2. THEORETICAL BACKGROUND

In this section, we first present the networkwide variable definitions, then conceptually describe the principal relations comprising each set (system) of models, as well as the analytical steps involved in the derivation of the models from a starting functional form relating any two variables.

4.2.1. Definition of Network Traffic Variables

The three fundamental traffic variables, speed, concentration, and flow have been generalized in previous work to the network level (Mahmassani *et al.*, 1984). All three are defined as average quantities taken over all vehicles in the network over some observation period τ. Average speed V (in mph) is defined as the ratio of total vehicle-miles to total vehicle-hours in the network during the period τ, i.e. $\Sigma x_i/\Sigma t_i$, where x_i and t_i are the distance travelled and the time spent in the network, respectively, by vehicle i, and the summations are taken over all vehicles circulating in the network during the observation period.

The average concentration K (in vehicles per lane-mile), for the same period, is the time average of the number of vehicles per unit lane-length in the system. Letting N(t) denote the number of vehicles in the system at time t, and L the total lane-miles of roadway, then

$$K = (1/t) \int_{t_0}^{t_0+t} [N(t) / L] \, dt,$$

where t_0 is the beginning of the observation period. Effectively, K can be simply calculated by dividing the total vehicle-hours (during t) by tL, where L is the total lane-miles of roadway. One of the key advantages of using simulation experiments is that one can maintain a constant concentration level in the network, by keeping the number of circulating vehicles constant over the period t. This strategy has been adopted in all the experiments discussed herein. The concentration is then simply equal to the known total number of vehicles in the network divided by L.

The average network flow Q is interpreted as the average number of vehicles that pass through a random point of the network, and given by $(\Sigma\, l_i\, q_i) / (\Sigma\, l_i)$, where q_i and l_i respectively denote the average flow (during t) and the length of link i, and the summations are taken over all network links.

Another key variable of interest is f_s, the average fraction of stopped vehicles over the observation period. It is an important descriptor of the productivity of an urban traffic network, and is suggested by the dichotomization of traffic in the network into moving and stopped vehicles. The variation of f_s with the prevailing network concentration K is one of the principal relations considered here, as seen later in this section. Other two-fluid variables of interest are the average running time T_r, stopped time T_s and total trip time T, all per unit distance; of course, $T=T_r+T_s$. The main result of the two-fluid theory is a model relating T_r to T (or T_s), that is invoked in all the following derivations. A detailed presentation of the two-fluid model assumptions and derivations will not be given here as it can be found elsewhere (Herman and Prigogine, 1979; Herman and Ardekani, 1984); the results necessary for the derivations in this chapter are presented as needed. Note one pertinent result used in our simulations to calculate f_s, namely that $f_s = T_s/T$, derived mathematically by Herman and Ardekani (1984) under steady-state conditions, and verified numerically in earlier simulation experiments.

4.2.2. The Traffic Models and Their Interrelation

The characterization of traffic network performance is provided by a set of relations among the variables V,K,Q and f_s, in addition to the two-fluid relation between T, T_r and T_s, which has the following form:

$$T_r = T_m^{[1/n+1]} T^{[n/n+1]} \qquad (1)$$

The parameter T_m is simply the inverse of V_m, interpreted as the average maximum running speed in the network (i.e. without any stopping), whereas the parameter n captures the sensitivity of running speed to the fraction of vehicles stopped in the network, and can serve as an indicator of quality of traffic service in a network (Herman and Ardekani, 1984).

The other relations comprising each model system are:

$$V = \mathbf{v}(K) \qquad\qquad f_s = \mathbf{f}(K)$$

$$Q = \mathbf{q_1}(K), \qquad \text{and} \qquad Q = \mathbf{q_2}(V).$$

Given a mathematical form for any one of the above relations, the functional forms of the remaining relations can be analytically derived by invoking one, or both, of the following:

1) the identity Q=KV, known to hold for individual roads, has been numerically verified, in earlier simulation work (Mahmassani *et al.*, 1984), to hold at the network level as well. Further evidence to this effect is also offered in this chapter for a more extensive set of simulations.

2) The two-fluid model, which effectively provides a relation between f_s and V, since T=1/V (by definition) and $f_s = T_s / T$. In this work, we assume the two-fluid model to hold, though comments about the range of applicability of this assumption are also given on the basis of the simulation results.

In each of the three model systems discussed in this paper, a different functional form will be postulated for either the speed-concentration relation **v**(K) , or for the fraction of vehicles stopped-concentration relation **f**(K). The functional form for the other relation is then derived from the postulated model, by invoking the two-fluid postulate, which states that:

$$V_r = V_m (1-f_s)^n , \qquad (2)$$

thereby specifying the dependence between V_r, the average speed of the moving vehicles in the network, and the average fraction of cars that are moving, over the observation period, $f_r = 1-f_s$.

The functions **v**(K), **f**(K), $\mathbf{q_1}$(K) and $\mathbf{q_2}$(V) are different in each of the three model systems, as they depend on the starting assumption regarding either **v**(K) or **f**(K). Then, since $V = V_r f_r = V_r (1-f_s)$ (by definition of the corresponding averages), we obtain, using Eq. 2:

$$V = V_m (1-f_s)^{n+1}. \qquad (3)$$

Then, if **f**(K) is given, substituting it for f_s in Eq. 3 yields **v**(K), as follows:

$$V = V_m (1-\mathbf{f}(K))^{n+1} . \qquad (4)$$

Similarly, if **v**(K) is given, substituting it for V in Eq. 3 allows the derivation of a functional form for **f**(K):

$$\mathbf{v}(K) = V_m (1-f_s)^{n+1}, \qquad (5)$$

which can be rewritten as:

$$f_s = 1 - (\mathbf{v}(K)/V_m)^{[1/n+1]} . \qquad (6)$$

The functions $\mathbf{q_1}$(K) and $\mathbf{q_2}$(V) can then be easily obtained by substituting **v**(K) into the identity Q = KV. In the first set of models discussed in this chapter, Ardekani and

Herman's (1987) postulated form for **f**(K) provides the starting point. The other two model systems are based on two alternative functional forms for **v**(K), both of which have been used in conjunction with individual streets and highways, as discussed in section 4.4. Before discussing the specific functional forms for each of the three model systems, the simulation experiments conducted to calibrate the models are described.

4.3. THE SIMULATION EXPERIMENTS

In this section, we first present a general description of the simulation experiments, including the common basic network configuration used. Next, we give pertinent information about the individual simulation runs. Additional detail on all these aspects can be found in Williams (1986).

4.3.1. General Description

A series of six simulation runs was performed to evaluate the network flow models of interest; several simulation experiments were also conducted to examine the network's performance in response to the various network features of interest, described in the next chapter. Each of these runs used the same basic network configuration and features, but involved a different vehicular concentration level. As noted earlier, the network is treated as a closed system, with a fixed number of vehicles maintaining a constant concentration throughout the observation (accumulation) period. The circulation of the vehicles is controlled by an elaborate set of detailed microscopic rules (governing car-following, queue discharge, lane switching, gap acceptance, and so on) embedded in the NETSIM package (PMM, 1973; FHWA, 1980). Note that some modification of the code was necessary in order to keep track of stopped versus running time in a manner that is consistent with the two-fluid definitions of these variables (see Williams [1986] for additional detail on these modifications).

The basic network configuration consists of 25 nodes, arranged in a 5-node by 5-node square, connected by two-way, four lane streets forming a regular, CBD-like grid, as shown in Fig. 4.1. In designing this network, we sought a degree of regularity that would not interfere nor mask the effect of the principal features investigated in the simulations. Each two-way street is represented by a pair of directed (one-way) links (80 in all). All links used in these simulations are 400-ft long, with no special geometric features. Vehicles are injected into the network via 12 additional entry links placed around the perimeter, three to a side, with each entry link connecting a source node (not part of the 25 network nodes, and numbered 801-812 in Fig. 4.1) to a non-corner boundary node.

Each network node corresponds to a signalized intersection, except for the four corner nodes, which are unsignalized since they present no conflicts to the drivers. Timing of signals at the interior nodes followed a two-phase timing scheme with a 50-50 split and no protected turning movements. Three-phase signals were used at the boundary nodes, providing a protected left turn for vehicles reentering the interior of the system, with the green time nearly equally allocated to vehicles leaving the interior of the network and those reentering it from the boundary. Two-way progression at the mean desired speed (35 mph in this case) was provided along the interior arterials by using 40-sec cycle length with single alternate operation.

Once the desired number of vehicles entered the network, they were not allowed to exit (no sink nodes were specified),thereby maintaining the concentration constant. At the interior intersections, 10% of the traffic turned left, 15% turned right, and the remaining 75% continued straight through. When vehicles reached a boundary node, they split equally between left and right turns, except at the corner nodes, where no choice is available. These traffic movements were used throughout the work reported herein. However, it should be recognized that the circulation pattern can itself be an important determinant of the network's performance. Results to that effect, using different turning percentages, are reported in Mahmassani et al. (1984).

Essentially, each simulation run yields one data point, consisting of values of the desired network-level average quantitites taken over the observation period (generally about

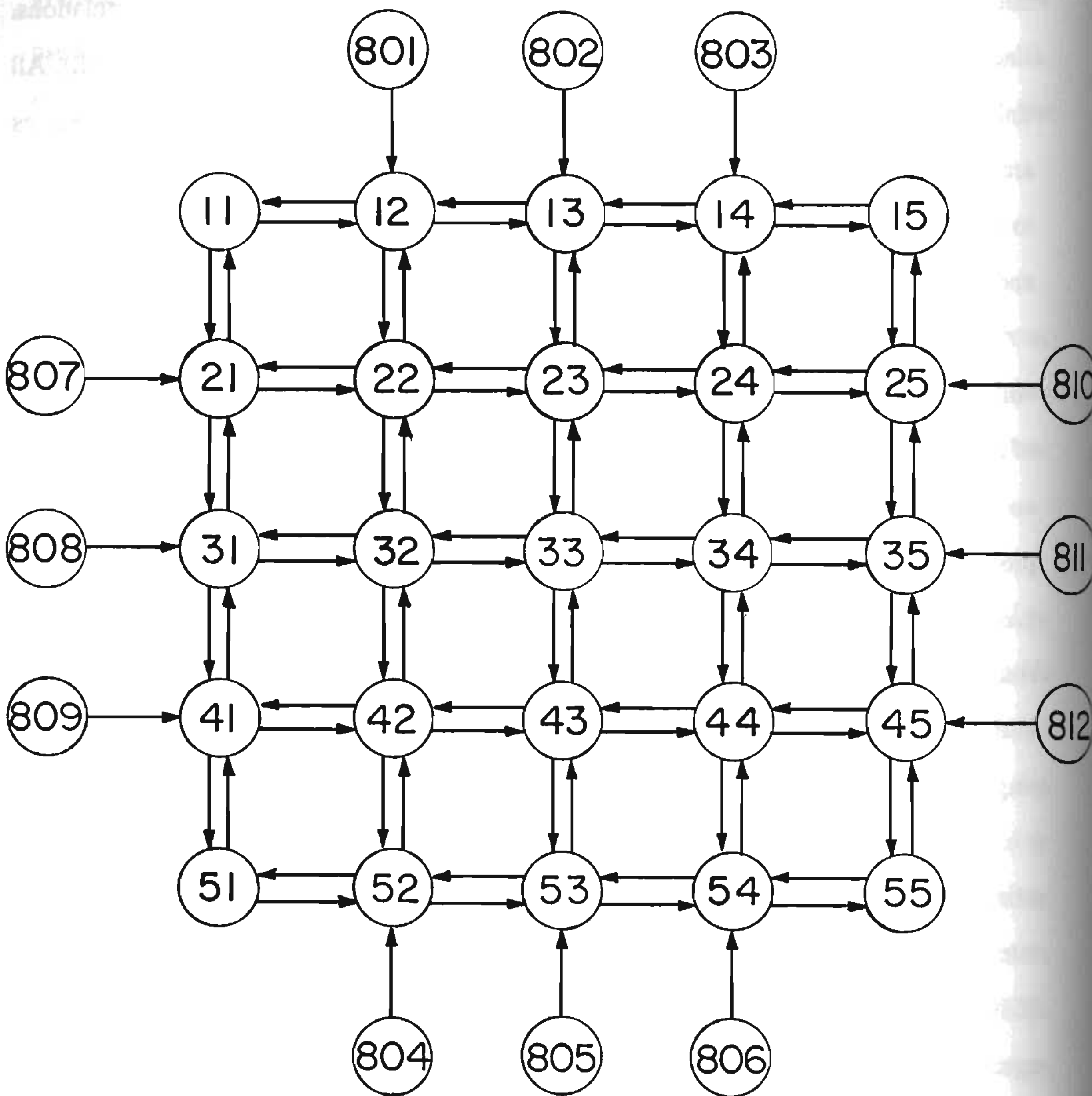


FIGURE 4.1: NETWORK CONFIGURATION. NOTE THAT NODES 801 TO 812 ARE SOURCE NODES AND AS SUCH ARE NOT PART OF THE CIRCULATION NETWORK.

15 minutes of accumulation, following an initial startup and loading period ranging from 5 to 15 minutes, as described below. The network concentrations in these runs are approximately 10, 20, 40, 60, 80 and 100 vehicles/lane-mile (the actual values, determined by simulation input practicalities, are 9.90, 19.80, 41.58, 61.38, 81.18, and 100.65 vehicles/lane-mile), representing a range from very light to extremely heavy traffic. For comparison purposes, the highest concentration observed in field work by Ardekani (1984) was approximately 30 vehicles/lane-mile in the Austin CBD during heavy traffic conditions. For all practical purposes, network concentrations in excess of 50 vehicles/lane-mile can be considered as very high, and are rather unlikely in actual operations. Therefore, the simulation results allow us to explore network behavior over a wide spectrum of concentrations, including conditions near saturation, which may not be easy to observe in actual street networks. The specifics of the simulation experiments conducted to examine the effect of network features on system performance are discussed in conjunction with the corresponding factor of interest.

4.3.2. Individual Runs

Individual runs consisted of three periods: the loading, network stabilization, and observation (or accumulation) periods. The first, the loading period, is necessary since each run begins with an empty network. As mentioned earlier, vehicles were loaded via the entry nodes (numbered 801-812 in figure 4.1), and the time required for this period depended on the desired network concentration for the particular run and the capacity of the signalized intersection where the vehicles actually entered the network (nodes 12-14, 21, 25, 31, 35, 41, 45, and 52-54). The loading periods in the simulations reported in this and the folowing chapters ranged from five to thirty-five minutes.

The second period, network stabilization, allows the vehicles to distribute themselves over the network once no additional vehicles are generated. Two primary methods were used to determine network stability: (1) flow, averaged separately over the boundary and interior links, and (2) the fraction of vehicles (and, thus, the relative concentration) on the boundary and the interior of the network. Average flows, which were found by averaging the one-minute vehicle discharges from each link separately for the boundary and interior links, are shown in Fig. 4.2 for an illustrative 35-minute run. During the loading period (first five minutes), the average flow for both the interior and boundary links rapidly increases. The flow rates continue to increase for a couple of minutes after the loading period as the vehicles queued on the entry links enter the network. However, after eight minutes, (three minutes after the end of network loading), flow fluctuations, which are still present, occur around apparently stable values. While Fig. 4.2 shows the average flows for the interior and boundary links to be about the same, this was not the case in all runs, although the largest difference was less than 15 percent. The network was considered stable when both the average boundary and interior flows stabilized about particular values for each run (not necessarily the same values for both).

The second criterion for network stability is based on the distribution of vehicles over the boundary and interior links. Since the vehicles were not allowed to turn onto the boundary as they entered the network during loading, none of the vehicles in the network were on the boundary during the first couple of minutes of simulation. However, after circulating for eight minutes (three minutes after loading), according the the rules outlined above, the respective fraction of vehicles on the boundary and interior links appeared to stabilize around certain values. Results from a typical twenty-minute run are shown in Fig. 4.3, where the vehicle fractions appear to stabilize around 0.4 for the boundary links, and 0.6 for the interior links after the first eight minutes of simulation. Interestingly, these particular vehicle fractions closely match the ratios of the lengths of the boundary and the interior links (40 and 60%, respectively), resulting in about the same concentration on both network components. The network was considered "stable" when the vehicle fractions fluctuated about constant values (not necessarily resulting in the same concentration for the boundary and interior). However, the vehicle fractions did not deviate substantially from 0.6 and 0.4 across the various simulation runs considered in this study.

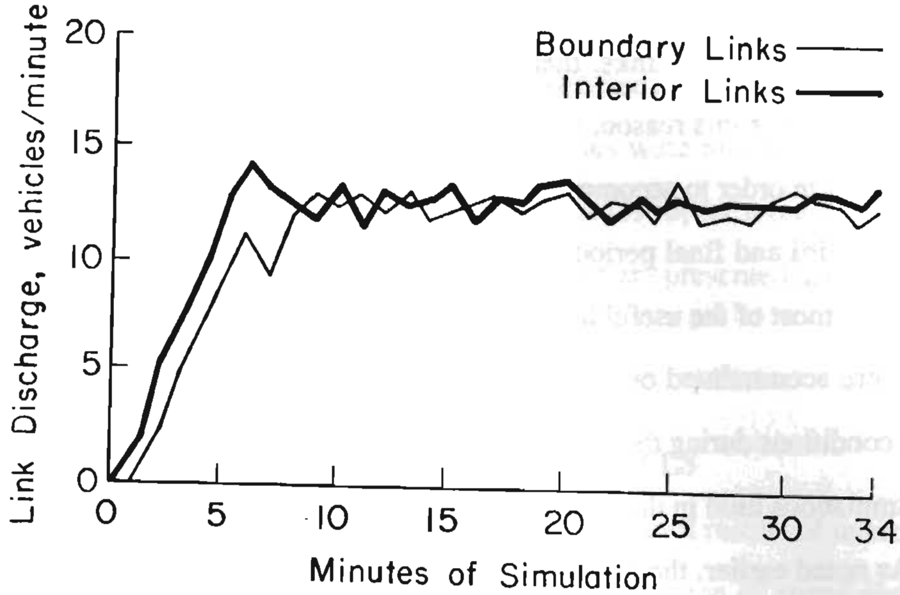


Figure 4.2. Link Discharge for Interior and Boundary Links

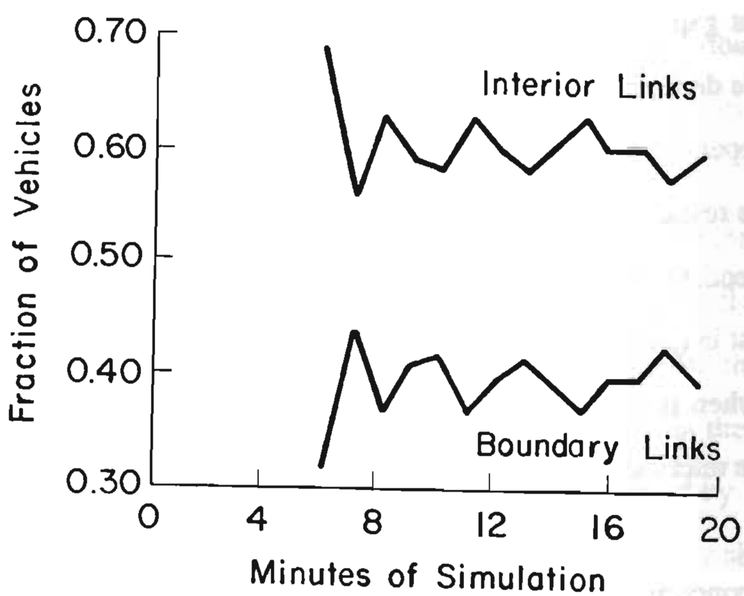


Figure 4.3. Fraction of Vehicles on Boundary and Interior Links

A period of three minutes after the loading stage was found to be adequate for network stabilization under a wide range of concentrations, so long as long queues were not built up on the entry links, thereby requiring long times to clear after the network loading period. For this reason, vehicle entry was spread over long loading times at high concentrations in order to accommodate the large number of vehicles entering the network.

The third and final period of each individual run is the observation period per se, during which most of the useful information is generated. All the network-level variables of interest were accumulated or averaged only over this third period, thereby excluding the transient conditions during the loading and stabilization periods. The duration of this period for the simulations used in this investigation ranged between 12 and 15 minutes.

As noted earlier, the simulation package used in these experiments treats many of the microscopic processes stochastically. In particular, a driver characteristic code is assigned randomly (from a pre-specified, discretized approximately normal distribution) to each vehicle as it enters the network, and determines the degree to which each vehicle (driver) is "aggressive" or "timid", as reflected in the parameter values affecting decision situations such as gap acceptance and reaction times. The asssignment of the driver characteristic code depends on a string of random numbers generated within the program, which, in turn, depends on the random number seed selected by the package user. In the early stages of the research, several otherwise identical runs were made, varying only the random number seed, to determine the variability of the networkwide time (and ensemble) averages of interest in this investigation. The results clearly confirmed that the variability in these averages (when taken over an observation period of 12 to 15 minutes) due to the stochasticity in the microscopic processes is insignificant and negligible for the purpose of this analysis.

Another concern is whether individual vehicles truly experience the traffic conditions in the entire network, or get trapped in a small sector of it. The turning movement rules, which govern the circulation pattern, were specified so as to allow each vehicle to circulate throughout the network. It was thus desirable to verify that this was indeed taking place, and that our experimental set-up was not leading to pathological behavior. This was accomplished by examining the detailed itineraries followed by several individual vehicles; the results confirmed that vehicles were allowed to circulate in the network, with no evidence of pathologies (see Williams [1986] for more detail).

In the next section, the results of the simulations are presented and used to address the network-level traffic relations of concern to this study.

4.4. ANALYSIS OF TRAFFIC NETWORK MODELS

Figures 4.4 through 4.7 present the basic network-level results of interest from the above runs; Figs. 4.4 and 4.5 respectively depict the variation of speed and flow with concentration, Fig. 4.6 depicts the corresponding variation of speed with flow in these runs, while Fig. 4.7 depicts the variation of the fraction of vehicles stopped with concentration. The general similarity of the K-V-Q patterns to their characteristic counterparts for individual road sections is striking; this observation forms the basis for the functional forms explored in this section. The results of the simulations are first examined to verify that the Q=KV identity holds for the network flow under consideration.

4.4.1. Verification of the Q=KV Identity

In earlier simulation experiments, the basic traffic flow identity relating speed, flow, and concentration was found to hold at the network level as well (Mahmassani *et al.*, 1984). Further evidence to that effect is also offered here for the new simulations (including those used in the analysis of network performance in the next chapter). The average flow Q (in vehicles/lane-hour) for each run is estimated by averaging, over the accumulation period, the minute-by-minute average flows taken over all links, based on the number of vehicles discharged from the downstream end of each link (each minute). Each link's discharge is first divided by two to obtain the flow on a per lane basis. Note with regard to the expression given earlier in this section for Q that the link length l_i is the same

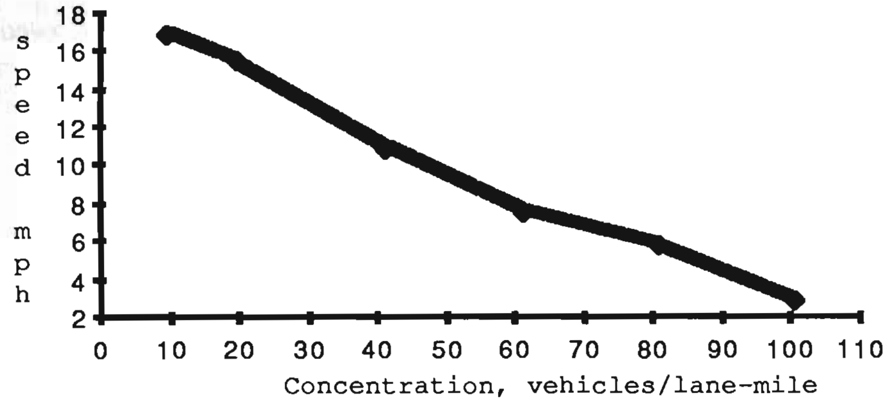


FIGURE 4.4. Plot of Speed vs. Concentration for Simulation Experiments

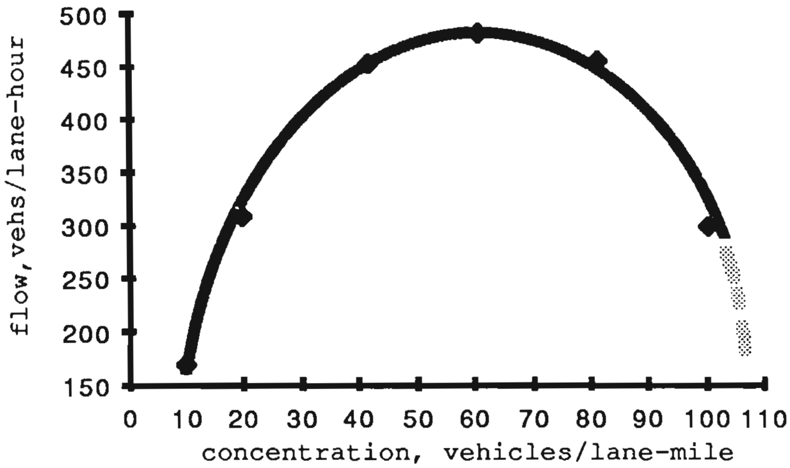


FIGURE 4.5. Plot of Flow vs. Concentration for Simulation Experiments

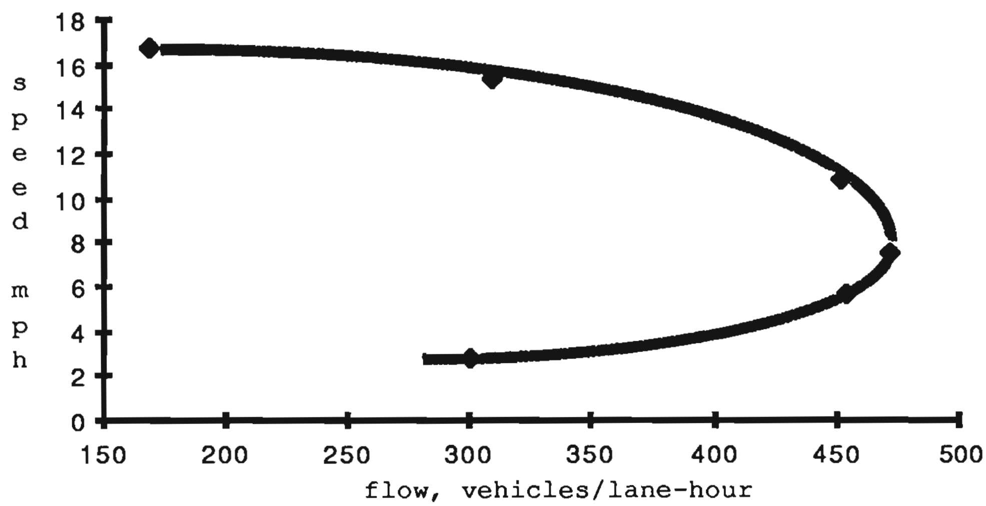


FIGURE 4.6. Plot of Speed vs. Flow for Simulation Experiments

across all links in these experiments. The other network variables, K and V, are obtained as described earlier in this section. The product KV can then be compared to the independently estimated Q.

The results of this comparison for the six runs are shown in Table 4.1, clearly illustrating the closeness of the two quantities. The small numerical discrepancies are mostly due to the one-minute discretization used in estimating Q, whereas V is determined from quantities accumulated over smaller time steps. The largest differences occur at the high concentration levels, where network operation is least stable. In addition, the independently calculated Q is plotted in Fig. 4.8 against the corresponding product KV for all the simulations conducted in connection with this study, clearly indicating a linear trend with a slope of 1. These results, coupled with earlier similar results (Mahmassani *et al.*, 1984), further confirm the validity of Q=KV at the network level as well.

Next, the three model systems are presented and discussed on the basis of their calibrated performance against the observations generated by the simulations.

4.4.2. Network Flow Models

The first set of models discussed is that proposed by Ardekani and Herman (1987), which starts with a postulated functional form for the f_s-K function **f**(K). The starting point for the other two sets of models is a postulated functional form for the network-level speed-concentration relation **v**(K). Given the similarity of the plots in Figs. 4.4 through 4.6 to the patterns encountered in individual road facilities, two well-known functional forms for the speed-concentration relation at the link level are considered here for the network-level variables. The first is Greenshields' widely utilized linear speed-concentration relation, while the second is the nonlinear 'bell-shaped' function, originally proposed by Drake, Schofer and May (1967) for arterials (see Gerlough and Huber (1975) for an overview of these models).

4.4.2.1. Model System 1

The starting point for this model system is the following postulated functional form for **f**(K), which specifies the f_s-K relation:

$$f_s = f_{s,min} + (1-f_{s,min})\,(K/K_j)^{\pi} \qquad (7)$$

where $f_{s,min}$ is a parameter intended as the minimum fraction of vehicles stopped in the network,

K_j is a parameter intended as the "jam" concentration, at which the network is effectively saturated, and

π is a parameter that determines the fraction of vehicles stopped at a given partial concentration (K/K_j), and could serve as a measure of the quality of service in a network.

The principal assumptions of this relation are reflected in its boundary conditions. In particular, it recognizes that in an urban street network, even under very low concentrations, there is some non-zero $f_{s,min}$ fraction of stoppage that is inevitable in the network (unless traffic control is fully responsive, unlike any currently in operation, or drivers do not obey traffic laws), as suggested by simulation results (see Fig. 4.7, for example) and field studies (Ardekani and Herman, 1987). On the other hand, f_s goes to 1, meaning that all vehicles are stopped, as K goes to K_j. Thus, Eq. 7 states that f_s is an increasing function of K, that varies in the range from $f_{s,min}$ to 1 as K goes from zero to K_j.

Following the steps presented in section 4.2.2, **v**(K), which specifies the speed-concentration relation, can then be found by substituting Eq. 7 into Eq. 4, yielding:

$$V = V_m\,(1-f_{s,min})^{n+1}\,(1 - (K/K_j)^{\pi})^{n+1}. \qquad (8)$$

The flow-concentration relation $\mathbf{q_1}(K)$ is then easily found using Q=KV, as shown in the previous section:

$$Q = KV_m\,(1-f_{s,min})^{n+1}\,(1 - (K/K_j)^{\pi})^{n+1}. \qquad (9)$$

Model Estimation. The estimation of the parameters of Eq. 7 can be performed by rewriting it as $f_s = a + bK^{\pi}$, where $a = f_{s,min}$ and $b = (1-f_{s,min})/K_j^{\pi}$. Nonlinear least squares estimates can then be obtained for a, b and π, from which the original parameters can be recovered, yielding for the set of observations generated from the simulation experiments estimated values of 0.187, 134.12 vehicles/lane-mile, and 0.208 for $f_{s,min}$, K_j, and π, respectively. Figure 4.9 depicts the curve corresponding to the estimated parameter values, and also includes the points observed in the simulation experiments for comparison purposes. It appears that the estimated model somewhat overestimates $f_{s,min}$. More seriously, between (the practically meaningful) concentrations of 20 and 80 vehicles/lane-mile, it predicts the opposite concavity than that suggested by the observed values.

The ability of this model to describe the speed-concentration pattern can be assessed by comparing the resulting Eq. 8 to the observed simulation results. This requires estimates of the two-fluid parameters n and T_m, which can be obtained by ordinary least squares estimation of the linear regression equation obtained by taking the natural logarithms of both sides of Eq. 1 (Herman and Ardekani, 1984):

$$\ln T_r = (1/n+1) \ln T_m + (n/n+1) \ln T. \quad (10)$$

Using all six data points, the estimates of the two-fluid parameters n and T_m were found to be 1.051 and 2.692 min/mile, respectively. However, it is important to recognize that this analysis assumes that the two-fluid model continues to hold at very high concentrations, well in excess of those encountered in previous field studies. Moreover, the reliability of the simulation at very high concentrations, where conditions are inherently unstable, cannot be taken for granted. The plot of the $\ln T_r$ vs. $\ln T$ trend (linear under the two-fluid assumptions) for the six observed points (Fig. 4.10) suggests that the sixth point (representing the highest concentration examined) deviates from the linear trend established by the other five points. However, the simulation results are clearly in agreement with the two-fluid assumptions over the range of practically meaningful concentrations. Hence, n and T_m were re-estimated, omitting the sixth point, yielding 2.349 and 1.809 min/mile, respectively.

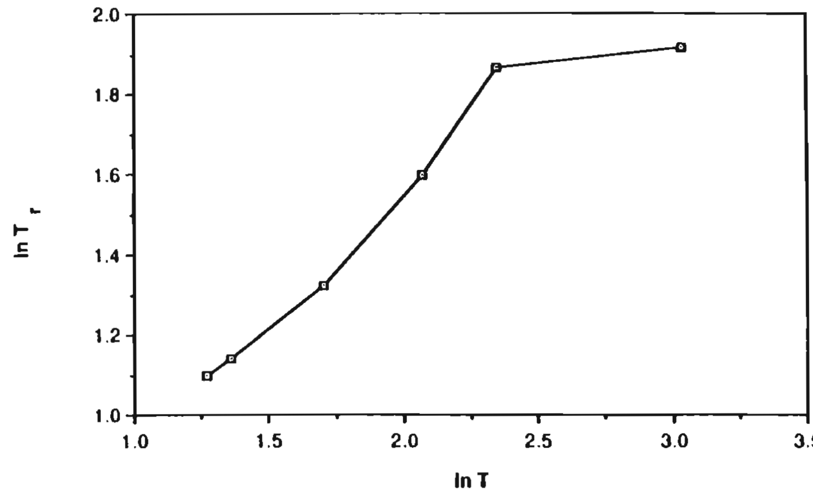


Figure 4.10. Plot of ln T_r vs. ln T for Simulation Experiments.

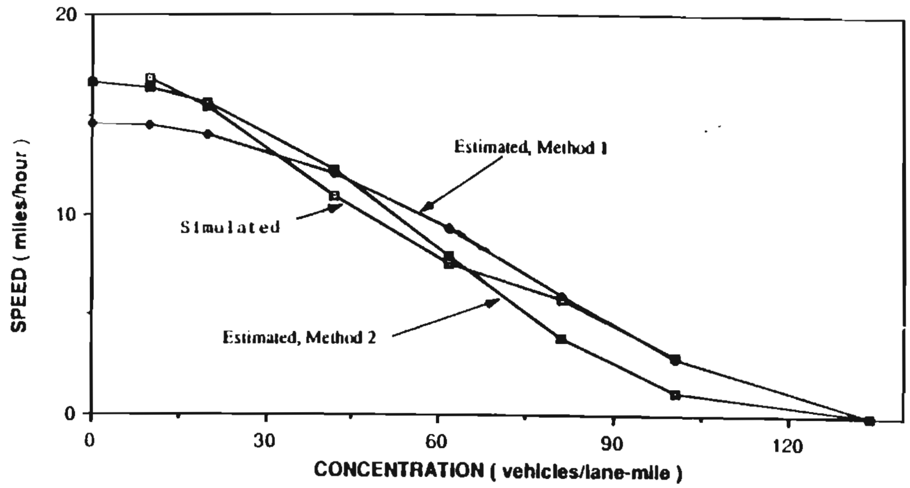


Figure 4.11. Comparison of V-K Model Derived in Model System 1 with Observed Simulation Results.

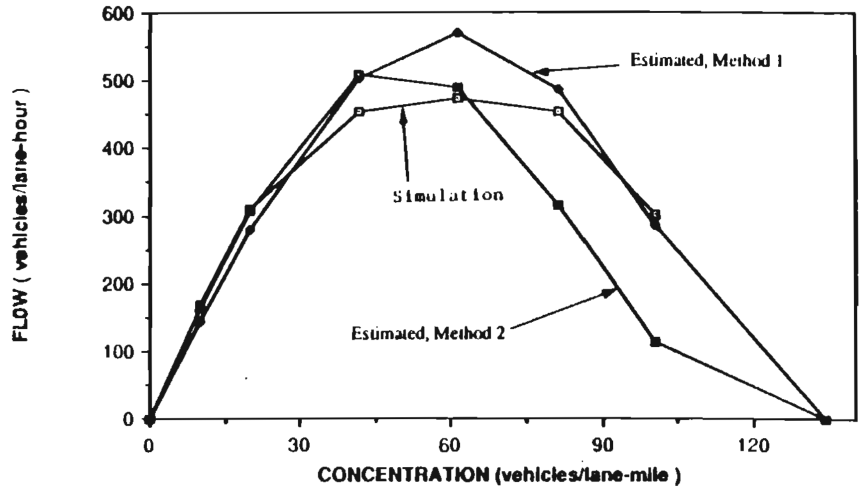


Figure 4.12. Comparison of Q-K Model Derived in Model System 1 with Observed Simulation Results.

models). Greenshields' function, which forms the basis of this set of models, has the following well-known form:

$$V = V_f (1 - K/K_j), \tag{11}$$

where V_f and K_j are parameters to be estimated, and respectively interpreted as the free mean speed, which is that experienced when interference from other vehicles in the traffic stream is virtually nonexistent (though vehicles must still obey traffic controls in the system), and the 'jam' concentration, defined earlier.

Note that the V_f parameter is distinct from the two-fluid parameter V_m (which is equal to $1/T_m$), the average maximum running speed, which occurs when $f_s = 0$, i.e. when no vehicles are stopped in the network. As stated in the previous subsection, this condition does not occur in existing urban traffic systems over any meaningful length of time, even at very low concentrations, due to the presence of the traffic control system. Therefore, because the average travel time corresponding to the free mean speed will exceed the minimum running time T_m by some usually non-zero stopped time, we will always have $V_f \leq V_m$, and in most cases $V_f < V_m$. Note that in the first model system discussed above, $V_f = V_m (1-f_{s,min})^{n+1}$ (by setting K=0 in Eq. 8).

The functional form for f(K), the f_s-K relation that is compatible with the above linear V-K model and the two-fluid model, is found following the steps presented earlier. Namely, by substituting Eq. 11 into Eq. 6, we obtain:

$$f_s = 1 - [(V_f/V_m)(1 - K/K_j)]^{1/n+1}. \tag{12}$$

This function has similar boundary conditions as Ardekani and Herman's, since $f_s=1-(V_f/V_m)^{1/n+1}$ for K=0, and $f_s = 0$ for $K=K_j$. However, here the minimum fraction of vehicles stopped, $f_{s,min}$, is explicitly stated in terms of the speed parameters V_f and V_m. For instance, it becomes clear in this expression that if $V_f = V_m$, then $f_s = 0$ at $K = 0$.

Finally, the flow-concentration relation is again obtained by applying Q = KV:

$$Q = V_f (K - K^2/K_j). \tag{13}$$

Model Estimation. Least squares estimates of the model parameters can be easily obtained since Eq. 11 is linear in K. Using the observed simulation data of Fig. 4.4, the estimated values are 18.02 mph and 116.3 vehicles/lane-mile for V_f and K_j respectively. The resulting calibrated functions, along with the observed values, are plotted in Figs. 4.13 and 4.14, for Eqs. 11 and 13 respectively, exhibiting in both cases a rather close fit to the data. Moreover, the free mean speed does not appear to be underpredicted; however, because Eq. 11 ignores the slight nonlinear trend exhibited by the K-V data, the speed is slightly overpredicted in the middle concentration range, resulting in an overprediction of the flow in the same range (Fig. 4.14).

Note that there is no need to exclude the sixth data point, for the highest K (as in 'method 2' in the previous section), in estimating the parameters of Eq. 11, because the postulated speed-concentration and resulting flow-concentration relations appear to be consistent with the pattern exhibited by the data, including the range corresponding to that high concentration point. The main justification for excluding that point in the previous section is that the discrepancy appears with regard to the two-fluid model assumption's applicability at higher concentrations. Since v(K) has been specified directly here, this assumption has no direct bearing on the parameter estimates presented so far in this subsection. This is however not the case when Eq. 12 is evaluated for the f_s-K relation, since this function's parameters are not estimated directly on the f_s-K data, but from the two-fluid and V-K models separately (just as the parameters of the V-K relation in the first model system were not directly estimated using the corresponding V-K data), as discussed hereafter.

Assessment of the f_s-K function in this model system requires the estimation of the two-fluid parameters, n and T_m. Those found by methods 1 and 2 in the previous section are both used here. The results are plotted, along with the observed data, in Fig. 4.15. If method 2 is used in the two-fluid estimation, the resulting f_s-K curve provides a much better fit to the data, up to a concentration of about 80 vehicles/lane-mile, than that obtained

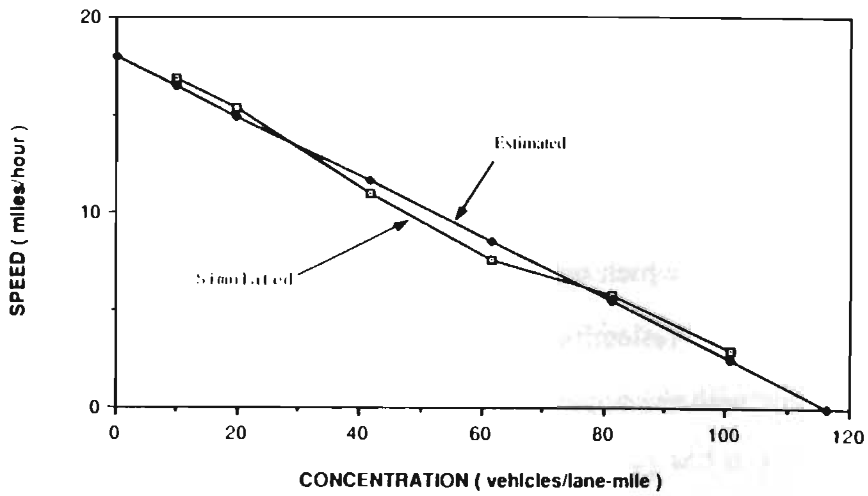


Figure 4.13. Comparison of Linear V-K Model with Observed Simulation Results.

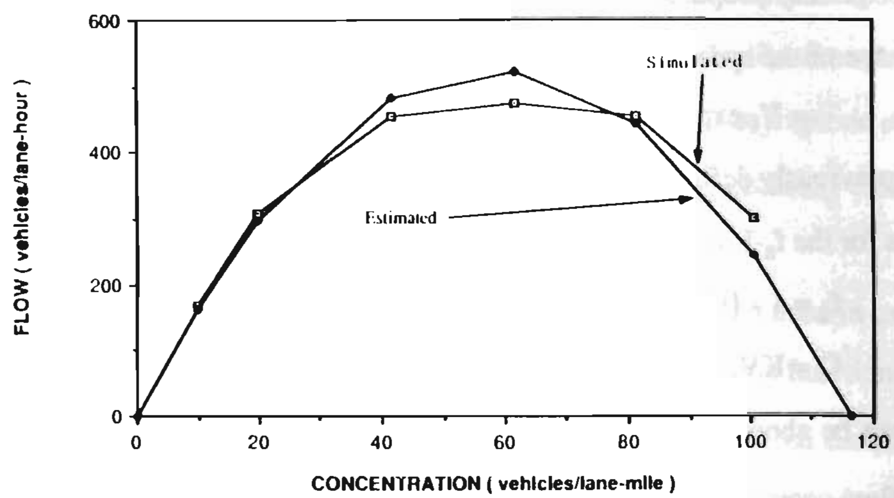


Figure 4.14. Comparison of Q-K Model Derived from Linear V-K Relation with Observed Simulation Results.

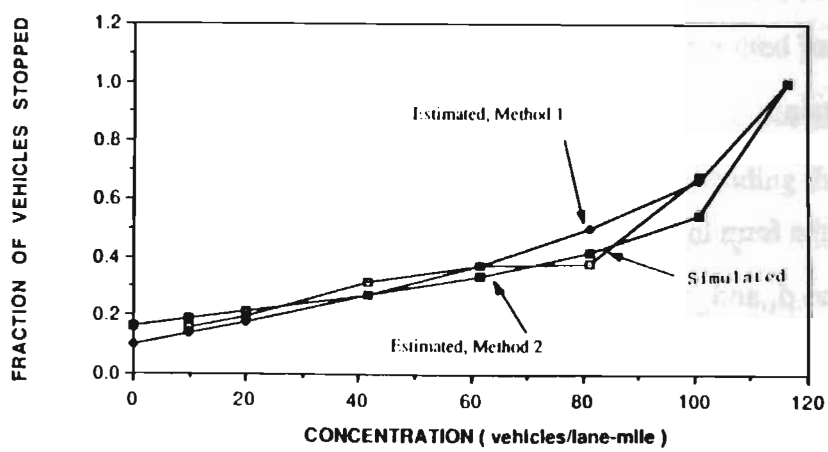


Figure 4.15. Comparison of f_s-K Model Derived from Linear V-K Relation with Observed Simulation Results.

by method 1. As with the first model, this model's concavity is opposite to that exhibited by the data in the middle range of concentrations. What is interesting however is that this model's fit to the f_s-K data, with which it was not directly calibrated, is almost as good as that obtained with Eq. 7, which was directly calibrated using that data. Actually, neither model seems capable of reflecting the concavity of the data over the middle range of concentrations, nor with the apparent inflection point prior to the higher concentration values.

4.4.2.3. Model System 3: Nonlinear V-K Relation.

The second postulated functional form for **v**(K) is a nonlinear 'bell-shaped' function, originally proposed by Drake *et al.* (1967) for arterials, that could capture the apparent shape of the speed-concentration data in Fig. 4.4:

$$V = V_f \exp[-\alpha\, (K/K_m)^d] \tag{14}$$

where V_f (previously defined), K_m, α, and d are parameters to be estimated. The resulting expressions for the f_s-K and Q-K models are then obtained as shown previously, yielding:

$$f_s = 1 - \{(V_f/V_m) \exp[-\alpha\, (K/K_m)^d]\}^{1/n+1}, \text{ and} \tag{15}$$

$$Q = KV_f \exp[-\alpha\, (K/K_m)^d]. \tag{16}$$

It can be shown, by solving $dQ/dK = 0$, that K_m is the concentration at which maximum flow occurs in the network.

Model Estimation Results. Estimation of the parameters of Eq. 14 can be accomplished by nonlinear least squares after some manipulation. Taking the natural logarithms of both sides of the equation, and rearranging the last term on the right-hand side, we obtain:

$$\ln V = \ln V_f - (\alpha / K_m^d)\, K^d, \tag{17}$$

which is of the form $\ln V = c_0 + c_1 K^d$. Nonlinear least squares estimates can then be found for c_0, c_1 and d, and it is then possible to recover $V_f = \exp(c_0)$. But, since $c_1 = (\alpha/K_m^d)$, and only c_1 and d are known, it is not possible to obtain unique values of α and K_m. However, this is not particularly problematic from the standpoint of the model's performance, since only the value of c_1 is necessary. Using the observed simulation results, the parameter estimates are 1.49, 17.95 miles/hour, and 1.83×10^{-3} for d, V_f and c_1 respectively. In order to have an idea of the approximate magnitude of α, the property that K_m represents the concentration at maximum flow can be used. Noting from Fig. 4.5 that the maximum flow occurs for a concentration between ~60 and ~65 vehicles/lane-mile, a range of α can be calculated given the estimated values of c_1 and d, yielding a range from ~0.81 (for K_m=60) to ~0.91 (for K_m=65).

Figure 4.16 shows that this model provides a close fit to the observed speed-concentration data throughout its range. The resulting flow-concentration model, given by Eq. 16, is plotted in Fig. 4.17, along with the observed points. Unlike the relation derived from the linear V-K model, Eq. 16 is not symmetric, and exhibits asymptotic decay of flow as the concentration increases. The model does however provide generally close agreement with the observed data.

The estimated values of the two-fluid parameters are again used in examining the descriptive ability of the corresponding f_s-K relation (Eq. 15). This function is plotted in Fig. 4.18, for the two sets of two-fluid parameter estimates presented earlier. The curve with the parameters found by method 2 (i.e. excluding the last data point) provides much closer agreement with the corresponding five points than that using the method 1 estimates, as the latter seem to be noticeably influenced by the highest concentration point. Unlike the previous model, maximum f_s (equal to 1 for plausible signs of the parameters) is approached asymptotically . Whether this behavior is 'correct' or not is difficult to determine, even in a simulated system, given that it would require loading the network to very unrealistic levels. Moreover, the performance of the derived f_s-K model does not depend only on the postulated V-K function, but also on the two-fluid model invoked in its derivation. As seen earlier, the simulation data appeared to significantly depart from the two-fluid trend at higher concentrations. Nevertheless, Eq. 15 seems to capture the

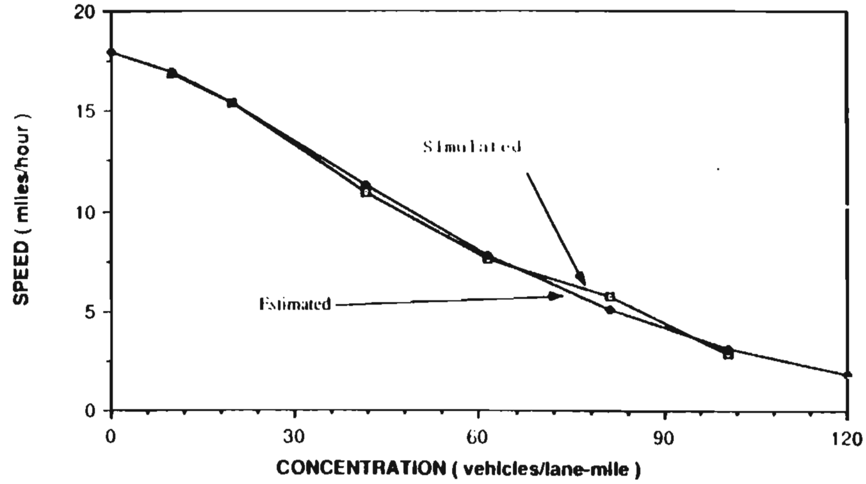


Figure 4.16. Comparison of Bell-Shaped V-K Model with Observed Simulation Results.

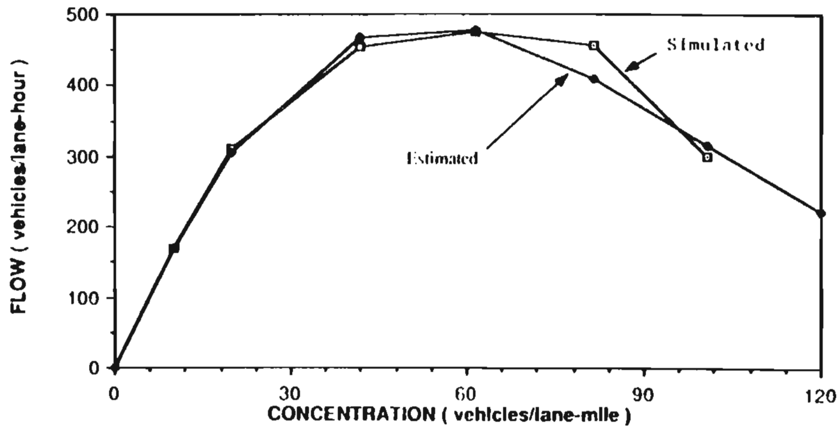


Figure 4.17. Comparison of Q-K Model Derived from Bell-Shaped V-K Relation with Observed Simulation Results.

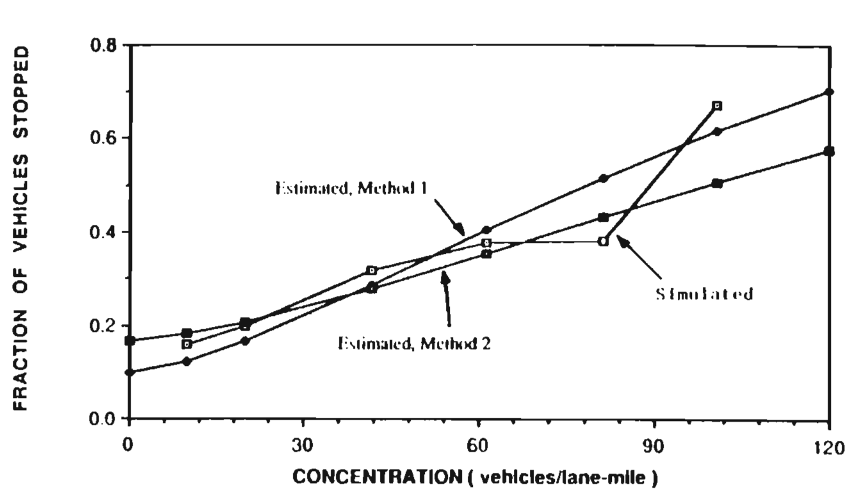


Figure 4.18. Comparison of f_s-K Model Derived from Bell-Shaped V-K Relation with Observed Simulation Results.

network's performance quite well for the range of concentrations that are likely to be encountered in actual urban traffic networks.

4.5. SUMMARY AND CONCLUDING COMMENTS

The principal conclusion from the work presented in this chapter is that it is possible to characterize traffic flow in urban street networks using relatively simple macroscopic models relating the principal networkwide traffic variables. Furthermore, it is remarkable that these relations appear to be not unlike those that have been established at the individual facility level. As illustrated in this paper, the characteristic shape of the fundamental traffic relations encountered for highways and arterials, and upon which traffic engineering procedures have been built, seem to be present at the network level as well, despite the complex interactions that take place in urban street networks.

In terms of the relative performance of the three sets of models presented in this chapter, all three seem to provide a more or less reasonable approximation of the patterns exhibited by the simulated data, at least over certain ranges of concentration. Model systems 2 and 3 provided a much better fit to the speed-concentration data (relative to the derived model in the first system), over the full range of observations; however, this is to be expected since they were calibrated using that data. Generally, the relation between the fraction of vehicles stopped and concentration remains the most problematic in terms of finding a model that is sufficiently convincing over the full concentration spectrum. However, the performance of the derived models was quite good for concentrations up to about 80 vehicles/lane-mile, already well in excess of realistic values that might be encountered in actual networks. Generally, the nonlinear v(K) function yielded a better set of models in terms of agreement with the simulation results. However, the linear approximation for that same relation, on which the second set of models was based, might offer the advantage of familiarity and ease of comparability.

Still focusing on the less than 80 vehicles/lane-mile concentration range, the results in this chapter provide another demonstration of the two-fluid model's validity. This model served as the principal theoretical bridge between the postulated v(K) function and the corresponding f(K) model. In both the second and third sets of models, the derived model performed quite well against the observed data, even though it was not directly calibrated using that data. However, the behavior of the various network variables at very high concentrations remains to be understood. Based on the simulation results presented here, the applicability of the derived functions at these levels decreased markedly, suggesting that the two-fluid assumptions might not directly extend to these high concentration levels. For instance, it is possible that the relation between $\ln T_r$ and $\ln T$ may be nonlinear over the full spectrum of concentrations, though it can be treated as effectively linear over most of the practically meaningful range of concentrations.

Naturally, since the results presented here are based on a limited set of simulated experiments, they must be interpreted in an exploratory sense aimed at stimulating further inquiry into this problem area. These simulations have been performed for a relatively small network, under controlled conditions, and subject to the microscopic rules embedded in the NETSIM package. However, the results appear to correspond well with the limited field studies conducted to date. Simulations of the kind described here have been very useful in supporting the continued development of network-level traffic flow relations and the investigation of their properties. While simulations in larger, more elaborate networks will undoubtedly provide useful information to advance knowledge on this topic, it is essential to obtain networkwide data on the operation of actual urban traffic systems. The cost and the scale of the problem may appear discouraging; however, technological developments in remote-sensing, telecommunications, opto-electronics, among others, offer challenging opportunities for learning about the workings of traffic in urban areas.

To more completetly characterize the performance of urban traffic networks, it is desirable to the understand the sensitivity of the above macroscopic relations and parameters to various network features. Indications along these lines are presented in the next chapter, with regard to the effect of network link width and length, of traffic timing coordination, and of the intensity of interfering events. In that analysis, and given the results seen in this chapter, parameters will be estimated only for the set of models based on the nonlinear form for v(K).

CHAPTER 5

PERFORMANCE OF TRAFFIC NETWORKS: EFFECT OF SYSTEM FEATURES

5.1. INTRODUCTION

The comparison of the three sets of network-level models was conducted in the previous chapter using simulations performed under one set of network conditions. In this chapter, we continue the investigation of a traffic network's overall performance, focussing on the factors that significantly affect the quality of service that it offers. This is accomplished by systematically varying the specified network conditions and examining the response in terms of the network-level relations and their parameters. As concluded in the previous chapter, the model system derived from the nonlinear "bell-shaped" curve for the speed-concentration relation (model system 3) as well as the two-fluid model parameters are used in this chapter to describe the network's response to the various factors.

Three broad categories of factors can be thought to affect a traffic network's performance: 1) topological features, 2) traffic control and 3) traffic activity patterns. These are discussed in turn hereafter, along with the corresponding factor levels considered in this investigation.

The first category includes the configuration of the network (e.g. grid, radial, irregular) and the physical characteristics of its components (e.g. street width, block length, and other geometric design chatracteristics). Only a grid network is used in this study, with street links varying in width from two to six lanes and in length from 400 to 1000 feet.

The factors in the traffic control category primarily include intersection control, signal timing and coordination schemes, as well as movement restrictions, speed control, and the battery of actions available to traffic engineers in controlling flow. The effects of signal timing on network performance are examined in this chapter for three alternative control strategies: 1) coordination for "perfect" progression, 2) simultaneous operation (all traffic signals begin their north-south green phase simultaneously, then begin their eastwest green phase together), and 3) random offsets, when the traffic signal cycles begin at randomly designated times. Of course, many other network factors can be included in traffic control, but are not investigated in this work, such as one-way street designation, speed limits, turn prohibitions, and protected turns at intersections.

The third category, referred to as traffic activity patterns, broadly encompasses the circulation patterns of vehicles in the network, resulting from trip routing and timing decisions, as well as the frequency and intensity of interfering occurrences (e.g. temporary lane blockages due passenger or package pick-up or delivery, parking manoeuvres) associated with a particular adjoining land use and activity pattern. The behavioral characteristics of motorists would also be included in this category. Indirectly, land use is an important element to the extent that it generates trip patterns and affects interfering activity. In future work, this category should be separated into two types of factors: 1) the more global trip patterns, resulting from the spatial distribution of residences, businesses, economic and social activities, and 2) the more local effect of interfering events, directly reflecting the intensity of urban activities adjoining the street network. The first type would then provide the currently missing linkage between the evolving spatial patterns in the urban area and their effect on the traffic system's performance. In this chapter, the investigation is limited to the effect of the interfering activities, which can be modelled as temporary blockages of the right lane, arising with different frequencies and over different durations. The "short-term" events capability of the NETSIM package is used for this purpose .

Clearly, this is not intended as a comprehensive or exhaustive assessment of the factors determining network performance, but as illustrative of the phenomena affecting traffic network performance, and of the direction that future work can take in seeking a further understanding of these phenomena within the framework proposed in this study. It does however begin to give answers about the relative magnitude of benefits in traffic service quality that could be reasonably expected under different types of improvements.

The general approach for these comparisons consists of performing a series of simulation runs for each level of the experimental factor under consideration. Each run corresponds to a concentration level that is maintained constant throughout the simulation, as before. For the topological factors of link width and block length, as well as for the traffic signal timing coordination factor, each series consisted of five simulation runs, with target concentrations of 8, 20, 40, 60 and 80 vehicles/lane-mile; higher concentrations were not included in light of the results of the previous chapter. For the traffic interfering activities factor, only three simulation runs were performed, with target concentrations of 20, 40 and 60 vehicles/lane-mile, as discussed in section 5.4. The actual concentrations, though quite close to the target values, do not exactly match them due to the manner in which vehicles are loaded onto the network, as described in the previous chapter. Basically, the network configuration and all the simulation details for this investigation are the same as those used for the analysis of chapter 4, with the exception of those elements corresponding to the particular factor under investigation. These are detailed in conjunction with the discussion of the particular experimental factor under consideration. The next section addresses the network topology factors, followed by the traffic control factor in section 5.3. The effect of traffic interfering activities is examined in section 5.4, and concluding comments are presented in section 5.5.

5.2. NETWORK TOPOLOGY

We first examine the effect of the block length on network performance, followed by that of the street's number of lanes.

5.2.1. Link Length.

The variation of block length is generally not an option that a traffic engineer can turn to in trying to solve traffic problems in an existing system. However, examination of this factor yields insight into the effects on network traffic performance of the frequency of intersection interference (such as turning traffic and the associated delays incurred by both the through and turning traffic). Thus the general results pertaining to this factor would be of practical interest when considering a grade separation at a busy intersection, and the resulting insights useful in comparisons of network operations in cities with different characteristics as well as in the planning of new or rapidly developing communities.

Three levels for the block length factor are considered: 400, 700 and 1000 feet, with the base or reference case being the 400 ft. length. In the simulation experiments, all other factors were unchanged. The resulting network-level V-K, Q-K, and f_s-K relations for these conditions are shown in Figs. 5.1, 5.2, and 5.3 respectively. The principal effect of link length on the network V-K relation is, as expected, higher average network speed for longer links, particularly at the lower concentrations. Longer links, resulting in greater distances between traffic signals, provide the opportunity to spend a higher fraction of time at or close to one's desired speed; however, at higher concentrations, this ability is offset by increasing congestion, thereby greatly reducing the influence of link length. The f_s-k curves in Fig. 5.3 indicate a general trend of decreasing average fraction of vehicles stopped, at a given concentration, with increasing block length. The improvement due to longer blocks is actually more dramatic at higher concentrations, with the fraction of time stopped reduced by about 30% as block length is increased from 400 to 700 ft., and by an additional 20% as it is increased to 1000 ft. (at the highest concentration level considered in this analysis).

Estimates of the parameters of the nonlinear **v**(K) (Eq. 14 in chapter 4) are shown in Table 5.1, and those of the two-fluid n and T_m are given in Table 5.2. These values confirm that the free mean speed, V_f, increases with the link length, along with the parameter α (note however that the values of α shown assume a constant K_m across all cases, because these two parameters cannot be independently identified, as discussed in the previous chapter), while d exhibits no particular pattern. Both of the two-fluid parameters exhibit clear patterns, with n increasing sharply, while T_m decreases, with link length. The behavior of T_m is expected and conforms to intuition; that of n is more subtle. The

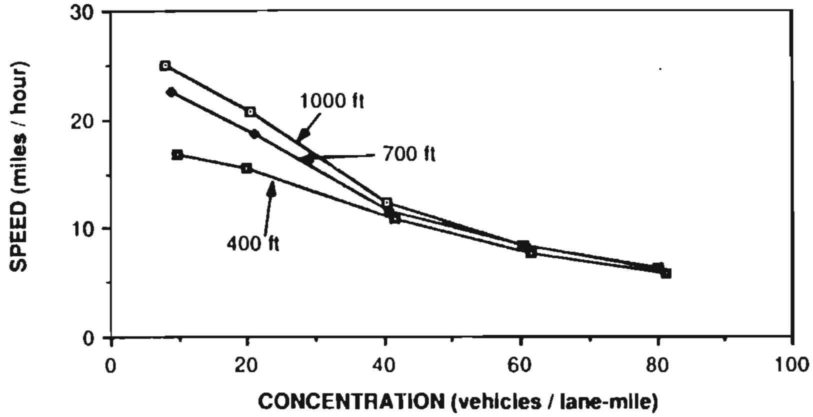


Figure 5.1. Effect of Block Length: Speed-Concentration Relation

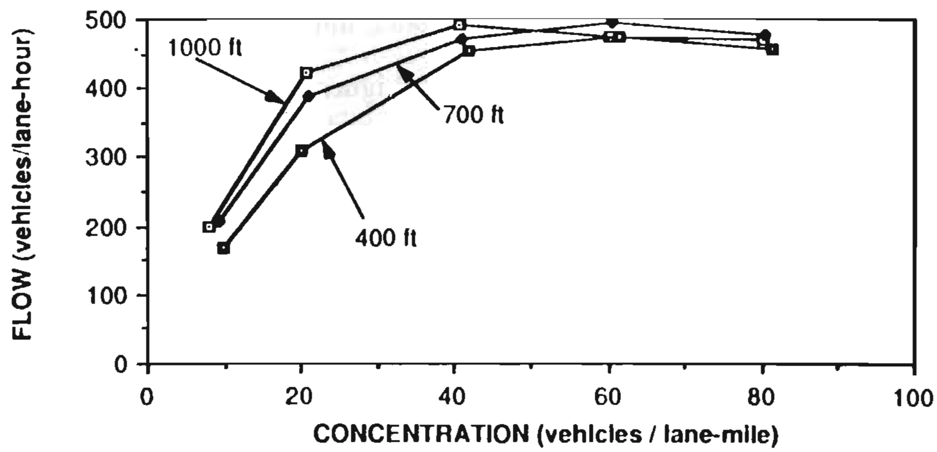


Figure 5.2. Effect of Block Length: Flow-Concentration Relation

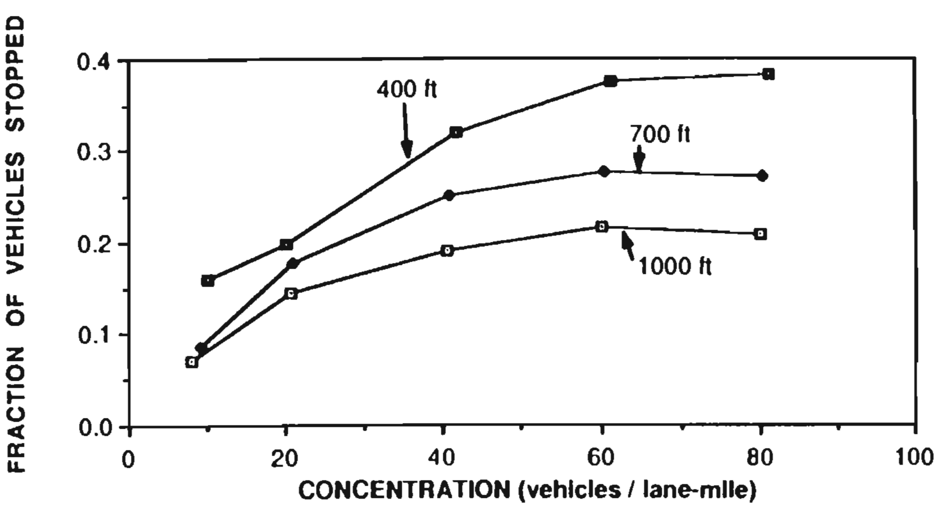


Figure 5.3. Effect of Block Length: f_s-K Relation

Table 5.1 Parameter Estimates for Speed-Concentration Model: Effect of Block Length

Block Length (ft)	V_f (mph)	d	$\alpha^{\dagger}$
400	18.584	1.225	0.829
700	27.892	0.914	1.173
1000	29.598	0.979	1.204

†Note: Assumes that K_m = 60 vch/lane - mile

Table 5.2 Two-Fluid Parameter Estimates: Effect of Block Length

Block Length (ft)	n	T_m (min/mile)
400	2.349	1.809
700	4.876	1.195
1000	8.148	0.878

Table 5.3 Parameter Estimates for Speed-Concentration Model: Effect of Link Width

Link Width Number of Lanes	V_f (mph)	d	$\alpha^{\dagger}$
400-ft. Links			
1	18.210	1.232	0.733
2	18.584	1.225	0.829
3	18.297	1.334	0.863
1000-ft. Links			
1	27.708	1.025	1.037
2	29.598	0.979	1.204

†Note: Assumes that K_m = 60 veh/lane - mile

parameter n can be thought of as the approximate slope of the T_s vs. T trend; higher n values then indicate that average (total) trip time T is increasing with concentration at a faster rate than average stopped time Ts. This occurs in longer blocks because the average running time T_r (also the difference between T and T_s) is more sensitive to congestion, since longer block distances allow higher average running speeds at low concentration.

5.2.2. Link Width (Number of Lanes).

This factor has more direct traffic engineering applicability, as the number of travel lanes on city streets can be manipulated through changes in parking regulations, restriping or low-cost physical improvements. Parking may be eliminated or allowed along city streets, increasing or decreasing, respectively, the number of travel lanes by one for each side of the street affected. In addition, parallel parking may be converted to angle parking, or vice versa, generally reducing or increasing the number of travel lanes by one; (conversion from parallel to angle parking along a street also has serious safety considerations, which may override traffic flow considerations). Of course, lanes may be added or dropped by construction, generally at the expense or in favor of sidewalks and other pedestrian amenities.

The effect of link width on the network parameters is examined by considering link widths of one, two, and three lanes, for a network with 400-ft. block lengths, and widths of one and two lanes for a network with 1000-ft. block lengths. Only two link widths were considered for the latter network primarily because of computational cost considerations, given the very large number of vehicles needed to maintain the desired vehicular concentration. It is important to note in this regard that the comparisons in this section are made, as before, at constant concentration; thus, when a lane is added in the simulation, the number of vehicles is increased accordingly. Of course, when a lane is added in the field, there is generally no immediate corresponding increase in concentration, resulting in an improvement in service quality. What is examined in this analysis then is the effect of link width in terms of the greater freedom to maneuver when more lanes are available, and the associated possibility of avoiding left-turning vehicles waiting for a gap in the opposing stream, etc.

The V-K, Q-K, and f_s-K relations for the three widths associated with the 400-ft. network are shown in Figs. 5.4, 5.5, and 5.6, respectively. As expected, the primary differences are manifested in the Q-K relation (Fig. 5.5), with average network flow increasing dramatically with the number of lanes, due to the larger number of vehicles required to maintain constant concentration. However, the differences are minimal on a per lane basis. The other two relations do not exhibit dramatic differences across link widths. An improvement in average speed can be noted at higher concentrations for the wider links, for the reasons mentioned above; for these same reasons, virtually no differences exist at low concentration, since interference from other vehicles is very limited then. Plots of the above relations for the network with 1000-ft. links, shown in Figs. 5.7, 5.8 an 5.9, in the same order as before, reveal similar results to the case with 400-ft. links, suggesting that the qualitative effect of number of lanes seems to be independent of link length.

The parameter estimates are presented in Tables 5.3 and 5.4, confirming that V_f does not vary significantly across the three cases. While the estimates of n (in table 5.4) exhibit little change between one and two lanes, it drops noticeably when the widths are increased to three lanes (for the 400-ft. network). This is due to the slower deterioration of the average running time with increasing concentration when drivers enjoy the greater freedom afforded by the greater number of lanes. It is nonetheless notable that this benefit seems to follow a threshold pattern, with meaningful improvement appearing in thetransition from two to three lanes. Similarly, T_m increases when the link widths are increased to three lanes, but shows little change between one and two lanes.

5.3. TRAFFIC CONTROL

Three levels of traffic signal control are examined: 1) traffic signal interconnection to provide "perfect" progression, with the signal turning to green as it is reached by the

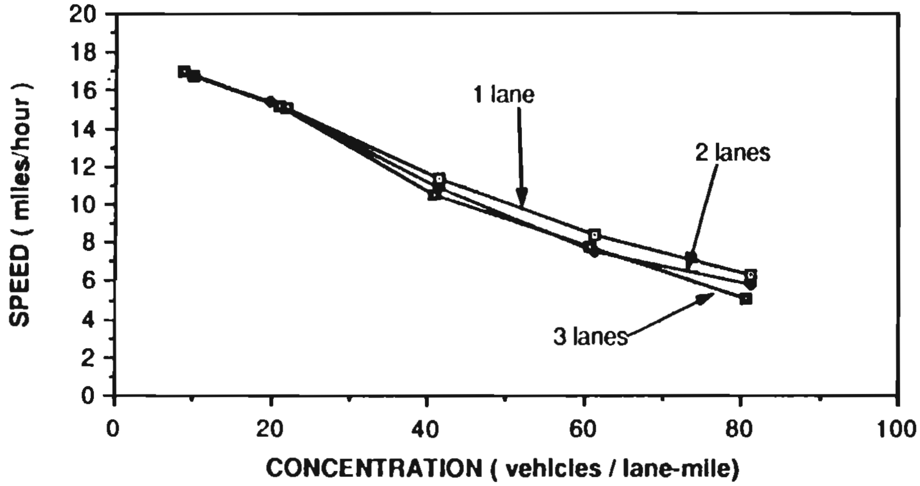


Figure 5.4. Effect of Link Width, 400-ft. Links: Speed-Concentration Relation

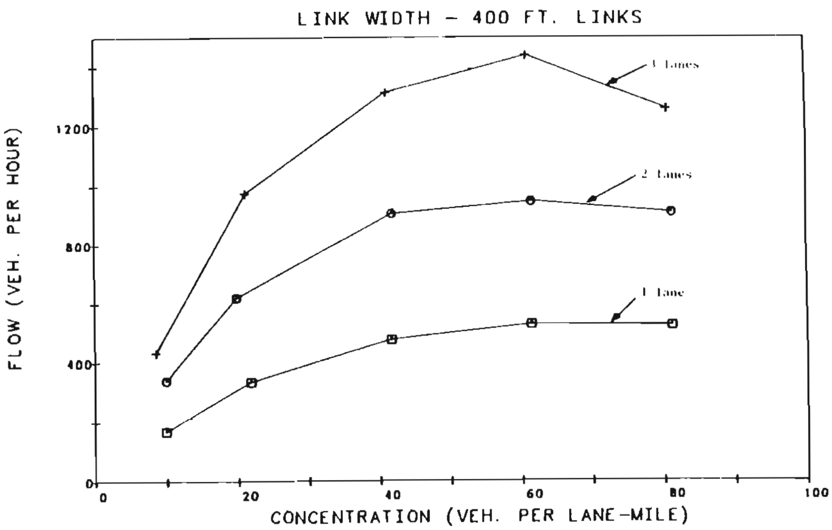


Figure 5.5. Effect of Link Width, 400-ft. Links: Flow-Concentration Relation

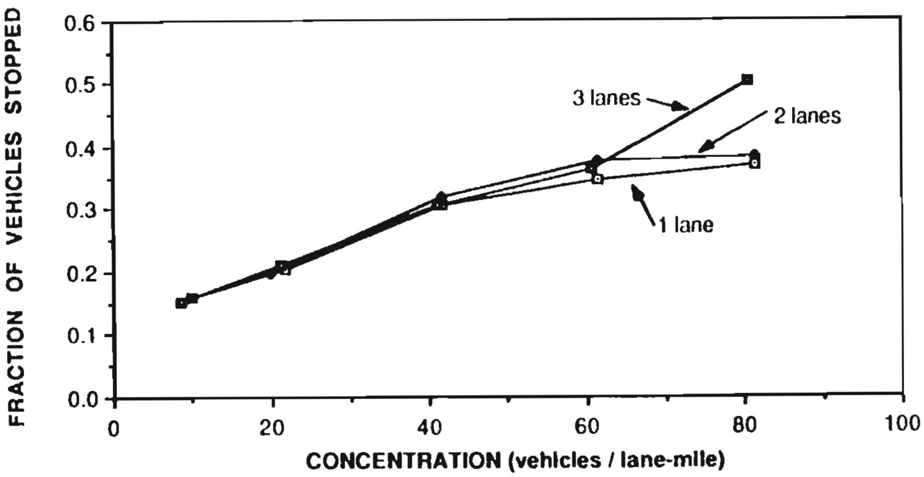


Figure 5.6. Effect of Link Width, 400-ft. Links: f_s-K Relation

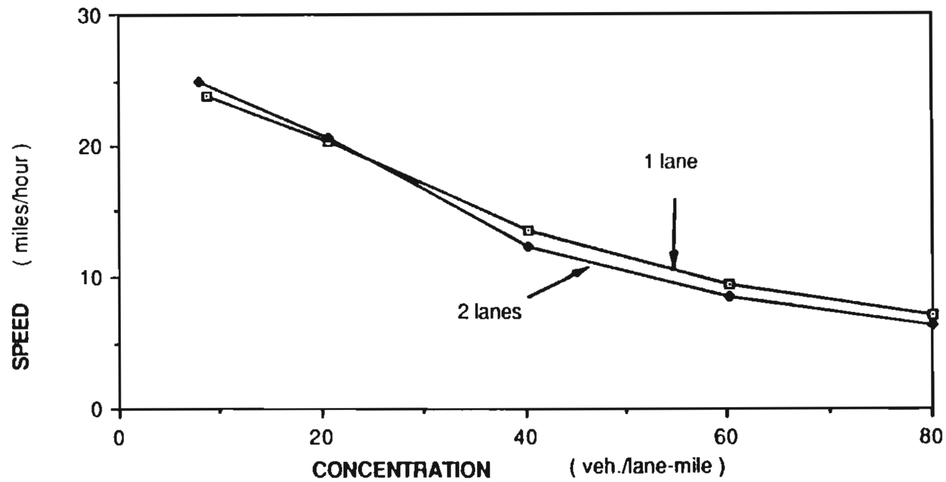


Figure 5.7. Effect of Link Width, 1000-ft. Links: Speed-Concentration Relation

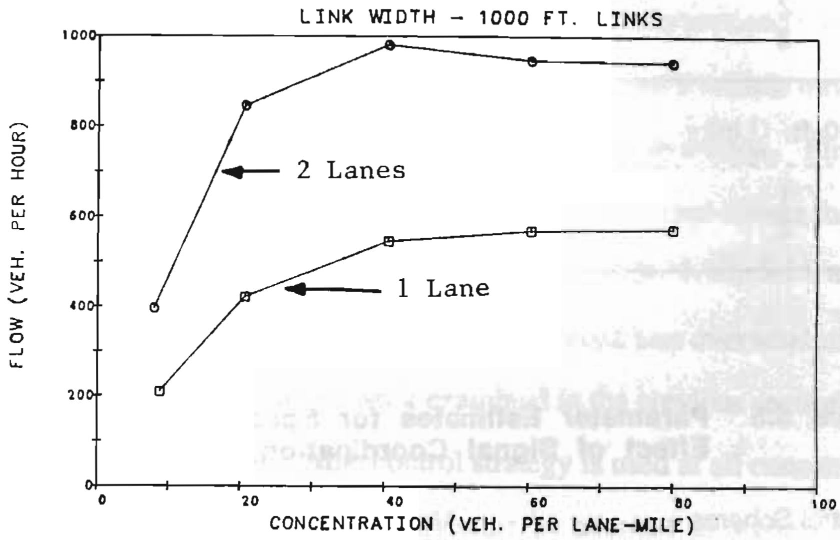


Figure 5.8. Effect of Link Width, 1000-ft. Links: Flow-Concentration Relation

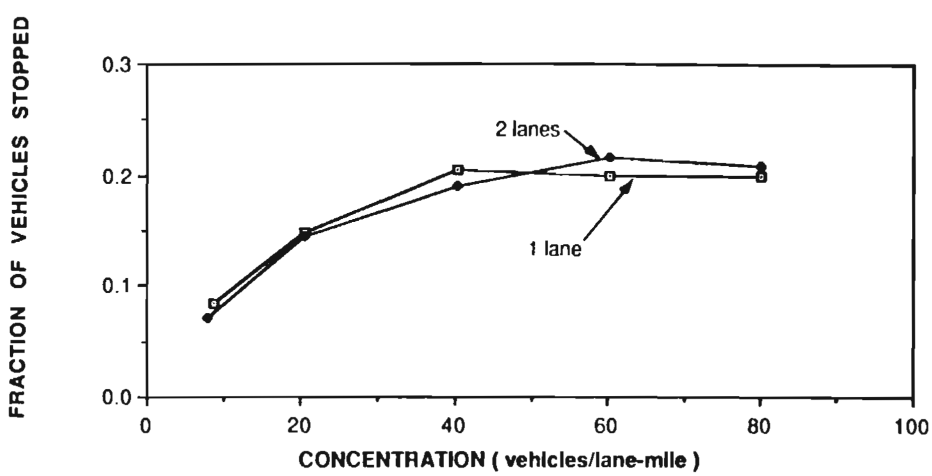


Figure 5.9. Effect of Link Width, 1000-ft. Links: f_s-K Relation

Table 5.4 Two-Fluid Parameter Estimates: Effect of Link Width

Link Width Number of Lanes	n	T_m (min/mile)
400-ft. Links		
1	2.328	1.833
2	2.349	1.809
3	1.897	2.365
1000-ft. Links		
1	8.963	0.708
2	8.148	0.878

Table 5.5 Parameter Estimates for Speed-Concentration Model: Effect of Signal Coordination

Control Scheme	V_f(mph)	d	α
No Short-Term Events Specified			
Single Alternate	29.598	0.979	1.204
Random Offset	28.541	1.005	1.170
Simultaneous	25.128	1.175	1.030
With Short-Term Events Present			
Single Alternate	31.170	0.860	1.284
Random Offsets	30.086	0.881	1.258
Simultaneous	25.345	1.100	1.062

incoming platoon from the upstream intersection; 2) simultaneous operation, with all signal cycle beginning at the same time; and 3) isolated operation, with the signal cycles beginning at randomly selected times. Because the average desired speed of the vehicles in the network is 35 miles/hour, a network using 1000-ft. links was necessary in order to provide "perfect" progression over a reasonable distance through the network. Since a 40-second cycle length is used throughout this simulation, 1000-ft. links provide near-perfect progression using a single alternate offset scheme. Simultaneous operation was achieved by assigning identical offsets to all twenty-five traffic signals. These offsets yield the opposite signal indication at the immediately downstream intersection from that obtained under the single alternate progression, thereby guaranteeing that a vehicle travelling straight through at the average desired speed will stop at each intersection. Finally, isolated uncoordinated operation was simulated by using randomly selected offsets for each signal.

The V-K, Q-K, and f_s-K results are shown in Figs. 5.10, 5.11 and 5.12. The differences due to traffic control initially appear to be much less overwhelming than those due to the geometric features of the network examined in the previous section. Moreover, it is difficult to generalize when the same control strategy is used at all concentration levels. The nature of traffic control is such that it must be adapted to prevailing conditions, including concentration. Nevertheless, the single alternate scheme that provides progression along major directions seems to yield the highest average speed, and the lowest fraction of vehicles stopped, particularly at lower concentration, as expected. As concentration increases, the superiority of the progressive scheme becomes less evident. It should be noted however that the differences among the three timing strategies observed for concentrations up to ~40 vehicles/lane-mile are quite meaningful. Keeping in mind that the peak observed concentrations in field work in Austin and Dallas do not exceed ~30 vehicles/lane-mile, the practical significance of the above conclusions should not be dismissed on grounds that all schemes are equally bad under very heavy congestion; an average network concentration of ~25 vehs/lane-mile already describes a heavily congested situation for most motorists circulating through the network.

In addition, the effect of traffic signal timing coordination was tested in conjunction with the "interfering traffic activities" factor, by conducting an additional series of simulation runs for each signal control strategy, but with the activation of the "short-term events" option in the simulation program. As described in more detail in the next section, this option provides an "activity level" which simulates interference to moving vehicles, in addition to that caused by the other moving vehicles, through the temporary blockage of a link's right lane. Introduction of these events, resulting in increased congestion (particularly at the higher concentration levels), further impedes the moving traffic, restricting the vehicles' ability to move at their desired speed (such as in the progression band). The short-term events were specified to occur stochastically, with a mean frequency of 30 times per hour (mean rate), and mean duration of 25 seconds each; this activity level is denoted AL1 in the next section, where the effects of the activity levels are specifically addressed). The results of these simulations with the "short-term events" option are shown in Figs. 5.13, 5.14 and 5.15, indicating the same general conclusions as in the case without interfering activities. The improvement due to progression timing is somewhat reduced in the presence of these short term blockages, since these further contribute to congestion in the system. Nevertheless, improvement at lower concentration can be of practical significance.

The parameter estimates are presented, for both cases (with and without short term events), in Tables 5.5 and 5.6, confirming the above comments regarding the effect of timing on V_f. The rank order of the magnitudes of the two-fluid parameter T_m for the three cases parallels that of V_f. On the other hand, the two-fluid parameter n does not exhibit a similarly consistent pattern. It is interesting to note that n increased as the timing went from random (isolated operation) to coordinated progression. While overall service quality is better under coordination, based on the other indicators discussed here, the increase in n is

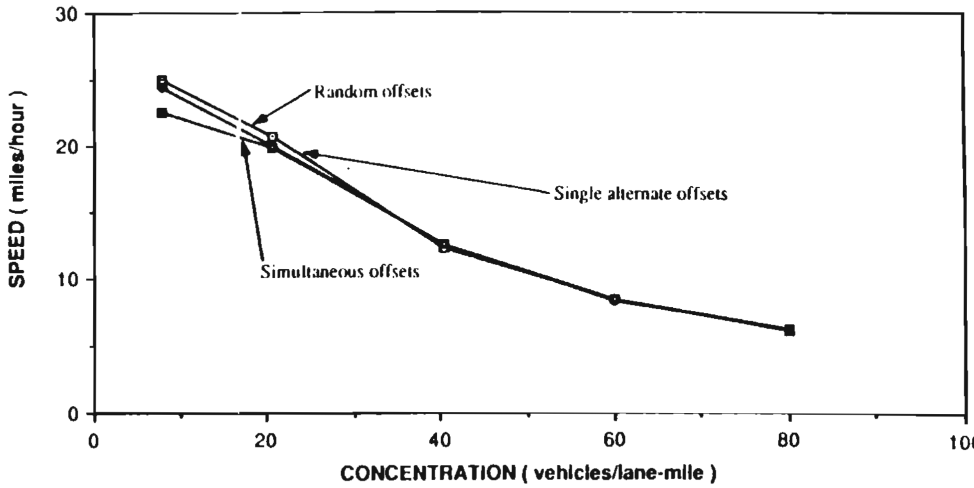


Figure 5.10. Effect of Traffic Control: Speed-Concentration Relation

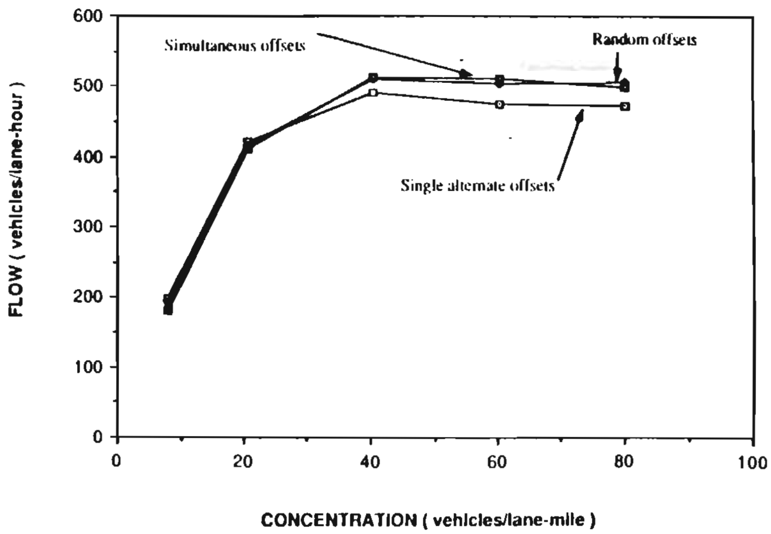


Figure 5.11. Effect of Traffic Control: Flow-Concentration Relation

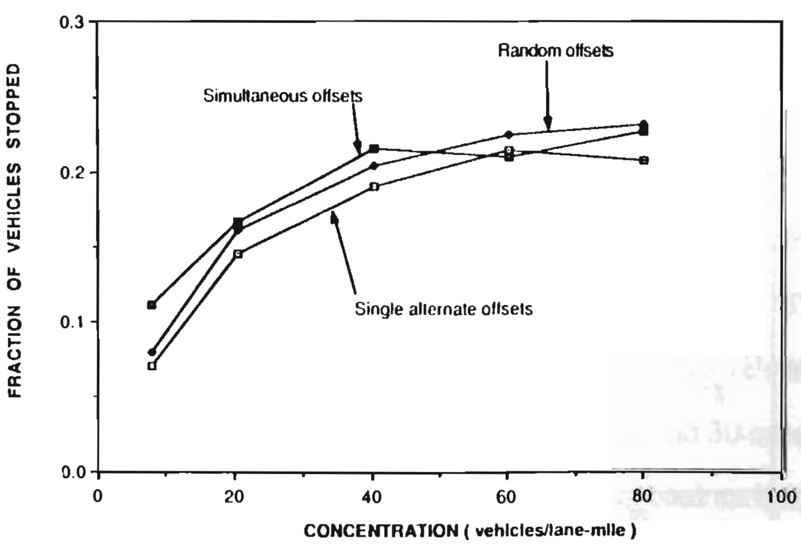


Figure 5.12. Effect of Traffic Control: f_s-K Relation

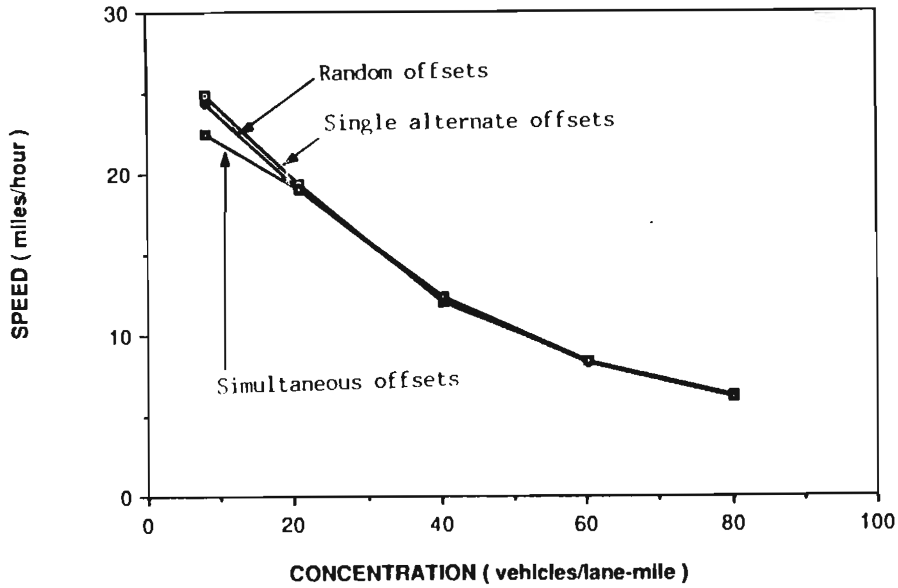


Figure 5.13. Effect of Traffic Control (with Traffic Interfering Activities): Speed-Concentration Relation

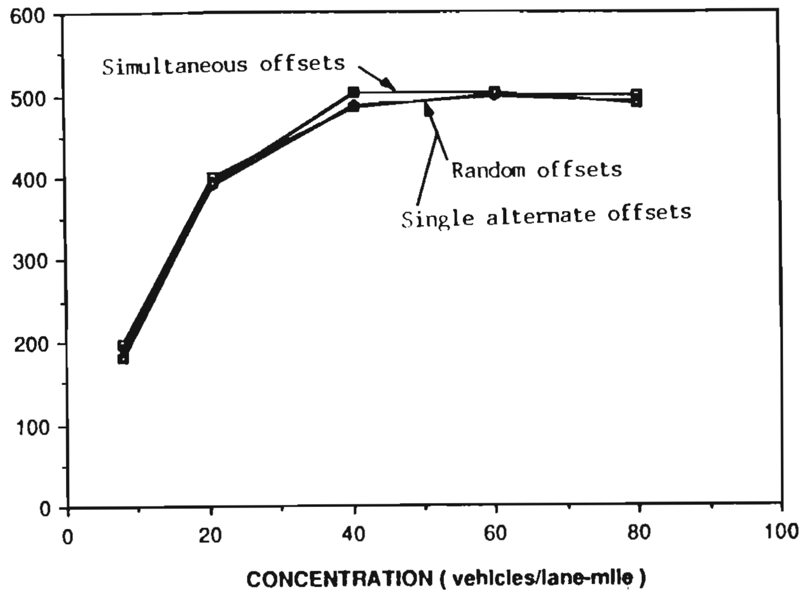


Figure 5.14. Effect of Traffic Control (with Traffic Interfering Activities): Flow-Concentration Relation

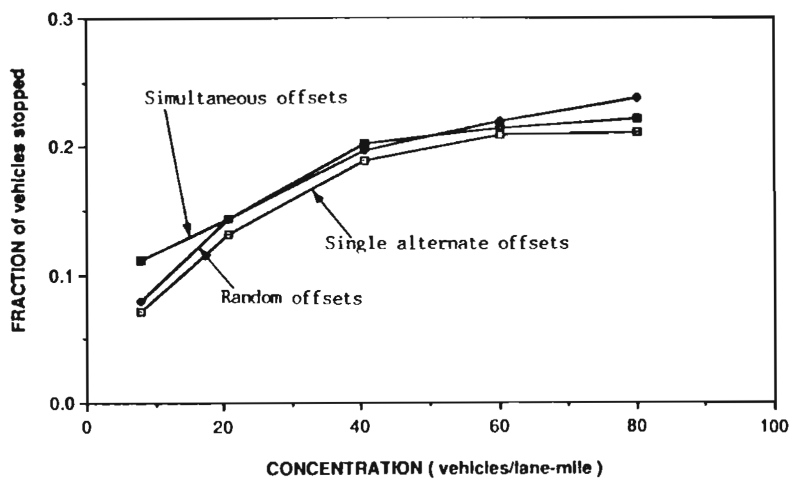


Figure 5.15. Effect of Traffic Control (with Traffic Interfering Activities): f_s- K Relation

due to the same type of phenomenon that resulted in a higher n for longer block length, discussed in the previous section. Essentially, progression allows drivers to achieve higher running speeds when moving, and therefore makes the average running speed relatively more sensitive to increasing concentration.

5.4 INTERFERING ACTIVITY LEVEL

In this section, we consider the effect of differing levels of traffic-interfering activity in the test network on the quality of traffic service. The "activity level", for the purpose of this study, refers to the degree of interference experienced by moving vehicles. As noted earlier in this chapter, such interference may be due to intralink perturbations such as vehicles stopping for loading and unloading, pedestrian activity (other than at intersections), and many other similar activities which are an inherent feature of a city street network. By examining the effect of this factor on the parameters of the traffic models of interest, we can gain insight into the determinants of these parameters, and thus enhance their interpretability and usefulness, particularly in the context of comparisons between cities with different characteristics. In our simulations, this factor is modelled using the "short-term events" option of the simulation package, where the user may specify the mean frequency and mean duration of these events. These are then generated, in the simulation, stochastically and independently as blockages in the right lane in the center of the link. The average fraction of time that the right lane is blocked can be found by dividing the mean event duration by the inter-event time (the inverse of the event frequency).

In the experiments described herein, short-term events are identically specified for all 80 links in the test network at the four levels shown in Table 5.7. The first, denoted AL0, is the base case, with no events specified. The other three levels are denoted AL1, AL2 and AL3, and correspond to mean frequencies of 30, 72 and 30 events per hour, and mean durations of 25, 25 and 60 seconds, respectively. As noted earlier, only three runs were conducted at each level of this factor, with target concentrations of 20, 40 and 60 vehicles/lane-mile.

The estimates of the network level model parameters of interest are presented in Table 5.7 (for the nonlinear v(K) relation) and Table 5.8 (for the two-fluid parameters), for each of the activity levels under consideration. All parameters exhibit clear trends in response to the activity levels. However, we have to be cautious in interpreting these parameters, particularly those of the nonlinear v(K) relation, because we have only three observations per factor level. Generally, it appears that increasing activity levels and the accompanying impedence to moving vehicles results in greater sensitivity to the prevailing concentration level. This is manifested in the value of the parameter d in the v(K) model, which decreases with increasing interfering events, indicating a more rapid rate of deterioration of the average speed in the range of concentrations $K<Km$ (which is where most of our simulated observations were). Similarly, the two-fluid parameter n increased with increasing activities, reflecting the greater sensitivity of running time.

On the other hand, the free-flow speed V_f increased, while the minimum running time T_m decreased, with increasing activity levels; this behavior appears to be counterintuitive, and is more of a statistical aberration than a reflection of a true phenomenon. Both values represent some condition under $K=0$, and are estimated by extrapolation from the points corresponding to concentration values in the range of 20 to 60 vehicles/lane-mile. Thus, for cases where the effect of concentration on average running speed was large, such as in the presence of high activity levels, lower intercepts (for $K=0$) were predicted than in the case where traffic interfering activities were not specified.

In all cases, it appears that parameters are more responsive to changes in mean event duration than in the mean frequency of these short term events, for a given fraction of time that the right lane is blocked. The behavior of the parameters d and n with respect to the activity level are shown in Figs. 5.17 and 5.18, respectively, where each parameter's variation is shown as a function of mean event frequency, mean event duration, and mean

Table 5.6 **Two Fluid Model Parameter Estimates: Effect of Signal Coordination**

Control	n	T_m
No Short-Term Events Specified		
Single Alternate	8.148	0.878
Random Offset	7.441	0.866
Simultaneous	9.919	0.502
With Short-Term Events Specified		
Single Alternate	7.643	1.023
Random Offsets	6.746	1.078
Simultaneous	8.954	0.681

Table 5.7 **Two-Fluid Model Parameter Estimates: Effect of Activity Levels (Short Term Events)**

Activity Level	Mean Frequency (Event/hour)	Mean Duration (Sec)	Frac. of Time Lane is Blocked	n	T_m (min/mile)
0	0	0	0	1.831	2.011
1	30	25	0.208	2.259	1.822
2	72	25	0.500	2.583	1.804
3	30	60	0.500	2.700	1.764

Figure 5.16. Variation of Parameter d with the Activity Level

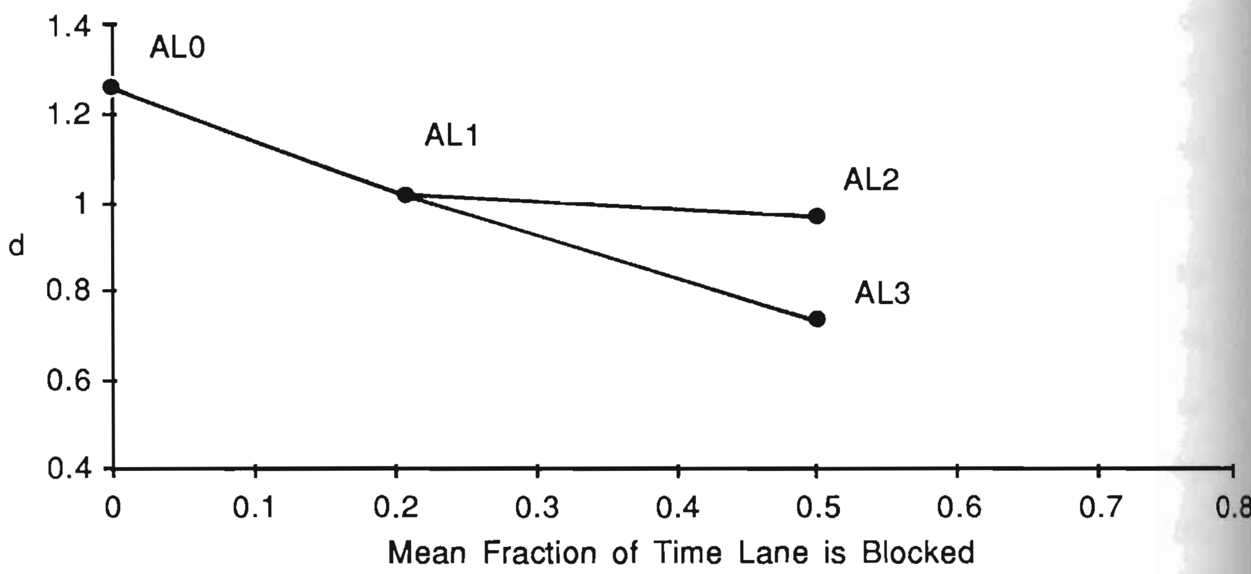
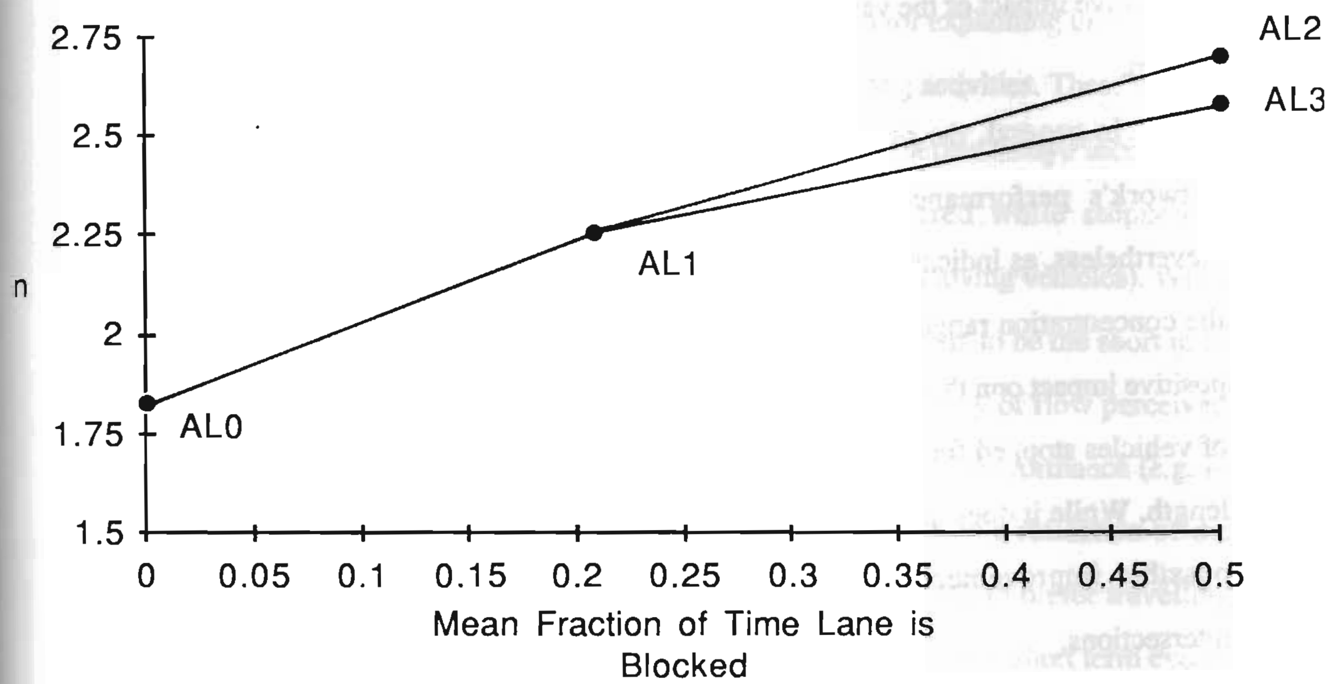


Figure 5.17. Variation of Parameter n with the Activity Level



fraction of time that the right lane is blocked. The greater influence exerted by the mean event duration is easily seen in these figures.

5.5. SUMMARY AND CONCLUDING COMMENTS

As noted earlier, the factors specifically examined in this study were not intended to be comprehensive, but mainly to explore the sensitivity of the network-level models and their parameters to selected examples of each of the three categories of factors thought to influence the network's performance. In this section, we review the principal parameters' response to these factors, and discuss how these parameters might be correlated, as well as the relative impact of the various factors considered on the quality of service offered by the system.

In general, the network's physical features seem to have a greater effect on the network's performance than traffic control through conventional signal timing. Nevertheless, as indicated in section 5.3, meaningful improvements could be achieved in the concentration ranges under which most actual networks operate. Overall, the greatest positive impact onn the level of service, defined as higher average speed and lower fraction of vehicles stopped for a given concentration level, was achieved by increasing the block length. While it does not seem practical to redesign entire CBD's, the results indicate the possible improvements that could result from reducing the required stoppage at intersections.

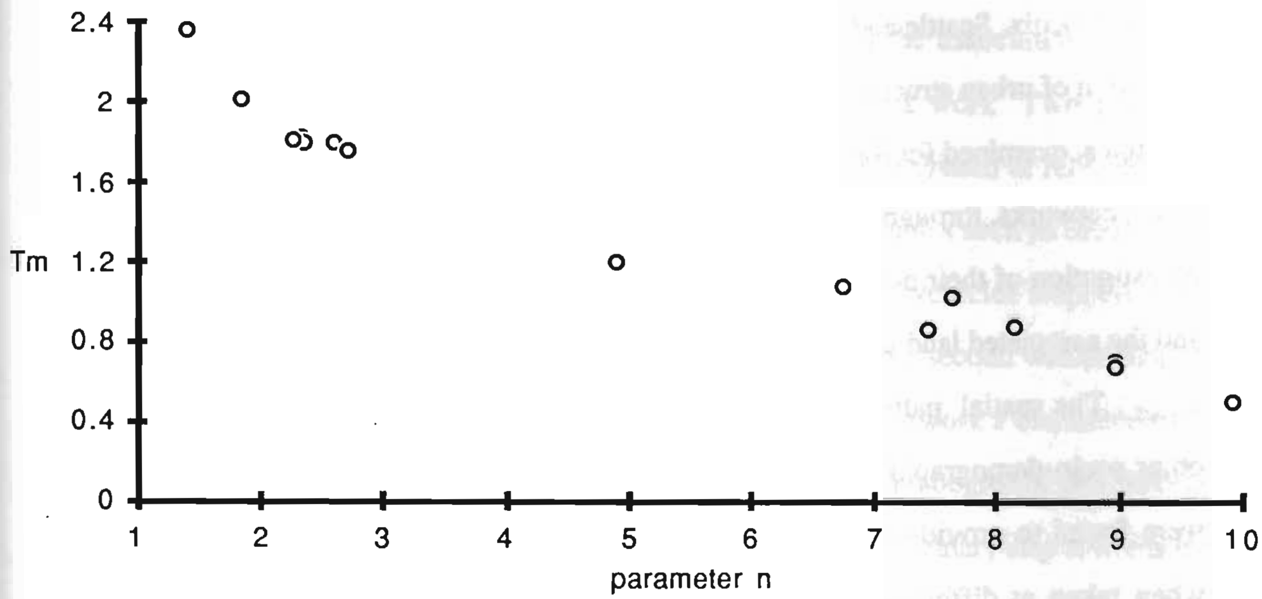
The effect of more lanes (keeping concentration constant) on average speed and fraction of vehicles stopped is not as dramatic; however, a better perceived quality of flow is seemingly reflected in a lower value of the two-fluid parameter n, due to the greater freedom to maneuver available to motorists. Regarding the latter parameter's behavior, n does not appear to constitute a reliable objective to guide network improvements. For instance, the above-mentioned clearcut overall improvement due to longer blocks is accompanied by an increase in n, because the higher running speeds possible in the longer links are correspondingly more susceptible to rapid deterioration due to intralink friction generated at higher concentrations. The same type of phenomenon is manifested in the results of the traffic control experiments, where the clearly beneficial effects of coordinated operation at lower concentrations were accompanied by an increase in n. On the other hand, for a given link length and coordination scheme, a lower n seems to correspond to improved traffic conditions, as seen when an extra lane was added to an existing link, or when the level of traffic interfering activities was reduced. This suggests caution in making comparisons on the basis of n among networks with different link sizes or signal coordination strategies.

It appears that the most direct and consistent mechanism for explaining the variation of the parameter n is the occurrence of short term traffic interfering activities. These directly interfere with the moving vehicles' running speed, and thus generate delays incurred in traversing the links themselves, as opposed to delays incurred while stopped at intersections (which do not affect the average running speed of the moving vehicles). While other factors naturally affect n, its most consistent determinant appears to be the short term events. From this standpoint, n might be a good measure of the quality of flow perceived by system users; this quality is not necessarily correlated with overall performance (e.g. in terms of average speed or fraction of vehicles stopped), but rather is a reflection of the quality of flow as vehicles are moving, since drivers can be expected to prefer travelling unimpeded at their desired speeds. To the extent that the intensity of these short term events reflect the character of life in a particular city, it can be expected that behavioral characteristics of the users as well as of the urban dwellers at large can greatly affect the performance of a traffic network. By the same token, measures to control these short term events, such as the provision of facilities for loading and unloading, better pedestrian channelization, and stricter enforcement of certain regulations (e.g. preventing illegal short-term parking) can have a significant impact on the overall quality of service, in terms of lower average speed for a given concentration level as well as greater ease of movement reflected by a lower value of n.

The interpretablity of any of the speed-concentration relation parameters, taken individually, appears limited, particularly if based on few observations, as in the tests for the effect of traffic-interfering activities. The greatest effect on the parameter d was that of the traffic control and activity level factors. However, it did not exhibit a generally consistent behavior in response to network changes.

In terms of the joint variation of these various parameters, no consistent patterns have emerged from this limited set of experiments. The clearest trend observed is the negative correlation between the values of the two-fluid parameters n and T_m, illustrated in Fig.5.18, which depicts all the points generated by the simulations discussed in this chapter. Fig. 5.18 seems to indicate that measures that reduce T_m, i.e. improve the maximum running speed in the network, tend to increase n, i.e. the sensitivity of the average running speed to increasing vehicular concentration. None of the other parameters appear to be as consistently related; rather, opposite trends are present between any two parameters over the spectrum of observations. Further understanding of how these parameters are related, and of the interaction among various network features in determining the response of these parameters, can only be accomplished through a more extensive systematic investigation.

Figure 5.18. Plot of the Two-Fluid Parameters n vs. Tm for All . Simulation Experiments



CHAPTER 6

CONCLUSION

Several useful and promising directions for exploring, characterizing and understanding urban development patterns, as they affect and interact with transport in cities, have resulted from this study. The work was pursued along three levels: 1) the long term (since the turn of the century) evolution of several infrastructural attributes of cities, tracked primarily for Austin and San Antonio, and, to a more limited extent, for eleven other cities (Atlanta, Boston, Chicago, Cincinatti, Denver, Los Angeles, Miami, New York, Phoenix, Seattle and St. Louis), providing a means for comparing these cities; 2) the evolution of urban structure, in terms of the spatial density patterns of population and other variables, examined for four case areas; and 3) the characterization of traffic flow quality in urban networks, through the development of macroscopic network-level relations, and the investigation of their dependence on the physical and operational features of the network and the associated land use.

The spatial patterns of population density , household automobile ownership and other socio-demographic variables, as a function of distance from the central city core, were found to provide a useful characterization of urban structure, and of its evolution when taken at different time intervals. The four case cities (Austin, Atlanta, Dallas, Phoenix) considered in conjunction with this analysis illustrated several important continuing trends observed over the past three decades. In particular, the density functions captured the spatial dispersion away from the traditional central urban cores into the suburbs and exurbs, as well as the simultaneous densification of existing suburbs to levels almost comparable to the central cores. Parallelling this pattern has been a continuing growth of average household automobile ownership, in all parts of the urban area, with a distinct spatial and temporal pattern that seems to be robust across the case areas considered, as well as within radial corridors in the one case that was so analysed (Austin). These results suggest that average trip distances are likely to further increase, and that congestion levels in the densifying "suburban" communities will reach levels typically associated with CBD traffic systems. On the other hand, denser suburban communities and the emergence of stronger suburban foci might offer opportunities for the provision of mass transit services with more acceptable service levels than has traditionally been possible in low-density suburbs.

However, suburban networks generally exhibit somewhat different characteristics from CBD systems, in terms of physical features, traffic control and adjoining land use and its generation of traffic-interfering activities. The implications of these differences on the performance of the traffic system, in terms of congestion and the associated quality of traffic service, were addressed by the third component of this work. Two principal objectives guided this part of the study. The first was to develop a system of relations that could comprehensively describe the joint behavior of traffic variables such as the average speed, flow and concentration, as well as the average fraction of vehicles stopped in the network, and the two-fluid stopped and running time variables. The second was to analyze the sensitivity of these relations and their parameters to the network's characteristics, particularly its topological features, prevailing control strategies and degree of interference from the adjoining land use. The latter objective is viewed as a preliminary step towards the more ambitious goal of relating the quality of traffic service and associated mobility level, to both network features and urban structure. Methodologically, both objectives were supported by microscopic simulation experiments, which provided useful macroscopic insights into network traffic behavior.

The principal conclusion from that investigation of traffic network performance is that it is possible to characterize traffic flow in urban street networks using relatively simple macroscopic models relating the principal networkwide traffic variables. Furthermore, it is remarkable that some of these relations closely parallel those that have been established at the individual facility level, despite the complex interactions that take place in a network. The analysis presented in Chapter 5 suggested that these relations and their parameters appear to follow systematic trends in response to the network features. The factors specifically examined in this study were not intended to be comprehensive, but illustrative of the phenomena affecting traffic network performance, and of the direction that future work can take in seeking a further understanding of these phenomena within the framework proposed in this study. However, the analysis yielded several useful indications about the relative magnitude of benefits from different types of improvements.

In general, the network's physical features seemed to have a greater effect on the network's performance than traffic control through conventional signal timing. Nevertheless, meaningful improvements could be achieved in the concentration ranges under which most actual networks operate. Overall, the greatest positive impact on the level of service, defined as higher average speed and lower fraction of vehicles stopped for a given concentration level, was achieved by increasing the block length. While it does not seem practical to redesign entire CBD's, the results indicate the possible improvements that could result from reducing the required stoppage at intersections. The effect of more lanes (keeping concentration constant) on average speed and fraction of vehicles stopped was not as dramatic; however, a better perceived quality of flow was reflected in a lower value of the two-fluid parameter n, due to the greater freedom to maneuver available to motorists.

Interesting insights were obtained regarding the effect of traffic interfering activities (short term events) typically present in an urban setting. Because the intensity of these short term events reflects the character of life in a particular city, it can be expected that behavioral characteristics of the users as well as of the urban dwellers at large can greatly affect the performance of a traffic network. By the same token, measures to control these short term events, such as the provision of facilities for loading and unloading, better pedestrian channelization, and stricter enforcement of certain regulations (e.g. preventing illegal short-term parking) could have a significant impact on the overall quality of service, in terms of lower average speed for a given concentration level as well as greater ease of movement reflected by a lower value of the two-fluid parameter n. A better understanding of these mechanisms would be useful in trying to understand and deal with the scenarios of deteriorating traffic conditions in densifying suburban communities. The private automobile has been an essential element in the development of these suburban development patterns, which initially seemed to offer congestion-free circulation. Rapid degradation of traffic conditions in these communities could have important consequences on the outcomes of the personal location, mobility and travel decisions of area residents, particularly with regard to automobile ownership and utilization.

6.1. FUTURE RESEARCH DIRECTIONS

Several opportunities for the continuation and expansion of the present research appear promising in terms of substantive and methodological contributions to the characterization of urban patterns, their evolution and the interaction with and effect on personal mobility and transportation. These opportunities exist in each of the three levels that formed the focus of the effort described herein, as well as in their integration into an overall operational framework that would provide the linkage between the evolution of the urban-wide attributes (considered in chapter 2), the spatial distribution of residents, firms and other activities, the resulting flow patterns and ultimately their effect on the performance of the transport network, given its features.

It is clear that much of the work in this report is of an exploratory nature. Thus the analysis of the attributes tracked for a small set of cities in chapter 2 can be enhanced through the consideration of additional descriptors and increasing the sample size, in order to gain greater confidence in the ability of some of these descriptors or combinations thereof to provide a robust characterization of urban evolution, and to subsequently support related predictions of future development.

Several directions for meaningful additional work are suggested by the results, presented in chapter 3, of the analysis of urban structure and its evolution, using the spatial density patterns of residence-based variables. First, it would be extremely valuable to consider the corresponding spatial patterns of the location of firms and businesses engaged in various professional, service or manufacturing activities. This can be accomplished using data sources such as industrial and business directories that give information at the firm level. Second, a larger set of case areas should be considered in order to further assess the robustness of some of the patterns seen here, as well as to possibly classify areas in terms of their respective spatial and temporal patterns. In particular, it would be interesting to examine areas such as San Francisco and Boston that have relatively more elaborate spatial patterns that have developed over longer time frames than many other U.S. cities. Third, it would be interesting to capture the characteristics of the transportation network serving the areas under consideration by examining the densities as a function of network distance or travel time instead of Euclidian distances from the CBD. Fourth, it is desirable to integrate the above spatial patterns into an approach to estimate flow patterns and aggregate travel descriptors (such as average trip lengths) resulting from the interaction between the spatial patterns of residence-based variables, and those of the location of firms and businesses. This would provide an important link to the study of the transport implications of evolving urban structures. Such an approach could additionaly incorporate the characteristics of the transport infrastructure.

Regarding the traffic network performance component of the present research, it has now reached a level of maturity which holds considerable promise for possibly very powerful results. A systematic investigation of an expanded list of factors would yield more definitive relations between the parameters of the traffic models and the network's features. A key factor to consider would be the flow (demand) pattern, in order to tie in with the spatial patterns of activities and residences discussed above. Methodologically, simulation experiments using a somewhat more realistic-looking network setting would be useful; interesting opportunities along these lines involve using a supercomputer to simulate larger networks, with more vehicles, over a longer period of time. In addition, combining the simulation-based work with carefully conceived observational field work could yield significant results.

REFERENCES

Adams, John S. (1986), "The Geographical Evolution of U.S. Urban Areas at the National and Local Scales", National Academy of Engineering Workshop on Future Infrastructures, Woods Hole, Massachusetts.

Anderson, John E. (1985), "The Changing Structure of a City: Temporal Changes in Cubic Spline Urban Density Patterns", *Journal of Regional Science* 25, pp. 413-427.

Anderson, John E. (1982), "Cubic Spline Urban Density Functions," *Journal of Urban Economics*, 12, 155-167.

Ardekani, Siamak (1984), "The Two-Fluid Characterization of Urban Traffic: Theory, Observation and Experiment", Ph.D. Dissertation, Department of Civil Engineering, The University of Texas at Austin .

Ardekani, Siamak, Dona, Edgar, Govind, Shekhar and Herman, Robert (1986), "The Dynamic Characterization of Cities", National Academy of Engineering Workshop on Future Infrastructures, Woods Hole, Massachusetts.

Ardekani, Siamak and Herman, R. (1982), "Quality of Traffic Service," Research Report 304-1, Center for Transportation Research, The University of Texas at Austin.

Ardekani, S., Torres-Verdin, V., and Herman, R. (1985), "The Two-Fluid Model and the Quality of Traffic in Mexico City," published in Spanish as "El Modelo Bifluido y La Calidad del Transito en La Ciudad de Mexico", *Revisto Ingenieria Civil,* Calegio de Ingenierus Civiles de Mexico Jan/Feb.

Brotchie, J., Newton, P., Hall, P. and Nijkamp, P. (1985), "Introduction", in *The Future of Urban Form: The Impact of New Technology*, Nichols Publishing Co.

City Directory of Austin (1959 to 1985), pub. R.L.Polk and Co., Dallas, Texas.

City Directory of San Antonio (1900 to 1985), pub. R.L.Polk and Co., Dallas, Texas.

Clark, Colin (1951), "Urban Population Densities", *Journal of the Royal Statistical Society,* A, 114, 490-496.

County and City Data Book (1949, 1956,1962, 1967, 1972, 1977 and 1983), U.S. Bureau of Census, Washington D.C .

Drake, J., Scofer, J. and May, A. D. (1967), "A Statistical Analysis of Speed-Density Hypotheses", *Proceedings of the Third International Symposium on Theory of Traffic Flow*, American Elsevier .

Frankena, Mark W. (1978), "A Bias in Estimating Urban Population Density Functions," *Journal of Urban Economics* 5, 35-45.

General Directory of the City of Austin (1900 to 1958), pub. Morrison and Fourmy, Gavleston, Texas.

Gerlough, D. L. and Huber, M. J. (1975), "Traffic Flow Theory: A Monograph", *TRB Special Report* 165, Washington, D.C. .

Govind, Shekhar, Robert Herman and C. Michael Walton (1986) "Characterizing the Evolution of Cities", working paper, Center for Transportation Research, University of Texas at Austin, August .

Herman, Robert and Montroll,Elliott W., (1972) "A Manner of Characterizing the Development of Countries", *Proceedings of the National Academy of Sciences*, Vol. 69, No. 10, October .

Herman R. and Prigogine, I. (1979), "A Two-Fluid Approach to Town Traffic," *Science* 204, pp. 148-151.

Herman, R. and Ardekani, S. (1984), "Characterizing Traffic Conditions in Urban Areas," *Transportation Science* 18, pp. 101-139.

Lakshmanar, T.R. and Chatterjee, L. (1985), "Technical Change and Metropolitan Adjustments: Some Policy and Analytical Implications", *Regional Science and Urban Economics*, 15.

Mahmassani, H.S., Williams, J.C., and Herman, R. (1984), "Investigation of Network-Level Traffic Flow Relationships: Some Simulation Results," *Transportation Research Record* 971, pp. 121-130 .

McDonald, John and Bowman, H. Woods (1976), "Some Tests of Alternative Urban Population Density Functions," *Journal of Urban Economics* 3, 242-252.

Muth, Richard (1969), *Cities and Housing, Chicago*, University of Chicago Press.

Nijkamp, P. and Schubert, U. (1985), "Urban Dynamics", in Brotchie, et al. (eds.) *The Future of Urban Form*, Nichols Publishing Co.

Peat, Marwick, Mitchell and Company (1973), *Network Flow Simulation for Urban Traffic Control System - Phase II*, Volumes 1-5; prepared for the Federal Highway Administration, U.S. Department of Transportation, Washington, D.C. .

Poirier, Dale J. (1973), "Piecewise Regression Using Cubic Splines", *Journal of the American Statistical Association* 68, 515-524.

Pressman, N. (1985), "Forces for Spatial Change", in Brotchie, et al (eds.), *The Future of Urban Form*, Nichols Publishing Co.

"Traffic Network Analysis with NETSIM - A User Guide" (1980), Implementation Package FHWA-IP-80-3, Federal Highway Administration, U.S. Department of Transportation, Washington, D.C. .

Williams, J. C., Mahmassani, H.S. and Herman, R.(1985), "Analysis of Traffic Network Flow Relations and Two-Fluid Model Parameter Sensitivity", *Transportation Research Record* 1005, pp. 95-106 .

Yeates, Maurice (1976), *The North American City*, Harper & Row, New York .