1-1-2007

Incorporating spatial characteristics in travel demand models

Valerian Kwigizile
University of Nevada, Las Vegas

INCORPORATING SPATIAL CHARACTERISTICS IN TRAVEL DEMAND MODELS

by

Valerian Kwigizile

Bachelor of Science in Civil Engineering
University of Dar Es Salaam, Tanzania
2001

Masters of Science in Civil Engineering
Florida State University, USA
2004

A dissertation submitted in partial fulfillment
of the requirements for the

Doctor of Philosophy Degree in Civil and Environmental Engineering
Department of Civil and Environmental Engineering
Howard R. Hughes College of Engineering

Graduate College
University of Nevada, Las Vegas
December 2007

UMI Number: 3302357

UNLV
UNIVERSITY OF NEVADA LAS VEGAS

Dissertation Approval
The Graduate College
University of Nevada, Las Vegas

17 OCTOBER, 20 07

The Dissertation prepared by

VALERIAN KWIGIZILE

Entitled

INCORPORATING SPATIAL CHARACTERISTICS IN TRAVEL DEMAND MODELS

is approved in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Examination Committee Chair

Dean of the Graduate College

Examination Committee Member

Examination Committee Member

Graduate College Faculty Representative

Examination Committee Member

1017-52

ABSTRACT

Incorporating Spatial Characteristics in Travel Demand Models

by

Valerian Kwigizile

Dr. Hualiang (Harry) Teng, Examination Committee Chair
Assistant Professor
University of Nevada, Las Vegas

The goal of this study was to address one of the major weaknesses of the ubiquitous four-step procedure for travel demand modeling: omission of spatial interactions between the variables. While contiguity of the analysis zones is commonly used to define spatial interaction of the variables in spatial analysis, it might not capture the interactions of travel demand variables. In this study, the efficacies of four alternative methods for defining spatial relationships: contiguity, separation, a combination of contiguity and separation, and economic linkages (accessibility), were evaluated. The home-based-work (HBW) spatial models and non-spatial models for trip attraction, and trip production were developed. For the destination choice, the spatial models were developed by using separation and accessibility alternatives for defining spatial relationship. Comparison of the trip attraction models indicated that the model estimated using the separation spatial relationship had the best fit. Furthermore, comparison of the best spatial model and the non spatial model indicated that the spatial model outperforms the non spatial model by increasing the prediction accuracy by 14%. For the trip production model, the results

indicated that the spatial variable is unnecessary. For destination choice, the spatial model developed using separation spatial relationship was found to be the best based on statistical tests. To compare the spatial model and the non-spatial model, the forecasted alternative destination shares were used. The results indicated that the difference between the forecasted alternative shares by using spatial and non-spatial models is small when there is a small percentage increase in casino/hotel and retail jobs. In order to use the developed destination choice models for long-term forecasting, additional variables such as housing location should be included. Also, since the design of the analysis zones used in this study may not be optimal, an attempt to design new analysis zones through a careful aggregation process in which homogeneity is carefully controlled, is recommended.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

First and foremost, I want to thank my doctoral advisor, Dr. Hualiang (Harry) Teng, who agreed to guide me as I struggled to address one of the major weaknesses of the widely used four-step travel demand modeling approach through advanced econometrics. He was always available to render the support I needed. His advice nourished and improved my dissertation from day to day. I also want to thank my advisory committee members: Prof. Edward Neumann, Dr. Mohamed Kaseko, Dr. David James, and Dr. Sandra Catlin for taking their valuable time to read my work and provide their comments.

Second, I want to thank my family for their moral support and encouragement they provided throughout my studies. Special thanks should go to my wife, Liberatha, and my son, Ivan, for their understanding and unwavering faith in me. They provided all social and moral assistance I needed for my research. I would like to thank my mother, Martha, who always believed in my academic capabilities and gave me words of courage. Also, my father who passed away twenty years prior to my defense established the foundation of my academic success. I also want to thank my siblings for their support.

Lastly but not least, I want to thank the UNLV Graduate College for awarding me summer scholarships for two years in a row. This enabled me to concentrate on my research and hence expedited its completion. I also want to thank my fellow students at the UNLV Transportation Research Center (Spring 2005—Fall 2007) for their support.

CHAPTER 1

BACKGROUND AND RESEARCH MOTIVATION

1.1. Background

Travel demand forecasting models are used by Metropolitan Planning Organizations (MPO) for developing design traffic for preliminary engineering and final design of highway improvements. They are also used for various planning studies such as air quality conformity analysis, major investment studies, congestion management system studies, and long range plan alternatives analysis. Furthermore, the models are used to assess the impacts of socioeconomics, demographics, land-use, and transportation system changes on the performance of the transportation system.

Travel demand analysis and forecasting have developed rapidly over the past three decades. The 1950's rapid increase in car use resulted in major investments in new road infrastructures. This increase called for developing aggregate trip-based models to predict traffic flows between different zones. The developed models followed a four-step procedure: trip generation, trip distribution, mode choice and choice of the route. Weiner (1999) provides the detailed historical overview of the travel demand analysis and the urban transportation planning in the United States. In the 1970's, the focus shifted to regional planning, which include the travel needs of individual persons. As a result, disaggregate trip-based travel demand models, also known as discrete choice models, were developed. For example, Ruiter and Ben-Akiva (1978) developed the disaggregate travel demand models for the San Francisco Bay area. However, similar to aggregate models, disaggregate trip-based models analyze each trip independent of other trips made by the same individual. The major weakness of both the aggregate and disaggregate trip-based models is their focus on individual trips, ignoring the spatial and temporal relationship between all trips and activities completed by an individual as well as other individuals. This implies that the models may produce poor forecasts in the cases where it is important to measure how individual trips relate to each other, both spatially and temporally.

To improve travel demand forecasting accuracy of the widely used traditional four-step procedure, there have been several attempts to identify and address its weaknesses. Together with such efforts, new rule-based models such as activity-based travel demand models and Bayesian networks have been under investigation. Bowman (2000) summarizes the most important elements of activity-based travel theory in two basic ideas. First, the demand for travel is derived from the demand for activities. Secondly, humans face temporal-spatial constraints, functioning in different locations at different points in time by experiencing the time and cost of movement between locations. A linkage exists between activities, locations, times and individuals (McNally, 1996). While new modeling approaches have been attempted, most planning agencies are still using the traditional four-step procedure models for travel demand forecasting because of institutional requirements and financial limitations. Therefore, it remains critical and imperative to address the weaknesses of the four-step procedure model with goals of improving its forecasting accuracy. One of the weaknesses of the procedure is its omission of the possible spatial autocorrelation present in travel demand variables. This is the main focus of this research.

The four-step procedure model consists of four steps: trip generation, trip distribution, mode choice and finally trip assignment. The purpose of trip generation is to determine the number of vehicle- or person-trips to and from the Traffic Analysis Zones (TAZ) under consideration. Trip generation models consist of two types of models: trip-production models and trip attraction models. Trip production is defined as the home end of home-based (HB) trips or as the origin of a non-home-based (NHB) trip while trip attraction is defined as the non-home end of a HB trip or the destination of a NHB trip (Ortŭzar and Willumsen, 2001). The models are usually estimated for different trip purposes such as home-based-work (HBW), home-based-other (HBO) and non-home-based (NHB). The common factors considered for trip generation models are available automobiles per household, income, household size, number of job opportunities available and residential density. The multiple regression analysis is one of the commonly used methods for modeling trip generation. It can be used with both aggregate (zonal) and disaggregate (household and personal) data to estimate trips generated. Cross-classification (category analysis) is another common approach for modeling trip generation.

Trip distribution is the second step in which trip productions and trip attractions for each zonal pair are linked. Destinations for each trip are determined in order to produce origin-destination (O-D) tables. Therefore, trip distribution is essentially about destination choice from which a trip matrix (or trip table) is generated for each trip purpose (McNally, 2000). Gravity models are commonly used in trip distribution and are functions of activity system attributes (indirectly through the generated productions and attractions) and network attributes (typically, inter-zonal travel costs). Due to complexity of interaction between origin and destination, the distribution models have generally been reinterpreted in terms of discrete choice theory; and statistically correct estimation methods are generally used (Hensher and Button, 2000).

Mode choice is the third step used to split the total zone-to-zone trips using each available mode between each zone pair. Very often, the mode choice models are based on logit formulation and a simple multinomial logit technique is one of the methods used for estimation. The last step is traffic assignment in which the distribution of traffic in a network considering a demand between locations and the transport supply of the network is performed. The person-trips are converted to vehicle-trips by applying occupancy rates prior to network assignment. Generally, the assignment methods are looking for a way to model the distribution of traffic in a network according to a set of constraints, notably related to transport capacity, time and cost (Rodrigue, *et al.*, 2006). Figure 1.1 is a schematic presentation of the travel demand four-step procedure.

1.2. Research Motivation

Modeling of trip generation and destination choice is based on observations made at different locations. For trip generation models, the zonal trip totals are used as response variables, while for destination choice models, the zones are used as origins and alternative destinations. Since the zones are geographical entities, their spatial relationship (relative locations) require consideration in the models. Specifically, this research is inspired by the theory that zonal trip generation totals are spatially distributed and hence affect each other.

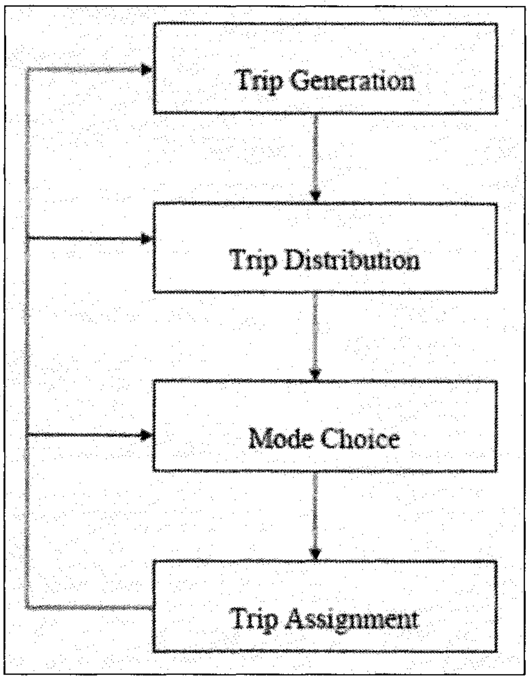


Figure 1.1. Sequence of the four-step procedure

Most linear regression trip generation models of the traditional four-step procedure do not incorporate exclusive variables to account for spatial autocorrelation possibly present in travel demand observations. If the observations of the variables such as zonal trip totals are spatially correlated, they constrain the possible analyses which can be applied to those observations and influence the final conclusions that can be reached. A fundamental geographical concept is that nearby observations often share similarities than observations which are far apart. This leads to a concept of presence of spatial dependence in zonal trip total observations. It simply suggests that characteristics of proximal locations (zones) might be spatially correlated, positively or negatively.

Significant positive spatial correlation means that similar observations are clustered while significant negative spatial correlation indicates that neighboring observations are more dissimilar. Spatial correlation means that whatever is contributing to trip generation totals in one zone also causes similar observation in nearby zones. For example, trip generation totals in nearby zones might be similar due to factors such as socioeconomic status and type of land-use. Spatial dependence in zonal trip generation total observations leads to spatial autocorrelation problem in statistics, similar to temporal autocorrelation in time series data. Spatial autocorrelation violates standard statistical assumptions of independence among the observations. This calls for investigating ways of incorporating the effects of spatial autocorrelation possibly present in travel demand variables. The fundamental question is how to quantify spatial relationship of the observations.

Spatial location of travel demand variables can be considered in step one of the four-step procedure, which is trip generation, and in step two, which is trip distribution, or destination choice. For step one, consider an urban area divided in ten zones as shown in Figure 1.2. It shows clearly that zonal aggregated trip totals are observations taken at different geographical locations.

There are two possible regression analysis approaches for estimating the impact of land-use, socioeconomics and demographic variables on the zonal trip totals. In the first approach, which is commonly used, planners estimate linear models without explicitly considering spatial location of the observations which might be inducing spatial dependence in the observations. However, if there is spatial dependence, the models estimated with this approach are prone to poorly estimating the observed values in some locations.

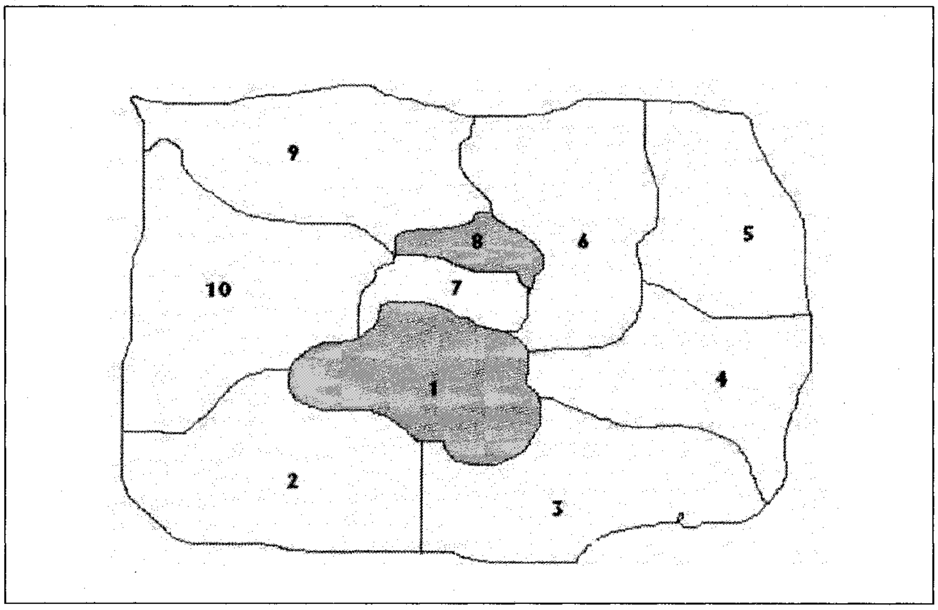


Figure 1.2. Example of zoned urban area

In the presence of spatial autocorrelation, the second approach, which is using spatial models, is necessary in order to account for the variation between observed quantities at different locations (Haining, 2003). Spatial dependence in a collection of sample data observations such as number of trips per zone refers to the fact that one observation associated with a specific location is spatially related to another observation at another location. Information loss occurs when two spatially dependent observations are made. Also, spatial dependence of the observations leads into wrong conclusions since it makes the test statistics invalid.

Consider a linear model of zonal trips, Y_i, and explanatory variable, X_i, which can be written as:

$$Y_i = \alpha + \beta X_i + \varepsilon_i, \tag{1.1}$$

in which ε_i is the random error. The test statistic for the hypothesis testing of the coefficient, β, can be given as:

$$t = \frac{\hat{\beta}}{\sqrt{Var(\hat{\beta})}}, \tag{1.2}$$

in which $\hat{\beta}$ is the estimated coefficient, and $Var(\hat{\beta})$ is the variance of the estimated coefficient. Standard statistical textbooks (for example Pindyck and Rubinfeld (1998), Gujarati (2003) and Greene (1997)) have shown that the variance of the estimated coefficient can be computed as follows:

$$Var(\hat{\beta}) = E\left[(c_1\varepsilon_1)^2 + 2(c_1c_2\varepsilon_1\varepsilon_2) + (c_2\varepsilon_2)^2 + 2(c_2c_3\varepsilon_2\varepsilon_3) + \ldots + (c_N\varepsilon_N)^2\right], \tag{1.3}$$

in which $c_i = \frac{1}{\sum_{i=1}^{N}(X_i - \overline{X})^2}$. Under the assumption that the observations are not correlated, that is $E(\varepsilon_i\varepsilon_j) = 0$ for $\neq$, the variance can be estimated as:

$$Var(\hat{\beta}) = c_i\sigma^2, \tag{1.4}$$

in which $\sigma^2 = var(\varepsilon_i^2)$. However, if spatial correlation exists in the observations, $E(\varepsilon_i\varepsilon_j) \neq 0$ and the variance can be estimated as follows:

$$Var(\hat{\beta}) = c_i\sigma^2 + 2\sum_{i,j} c_ic_j.Cov(\varepsilon_i, \varepsilon_j), \tag{1.5}$$

It can be clearly seen from Equation (1.5) that if there is positive spatial correlation and the ordinary estimation method which uses Equation (1.4) is applied, the variance will be underestimated. The effect of this underestimation is the overestimation of the test statistic in Equation (1.2). With spatial dependence in the observations, the risk of committing a type I error is increased, meaning that the probability of rejecting the null hypothesis when it is true is greater than the nominal value selected for the hypothesis test.

For the destination choice models, the choice set is based on Traffic Analysis Zones (TAZ). Since the TAZs are geographical units, their spatial relationships may have an impact on individuals' destination choices. It is equally important to consider the individuals' socioeconomics, demographics and spatial relationship of origin and destination to more accurately study their destination choice decision making. Taste variation of destination zones may result not only from individuals' and land-use characteristics, but also spatial relationships of the origin and destination. The spatial relationship may further depend on zonal characteristics such as shape and size. Therefore, ways to quantify spatial dependence in discrete choice models of destination choice requires exploration.

1.3. Problem Statement

The accuracy of the forecasts which are produced by a transport demand model depends both upon the accuracy with which its inputs (or "planning variables") can be forecast and on the analyst's knowledge of how the coefficients and parameters of the model vary through time and space. Time variation can be captured by developing time series models. Although spatial relationships of the analysis zones might have an effect on trip generation and destination choice decisions, they are not investigated explicitly in most of the existing four-step models. In a recent study, Gamas *et al.* (2005) estimated trip generation in Mexico City using spatial effects and urban densities. The coefficients for work, shopping and school trip generation models were estimated. To correct the bias in the coefficients estimated, spatial regression was used instead of the traditional ordinary least squares (OLS) estimation method. The models were run for both trip productions and trip attractions for all trip types. The spatial relationship of the observations was defined by using a contiguity matrix in which a value of one was assigned to neighboring zones and zero otherwise. Most models showed significant spatial autocorrelation, indicating that if they had not been specified accordingly, the coefficients would have been biased.

Using contiguity of the zones as the major criterion for defining spatial relationship of the observations may not explain spatial relationship correctly. Different definitions of spatial relationship may produce different model results. For example, the shape and size of the zones may have an impact on their spatial relationship when the two zones are barely contiguous but extend to opposite directions. Difference in the size of the zones may have a similar effect of reducing the spatial relationship of the zones. Such situations may invalidate the assumption that the zones are spatially neighbors. Figure 1.2 illustrates this by showing the possible effect of a combination of the sizes and shapes of the zones in defining their spatial relationships. For example, if using binary contiguity criterion only to define the spatial relationship, zones 1 and 8 would be deemed spatially unrelated while zones 1 and 6 would be deemed spatially related. However, zone 8 may be actually more spatially close to zone 1 than zone 6 is close to zone 1. Similarly, if the analysis zones are separated by man-made or natural features such as a freeway or a river, it is possible to have unrelated zones defined as spatially related.

Also, home-based work trips are normally viewed as trips involving long–term decisions such as job location and residential location. Since most job locations are fixed, workers have no alternative choices and therefore do not really make a destination choice for such trips. Some past studies that investigated the effect of spatial location in destination choice models concluded that work trips are fixed in space and therefore it is not possible to apply probabilistic models to explain this type of activity (e.g. Hamadou, *et al.*, 2004). However, this assumption may not be applicable to all urban areas, especially those in which numerous comparable job opportunities exist and are distributed over the area. Spatial relationship of the origin and the work destination may have an impact in the long-term decision of where to work.

The Las Vegas valley is a unique urban area in which the major employment industry (hotel/casino) provides numerous comparable job opportunities around the area. Based on the Las Vegas land-use data, hotel and casino employment constitute 30% of all employment opportunities available in the valley. Apart from the resort corridor, popularly known as "the strip", several comparable hotel/casino employment opportunities are available and distributed in other areas around the valley. Consequently, more than 50% of jobs are available from a combination of hotel/casino and retail employment opportunities. The retail job opportunities are also spatially distributed over the valley. Therefore, a high percentage of workers in the valley have wide choices of comparable employment opportunities at different locations. Therefore, spatial location of the job relative to the residence of the worker may be an important factor in estimating long-term trip generation as well as destination choice models for the valley. Alternative methods of incorporating the spatial relationship of origin and destination for urban areas with setting like Las Vegas valley require examination. Most models developed assume that the work location is fixed and therefore the choices for residence are made. These models are called residential choice models. Conversely, the assumption in this study is that individuals with fixed residences face a wide set of choices for working locations. Therefore, commuters make long-term travel and destination choice decisions by considering their residential location relative to their potential work destinations(s).

1.4. Research Objectives

The general objective of this research was to investigate the effect of spatial relationship of the observations for trip generation and destination choice models of the four-step procedure. More specifically, the objectives can be grouped into four categories:

- To test the presence of spatial autocorrelation in travel demand variables for trip generation and destination choice models.
- To develop a methodology for defining a spatial variable necessary to account for spatial autocorrelation and ways to incorporate it in the spatial model.
- To evaluate the efficiency of alternative methods for defining spatial relationship of the observations in trip generation and destination choice models.
- To compare the estimated spatial models for trip generation and destination choice with corresponding non-spatial models.

To fulfill the objectives, the home-based-work trips as reported in the 1996 Las Vegas Household Travel Survey conducted by the Regional Transportation Commission of Southern Nevada (RTC) were used to investigate the effect of spatial location of travel demand variables in long-term (work trip) trip-making and destination choice decisions. In the first part, trip generation linear regression models were developed by explicitly incorporating spatial relationship of observations in order to investigate its effect on the number of trips generated. Alternative approaches for quantifying the spatial relationship were examined. In the second part, the individual demographics and socioeconomics, as well as the characteristics and spatial relationship of the origins and destinations were incorporated in the discrete choice models to investigate their effects on destination choice decision making process. Alternative approaches for quantifying the spatial relationship were also examined. Results of the spatial models were compared to those from the non-spatial models.

1.5. Significance of the Research

Most United States metropolitan planning organizations (MPO), including the Regional Transportation Commission of Southern Nevada (RTC), use the four-step procedure model for forecasting future traffic necessary for preliminary engineering and roadway design of highway improvements. The forecasts are also used for air conformity analysis, major investment studies, congestion management system studies, and long range plan alternatives. The accuracy of the forecasts produced by these models is of paramount importance to planners for the benefit of roadway network users. Understanding the effect of spatial relationship of the observations is required to improve the forecasting accuracy of the models. More importantly, identification of the most effective method for quantifying spatial relationship is needed. The results of this study can be used to explain the effect of spatial relationship of observations in explaining variability. In addition, the results of the study can establish basis for similar studies in urban areas with numerous jobs distributed over the area.

The developed spatial trip generation and destination choice models are expected to be used by transportation planners to forecast future travel demand more accurately. The new models can be integrated in the existing modeling platforms such as TransCAD used by the Regional Transportation Commission of Southern Nevada (RTC) to replace the existing non-spatial models. TransCAD is a package manufactured by Caliper Corporation that fully integrates Geographic Information Systems (GIS) with planning modeling and logistics application.

1.6. Scope and Limitations of the Research

Although the 1996 Las Vegas Household Travel Survey data contains different trip types, this study focused on home-based work (HBW) trips to fulfill the research goals. Aggregated models for trip attraction and trip production were used due to their simplicity in estimation as well as interpretation of the results. In addition, the disaggregate destination choice model was used to investigate the effect of spatial location on individuals' work location decision making. It should also be noted that spatial models use zonal neighborhood information to investigate the effect of spatial autocorrelation. This information is subject to change due to administrative and planning requirements. This may affect forecasting capability of such models since the weight matrix defining neighborhood of the zones was based on the zonal structure. The matrix used for forecasting would be misleading if the urban area is restructured by redefining the boundaries of the zones. The districts created by the Regional Transportation

Commission of Southern Nevada (RTC) by aggregating homogeneous Traffic Analysis Zones (TAZs) were used in this study. The districts may not be the optimal zones designed to eliminate the modifiable areal unit problem (MAUP). In order to create structure-insensitive spatial variables, it is important to have as much control over the configuration of the areal units as possible.

1.7. Dissertation Outline

Chapter two of this dissertation provides a review of travel demand models. The modeling approaches for trip generation and destination choice are presented together with their strengths and weaknesses. Past modeling efforts attempting to address the weaknesses are also summarized. Chapter three details the methodology followed in this study. Specifications of the spatial trip generation and destination choice models are given. The alternative methods for quantifying spatial relationship of the observations are also presented. Data collection for this study is presented, followed by selection of the variables for modeling. The tests for spatial autocorrelation in the selected variables are described followed by model estimation methods. A detailed method for interpreting the results is given followed by a presentation of methods used for comparing different models. The chapter is concluded by outlining the methodology for creating Origin-Destination (O-D) matrices using the results of the multinomial logit model of destination choice. Chapter four details the results obtained by implementing the methodologies explained in Chapter three. Chapter five provides the conclusions and recommendations for future work.

CHAPTER 2

REVIEW OF TRAVEL DEMAND MODELING

2.1. Overview of Travel Demand Modeling

Typically, models developed in the travel demand modeling system integrate insights from the travelers' psychology with neo-classical economic theory. They relate socio-economic travel demand attributes and level of service or supply variables. Microeconomic demand theory, which is concerned with the interaction between buyers (in this case travelers) and suppliers (in this case the set of all attributes of transportation that have a bearing on the quantity and nature of transport activities that actually take place), underlies travel demand modeling. However, direct application of the microeconomic demand theory to transportation faces methodological difficulties such as specification of appropriate demand models and experimentation needed to validate these models. Individual travel behavior is subject to many more uncertainties than other consumption activities. Factors that affect the demand for and the supply of transportation are numerous and mutually dependent. The demand for travel is derived and takes place over space, while transport supply is a service and not a good. Therefore, in modeling travel demand, the socio-economic travel demand attributes and level of service or supply variables are used simultaneously.

Demand for travel includes many choices made by individuals for trip making such as how many trips to make during a given period of time, what destination to choose, what mode of travel to use, and which route to take for the trip. Modeling requires a clear understanding of whether these choices are made independently of others, and if they are dependent, whether they are made simultaneously or sequentially in some order. It is not possible to include all variables that can possibly influence the individual choices, and thus necessitate a model that can incorporate the stochastic nature of choice probabilities to take into account the missing variables. Stochastic choice probabilities can be measured at an aggregate market level, indicating for each alternative the probability that it is the choice of an individual chosen at random from the population when the sampling is defined over a given period of time during which it is postulated that a choice will be made (Kanafani 1983). At the disaggregate level, the choice probability of an alternative defines the number of times a choice is made if the individual faces the same choice environment.

2.2. Travel Demand Modeling Approaches

There are two basic approaches for modeling travel demand: the structured (sequenced) choice model approach, and the direct approach. Both of these approaches can be applied at the individual level (disaggregate) and the market level (aggregate). In the aggregate level, data items represent an average over a group of travelers while at disaggregate levels, the data is collected specifically for a single individual. For aggregation, the area in consideration is initially divided into zones, popularly known as Traffic Analysis Zones (TAZ).

In the sequenced choice model approach, the choice processes are modeled explicitly in order to forecast the number of trips made by an individual during a given period of time for specific purpose, from origin to destination, by specific mode, route, and at a specific time of the day. The typical sequence used is estimating the number of trips generated (*trip generation*), distributing the trips generated among the available destinations (*trip distribution*), distributing the trips among the mode alternatives available (*mode choice*), and finally distributing all trips among the available routes (*assignment*). This method is commonly referred to as the four-step procedure or Urban Transportation Planning System (UTPS).

In the direct approach, the concepts of microeconomic demand modeling are applied to derive the number of trips demanded by individuals as a function of demand and supply characteristics in a single function. Between the sequenced and direct approaches, the first two steps of the traditional four-step procedure is the focus of this research. The following sections detail modeling approaches for step one (trip generation) and two (trip distribution or destination choice). Previous efforts for addressing weaknesses of the modeling approaches for the two steps are also documented.

2.3. Classification of Trips

Ortŭzar and Willumsen (2001) categorize trips into Home-based (HB) and Non-home-based (NHB) trips. The home-based (HB) trip is the one where the home of the trip maker is either the origin or the destination of the trip, while, the non-home-based trip is the one where neither end nor origin of the trip is the home of a traveler. Trip production is defined as the home end of an HB trip or as the origin of a NHB trip, while trip attraction is defined as the non-home end of an HB trip or the destination of a NHB. Figure 2.1 which is reproduced from Ortúzar and Willumsen (2001) shows the difference between trip attraction and trip production as related to origin and destinations.

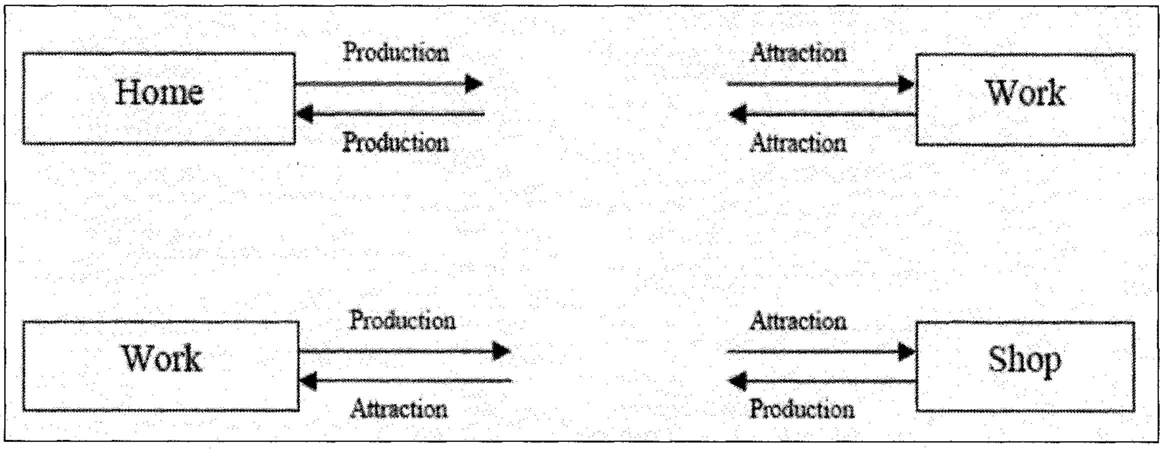


Figure 2.1. HB and NHB trip attraction and trip production

Furthermore, trips can be classified in three major ways: purpose, time of the day and person-type. Trips by purpose are further divided into two categories, which are compulsory (or mandatory) trips and discretionary (or optional) trips. Compulsory trips involve work and trips to school or college, while discretionary trips involve shopping, social and recreational activities and all trips not associated with work or school. For analytical purposes, non-home-based (NHB) trips are normally not separated because they only amount to 15 – 20% of all trips. Classification of trips by time of the day proportions trips according to peak and off-peak period trips. The proportion of journeys by different purposes usually varies greatly with time of the day. Classification of trips by person-type views individual travel behavior as heavily dependent on the socioeconomic attributes of individuals. Three major categories of trip classification by person-type are:

- by income level,

- car ownership, and
- household size and structure.

2.4. Models for Analysis of Trip Generation

Trip generation is the first and basic step in the conventional four-step procedure for travel demand forecasting. It is the step in which models forecasting the number of trips resulting from land-use of a specific area (zone) as well as demographic and socio-economic characteristics of the travelers are developed. Specifically, it predicts the number of trips originating or destined for a particular area (zone). Analysis in the main trip generation is focused on residences and is considered a function of the social and economic attributes of a household. Additionally, trip generation is particular to land-uses of different zones. The zones also comprise destinations of trips, thus called trip attractors. The analysis of trip attraction focuses mainly on non-residential land-uses.

The trip generation model estimates the number of motorized person-trips to and from each zone in the area of study. By looking at the number of trips aggregated over all modes, destinations, and routes, the implied assumption is these numbers represent an equilibrium between the demand for transportation and the supply conditions prevailing in the transportation system at the time observations are made (Kanafani 1983). Forecasting trip patterns using trip generation models imply that either the demand for transport is inelastic with respect to supply conditions, or the supply conditions will not change significantly between the time of analysis and the time of forecast. Supply conditions refer to all attributes of transportation, such as travel time, operating costs and delays in transit, that have an influence on the quantity and nature of transport activities that actually take place.

Trip production analysis predicts all trips generated or produced at an inhabited or activity location, often aggregated over a group of households or individuals in a zone. Key factors for trip productions include car availability (ownership), income, household structure, family size, residential density, accessibility, and value of land. Trip attraction analysis predicts all trips attracted to specific areas (zones) where the trip purpose can be fulfilled. Key factors affecting trip attractions include zonal employment in different sectors and measure of accessibility.

There are two major approaches for modeling trip generation: regression analysis and cross-classification (or category) analysis. Regression analysis involves setting up models, usually linear, that relate the trips generated to relevant explanatory variables such as socioeconomic characteristics of trip makers. The major drawback of linear regression analysis is the possibility of non-linearities among the independent variables. However, this can be overcome by transforming the variables into a suitable form. In cross-classification, all households are grouped into homogeneous groups on the basis of socioeconomic characteristics. The number of trip productions per household for a given trip purpose is estimated as a function of household attributes. One of the advantages of cross-classification is that there is no need for an assumption of linearity between independent and dependent variables. However, the approach requires detailed data to construct and predict trip generations.

2.4.1. Multiple Regression Models for Trip Generation

In regression models, the dependent variable is assumed to be a function of a series of independent variables. The multiple regression model is of the form shown in Equation (2.1).

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + \varepsilon_i \tag{2.1}$$

where

$\beta_0, \beta_1, \ldots, \beta_k$ are regression coefficients,

$x_{1i}, x_{2i}, \ldots, x_{ki}$ are k independent (explanatory) variables

Y_i is the dependent (response) variable

ε_i is the stochastic error term

$i = 1, 2, \ldots,$ and N: indexes N sample observations.

The parameters of Equation (2.1) can be estimated by two methods: (1) Ordinary Least Squares (OLS) and (2) Maximum Likelihood (ML). The method of OLS is used often, primarily because it is intuitively appealing and mathematically much simpler than the method of Maximum Likelihood. Gujarati (2003) mentions OLS assumptions made regarding Equation (2.1) as follows:

i. The regression model is linear in parameters,

ii. Explanatory variables (x_i) are non-random,

iii. Zero mean value of the stochastic error, ε_i, that is, $E(\varepsilon_i / x_i) = 0$,

iv. Homoskedasticity or equal variance of stochastic error term, ε_i, that is, $E(\varepsilon_i^2 / x_i) = \operatorname{var}(\varepsilon_i / x_i) = \sigma^2$,

v. No autocorrelation between the error terms, that is, $\operatorname{cov}(\varepsilon_i, \varepsilon_j) = 0$,

vi. Zero covariance between error term ε_i and explanatory variable x_i; that is, $\operatorname{cov}(\varepsilon_i, x_i) = E(\varepsilon_i x_i) = 0$,

vii. The number of observation, N, is greater than the number of parameters to be estimated,

viii. The x_i values in a given sample are not the same,

ix. The regression model is correctly specified, and

x. There is no perfect multicolinearity among the explanatory variables.

Regression analysis is concerned with estimating the parameters of Equation (2.1). Consider a regression model with three variables: two independent variables and one dependent variable as shown in Equation (2.2).

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i \tag{2.2}$$

The coefficient β_1 measures the change in dependent variable Y resulting from unit change in x_1 given that the variable x_2 is held constant, while the coefficient β_2 measures the change in Y associated with a unit change in x_2 given that the variable x_1 is held constant. The objective of least-square estimation is to find the values of β_0, β_1 and β_2 which minimize the sum of the squared deviations of the observations from the fitted line. Mathematically, the least square criterion can be represented as:

$$\text{Minimize} \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2 \tag{2.3}$$

where $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i}$

Assuming that there are more observations than the parameters to be estimated, three parameters in this case, and that the underlying equations are independent, standard textbooks (Pindyck and Rubinfeld 1998, Gujarati 2003, Greene 1997) show that

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x}_1 + \hat{\beta}_2 \bar{x}_2 \tag{2.4}$$

where

$$\bar{Y} = \frac{\sum_{i=1}^{N} Y_i}{N}, \quad \bar{x}_1 = \frac{\sum_{i=1}^{N} x_{1i}}{N} \text{ and } \bar{x}_2 = \frac{\sum_{i=1}^{N} x_{2i}}{N}$$

Further computations and simplifications indicate that

$$\hat{\beta}_1 = \frac{\left(\sum x_{1i} y_i\right)\left(\sum x_{2i}^2\right) - \left(\sum x_{2i} y_i\right)\left(\sum x_{1i} x_{2i}\right)}{\left(\sum x_{1i}^2\right)\left(\sum x_{2i}^2\right) - \left(\sum x_{1i} x_{2i}\right)^2} \tag{2.5}$$

$$\hat{\beta}_2 = \frac{\left(\sum x_{2i} y_i\right)\left(\sum x_{1i}^2\right) - \left(\sum x_{1i} y_i\right)\left(\sum x_{2i} x_{1i}\right)}{\left(\sum x_{1i}^2\right)\left(\sum x_{2i}^2\right) - \left(\sum x_{1i} x_{2i}\right)^2} \tag{2.6}$$

In order to statistically determine if the dependent variable is at all related to the explanatory variable, a "zero" null hypothesis test is conducted. If the errors are normally distributed, a *t*-ratio is computed as follows:

$$\frac{\beta}{se(\hat{\beta})} \sim t_{N-k} \tag{2.7}$$

where $se(\hat{\beta})$ is the estimated standard error of estimator.

To measure the proportion of variation in the dependent variable that is explained by the multiple regression equation, the multiple coefficient of determination, R^2, is determined as follows:

$$R^2 = \frac{\hat{\beta}_1 \quad y_i x_{1i} + \hat{\beta}_2 \quad y_i x_{2i}}{\sum y_i^2} \tag{2.8}$$

The value of R^2 lies between 0 and 1. The fit of the model is said to be better as the value is closer to 1. It should be noted that increasing the number of independent variables in a model can never decrease the coefficient of determination, but will likely increase it. Therefore, when comparing the goodness-of-fit of two models with the same dependent variable, an adjusted coefficient of determination should be used. This takes into account the number of independent variables present in the model. It can be computed as follows

$$\overline{R}^2 = 1-(1-R^2)\left(\frac{N-1}{N-k}\right) \tag{2.9}$$

where $\overline{R}^2$ is the adjusted value.

In order to find out if, jointly, the explanatory variables do explain the variation of the dependent variable about its mean, the F statistic is used. It is basically used to test the significance of the coefficient of determination, R^2, and is calculated as follows

$$F_{k-1,N-k} = \left(\frac{R^2}{1-R^2}\right)\left(\frac{N-k}{k-1}\right) \tag{2.10}$$

However, it should be noted that, if the independent variables are highly correlated with each other, the F test of the significance of a regression equation may allow for rejection of the null hypothesis even though none of the regression coefficients are found significant according to their individual t-ratio tests. It is therefore important to make sure that independent variables used in a multiple regression equation are not highly correlated.

2.4.2. Cross-Classification (Category analysis)

Rather than grouping households spatially (i.e., by zones) as in regression models, cross-classification analysis groups individual households according to common socioeconomic characteristics (auto-ownership level, income, household size, etc.) to create relatively homogeneous groups (Meyer and Miller 2001). The basic assumption of this method is that trip generation rates are relatively stable over time for certain household stratifications. Suppose $t_p(j)$ is the average number of trips with purpose p and made by household of type j. The standard method for computing the rate for each cell is to group households in the calibration data to the individual cell groupings and sum the observed trips $T_p(j)$ by group purpose. To obtain the rate for each cell, the total number of trips in that cell is divided by its number of households as shown in Equation (2.11). Table 2.1 presents an example of cross-classification analysis.

$$t_p(j) = \frac{T_p(j)}{H(j)} \tag{2.11}$$

2.4.3. Previous Efforts in Modeling Trip Generation

Trip generation is a key component of any transportation demand modeling system. Both regression analysis models and cross-classification methods have been applied by different transportation planning agencies as well as researchers. However, the developed models do not address the issue of potential spatial autocorrelation that may produce inefficient estimates. Researchers and transportation planning agencies often use linear regression models due to their ability to test a number of variables thought to affect tripmaking behavior and ability to statistically select those which are proven more important. Effects of changes in individual and household travel behavior are assessed using coefficients of regression models, or by cross-classification of households and individuals on a few variables such as income, household size, number of workers in a household, and car ownership. The following are selected previous efforts to address some weaknesses of trip generation models.

Table 2.1. Example of cross-classification analysis

a) Number of households (HH) and trips made

	Available Vehicles					
	0		1		2+	
HH size	No. of HH	Trips	No. of HH	Trips	No. of HH	Trips
1	354	531	236	689	387	875
2	1324	2164	1132	1965	1654	2897
3+	1256	1987	987	1356	1124	1978

b) Households trip rates

	Available Vehicles		
HH size	0	1	2+
1	1.50	2.92	2.26
2	1.63	1.74	1.75
3+	1.58	1.37	1.76

c) Forecasted number of households in one zone

	Available Vehicles		
HH size	0	1	2+
1	24	42	8
2	10	51	107
3+	14	48	467

d) Forecasted number of trips from this zone

	Available Vehicles		
HH size	0	1	2+
1	36	123	18
2	16	89	187
3+	22	66	822
Total	74	277	1027

White (1976) examined the residual distributions in Ordinary Least Squares (OLS) household-based trip generation models using data obtained from Western Midlands Rural Travel Survey (1972). An OLS trip generation model was developed from household data and the residual distribution obtained from a typical estimated trip equation was examined in order to assess conformity with the assumptions of ordinary least squares. It was shown that the residual distributions obtained from typical OLS household-based trip generation models do not comply with the asssumptions inherent in the use of ordinary least squares. In particular, the author suggested that heteroskedasticity is likely to occur and therefore invalidate the estimation of coefficient standard errors by OLS procedure. Heteroskedasticity refers to violation of the assumption that error term has a constant variance. However, the author did not investigate the effect of spatial dependence.

Stopher and McDonald (1983) describes an alternative methodology for calibrating cross-classification models, namely multiple classification analysis (MCA). This technique was demonstrated to be able to overcome most of the disavantages normally associated with standard cross-classification calibration techniques. The method was based on analysis of variance (ANOVA), which provides a structured procedure for choosing among alternative independent variables and alternative groupings of the values of each independent variable. This procedure was contrasted with standard procedure for cross-classification that estimates cell values by obtaining the average value of the dependent variable (e.g., trip rate) for those samples that fall in the cell and are unable to use any information from any other cell. The process of selecting independent variables and selecting groupings of the chosen variables by ANOVA was illustrated with a case study. The degree to which there is statistical information provided to guide the analyst's judgment was shown. The results showed that the best household grouping is one that combines two- and three-person households. Also, it was shown that the MCA procedure allows trip rates to be computed for some cells that are empty of data, and it removes some possibly spurious rates that arise in the conventional method from small sample problems in some cells. The authors concluded that MCA provides a strong methodology for cross-classification modeling and that the procedure is effective in surmounting most of the drawbacks of conventional estimation of such models. However, as it is the case for conventional cross-classification analysis, the MCA procedure does not explicitly incorporate spatial dependence that might be present.

Said and Young (1989) discussed the General Linear Model (GLM) framework as an alternative statistical method for estimating work trip rates for households in Kuwait. The authors suggested that the framework provided a flexible range of statistical models for respresenting the dependence of mean household trip rates on explanatory variables of interest and for selecting the distribution of trip rates of households within individual classification cells. Seven different household major groups were identified from the 1985 census for Kuwait. One of these groups, Kuwaiti households living in villas, was used for some illustrative GLM analysis in which the results of an extensive home interview survey conducted in 1988 were utilized. The analysis showed that work trip rates of this household group were influenced by car ownership, household size, and an interactive effect of these two variables. However, none of spatial variables was considered in the model developed.

Jacobson (1982) proposed two models as alternatives to the ordinary least square model: an integer-dependent variable model and an error-component model of a time-series of cross-sections. The only difference across the proposed alternative models occured in the assumption regarding the distribution of random disturbances. It was hypothesized that significant differences would result in the model forecasts if the appropriate *a priori* distributional assumption were chosen for the disturbances, without changes in the explanatory variables. The data used for this analysis was a 23-day diary of shopping travel by able bodied elderly individuals in Lawrence, Massachusets. The findings suggested that, when models are developed that consider explicitly the discrete nature of the daily trip generation variable (i. e., the number of trips taken by an individual on a given day), forecasts which are not significantly different from the ordinary least squares forecasts were obtained. However, spatial dependence was not investigated in this work.

Doubleday (1977) attempted to assess the temporal stability of one type of person trip generation model (the category analysis model with the individual as the behavioral unit) by conducting comparative studies based on the Reading Travel Surveys of 1962 and 1971 data. Categories were defined with respect to a subset of the variables: employment status, socio-economic group, household structure, car availability and household car ownership. Particular attention was given to optional trips made by persons in certain employment status groups, notably housewives and retired persons. The differences in trip rates between the two years were tested for statistical significance. The temporal variation of trip rates were found to be dependent to a certain extent on the scheme of categorization adopted. However, due to methodological and data limitations, the study concluded tentatively that the trip rates of certain groups in the population are susceptible to variation in response to changes in the level of accessibility. The author did not assess the possible effect of spatial variation.

Meurs (1990) presented and estimated a number of models describing the correlation of trip making over time. Unobserved heterogeneity was taken into account using random effects. The author pointed out that one of the reasons for potential problems may be the omission of variables in the model. If the omitted variables are correlated with the included explanatory variables, model coefficients will be biased. Part of the effects of the omitted variables will be captured by the coefficients associated with the included variables. Furthermore, the author pointed out important omitted variables often as life-styles, tastes and preferences, and sometimes residential and spatial characteristics. The basic models considered in this study were the serial correlation and the state-dependence models. Trip making in total and by transit was best described by using state-dependence models, while trip making by car was described by a model with lagged exogeneous variables. The generalized methods of moments were used for estimation of the models because it is asymptotically efficient and does not require assumptions about initial conditions. However, the author concentrated on temporal variations only.

2.5. Models for Trip Distribution

Trip distribution modeling is an important step for any travel demand model system because trip makers in an urban area normally face a number of destinations for trips of different purposes. There are two types of trip distribution processes (Kanafani 1983): long-term process, and short-term process. Distribution of home-to-work trips is an example of long-term process while convenience shopping trips is an example of short-term trip distribution process. The first is a process that is stable and changes only in the long run, either by the change of residential location or of employment. Very often, the destination for work is defined by location of the work. For example, work destination for an instructor is fixed and defined by the location of the school. However, if a worker faces similar work opportunities at different locations, choice of working location is subjected to consideration of specific attributes of all available alternatives. The second process is more random in its nature, for it is possible that a trip maker may change the destination of even a regular shopping trip, from day to day. The first process is the focus of this research.

In the trip distribution step, models for forecasting destination choices are developed. There are three categories of trip distribution models: (1) physical models of spatial interaction, (2) choice models, and (3) origin-destination demand models. Each of these categories of models is introduced below.

2.5.1. Physical Models of Spatial Interaction

In the physical models of spatial interaction, gravity models are typically used. They start with assumptions about group trip-making behavior and the way this is influenced by external factors such as total trip ends and distance traveled. They were originally generated from an analogy with Newton's gravitational law. Casey (1955) suggested an approach to synthesize shopping trips and catchment areas between towns in a region using the following functional form:

$$T_{ij} = \frac{\alpha P_i P_j}{d_{ij}^2} \tag{2.12}$$

where P_i and P_j are the populations of the towns of origin and destination, and d_{ij} is the distance between the two towns while α is a proportionality factor. Decreasing functions were further assumed that they could better model the effect of distance apart by representing the disincentive to travel as cost increases. The modified gravity model takes the following form (Meyer and Miller 2001):

$$T_{ij} = \frac{P_i A_j f_{ij} k_{ij}}{\sum_{\text{all } j,\, j \neq i} A_j f_{ij} k_{ij}} \qquad (2.13)$$

where

P_i = total number of trips produced in zone i

A_j = total number of trips attracted to zone j

f_{ij} = friction factor (deterrence function)

k_{ij} = adjustment factor for used when there are pairs of zones which have a special relationship in terms of trip making.

Popular versions of the 'deterrence function' include:

Exponential function: $f_{ij} = e^{-\beta C_{ij}}$ (2.14)

Power function: $f_{ij} = C_{ij}^{-\phi}$ (2.15)

Combined function: $f_{ij} = C_{ij}^{\phi} e^{-\beta C_{ij}}$ (2.16)

where

C_{ij} is the generalized cost between origin i and destination j

ϕ and β are calibration parameters to be estimated

Calibration of the gravity model involves estimating the parameters such that the model closely reproduces the base year trip patterns. For example, the initial value for β can be 'borrowed' or guessed and the gravity model run repeatedly to produce trip patterns that are compared to the base year trip patterns until they are close. Although the gravity model is widely used for trip distribution, it has its own shortcomings. Butler (1972) mentioned one of the theoretical problems of the gravity models as the lack of theoretical base. The model deals with the theory of individual movement behavior at an aggregate level without having first aggregated behavior patterns. In addition to lack of ideal test of closeness of fit, the gravity models have may give less accurate long-term forecasts. A goodness of fit of the models is accepted when the calibrated trip distribution is statistically close to the observed trip distribution for the base year. However, comparison of trip length frequency distribution does not necessarily provide enough information for assessing how well the trip interchanges match (Duffus et al. 1987).

2.5.2. Choice Models of Trip Distribution

In choice models, the choice of destination is made on the basis of a comparison of the attributes of all the available alternative destinations. The socioeconomic characteristics of a traveler are also considered in the model. Models based on the principle of individual utility maximization are typically used. Trips between origin and destination are obtained by multiplying the choice probabilities with total trip originations.

In order to represent the attractiveness of the alternatives available, utility functions are defined for each alternative. These functions consist of variables representing attributes of the alternative and of the decision maker (traveler). Generally, they consist two parts:

- The systematic, observable utility that is similar to the conventional microeconomic utility functions, and
- A random term that is intended to capture such effects as variations in perceptions and tastes of individual trip-makers, misspecification of the utility function by the analyst and measurement errors on the part of the analyst (Manski 1973).

The form of the utility function of an alternative j to an individual n, U_{jn}, is assumed as follows:

$$U_{jn} = V_{jn} + \varepsilon_{jn} \tag{2.17}$$

where

V_{jn} = deterministic (systematic) component, and

ε_{jn} = an additive random component.

The deterministic (systematic) component is assumed to be a function of the attributes of the alternative, X_j, and the characteristics of individual, S_n. From empirical investigation and from behavioral postulates, it is possible to specify the form of the V_{jn} function and to select the variables to include in it. The empirical observation of the random component of the choice function is less practical, for it requires the observation of an individual on repeated occasions, under experimentally controlled conditions, in order to observe the variability of perception and behavior in the face of a choice function (Kanafani 1983). Therefore, statistical assumptions are made regarding the distributional nature of the random component, ε_{jn}.

2.5.3. Types of Choice Models for Trip Distribution

One assumption concerning the distribution of the random term, ε_{jn}, is that they are each independently and identically distributed (IID) with a Gumbel Type I distribution whose cumulative distribution function is given as:

$$F_e(x) = e^{-\varphi e^{-x}}; \quad \varphi \quad 0; \tag{2.18}$$

Hensher and Johnson (1981), Maddala (1983), Ben-Akiva and Lerman (1985), Train (1986), Johnson and Kotz (1970) and Cramer (1991) have shown that if the random components of the utility function are identically and independently distributed (IID) according to a Gumbel distribution, the probability that individual n will choose alternative i can be computed as shown in Equation (2.19). Absence of autocorrelation and heteroskedasticity are also assumed.

$$P_n(j) = \frac{e^{V_{jn}}}{\sum_{p \in C_n} e^{V_{pn}}} \tag{2.19}$$

where j = 1, 2,…,J is the number of alternatives and C_n is the choice set of alternatives available for individual n. The functional form of Equation (2.19) is a well-known form of the logit model referred to as the multinomial logit (MNL) model. The probability of individual n to choose alternative 1 given only two alternative choices can be represented as:

$$P_n(1) = \frac{e^{V_{1n}}}{e^{V_{1n}} + e^{V_{2n}}} \tag{2.20}$$

and is known as Binary Logit model.

Another possible assumption on the random component of the utility function is that they are multivariate normally distributed. Generally, a vector-valued random variable **X** = (X_1, X_2, …, X_k) is multivariate normally (MVN) distributed with mean **m** = (m_1, m_2, …, m_k) and nonsingular matrix

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \cdots & \sigma_{1k}^2 \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ \sigma_{k1}^2 & \sigma_{k2}^2 & \cdots & \sigma_{kk}^2 \end{bmatrix}$$

if its density function is

$$\Phi(x \backslash m, \Sigma) = (2\pi \backslash \Sigma \backslash)^{-k/2} \exp\left[-(x-m)\Sigma^{-1}(x-m)^T / 2\right] \quad (2.21)$$

The multivariate normal distribution assumption leads to the multinomial probit (MNP) model. In the MNP model, the variances of the error terms can be different and the error terms may be correlated. Unfortunately, the choice function of a MNP model cannot be easily written in a closed form, except for the case of two alternatives, and thus must be evaluated numerically (Daganzo, 1979). Suppose ε_{in} and ε_{jn} are random components of utility functions for alternatives i and j, and are both normal with zero means and variances σ_i^2 and σ_j^2, respectively, with covariance σ_{ij}. The term ε_{in} - ε_{jn} is also normally distributed with mean zero and variance $\sigma_i^2 + \sigma_j^2 - 2\sigma_{ij} = \sigma^2$. Ben-Akiva and Lerman (1985) used this result to show that the probability of individual n to choose alternative i can be computed as:

$$P_n(i) = \Phi\left(\frac{V_{in} - V_{jn}}{\sigma}\right) \quad (2.22)$$

where $\Phi(\ )$ denotes the standardized cumulative normal distribution. This model is called Binary Probit model.

The results from binary logit and probit model are not much different when the independence of utility is assumed. DeDonnea (1971) compared the results of binary logit and probit models of mode choice and found that they are close. Since the predictions are so close, the choice between these two models should always be made on the basis of whether the independence assumption can be made or not. In general, the logit model is preferable because it approximates a normal distribution quite well and it is analytically convenient.

2.5.4. Origin-Destination Demand Models

In the origin-destination demand models, the socioeconomic demand and supply system variables are multiplicative. Unconstrained gravity models are typically used to generate the total traffic flows rather than choice probabilities. Both trip production and trip attraction variables are included in the model.

2.5.5. Previous Efforts in Modeling Destination Choices

There have been several efforts of developing models for trip distribution in order to overcome the weaknesses of conventional models for trip distribution. The models developed include developing least squares estimation models and utility maximization models. Wansbeek (1977) analyzed the estimation of the parameters of a standard trip distribution model by means of the ordinary least squares (OLS) method. In the context of the linearized trip distribution model, simple formulae were derived for the least-squares estimators of the parameters and their covariance matrix using a generalized inverse to solve the normal equations. An extension to the case of a non-linear distance function was given. However, the methods suggested did not specify explicit ways of incorporating spatial behavior in models of trip distribution.

Ashtakala (1985) developed a generalized power model for trip distribution, based on the the concepts of conventional trip distribution, linear regression analysis and power transformation on the independent variable. In the proposed model, the relationship between trips from an origin to a destination (dependent variable) and attraction of the destination (independent variable) were taken into consideration. Regression analysis was used to determine the parameters of the relationship which was optimized by varying the power parameter. Regression parameters and the optimum power parameters were used in formulating the trip distribution model. This model was termed as Generalized Power model for distribution. The model was developed by using cordon origin-destination (OD) survey data of Red Deer in Alberta. Evaluation of the model's performance and application of power transformations indicated to have strengthened the model for trip distribution.

Kitamura (1984) developed a model of destination choice employing "prospective utility" of a destination zone as its attraction measure. The focus of the study was mainly on the effect of trip chaining on destination choice and on the adequacy of the conventional assumption that this linkage effect can be ignored and trips can be separately and independently analyzed. The prospective utility accounted for dependency of destination choice and therefore made possible relevant treatment of interdependent choices in a trip chain. A parameter was included in the model to represent the magnitude of the future dependency. The value of this parameter was estimated in the empirical analysis and was concluded to be significantly different from zero. It was also found that the estimates of travel time and zonal attribute coefficients differed substantially when this parameter was excluded from the model. The study demonstrated the effect of exclusion of important variables in trip distribution models.

Southworth (1981) calibrated multinomial logit models of mode and destination choice for a sample of car-owning households in the West Yorkshire Region of England. The models were calibrated using a disaggregate database. Previous practices such as conventional trip distribution methods were extended to investigate the effects on model parameter values of socio-economic standardization of trip making households for the journeys to work, shop and recreation. The models were calibrated at regional scale. All the model runs showed statistically significant results on the basis of maximum likelihood fitting criterion used. However, a number of model coefficients associated with important travel cost variables had large standard errors, preventing many comparisons across socio-economic groupings. However, the author did not investigate the effect of spatial characteristics of both travelers and their origin and destination location, and ways of incorporating them in multinomial logit model.

D'Juran (1995) developed a combined fratar-gravity model for trip distribution with the objective of investigating its implementation in a three-dimensional modeling specification in order to combine the virtues of these two proven modeling techniques. In the fratar model, information on the likely growth in the number of trips originating and/or attracted to each zone is used to develop growth factors used to forecast future trips. The model proposed in that study grouped zones in an urban area into larger superzones for use in the Fratar module. Superzones were defined as amalgamations of the zones in the Greater Toronto area by geographic location and by household growth categories. In the first stage, the Fratar approach was applied to a small superzone trip matrix in order to produce a predicted superzone matrix. The superzone matrix produced was then used as a third dimension constraint in a 3-D trip distribution model. The model was compared to the conventional gravity model to assess its performance in both long-term and short-term prediction and was concluded to be only marginally better than the gravity model in predicting trip movements in the short term (5 years) period only. The model did require significantly more data input than the gravity model, and as a result, the use of the combined model proposed in this study might be inappropriate for some modeling applications.

In a recent Ph. D. dissertation, Limanond (2001) conducted a study on effects of neighborhood setting and intra-neighborhood location on shopping travel behavior of residents in traditional neighborhoods. The activity-based approach was used to investigate how travel decisions of traditional neighborhood (TN) residents vary spatially within neighborhood and across neighborhoods of different regional settings. The primary focus was on shopping travel decisions, specifically to mode and destination choices for home-based shopping tours, and household shopping tour generation. A nested-logit model was constructed to consider five dimensions of shopping travel decisions: household tour generation, participating party, shopping tour type, mode and destination choices. The choice set of the destination choice model was uniquely constructed to separately represent the effects of neighborhood accessibility (characterizing how well a residential location interconnects to stores within the neighborhood) and regional accessibility (characterizing how well a residential location can access outside-neighborhood opportunities within the region). The model was calibrated using travel data collected in three neighborhoods in the Puget Sound, Washington area. The results revealed that both neighborhood and regional accessibility had interrelated effects on the mode-destination choice decisions for home-based shopping only tours, resulting in spatial variations of the travel decisions both within neighborhood and across neighborhoods.

2.6. Spatial Models for Travel Demand Analysis

Spatial models are used for modeling data in which observations of the variables are themselves spatially dependent from each other. Spatial dependence in a collection of sample data observations refers to the fact that one observation associated with a location *i* depends on other observation at location *j*. Spatial models are designed to account for the variation between observed quantities at different locations (Haining, 2003). Variation can be represented through the mean or the correlation structure or a combination of both.

2.6.1. The rationale for incorporating spatial factors in travel demand models

Both space and time play an important role in travel demand behaviors—time being the most perceived variable. Planners have hypothesized that the spatial distribution of land use within a region influences the travel behavior of residents living in that region (Limanond and Niemeier, 2004). The concept of accessibility has been mostly used in past studies (e.g. Handy 1993, Kockelman 1997) to examine the effects of land use patterns on travel behavior. Kockelman (1997) concluded that, after demographics were controlled for, the measures of accessibility, land use mixing, and land use balance—computed for trip makers' home neighborhoods and at trip ends—proved to be highly statistically significant and influential in their impact on all measures of travel behavior. Handy (1993) suggested that the amount that a person travels is influenced by both the character of the particular community in which he or she lives and the spatial structure of the region of which that community is a part.

Srinivasan (2001) examined the effect of neighborhood characteristics such as land use, network, and accessibility-related characteristics quantified through the use of geographical information systems on travel behavior. The author suggested that such measures could be used in conjunction with detailed surveys of travel behavior to specify, calibrate, and use models that are more sensitive to the fine-grained spatial structure of neighborhoods and transportation corridors in metropolitan areas. The conclusion of the study was that spatial characteristics affect travel behavior even on the relatively (spatially) restricted non work tour and could be potentially useful for transportation planning. Bento *et al* (2005) examined the effects of urban form and public transit supply on the commute mode choice and annual vehicle miles traveled (VMT). In order to establish the relationship between urban form and travel demand, the author suggested that the spatial distribution of firms (and associated wage gradient) and the set of possible employment locations affect commute lengths by affecting where households choose to live and where they choose to work. The results of this study indicated that individual measures of urban form and public transit supply have a small but statistically significant effect on travel demand. Ghaeli and Hutchinson (1998) described analyses of the intraregional differences in travel behavior in the greater Toronto area. The travel characteristics of residents living in inner suburban stable and growing areas, outer suburban stable and growing areas, and downtown areas were compared. Two geographical (spatial) scales: municipality and zones within municipality were used for analysis. The results of this study indicated that the average household characteristics in stable and growing sections of the new suburban areas were quite different. In addition, significant difference existed in the household trip production rates for households located in the low- and high-growth areas of both the stable suburbs and the growing suburbs. Recently, Naess (2006) conducted a comprehensive research study in Copenhagen Metropolitan Area with the objective of identifying how spatial planning in urban areas can be used to influence the amount of travel and the proportions carried out by different modes of conveyance. The author hypothesized that the travel between different destinations is influenced on the one hand by the reason people may have for going to a particular place, and on the other hand by the discomfort involved when traveling to this location. One way of measuring the discomfort is by using the distance or accessibility between the origin and the destination. There are mutual influences between the urban structural situation of the dwelling (location relative to various centers and facilities, and local transport infrastructure) and individual and household characteristics. Certain socio-economic characteristics and attitudes (e.g. car ownership and transport attitude) may themselves be influenced by the urban structural situation. The author concluded that urban structure, in addition to its direct effects, may influence activity participation and travel behavior indirectly via car ownership, transport attitudes and some other variables.

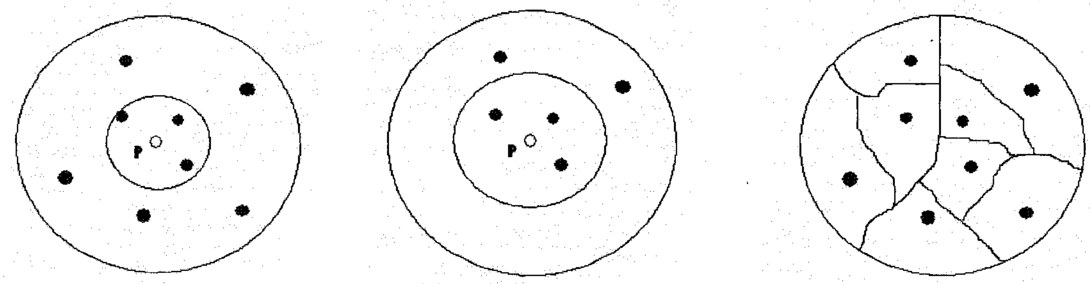
Based on a sample literature review summarized above, it can be clearly seen that the effect of spatial variables in travel demand is a research question which has not been answered fully. Urban structure is viewed as one of the important variables affecting travel behavior—both directly and indirectly. The literature suggests that, although travelers may not consider the urban structure as an explicit variable, they consider it through other variables such as accessibility and travel distance between their origin and possible destinations. One study (Bento *et al* (2005)) indicated that spatial distribution of firms and the set of possible employment locations affect travel behavior. Given that the individuals have chosen where to live and have a choice between different working locations, their decision may therefore be influenced by the spatial distribution of possible employment locations.

2.6.2. Quantifying Location in the Model

In the spatial models, location is quantified by introducing an additional variable with spatial connectivity of the observations. Connectivity is established through a spatial matrix defining neighborhoods of the observations. In general, the criteria used for defining neighborhoods include (Haining, 2003):

- *Straight line distance*: each point is linked to all other points that are within a specified distance
- *Nearest Neighbors*: each point is linked to its k ($k = 1, 2, 3,\ldots$) nearest neighbors.
- *Delaunay triangulation*: all points with a shared edge in a Dirichlet partitioning of the area are linked.

Figure 2.2 shows the details of the criteria for defining neighborhoods. In Figure 2.2a, all points within a specified distance from point P are deemed neighbors of point P, while in Figure 2.2b, a specific distance is first defined and point P is deemed a neighbor of k majority points within the specified distance of P. For example, using the inner circle, only red points are deemed neighbors of point P, while using the outer circle, green points are deemed neighbors of point P. Finally, Figure 2.2c shows that all points sharing an edge are deemed neighbors.



a. Straight line distance b. Nearest neighbors c. Delaunay triangulation

Figure 2.2. Some criteria used for defining neighborhoods

Spatial relationship can be represented in the form of a binary contiguity matrix in which if there are n locations, the matrix developed will have n x n elements. If two objects i and j are to be defined as mutually linked then $W(i, j) = W(j, i) = 1$, otherwise it is assigned a value of 0. It should be noted that an object cannot be connected to itself, that is, $W(i, i) = 0$ for all i. For example, considering the example of a zoned urban area in Figure 1.2, the contiguity matrix W can be defined as follows:

$$W = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix} \tag{2.23}$$

Another option of defining the spatial weight matrix is to make W_{ij} a distance-based weight which is the inverse distance between locations *i* and *j* (1/d*ij*). It compares the sum of the cross-products of values at different locations, two at a time, weighted by the inverse of the distance between the locations.

2.6.3. Modifiable Areal Unit Problem in Spatial Data Analysis

The Modifiable Areal Unit Problem (MAUP) is a common problem in analysis of spatial data. The MAUP is a potential source of error that can affect spatial data analysis which utilizes aggregated data (Unwin, 1996). The MAUP consists of two major parts: scale effect and zonal aggregation effect. Scale effect refers to the variation that can occur when data from one scale of areal units is aggregated into bigger or disaggregated into smaller areal units. Understanding the scale effect is important for the analysis of land use-travel interaction (Kwan and Weber, 2007). The zonal or aggregation effect refers to variability of analytical or statistical results derived from data for the same region, but aggregated or partitioned in different ways, with the number of areal units in different partitioning schemes being the same (Wong, 1996).

There have been several attempts to address the MAUP effect in analysis of spatial data. Wong (1996) summarized these attempts into three potential approaches: data manipulation, technique-oriented, and error modeling approach. In the data manipulation approach, researchers argue that if the chosen zones can be justified as the best among all possible spatial partitioning, the MAUP vanishes. For technique-oriented approach, Robinson (1950) argued that weighting areal units by population size or number of observations can eliminate the scale effect in aggregated data. For the error modeling approach, a spatial regression model in which a component for individual observation and the average of attribute values from surrounding areal units of the observations are incorporated, can minimize the MAUP. However, none of the three approaches have been proven to eliminate the MAUP effect.

Furthermore, Wong (1996) recommended three categories of guidelines for analyzing spatial data from different scales: using data from the finest scale, reporting error from aggregation, and using techniques insensitive to scale changes. Using data from the individual or at the most disaggregated level may be the most appropriate approach. However, individual observations or data at a very high level of resolution may be too massive and difficult to represent by cartographic means. Also, the author recommended that reporting the scale-sensitivity of results may indicate how reliable the results can be. Khatib *et al* (2001) conducted a study to determine how different TAZ structures and different roadway network details affect the ability of a statewide transportation planning (STP) model to replicate annual daily traffic counts on the network. One of their findings was that a more detailed network would achieve better assignment results no matter which level of TAZs is used. It was further concluded that the effect of interaction between the zoning structure and the detail of network should be considered. This implies that, not only the zoning structure, but also the details of the attributes of the zones play an important role in minimizing the MAUP in spatial data analysis. Kwan and Weber (2007) used the activity-travel diary data set collected in Portland (Oregon, USA) to evaluate the properties of a distinct type of accessibility by employing frame-independent and scale-invariant methods that do not produce results that are dependent on particular sets of zones or spatial scale. The importance of spatial scale to individual accessibility patterns was examined by carrying out analysis on the relationship between access and explanatory characteristics (e.g. socioeconomics and demographics) at a range of spatial scales. The results indicated that accessibility measure is scale independent and invariant.

CHAPTER 3

RESEARCH METHODOLOGY

To accomplish the objectives of this research, the spatial models of trip generation and destination choice were specified. Alternative methods for quantifying spatial relationship of observations in these models were also identified; details of these specifications are provided in this chapter. Also, the details of the methodologies for data collection, selection of modeling variables, model estimation, interpretation and analysis of model results and model comparisons are provided. Finally, the chapter provides details of the methodology for creating the trip distribution Origin-Destination (O-D) matrix using the results of the spatial multinomial logit model (MNL) developed.

3.1. Specification of Spatial Models

In order to develop the spatial model, a variable necessary to account for spatial interactions between the observations was incorporated in the model. This variable uses spatial referencing associated with each data value (eg. trip attractions in different zones) that is specified within the geographical system under study. For a spatial variable, geographical location needs to be defined by establishing "spatial connectivity" between observations made at different locations. The following are the specifications for trip generation and destination choice spatial models.

3.1.1. Spatial Trip Generation Models

There are two common types of spatial regression models: spatial lag and spatial error models. In this research, the two models were specified and their estimations compared to identify the best model. In the spatial lag model, observation at one location, Y_j, was used as the independent variable to predict the observation at another location, Y_i. The spatial connectivity between the locations was established a priori by using the spatial weight matrix, W_{ij}. In this model, the observation at one location was tested for its effect on the adjacent observation at another location while taking into consideration spatial linkage of the observations. Equation (3.1) shows the general form of the spatial lag model.

$$Y_i = \beta_0 + \rho W_{ij} Y_j + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i \tag{3.1}$$

where

W_{ij} is the spatial weight (connectivity) matrix

ρ is a spatial parameter to be estimated

β_k are the coefficients to be estimated

Y_i is vector of observed trips at location i

Y_j is a vector of observations at location j

x_{ik} are explanatory variables

ε_i are random errors

In the spatial error model, the stochastic errors are assumed to be spatially autocorrelated. Connectivity among the stochastic errors was established through spatial weight matrix. Equation (3.2) shows the general specification of the spatial error model.

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \lambda W_{ij} \varepsilon_j + u_i \tag{3.2}$$

where

λ is a spatial parameter to be estimated

Y_i is vector of observations at location i

u_i are random errors

3.1.2. Spatial Destination Choice Model

In this research, the discrete spatial destination choice model for Las Vegas was specified and estimated. There are two approaches for developing discrete destination choice model. The first approach is through collapsing the alternative choices into two (dichotomy), hence developing a binary logit model. This is done by assigning a value of one to the chosen alternative while others are assigned a value of zero. The second approach is to keep all the alternative choices and therefore develop a multinomial logit model. However, when multinomial data are treated as dichotomous, not only do the expected confidence interval widths become greater, but the penalty in terms of larger sample size requirements for hypothesis testing can be severe (Bartfay and Donner, 2000). There are clear advantages in preserving multinomial data on the original scale rather than collapsing the data into a binary trait. Therefore, the usual multinomial logit model modified to incorporate spatial relationship of the alternatives was developed for destination choice. The model is estimated through application of maximization concepts of alternatives, from which the alternative with maximum utility is chosen (Ben-Akiva and Lerman, 1985).

To account for the spatial relationship, the utility function was modified to incorporate the spatial interaction term in order to minimize uncertainty resulting from spatial autocorrelation. Also, the socioeconomic variables were included in the utility function with alternative specific coefficients except for the base alternatives. The base alternatives were the two districts containing the resort corridor, popularly known as "the strip" and the downtown area. These districts contain most of the hotel and casinos. The function was defined as follows:

$$U_{jn} = \sum_{k=1}^{K} \beta_k x_{jnk} + \sum_{m=1}^{M} \gamma_{jm} S_{nm} + \rho Z_{ijn} + \varepsilon_{jn} \tag{3.3}$$

where

U_{jn} is the utility of alternative destination j to individual n

$k = 1, \ldots, K$ are the attributes of the alternatives

β_k is the coefficient of attribute k

x_{jnk} is the k^{th} attribute associated with alternative destination j and individual n

$m = 1, \ldots, M$ are the socioeconomic characteristics of the individuals

γ_{jm} is the alternative specific coefficient of socioeconomic characteristic m

S_{nm} is the m^{th} socioeconomic characteristic associated with individual n

ρ is the coefficient associated with spatial variable

Z_{ijn} is the spatial variable defined by separation or accessibility between origin (i) and destination (j) for individual n

ε_{jn} is the stochastic error term

The probability of individual n to choose alternative destination j is therefore computed as follows:

$$P_n(j) = \frac{e^{\sum_{k=1}^{K} \beta_k x_{jnk} + \sum_{m=1}^{M} \gamma_{jm} S_{nm} + \rho Z_{ijn}}}{\sum_{p \in C_n} e^{\sum_{k=1}^{K} \beta_k x_{pnk} + \sum_{m=1}^{M} \gamma_{jm} S_{nm} + \rho Z_{ipn}}}, \quad p \neq j \tag{3.4}$$

3.2. Methods for Quantifying Spatial Relationship

In order to quantify the spatial relationships of the zones, different $N \times N$ spatial matrices, in which N is the number of zones, were developed. Four alternative methods for defining the elements of the matrices were adopted: (a) Contiguity, (b) Separation, (c) Contiguity-Separation, and (d) Economic linkages (accessibility measure). For each alternative, the spatial models of trip generation were developed and compared to identify the best alternative for defining spatial relationship of the observations. Similarly, different models of destination choice were developed and compared. The following sections provide details of each method while Appendix A presents the matrices created for each method.

3.2.1. Contiguity

Contiguity matrix is the popular method for defining spatial relationship of geographical units such as TAZs. In this method, if two locations i and j are to be defined as mutually linked, then $W(i, j) = W(j, i) = 1$; otherwise it is assigned a value of zero. Because an object cannot be connected to itself, $W(i, i) = 0$ for all i.

3.2.2. Separation

Separation was measured as the inverse of the distance between the zones. The separation only was used in order to avoid omission of observations which are not contiguous to the specified observation, but are influencing it. Therefore, the influence of all other observations to a specified observation decayed as a function of their distance apart. Elements of the spatial weight matrix were generated as follows:

$$W_{ij} = D_{ij}^{-1} \quad (3.5)$$

where D_{ij} is the distance between zone i and zone j.

3.2.3. Contiguity-Separation

In this specification, the effect of separation between observations is combined with contiguity to explain the spatial interaction between geographical entities. Separation was measured as the inverse of the distance between the observations. Since the distance is measured between zonal centroids, combination of contiguity and distance was expected to indicate how "truly" the zones are spatially close. Elements of the spatial weight matrix were generated by multiplying separation and contiguity as follows:

$$W_{ij} = \mu_{ij} D_{ij}^{-1} \quad (3.6)$$

where $\mu_{ij} = 1$, if i is contiguous to j; otherwise $\mu_{ij} = 0$. D_{ij} is the distance between zone i and zone j.

3.2.4. Economic linkages (accessibility measure)

Although two zones might be contiguous, if there is no accessibility between them, the spatial relationship measured by contiguity only, might not represent the reality. Therefore, accessibility between two zones is another way of defining spatial relationships. Traditionally, it is measured as the product of employment opportunities and the inverse of their separation distance (Stewart 1958 and Koenig 1980). The difference between the methods of combining contiguity and separation and that of combining employment opportunities and separation is that, the latter introduces economic linkage of the two zones. The accessibility measure between two zones can be used as elements of the spatial weight matrix to define interaction between the two zones. This means, a zone is spatially related to another zone if the accessibility measure between them is higher compared to other zones. Mathematically, the spatial weight matrix can be written as follows:

$$W_{ij} = E_j D_{ij}^{-1} \tag{3.7}$$

in which E_j is the total employment in zone j ( )

In my research, instead of using the straight line distance between the observations, the separation distance was measured along the shortest and potential network route of individuals moving from one district to another. This reflected the typical distance traveled and the actual spatial separation of the observations. Free online mapping portals used for searching for driving directions were used to identify potential network routes. ArcGIS 9.0 software was then used to create the maps. Figure 3.1 illustrates the difference between straight line distance and network distance by showing the possible route from district 3 to district 15. While the "actual" separation distance along the travel route is 10.6 miles, the straight line distance is 6.28 miles: less by more than 40%. Therefore, using straight line distance may distort the effect of separation distance on spatial relationship. Appendix B presents the typical network routes for individuals originating from zone 3 to other zones.

3.3. Data Collection and Processing

The data used for this research included travel records of surveyed individual travelers, land-use attributes of their origin and destination districts, their demographic and socioeconomic characteristics, and geographic location of origin and destination of each trip. The data were collected by extracting relevant information from two databases maintained by the Regional Transportation Commission of Southern Nevada (RTC): (1) The 1996 Las Vegas Household Travel Survey database, and (2) The Land-use database.

3.3.1. The 1996 Las Vegas Household Travel Survey Database

RTC conducted a household travel survey between April 1996 and June 1996 to collect retrospective travel information from both residents and visitors (RTC, 1998). The survey consisted of a stratified random sample of households based on the number of vehicles available and household size. Random-digit dialing was used to develop the sample list of telephone numbers. A total of 3,738 households were recruited for the study but only 1,887 (approx. 50%) households provided complete information during the survey. Each household member of age 16 and above was asked to record each activity they undertook during the 24-hour period from 3.00 a.m on the diary day to 3.00 a.m on the following day. Nineteen (19) main activities were coded as well as a catch all "other" category. Using the 19 main activities, the resulting trips can be categorized into four trip purposes:

- Home-Based Work
- Home-Based non-Work
- Non-Home-Based Work, and
- Non-Home-Based Non-Work.

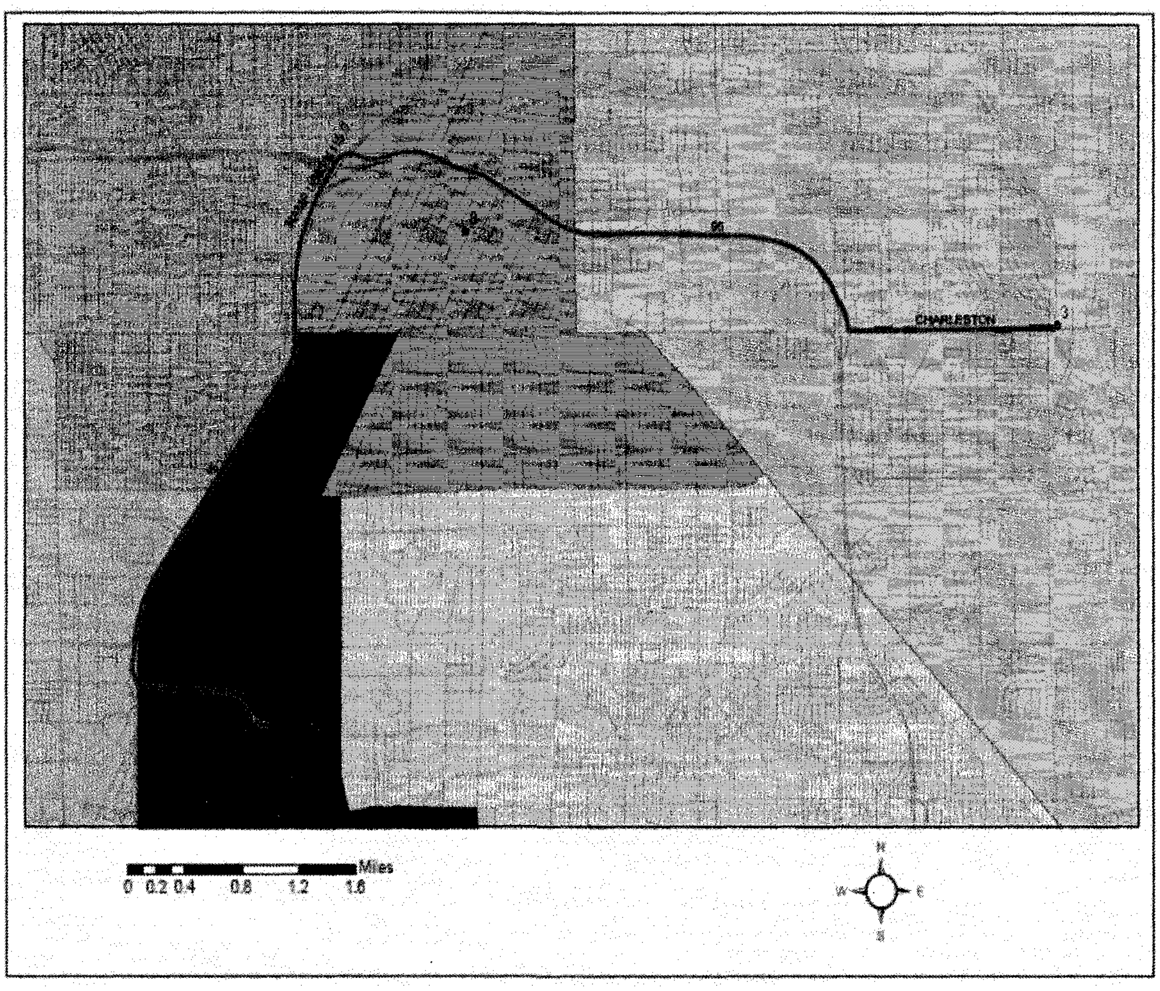


Figure 3.1. Travel route from district 3 to district 15

For each stop made during specified time period, the individuals were asked to record the correct addresses. This enabled geocoding the geographic location of their origins and destinations in the form of (x, y) coordinates. In addition, the diary was used to collect personal demographic and socioeconomic information such as age, gender, income, household size, number of vehicles available, etc. In this research, Home-Based Work (HBW) trips were identified and extracted from the database for developing models.

3.3.2. The Land-use Database

The RTC also maintains a database of land-use based on Traffic Analysis Zones (TAZ). TAZs are the basic geographical zonal units used for land-use and trip generation estimations. Their boundaries are made up of major streets, census tracts, natural or man made barriers like railroad tracks, or rivers. A general rule for the size of the TAZ per model area is one TAZ per 1000 in population. The RTC, local entities and consultants provided input to define the TAZ boundaries based on the criteria mentioned above. In 1996, the Las Vegas Valley was divided into more than 1000 TAZs. Often, several TAZs are combined to form larger zones called districts based on certain criteria. Districts often follow travel corridors, political jurisdictions and natural boundaries. The Las Vegas Valley has been divided up into 18 districts based on natural and man-made barriers (for example freeways and railroads). The same agencies involved in defining TAZs also provided input to define the district boundaries based on the criteria for defining districts. In this study, the TAZ and districts definitions have been verified with relevant RTC personnel. Figure 3.2 shows the Las Vegas Valley districts.

The information available in the land-use database includes, but not limited to, district population, area of the district, available dwelling units, occupied dwelling units, average household income, average household size, employment opportunities available in different job categories, etc. In this study, the variables relevant to travel behavior were extracted from the land-use database and associated with their corresponding Home-Based Work trips extracted from the 1996 Las Vegas Household Travel Survey database.

3.3.3. Database Merging and Processing

In order to merge the two databases, the 1996 Las Vegas Household Travel Survey data were geocoded into geographical information systems (GIS) using the ESRI ArcGIS 9.0 software. ESRI ArcGIS is an integrated collection of GIS software products for building a complete GIS. It enables users to deploy GIS functionality wherever it is needed—in desktops, servers, or custom applications; over the Web; or in the field (ESRI, 2007). The ArcGIS geodatabase framework allows accessing large volumes of geographic data stored in files and databases. It also allows extraction of specific data from the database. A map of the Las Vegas Valley districts with their land-use attributes was developed using the same software. The travel survey data was displayed on the same map in order to identify their location with relation to land-use data. Only 17 districts were used in this study because there were no households interviewed from district 18. The trips that originated from home to work (main or second job), volunteer work, and work-related trips were identified and categorized as HOME-BASED WORK trips. The number of trip attractions and trip productions for each district were obtained.

The two trip types were modeled differently since their explanatory variables are not the same. Total trips per district for both trip attraction and production were obtained by aggregating individual trips. For destination choice modeling, destination districts for all trips were identified and used to form the choice set. Each of the 17 districts was assumed as a possible alternative destination for work to all individuals and therefore, the size of the choice set was 17.

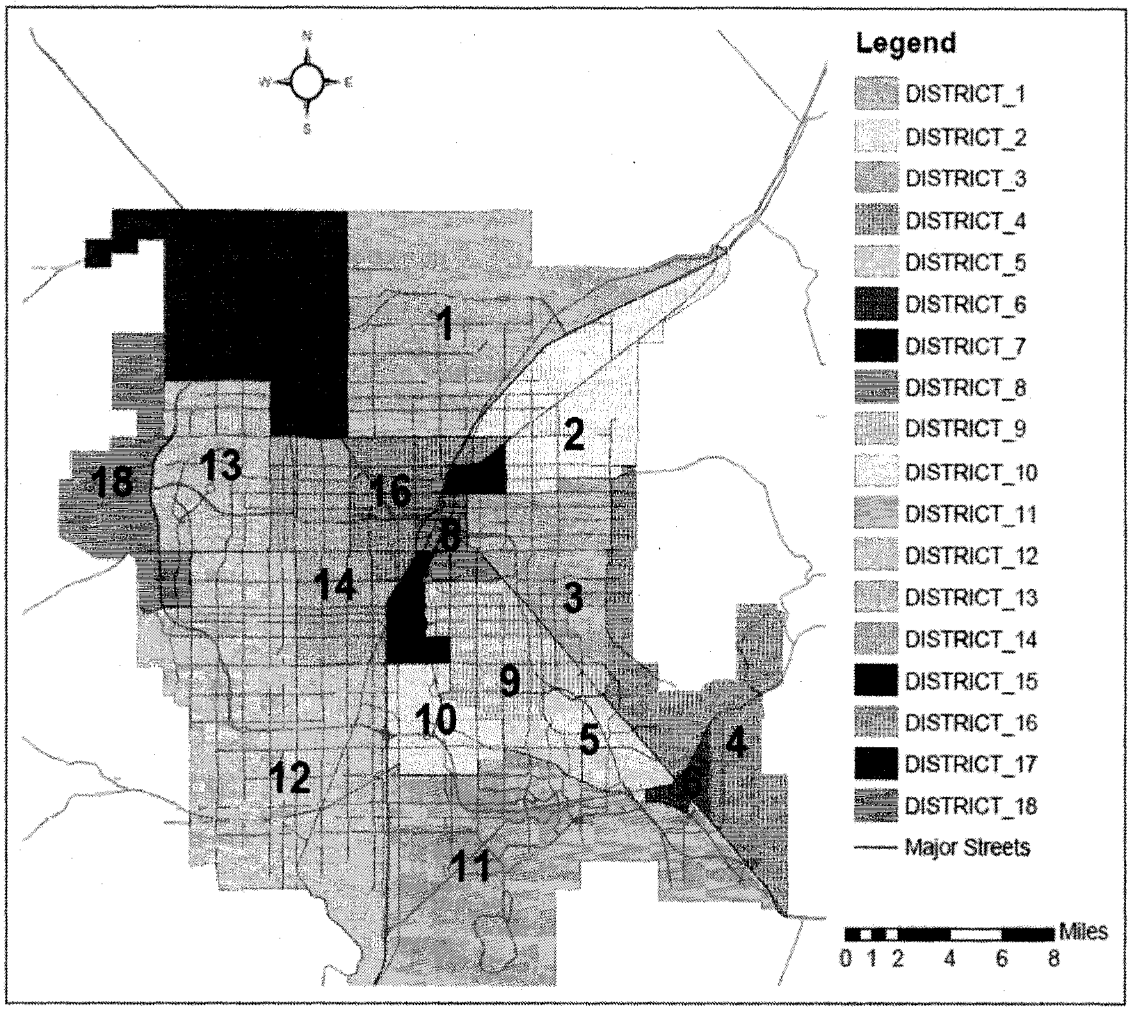


Figure 3.2. Las Vegas Valley districts

3.3.4. Data Expansion

It should be noted that only sampled households participated in the survey, and therefore the total number of trips obtained from the database is not the actual number of trips per district. As a consequence, the procedure for data expansion to reflect the actual estimates of number of trips was adopted. The aim of data expansion was to estimate factors to be applied to each survey observation so that the estimated number of trips would provide valid statistical estimates of the actual number of such trips under actual population estimates of the analysis year. The factors were developed through a two-tier process, in which individual households were categorized by available vehicles and household size. Sampled households in each category were expanded to estimate the 1996 total number of households in that category using the 1990 Census data as a base. The 1990 census data was used as a base because it contains the actual total number of households in 1990 for each category. For 1996, the household data available was the total number of households for all categories in each parcel-based planning area. The following are the details of developing expansion factors:

Let $S96_{(h,v)}$ and $T96_{(h,v)}$ be the sampled households and the total households in a category of households with size h and v vehicles available in 1996, respectively. The relationship between the two can be given as:

$$\left(S96_{(h,v)}\right)F = T96_{(h,v)}, \tag{3.8}$$

where F is the expansion factor . Also, it should be noted that, since $T96_{(h,v)}$ was not available, the use of the 1990 census data was warranted. If $T90_{(h,v)}$ is the he total households in a category of households with size h and v vehicles available in 1990 census, the 1996 and 1990 data for that category can be related as follows:

$$T96_{(h,v)} = GF.\left(T90_{(h,v)}\right), \tag{3.9}$$

in which GF is the growth factor. Combining Equations (3.8) and (3.9) yields the following relationship for the expansion factor:

$$F = GF.\left(\frac{T90_{(h,v)}}{S96_{(h,v)}}\right), \tag{3.10}$$

The growth factor for each district were derived and used to estimate the number of trips in each category as per 1996 Las Vegas population (Clark County, 1996). The factor for each district was used since the growth was not uniform over the valley. To derive the growth factor for each district, the map of the parcel-based planning areas was initially overlaid on a map of modeling districts using GIS. The parcel-based areas overlapping with each modeling district were identified and the growth factors were computed for these areas. If the modeling district contained more than one parcel-based area, the weighted average of the factors (weighted by households in each parcel-based area) was computed, otherwise, the parcel-based area factor was used as the growth factor for the modeling district. Figure 3.3 shows the growth factor levels for all districts while Appendix C shows the details of the parcel-based area growth factors.

3.4. Selection of Variables for Modeling

3.4.1. Trip generation Models

The response variable for trip attraction model was the number of trips attracted to a district, while for the trip production models the response variable was the number of trips produced from a district. Selection of explanatory variables was based on their reasonable relevance to the response variable and availability of such data in the databases. Also, the literature was reviewed to identify the variables used in similar existing trip generation models. Descriptive statistics and correlation tests were conducted in order to identify variables suitable for use in the model. In the presence of high correlation, the estimators are very likely to have large variances and covariances, making precise estimation difficult. Therefore, if two variables were highly correlated, only one of them was included in the model. The Pearson-Product-Moment Correlation coefficient was computed for variables following normal distributions, while the nonparametric (Spearman Rank Correlation) coefficient was used for the remaining variables.

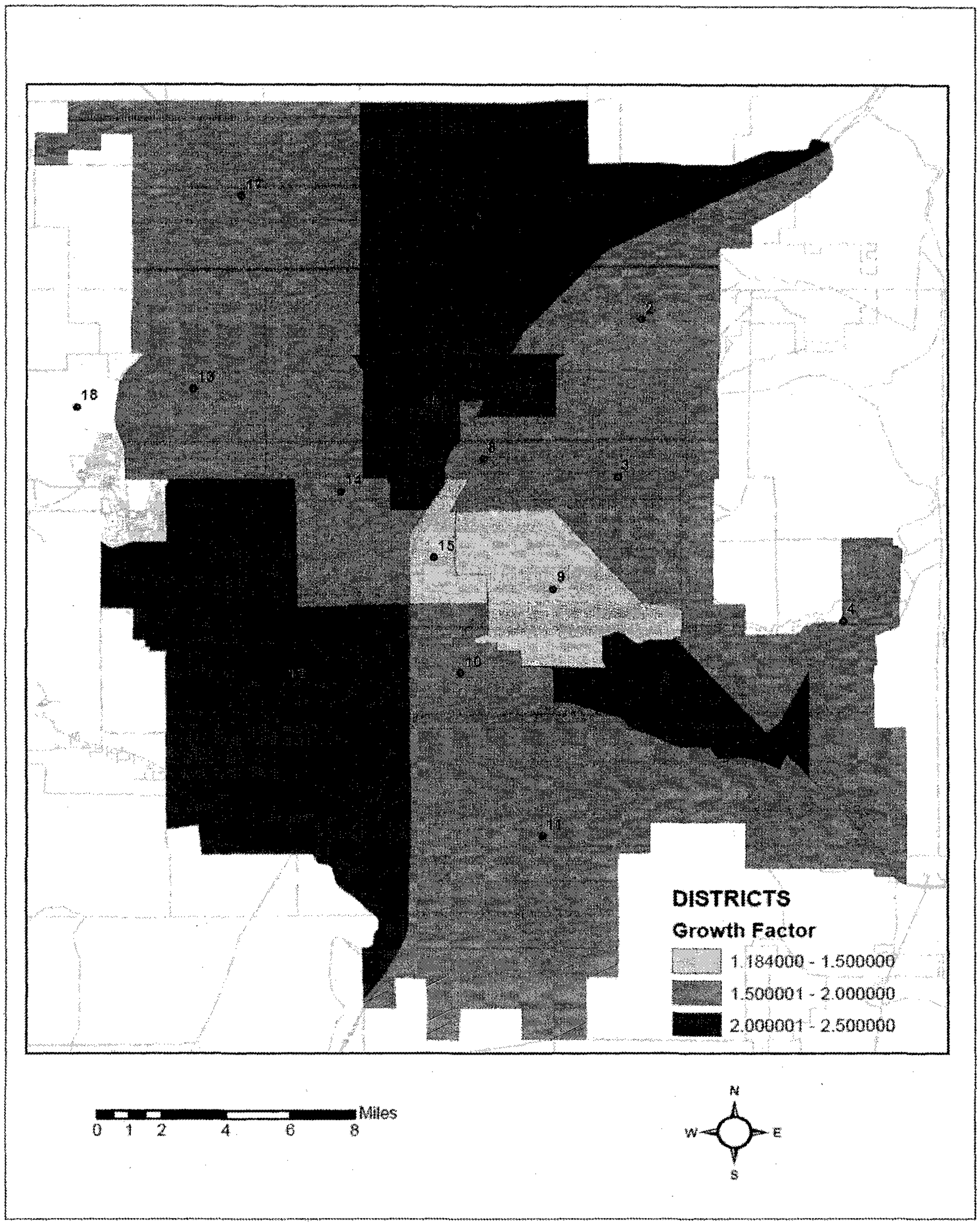


Figure 3.3. Growth factor levels of Las Vegas between 1990 and 1996

In order to satisfy the linear assumption, several transformations were tested to identify the suitable forms of variables. For the trip attraction model, the number of employment opportunities available at destination districts was used. Since the gaming industry constitutes the higher percentage of job opportunities, it was separated from other job opportunities. Table 3.1 shows the details of the variables for the trip attraction model with their description and codes used in the database.

Table 3.1. Variables for trip attraction model

Variable	**Description and Components of the Variable**	**Unit (Code)**
Gaming	Hotel/Casino Employment Opportunities	Total jobs per district
Non-Gaming	Regional Retail Employment Opportunities (*r_shop*)	Total jobs per district
	Community Shop Employment Opportunities (*c_shop*)	
	Other Retail Employment Opportunities (*other_ret*)	
	Other Non Retail Employment Opportunities (*other_non*)	
	Nellis Airforce Base Employment Opportunities (*nafb*)	
	McCarran Airport Employment Opportunities (*mia*)	
	UNLV Employment Opportunities (*unlv*)	
	Office Employment Opportunities (*other_office*)	
	Industrial Employment Opportunities (*indust*)	
Spatial	A variable quantifying spatial relationships of the observations	

The general equation for the trip attraction model was as follows:

$$
\begin{aligned}
(ATTRACTIONS)_i = {} & \beta_1 * (GAMING\ JOBS)_i \\
& + \beta_2 * (NON-GAMING\ JOBS)_i \\
& + \beta_3 * \left(\sum_j W_{ij} * (TRIPS\ ATTRACTED)_j \right), \\
& + CONSTANT
\end{aligned}
\tag{3.11}
$$

where $\beta's$ are marginal impact of the variables on trip attraction and W_{ij} is the matrix defining spatial relationship between observations at location i and j. For the trip production model, the variables selected included trip makers' socioeconomics and demographics. Table 3.2 shows the details of the variables for trip production model.

Table 3.2. Variables for trip production model

Variable	**Description**	**Unit (Code)**
poparea	Population per Area of a district	Population per Square Miles
hh_size	Average Household size	Persons
lowinc	Number of household with income less than \$17.5K	Households
mediuminc	Number of households with income between \$17.5K and \$47.5K	Households
highinc	Number of households with income greater than \$47.5K	Households

The general equation for the trip production model was as follows:

$$
\begin{aligned}
(PRODUCTIONS)_i = {} & \beta_1 * (POPULATION\ DENSITY)_i \\
& + \beta_2 * (AVERAGE\ HOUSEHOLDS)_i \\
& + \beta_3 * (NO.\ OF\ LOW\ INCOME\ HOUSEHOLDS)_i \\
& + \beta_4 * (NO.\ OF\ MEDIUM\ INCOME\ HOUSEHOLDS)_i\ , \\
& + \beta_5 * (NO.\ OF\ HIGH\ INCOME\ HOUSEHOLDS)_i \\
& + \beta_3 * \left(\sum_j W_{ij} * (TRIPS\ PRODUCED)_j \right) \\
& + CONSTANT
\end{aligned}
\tag{3.12}
$$

3.4.2. Destination Choice Models

Based on relevance and availability in the databases, the variables for the destination choice model were selected. Contrary to the gravity model which uses attraction and separation as the exclusive variables for trip distribution; socioeconomic and demographic characteristics of the trip makers were also used in this model. Table 3.3 shows the details of the variables chosen. The deterministic parts of the utility functions for all alternatives are shown in Equation (3.13). District eight and fifteen which contain the resort corridor and the downtown area were used as the base; therefore, individual specific variables (socioeconomics) were not included in the utility function of this alternative. The two districts contain most of the hotels and casinos. The alternative specific coefficients for the socioeconomics were specified in order to capture their specific effects on each alternative. It was also assumed that the propensity of the number of jobs available at the destinations was equal regardless of the destination, and therefore, generic coefficients were used for the variables quantifying job opportunities. Similar assumption was made for the spatial variable and the relative measure of distance with regard to the opportunity at CBD —the propensity is the same regardless of the destination.

Table 3.3. Variables for destination choice model

Variable	Description	Unit (Code)
CBD (cbd)	Distance from home to district *i* divided by distance from home to CBD/Resort, a relative measure with regard to the opportunity at CBD	-
Hotelre(H)	Available job opportunities in hotel and retail	Number of Employments
Otherjo (O)	Number of available jobs other than hotel and retail divided by area of the district	Employment Density
Vehhhworker (V)	Number of vehicles available divided by number of people with age16+ in a household	Vehicles per Persons
Spatial (S)	A variable quantifying spatial relationship between origin district and destination districts	
Age (A)	Age of trip maker	Years
Income (I)	Income level indicator for trip make: 0. Less than US$ 47,499 1. Greater than US$ 47,500	Indicator of income level

$$
\begin{aligned}
V_{1n} &= C_1 + \alpha_1 V_n + \beta_1 A_n + \delta_1 I_n + \phi(cbd)_1 + \gamma H_1 + \pi O_1 + \omega S_1 \\
V_{2n} &= C_2 + \alpha_2 V_n + \beta_2 A_n + \delta_2 I_n + \phi(cbd)_2 + \gamma H_2 + \pi O_2 + \omega S_2 \\
&\vdots \\
V_{7n} &= C_7 + \alpha_7 V_n + \beta_7 A_n + \delta_7 I_n + \phi(cbd)_7 + \gamma H_7 + \pi O_7 + \omega S_7 \\
V_{8n} &= \phi(cbd)_8 + \gamma H_8 + \pi O_8 + \omega S_8 \\
V_{9n} &= C_9 + \alpha_9 V_n + \beta_9 A_n + \delta_9 I_n + \phi(cbd)_9 + \gamma H_9 + \pi O_9 + \omega S_9 \\
&\vdots \\
V_{14n} &= C_{14} + \alpha_{14} V_n + \beta_{14} A_n + \delta_{14} I_n + \phi(cbd)_{14} + \gamma H_{14} + \pi O_{14} + \omega S_{14} \\
V_{15n} &= \phi(cbd)_{15} + \gamma H_{15} + \pi O_{15} + \omega S_{15} \\
V_{16n} &= C_{16} + \alpha_{16} V_n + \beta_{16} A_n + \delta_{16} I_n + \phi(cbd)_{16} + \gamma H_{16} + \pi O_{16} + \omega S_{16} \\
V_{17n} &= C_{17} + \alpha_{17} V_n + \beta_{17} A_n + \delta_{17} I_n + \phi(cbd)_{17} + \gamma H_{17} + \pi O_{17} + \omega S_{17}
\end{aligned}
\quad , \qquad (3.13)
$$

in which V_{jn} is the utility function for alternative *j* to individual *n*

C_j is the alternative specific constant

α_j is the alternative specific coefficient for vehicles available per workers

V_n is the number of vehicles available per workers for individual *n*

B_j is the alternative specific coefficient for the age of individual *n*

A_n is the age of individual n

δ_j is the alternative specific coefficient for the income of individual n

I_n is the income of individual n

ϕ, γ, π and ω are the generic coefficients for the relative measure of distance with regard to the opportunity at CBD, available job opportunities in hotel/casino and retail, available job opportunities in other industries, and spatial interaction of the origin and destination, respectively.

3.5. Testing for Spatial Autocorrelation in Selected Variables

In order to estimate spatial models, the variables were tested for spatial autocorrelation. There are two ways of checking for presence of autocorrelation in the variables: mapping the variables, and using analytical methods. When variables are mapped, one can identify the presence of any spatial patterns such as clustering. For example, when the income variable is mapped, one can identify whether high income or low income households form clusters. In a similar fashion, one can identify whether number of trips (attracted or produced) form clusters or spatial patterns. Analytical spatial correlation tests include the Moran-I test, the Geary's C test, and the Getis and Ord's G test. In this research, the analytical methods were used because they allow for computation of statistics necessary for drawing conclusions. The following are the details of the common spatial autocorrelation tests:

3.5.1. The Moran's I Test Statistic

The Moran-I test statistic is based on the value of the variable at any one location compared with the value at all other locations. The statistic is computed using the following formula:

$$I = \frac{N \quad {}_{i} \quad {}_{j} W_{i,j}(X_i - \bar{X})(X_j - \bar{X})}{(\sum_i \sum_j W_{i,j}) \sum_i (X_i - \bar{X})^2}, \tag{3.14}$$

where

I is the Moran-I test statistic

N is the number of observations

X_i is the observation at a particular location i

X_j is the observation at another location j

$\bar{X}$ is the mean of the variable

W_{ij} is a spatial weight matrix applied to the comparison between location i and location j

The expected value of I under the null hypothesis of no global spatial autocorrelation is given as:

$$E(I) = \frac{-1}{N-1}, \tag{3.15}$$

If the expected value $E(I)$ is less than the test statistic I, then the overall distribution of variable X can be seen as characterized by positive spatial autocorrelation and vice versa. The z-values, computed by subtracting expected value of I from calculated I and dividing the difference by the standard deviation of I, is used for inference and is computed as follows:

$$z_I = \frac{I - E(I)}{std(I)}, \tag{3.16}$$

where

$$std(I) = \sqrt{\left[\frac{N^2 \sum_{ij} W_{ij}^2 + 3\left(\sum_{ik} W_{ij}\right)^2 - N \sum_i \left(\sum_j W_{ij}\right)^2}{\left(N^2 - 1\right)\left(\sum_{ij} W_{ij}\right)^2}\right]}, \tag{3.17}$$

3.5.2. Geary's C Test Statistic

The Geary's C statistic is based on a comparison of related map values. The relationship of map values is defined using the spatial weights matrix, W_{ij}. The statistic is computed as follows:

$$C = \frac{(N-1)\sum_i \sum_j W_{ij}(X_i - X_j)^2}{2\left(\sum_i \sum_j W_{ij}\right)\sum_i (X_i - \bar{X})^2}, \tag{3.18}$$

The Geary's C statistic varies between 0 and 2. Under the null hypothesis of no global spatial autocorrelation, the expected value of Geray's C statistic is 1. Values greater than 1 indicate that the overall distribution of X is characterized by negative spatial autocorrelation, and vice versa. z-values, computed by subtracting expected value of I from calculated I and dividing the difference by the standard deviation of I, is used for inference, and is computed as follows:

$$z_C = \frac{C-1}{std(C)}, \tag{3.19}$$

3.5.3. Getis and Ord's G Test Statistic

The Getis and Ord G-Statistic measures the overall spatial clustering of values of X. It requires that the measured variable X is positively valued and has a natural origin (Getis and Ord, 1992, cited by Haining, 2003). Also, it requires the weights matrix to be composed of binary non-standardized elements. The statistic can be computed as follows:

$$G = \frac{W_{ij} X_i X_j}{\sum_i \sum_j X_i X_j}, i \neq j, \tag{3.20}$$

Under the null hypothesis of no global autocorrelation, the expected value of G is computed as follows:

$$E(G) = \frac{\sum_{i}\sum_{j} W_{ij}}{N(N-1)}, \tag{3.21}$$

Values larger than the expected value indicate that the overall distribution of X is characterized by positive spatial autocorrelation. The inference is based on z-values computed by subtracting the expected value of G from calculated G and dividing the difference by the standard deviation of G as follows:

$$z_G = \frac{G - E(G)}{std(G)}, \tag{3.22}$$

3.6. Model Estimation

3.6.1. Trip generation Models

For trip generation (attraction and production), the Ordinary Least Square (OLS) method of estimation was used to estimate non-spatial models while the Maximum Likelihood estimation (MLE) method was used for spatial model. OLS is a mathematical optimization technique which, when given a series of measured data such as zonal work trip totals, attempts to find a function which closely approximates the data (a "best fit"). It attempts to minimize the sum of the squares of the ordinate differences (called residuals) between points generated by the function and corresponding points in the data. MLE is a popular statistical method used to make inferences about parameters of the underlying probability distribution from a given data set. It is used to estimate the unknown parameters by fixing a set of data and then picking the parameters to the distribution that are most likely given the data. Non-spatial models for trip generation were estimated using Equation (2.1) while the spatial models were estimated using Equation (3.1) and/or Equation (3.2) which incorporated spatial characteristics.

The non-spatial and spatial models of trip attraction and trip production were estimated using the STATA® software. The software is a complete, integrated statistical package that provides most of the needed capabilities for data analysis, data management, and graphics. It was created in 1985 by StataCorp and is used as a statistical program by many businesses and academic institutions around the world. STATA®'s full range of capabilities includes data management, statistical analysis, graphics, simulations, and custom programming. The consistent, intuitive syntax of STATA® commands makes it straightforward to develop programs for complex or repetitious tasks (Hamiliton 2004). In recent years, STATA® has added many new features such as linear mixed models, balanced repeated replications, and multinomial probit. The software has become increasingly popular over the years because of its simplicity in design and usage of drop-down menus compared to many other statistical software packages (e.g. SAS®).

3.6.2. Destination Choice Models

For destination choice model, the classic MLE method was used to estimate the parameters of Equation (3.12). The BIerlaire Optimization toolbox for GEv Model Estimation (BIOGEME) software was used for estimation of the multinomial logit model of destination choice. This is a freeware designed for the maximum likelihood estimation of binary logit, multinomial logit, nested logit, and more complex models in the Generalized Extreme Value (GEV) family as well as mixtures of these models (e.g. mixed logit). The GEV is a family of models consistent with the random utility theory (Bierlaire 2003).

3.7. Interpretation and Analysis of the Results

3.7.1. Spatial Autocorrelation Tests

Spatial autocorrelation test results were interpreted by assessing the degree of spatial autocorrelation present in the observations. The Moran-I statistic varies between –1.0 and +1.0. Areas close together (adjacent) that have similar observations are expected to have a high Moran-I result. Positive Moran-I values indicate positive spatial autocorrelation, meaning that, similar observations are spatially clustered. Negative Moran-I values indicate negative spatial autocorrelation, meaning that, neighboring observations are dissimilar. For the Geary's C test, the expected value of C is one. If the calculated value is greater than one, it indicates negative spatial autocorrelation. If the calculated value is less than one, it indicates positive spatial autocorrelation. For the Getis and Ord's G, if the calculated value is greater than the expected value, it signifies positive spatial autocorrelation with a prevalence of high-valued clusters, while the smaller G value signifies positive spatial autocorrelation, but with a prevalence of low-valued clusters. The significance of the results was decided based on the p-values. The p-values are measures of how much evidence we have against the null hypotheses (H_0) that no global spatial autocorrelation is present in the observations. The smaller the p-value, the more evidence we have against H_0.

3.7.2. Trip Generation Models

For spatial trip attraction and trip production models, the estimated model parameters were assessed for correctness of sign (consistence with intuition), their level of significance in explaining the variability in the response variable, and the general goodness-of-fit of the model. Violation of intuitive expectations of the sign of the coefficient resulted in further investigation of the variable in order to decide whether it should be dropped from the model. It should be noted that "wrong sign" of the coefficients may result from computation errors, coefficient that do not differ from zero, or multicolinearity. For the latter reason, the correlation test results were used for investigation of the variables.

The level of significance of the parameters was assessed by observing the resulting test statistic, p-value and their 95% confidence interval. The coefficient with a test statistic greater than the critical value was interpreted as significantly different from zero and was expected to have a p-value less than 0.05 (at the 5% level of significance). The critical value is a cutoff value determining the boundary between those samples resulting in a test statistic that leads to rejecting the null hypothesis and those that lead to a decision not to reject the null hypothesis. Consequently, for the coefficient to be valid, the confidence interval must not span over zero. A coefficient with confidence interval spanning over zero indicates that it can actually be zero. Variables with coefficients meeting these criteria were concluded as significant variables in explaining the variability in the response variable. The results of coefficient estimation for destination choice model were interpreted in a similar fashion.

In order to assess the goodness-of-fit of the model, the adjusted R squared was used for models with the same dependent variable. This is because the adjusted R-squared value takes into account the number of explanatory variables present in the model.

3.7.3. Destination Choice Models

For the destination choice models, the goodness-of-fit was assessed by conducting a log-likelihood ratio test. The test compares two models provided the simpler model (in this case, non-spatial model) is a special case of the more complex model (in this case, the spatial model). Non-spatial model was deemed a special case of the spatial model because it eliminates the spatial term (restricts the coefficient of spatial term to zero).

3.8. Model Comparisons

To compare the trip generation models, the Residual Sum of Squares (RSS), the Akaike Information Criterion (AIC), and the Schwarz Information Criteria (SIC) were computed and compared. Furthermore, the final loglikelihood was used to compare the spatial models. Comparison of the destination choice models was made by first comparing the spatial models estimated using different definitions of spatial relationships. Finally, the best spatial model was compared to the non-spatial model. The final loglikelihood was also used to compare the models. The model with maximum final loglikelihood represented the best fit model. The following subsections provide the details for each comparison method.

3.8.1. Residual Sum of Squares (RSS)

The RSS quantifies the deviation of the estimated observations from the actual observations by summing their squares. Mathematically, it is computed as follows:

$$RSS = \sum_{i=1}^{N} \left(Y_i - \widehat{Y}_i\right)^2 \qquad (3.23)$$

where $\widehat{Y}_i$ is the estimated value for observation i and N is the number of observations.

3.8.2. Akaike Information Criterion (AIC)

The AIC is aimed at imposing a penalty for adding excessive independent variables in a model. The minimum AIC value represents a model with best goodness-of-fit. The AIC can be computed as follows:

$$AIC = e^{2k/N} \frac{\sum_{i=1}^{N} \left(Y_i - \hat{Y}_i\right)^2}{N} \tag{3.24}$$

where k is the number of parameters.

3.8.3. Schwarz Information Criterion (SIC)

The SIC is designed to impose a harsher penalty for addition of excessive independent variables than the AIC. Similar to the AIC, the lower the SIC, the better the model. Mathematically, it can be computed as follows:

$$SIC = N^{k/N} \frac{\sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2}{N}, \tag{3.25}$$

3.8.4. Adjusted Rho-Squared Value ($\bar{\rho}^2$)

The adjusted rho-squared value ($\bar{\rho}^2$) was used to compare the spatial multinomial models of destination choice. The model with highest $\bar{\rho}^2$ represented the apt model. Mathematically, it can be computed as follows:

$$\bar{\rho}^2 = 1 - \frac{L(\hat{\beta}) - K}{L(0)} \tag{3.26}$$

where $L(\hat{\beta})$ is the final loglikelihood, $L(0)$ is the initial loglikelihood, and K is the number of parameters estimated.

3.8.5. Likelihood Ratio Test

The best spatial model was compared to the non-spatial model by conducting the likelihood ratio test with the null hypothesis that the coefficient of a spatial variable is not different from zero. The test statistic was computed as follows:

$$-2\left(L(\widehat{\beta}_{NS}) - L(\widehat{\beta}_S)\right) \sim \chi^2_{\tau,0.05} \tag{3.27}$$

where $L(\widehat{\beta}_{NS})$ is the final loglikelihood of the non-spatial model, $L(\widehat{\beta}_S)$ is the final loglikelihood of the spatial model, and $\chi^2_{\tau,0.05}$ is the chi-squared critical value at 5% level of significance with τ degrees of freedom. τ is the number of restrictions, in this case, 1.

3.8.6. Forecasted Alternative Shares

The final comparison performed was between the short-term and long-term forecasts of destination shares. The comparison was made between the best spatial model and the non-spatial model. Three levels of increase in hotel/casino and retail jobs were used to forecast the destination shares and evaluate the difference between the two models.

3.9. Creating Origin-Destination Matrix Using MNL Results

The destination choice model shown in Equation (3.4) predicts the probability of individual n to choose destination j. However, for planning purposes, the aggregate probability of all individuals who choose destination j, is required in order to estimate the proportion of population from origin i choosing alternative destination j. The proportion is used to create the origin-destination (OD) matrix for the year of analysis. Suppose N_{ij} is the number of individuals from origin i choosing alternative j as their destination, the aggregate probability for this group, $P_{ij}^{N}(j)$ can be computed as shown in Equation (3.21).

$$P_{ij}^{N}(j) = \frac{\sum_{all\ N_{ij}} P_n(j)}{N_{ij}}, \tag{3.28}$$

Principally, the individual probabilities of selecting a specific destination are generated from trip attractions (job opportunities at the destination and spatial relativity of all possible destinations). The aggregate probability can be applied to trip productions estimated for the origin zone to create the O-D matrix. Conceptually, this makes the analysis similar to the gravity model (ITE 1992). Given *Pi* as the estimated trip productions from zone *i*, and *Aj* as the trip attractions for zone *j*, the number of trips between origin *i* and destination *j*, T_{ij} can be computed as shown in Equation (3.29). Table 3.4 shows the general form of the O-D matrix for the logit model with 17 alternative destinations, all available to each individual.

$$T_{ij} = \left(P_{ij}^{N}(j)\right) * Pi, \tag{3.29}$$

Table 3.4. General form of O-D matrix from logit model

		Attractions (at destination *j*)				
		1	**2**	**...**	**17**	***Pi***
Productions (at origin *i*)	**1**	$P_{1-1}^{N}(1) * P1$	$P_{1-2}^{N}(2) * P1$		$P_{1-17}^{N}(17) * P1$	$P1$
	2	$P_{2-1}^{N}(1) * P2$	$P_{2-2}^{N}(2) * P2$		$P_{2-17}^{N}(17) * P2$	$P2$
	. .					
	17	$P_{17-1}^{N}(1) * P17$	$P_{17-2}^{N}(2) * P17$		$P_{17-17}^{N}(17) * P17$	$P17$
	Aj	$A1$	$A2$		$A17$	$\sum Ai = \sum Pi = T$

CHAPTER 4

RESULTS AND DISCUSSIONS

A. TRIP GENERATION MODELS

4.1. Expansion Factors

Table 4.1 shows the initial expansion factors developed and used to expand the trip generation observations to reflect estimates of the Las Vegas valley population in 1996. These were computed by using the following equation which was developed in Section 3.3.4:

$$F = \left(\frac{T90_{(h,v)}}{S96_{(h,v)}} \right), \tag{4.1}$$

in which $S96_{(h,v)}$ are the sampled households in 1996 and $T90_{(h,v)}$ are the 1990 total households in a category of households with size h and v vehicles, respectively. The growth factors developed for each district as explained in Section 3.3.4 were applied to these initial factors to obtain final factors used to expand the sample.

Table 4.1. Initial expansion factors

HH Size (*h*)	**Available Vehicles per Household (*v*)**			
	0	1	2	3+
1	120.3	127.2	103.1	120.9
2	174.3	148.4	129.1	153.5
3	323.7	165.2	148.7	153.0
4+	194.2	220.6	162.0	164.2

4.2. Correlation Test Results

The correlation tests of the variables for trip attraction and trip production models were performed. The dependent variable for trip attraction model was the number of trips attracted to a district (*expattrac*) and the explanatory variables from the database were the number of gaming jobs per district and the number of non gaming jobs. Table 4.2 presents the results of the correlation test for trip attraction totals transformed by taking natural logarithm and the explanatory variables: gaming jobs and non-gaming jobs. The number in the bracket is the p-value associated with the correlation coefficient. The p-value less than 0.05 indicate that the variables are significantly correlated. The results indicated that the variables were not significantly correlated with a correlation coefficient of 0.0879 and a p-value of 0.7373.

It should be noted that when conducting a regression analysis, the assumption is that the residuals are normally distributed. One way to make it very likely to have normally distributed residuals, although not guaranteed, is to have a dependent variable that is normally distributed and predictors that are all normally distributed. Therefore, the variables were tested for normality, and the dependent variable was found to follow a distribution other than normal. To create linearity among the variables, several transformations and normality tests were performed. Close to normal and linear results were obtained for transformations by computing natural logarithm and/or by taking the inverse of the square root. Therefore, three types of models were estimated: using non-transformed number of trips, transformed by taking natural logarithm, and transformed by finding the inverse of the square root.

Table 4.2. Correlation coefficients for trip attraction model input variables

	expattraclog	*Gaming*	*Non-Gaming*
expattraclog	1.0000		
Gaming	0.5356 (0.0174)	1.0000	
Non-Gaming	0.8006 (0.0001)	0.0879 (0.7373)	1.0000

The dependent variable for the trip production model was the number of trips produced by a zone and the relevant explanatory variables were the population density, household size, and number of individuals with specific income level in a district. Table 4.3 presents the results of the correlation test for trip production totals transformed by finding the square root and the explanatory variables. Significant correlation coefficients are highlighted.

The results indicated that the variables *lowinc* and *mediuminc* are highly and significantly correlated to each other compared to other variables as shown in Table 4.3. The variable *mediuminc* is also highly and significantly correlated with *highinc*. While the *hh_size* variable was highly and significantly correlated with *highinc*, the *poparea* variable was highly and significantly correlated with *lowinc*. As a result, *poparea*, *mediuminc*, and *hh_size* variables were not included in the model.

Several transformations were performed and tested for linearity among the variables. Transformation indicated close to normal and linearity results for transformation by applying natural logarithm and for transformation by finding the square root. Three types of models were estimated: using non-transformed number of trips, transformed by taking natural logarithm, and transformed by finding the square root.

Table 4.3. Correlation Coefficients for all trip production input variables

	expproducsqrt	*lowinc*	*mediuminc*	*highinc*	*hh_size*	*poparea*
expproducsqrt	1.0000					
lowinc	0.4802 (0.0253)	1.0000				
mediuminc	0.8088 (0.0000)	0.6749 (0.0030)	1.0000			
highinc	0.8178 (0.0000)	0.1349 (0.6058)	0.6384 (0.0058)	1.0000		
hh_size	0.2513 (0.2842)	0.2393 (0.3549)	0.2288 (0.3770)	0.3239 (0.0248)	1.0000	
poparea	0.0488 (0.5511)	0.4243 (0.0896)	0.3913 (0.1203)	-0.0006 (0.9983)	0.3766 (0.1362)	1.0000

4.3. Trip Attraction Model Estimation Results

Different specifications for non-spatial model and spatial model were estimated using STATA® software. The Ordinary Least Square (OLS) method of estimation was used for the non-spatial model while the Maximum Likelihood Estimation (MLE) method was used for the spatial model. The following sub-sections give the details of estimation results.

4.3.1. Non-Spatial Models

The non-spatial trip attraction model of home-based work trips was estimated using non-transformed number of trips (Equation 4.2), number of trips transformed by taking natural logarithm (Equation 4.3), and number of trips transformed by finding the inverse of the square root (Equation 4.4). Since the non-transformed number of trips violates the OLS assumption of linearity, the results of the model estimated using non-transformed number of trips were excluded.

$$\begin{aligned}(ATTRACTIONS)_i = {} & \beta_1 * (GAMING\ JOBS)_i \\ & + \beta_2 * (NON-GAMING\ JOBS)_i\,, \\ & + CONSTANT\end{aligned} \tag{4.2}$$

$$(Ln(ATTRACTIONS))_i = \beta_1 * (GAMING\ JOBS)_i + \beta_2 * (NON-GAMING\ JOBS)_i + CONSTANT, \tag{4.3}$$

$$\left(\frac{1}{\sqrt{ATTRACTIONS}}\right)_i = \beta_1 * (GAMING\ JOBS)_i + \beta_2 * (NON-GAMING\ JOBS)_i + CONSTANT, \tag{4.4}$$

The models with transformed number of trips were compared by using the Adjusted-R squared value. It was found that the model estimated by using number of trips transformed by applying natural logarithm was better than the model using non-transformed trips in explaining variability in observed trips (Adj-R^2 of 0.84 against 0.73). Also, the t-statistics are consistently higher. Therefore, this model was chosen as the best fit for non-spatial trip attraction model. Table 4.4 shows a summary of the results of the selected non-spatial model of trip attraction. Full model results are shown in Appendix D.

The positive sign of the coefficients of all variables (*Gaming* and *Non_Gaming*) is consistent with intuition. It is expected that the higher the number of jobs available in gaming industry and other industries, the higher the number of trips attracted to a zone. Also, the positive constant indicates that the number of trips attracted to a specific zone is always positive. Mathematically, the resulting model can be written as:

$$(Ln(ATTRACTIONS))_i = 0.016 * (GAMING\ JOBS)_i + 0.033 * (NON-GAMING\ JOBS)_i + 8151.189, \tag{4.5}$$

Table 4.4. Non-Spatial Attraction Model using log transformed trips

	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
Gaming jobs	0.016	0.003	4.65	0.000	0.009	0.023
Non_Gaming jobs	0.033	0.004	7.54	0.000	0.023	0.042
_cons	8151.189	166.814	48.86	0.000	7793.409	8508.969
Number of obs: 17						
F(2, 14): 42.66						
Prob > F: 0.000						
R-squared: 0.8591						
Adjusted R-squared: 0.8389						

4.3.2. Spatial Models of Trip Attraction

To investigate the effect of location on number of trips attracted to a particular district, the variables were tested for spatial autocorrelation. The Moran's I, the Geary's C and the Getis & Ord's G statistics were estimated for all variables as well as three forms of the response variable. All three of the test results in Table 4.5 indicate that there is significant spatial autocorrelation in the number of trips attracted to zones transformed by applying natural logarithm. The results suggested incorporation of the effects of spatial location on trip attraction model.

Both spatial error and spatial lag models were estimated using the number of trips transformed by applying natural logarithm. Their final loglikelihoods were compared to identify the best fit model. Table 4.6a shows a summary of the spatial lag and spatial error trip attraction model results estimated using the binary contiguity spatial relationship. The results indicate that the spatial variable is not significant (p-value is greater than 0.05) at the 5% level of significance for both the spatial error and the spatial lag model. While the coefficients of the variables for the two models are similar, the final loglikelihood values indicate that the spatial lag model is slightly better than the spatial error model (loglikelihood of -121.85 against -121.87).

Table 4.5. Spatial autocorrelation statistics for trip attraction

MORAN's I					
Variables	**I**	**E(I)**	**sd(I)**	**z**	**p-value**
expattrac	0.289	-0.063	0.125	2.811	0.002
expattracsqrinv	0.225	-0.063	0.125	2.290	0.011
expattraclog	0.277	-0.063	0.129	2.635	0.004
Gaming	0.078	-0.063	0.066	2.114	0.017
Non_Gaming	-0.086	-0.063	0.125	-0.192	0.424
GEARY's C					
Variables	**c**	**E(c)**	**sd(c)**	**z**	**p-value**
expattrac	0.771	1.000	0.144	-1.589	0.056
expattracsqrinv	0.806	1.000	0.143	-1.359	0.087
expattraclog	0.751	1.000	0.135	-1.837	0.033
Gaming	1.026	1.000	0.217	0.119	0.453
Non_Gaming	1.107	1.000	0.145	0.741	0.229
GETIS & ORD's G					
Variables	**G**	**E(G)**	**sd(G)**	**z**	**p-value**
expattrac	0.459	0.309	0.039	3.892	0.000
expattracsqrinv	0.293	0.309	0.016	-0.989	0.161
expattraclog	0.317	0.309	0.003	2.679	0.004
Gaming	0.589	0.309	0.087	3.229	0.001
Non_Gaming	0.358	0.309	0.026	1.883	0.003

This model (spatial lag) can be written as follows:

$$\begin{aligned}\left(Ln(ATTRACTIONS)\right)_i &= 0.016*\left(GAMING\ JOBS\right)_i \\ &+0.032*\left(NON-GAMING\ JOBS\right)_i \\ &+0.002*\left(SPATIAL\right)_i \\ &+8083.894\end{aligned}, \qquad (4.6)$$

Table 4.6b shows a summary of the spatial lag and spatial error trip attraction model results estimated using the separation spatial relationship. The results indicate that the spatial variable is significant (p-value is less than 0.05) at the 5% level of significance for both the spatial error and the spatial lag model.

Table 4.6a. Spatial model of attraction using binary contiguity spatial relationship

	Spatial Lag Model				Spatial Error Model			
	Coef.	**Std. Err.**	**z**	**P>|z|**	**Coef.**	**Std. Err.**	**z**	**P>|z|**
Gaming	0.016	0.004	4.31	0.016	0.016	0.003	4.56	0.000
Non_Gaming	0.032	0.005	6.50	0.032	0.032	0.005	6.59	0.000
_cons	8083.894	347.954	23.23	0.000	8137.35	279.834	29.08	0.000
spatial	0.002	0.009	0.21	0.002	0.001	0.011	0.06	0.953
	Likelihood ratio test of rho=0: chi2(1) = 0.046 (0.830)				Likelihood ratio test of lambda=0: chi2(1) = 0.003 (0.953)			
	Number of obs = 17				Number of obs = 17			
	Log likelihood = -121.84752				Log likelihood = -121.86881			

While the coefficients of the variables for the two models in Table 4.6b are close, the final loglikelihood values indicate that the spatial lag model is slightly better than the spatial error model (loglikelihood of -118.91 against -119.00). This model (spatial lag) can be written as follows:

$$\begin{aligned}(Ln(ATTRACTIONS))_i = {} & 0.013*(GAMING\ JOBS)_i \\ & + 0.029*(NON-GAMING\ JOBS)_i \\ & + 0.059*(SPATIAL)_i \\ & + 7514.86\end{aligned}, \qquad (4.7)$$

Table 4.6b. Spatial model of attraction using separation spatial relationship

	Spatial Lag Model				Spatial Error Model			
	Coef.	**Std. Err.**	**z**	**P>|z|**	**Coef.**	**Std. Err.**	**z**	**P>|z|**
Gaming	0.013	0.003	4.56	0.000	0.013	0.003	4.62	0.000
Non_Gaming	0.029	0.004	7.89	0.000	0.028	0.004	7.84	0.000
_cons	7514.856	270.646	27.77	0.000	8063.099	136.273	59.17	0.000
spatial	0.059	0.022	2.66	0.008	0.068	0.025	2.67	0.007
	Likelihood ratio test of rho=0: chi2(1) = 5.927 (0.015)				Likelihood ratio test of lambda=0: chi2(1) = 5.735 (0.017)			
	Number of obs = 17				Number of obs = 17			
	Log likelihood = -118.907				Log likelihood = -119.003			

Table 4.6c shows a summary of the spatial lag and spatial error trip attraction model results estimated using the contiguity-separation spatial relationship. The results indicate that the spatial variable is significant (p-value is less than 0.05) at the 5% level of significance for both the spatial error and the spatial lag model. While the coefficients of the variables for the two models are close, the final loglikelihood values indicate that the spatial lag model is slightly better than the spatial error model (loglikelihood of -119.58 against -119.73). This model (spatial lag) can be written as follows:

$$\begin{aligned}(Ln(\mathit{ATTRACTIONS}))_i = {} & 0.012 * (\mathit{GAMING\ JOBS})_i \\ & + 0.028 * (\mathit{NON-GAMING\ JOBS})_i \\ & + 0.071 * (\mathit{SPATIAL})_i \\ & + 7863.54\end{aligned}, \qquad (4.8)$$

Table 4.6c. Spatial model of attraction using contiguity-separation spatial relationship

	Spatial Lag Model					Spatial Error Model			
	Coef.	**Std. Err.**	**z**	**P>|z|**		**Coef.**	**Std. Err.**	**z**	**P>|z|**
Gaming	0.012	0.003	3.88	0.000		0.013	0.003	4.10	0.000
Non_Gaming	0.028	0.004	7.34	0.000		0.029	0.004	7.35	0.000
_cons	7863.543	182.328	43.13	0.000		8543.65	230.40	37.08	0.000
spatial	0.071	0.031	2.29	0.022		0.079	0.035	2.26	0.024
	Likelihood ratio test of rho=0: chi2(1) = 4.573 (0.032)					Likelihood ratio test of lambda=0: chi2(1) = 4.293 (0.038)			
	Number of obs = 17					Number of obs = 17			
	Log likelihood = -119.58413					Log likelihood = -119.72389			

Table 4.6d shows a summary of the spatial lag and spatial error trip attraction model results estimated using the economic linkage (accessibility) spatial relationship. The results indicate that the spatial variable is significant (p-value is less than 0.05) at the 5% level of significance for both the spatial error and the spatial lag model. While the

coefficients of the variables for the two models are similar, the final loglikelihood values indicate that the spatial lag model is slightly better than the spatial error model (loglikelihood of -120.85 against -120.90). This model (spatial lag) can be written as follows:

$$\begin{aligned}(Ln(ATTRACTIONS))_i = 0.010 * (GAMING\ JOBS)_i \\ + 0.024 * (NON - GAMING\ JOBS)_i \\ + 0.005 * (SPATIAL)_i \\ + 8015.69\end{aligned}, \tag{4.9}$$

Table 4.6d. Spatial model of attraction using economic linkage (accessibility) spatial relationship

	Spatial Lag Model				Spatial Error Model			
	Coef.	**Std. Err.**	**z**	**P>|z|**	**Coef.**	**Std. Err.**	**z**	**P>|z|**
Gaming	0.010	0.005	2.15	0.031	0.011	0.005	2.19	0.029
Non_Gaming	0.024	0.007	3.32	0.001	0.024	0.007	3.26	0.001
_cons	8015.681	169.663	47.24	0.000	8054.397	158.467	50.83	0.000
spatial	0.005	0.003	1.47	0.141	0.005	0.003	1.53	0.125
	Likelihood ratio test of rho=0: chi2(1) = 2.042 (0.153)				Likelihood ratio test of lambda=0: chi2(1) = 1.943 (0.163)			
	Number of obs = 17				Number of obs = 17			
	Log likelihood = -120.84953				Log likelihood = -120.89914			

4.3.3. Comparison of Trip Attraction Models

The criteria for selecting the best method of quantifying spatial relationships were based on the goodness-of-fit of the models estimated using different spatial relationships. Table 4.7 presents the computed RSS, AIC, SIC and the final loglikelihood for the three spatial models. The results indicate that the model estimated using the spatial weight matrix defined with separation spatial relationship has the minimum RSS, AIC and SIC. In addition to minimizing RSS, AIC and SIC, it should be noted that the model estimated using contiguity-separation spatial relationship maximized the final loglikelihood. Figure 4.1 graphically compares the results. Appendix F contains the computations of RSS, AIC and SIC. Therefore, the spatial lag model estimated using separation spatial relationship was selected as the best spatial model for trip attraction. The coefficients estimated were compared to those estimated by using non-spatial model of Table 4.5.

Table 4.7. Comparison of the models for trip attraction

	Non Spatial	**Contiguity**	**Separation**	**Contiguity-Separation**	**Accessibility**
RSS	1.82	1.86	1.13	1.4	1.92
AIC	1.52	1.75	1.07	1.32	1.81
SIC	1.76	2.13	1.30	1.6	2.21
Log Likelihood	-	-121.85	-118.91	-119.58	-120.845

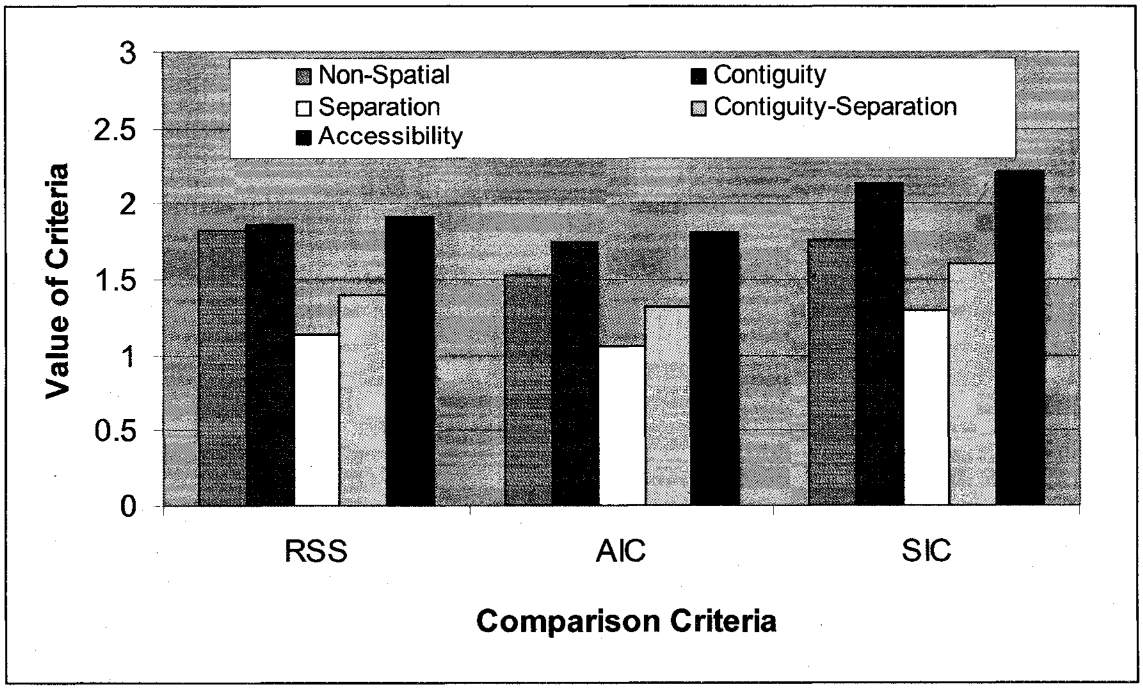


Figure 4.1. Comparison of the models of trip attraction

The comparison results showing the percentage change are presented in Table 4.8. Consistently, the comparison results indicated a decrease in the coefficients estimated with spatial trip attraction model. This suggests that non-spatial models of trip attraction overestimate the coefficients.

Table 4.8. Percentage changes in coefficient estimates for spatial models

Variable	Non Spatial Model		Spatial Model		% Change in coefficients
	Coefficient	**t-stat.**	**Coefficient**	**t-stat.**	
Gaming	0.016	4.65	0.013	4.56	-18.75
Non_Gaming	0.033	7.54	0.029	7.89	-12.12
_cons	8151.189	48.86	7514.86	27.77	-7.81
spatial	-	-	0.059	2.66	

In addition, the observed trips were compared to the modeled trips for both spatial and non-spatial trip attraction models. Figure 4.2a shows the bar-chart by district for spatial model while Figure 4.2b shows the bar-chart for non-spatial model. Figure 4.2c shows the linear comparison of modeled against observed trips for the spatial model while Figure 4.2d shows linear comparison for the non-spatial model. It can be clearly seen that the modeled trips are closer to the observed trips for the spatial model than for the non-spatial model. Figure 4.2e shows the percentage deviation of modeled trips from observed trips for both the spatial model and the non spatial model. The results indicates a smaller deviation for the spatial model, which suggests a much better prediction capability of the spatial model compared to the non-spatial model. Specifically, while the non-spatial model had an average absolute deviation of 31%, the spatial model had an average absolute deviation of only 17%—an increase of 14% in prediction accuracy.

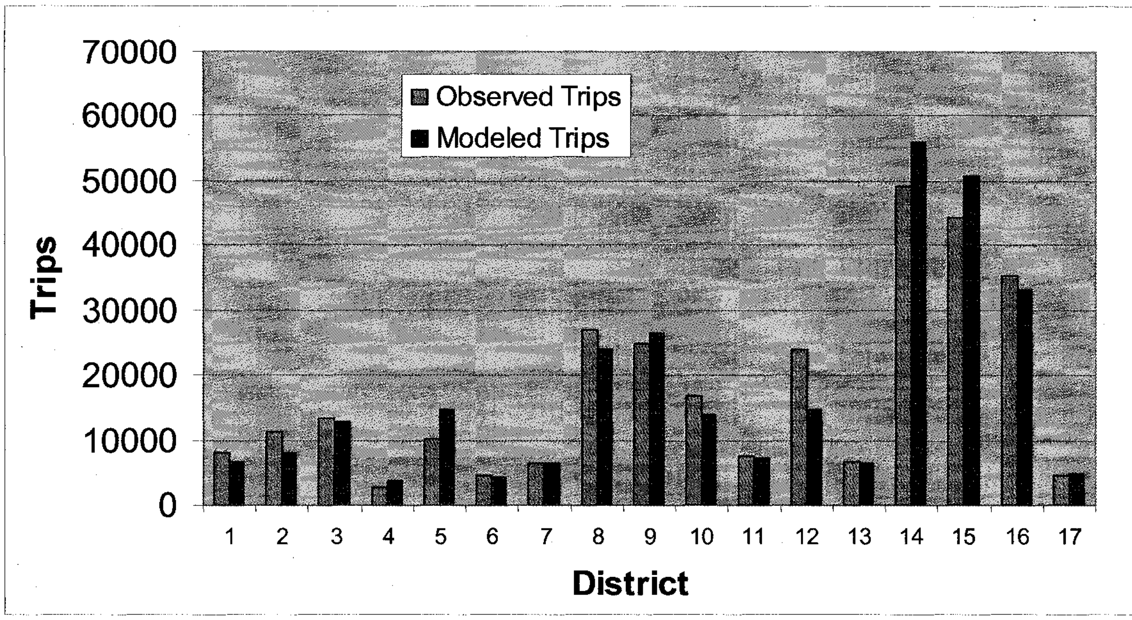


Figure 4.2a. Bar-chart comparison of observed trips against modeled trips for spatial model

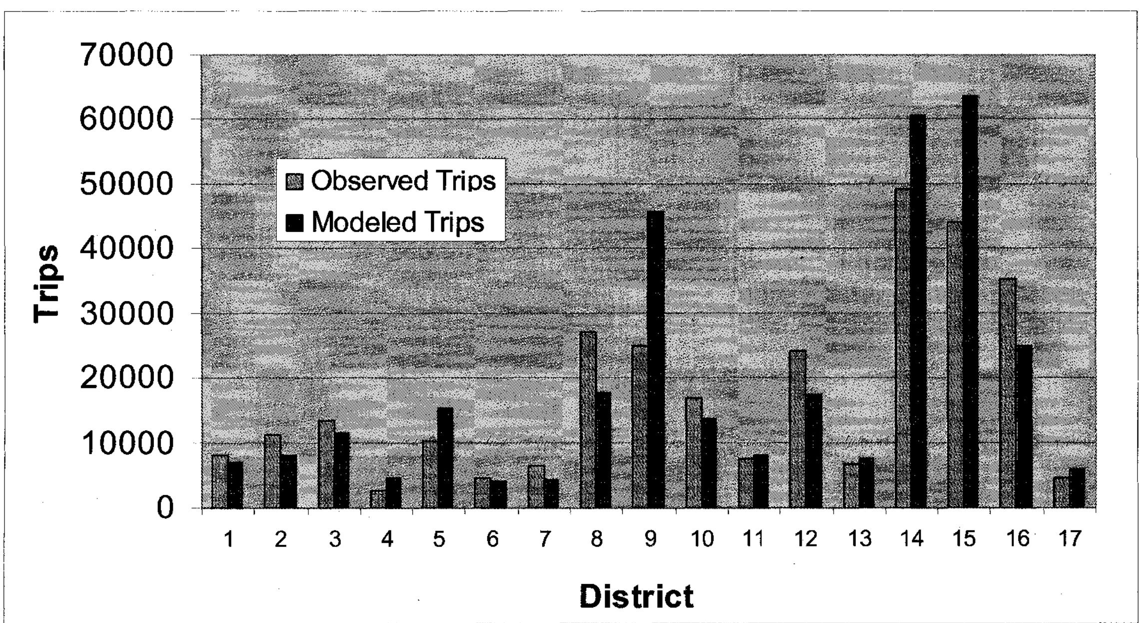


Figure 4.2b. Bar-chart comparison of observed trips against modeled trips for non-spatial model

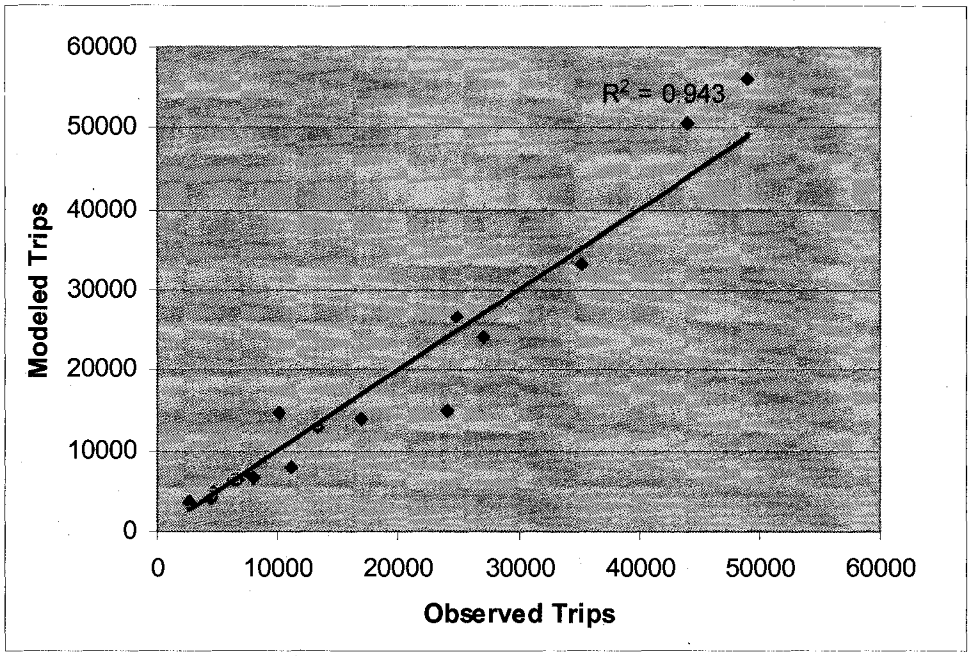


Figure 4.2c. Linear comparison of observed trips against modeled trips for spatial model

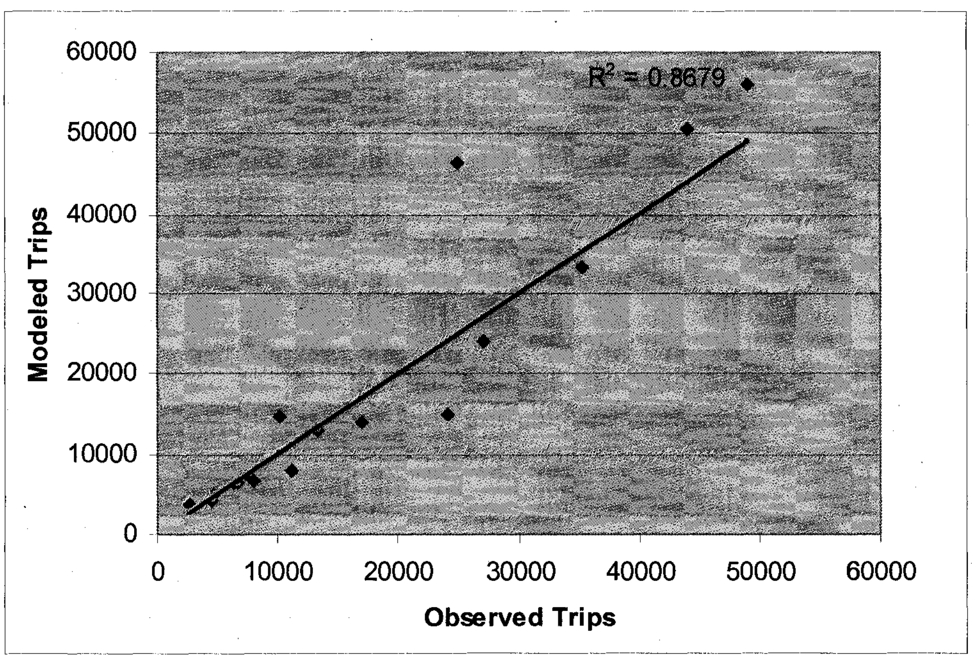


Figure 4.2d. Linear comparison of observed trips against modeled trips for non-spatial model

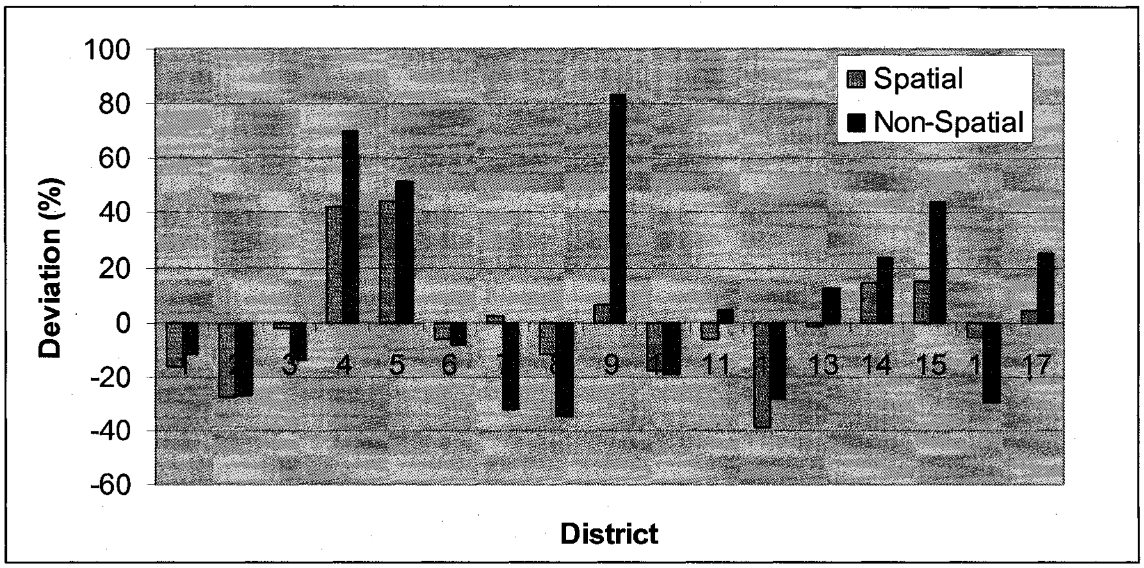


Figure 4.2e. Percentage deviation of modeled against observed trips for non-spatial model

Furthermore, the percentage deviations were plotted against observed trips to identify any presence of special pattern, for example higher deviations for zones with higher trip totals. Figure 4.3a shows the plot for the spatial model while Figure 4.2b shows the plot for non-spatial model. The results indicated no special pattern in deviations, suggesting that the models estimated using the number of trips transformed by taking natural log, have the correct functional form.

4.4. Trip Production Model Estimation Results

Similar to the trip attractions, different model specifications for non-spatial and spatial trip production models were estimated. The Ordinary Least Square (OLS) method of estimation was used for non-spatial model while the Maximum Likelihood Estimation (MLE) method was used for the spatial model. The following sub-sections present the details of estimation results.

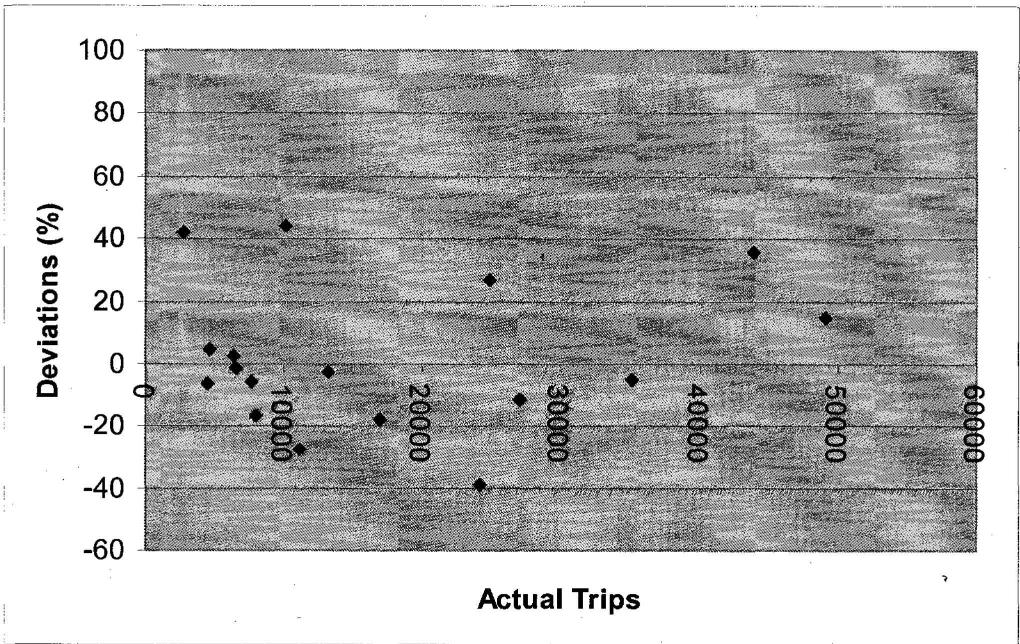


Figure 4.3a. Deviation against observed trips for the spatial model

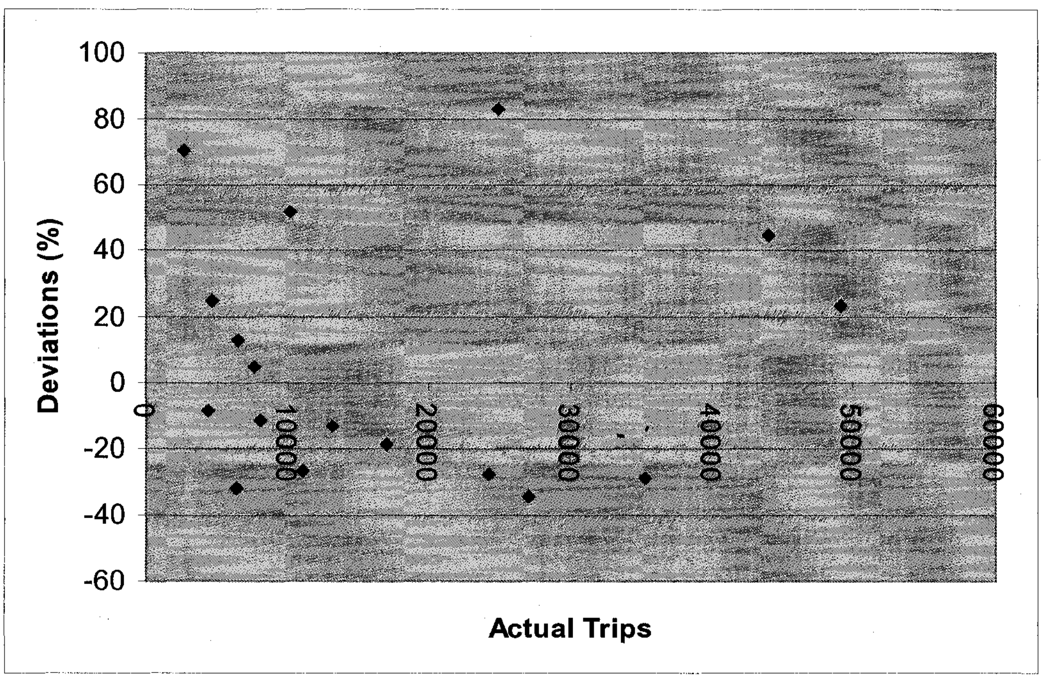


Figure 4.3b. Deviation against observed trips for the non-spatial model

4.4.1. Non-Spatial Models

The non-spatial trip production model was estimated using non-transformed number of trips (Equation 4.10), number of trips transformed by applying natural logarithm (Equation 4.11), and number of trips transformed by finding the square root (Equation 4.12). Since the non-transformed number of trips violates the linear assumption, the model results were excluded. The models with transformed number of trips were compared by using the Adjusted-R squared value.

$$\begin{aligned}(PRODUCTIONS)_i = \beta_1 * (NO.\ OF\ LOW\ INCOME\ HOUSEHOLDS)_i \\ + \beta_2 * (NO.\ OF\ HIGH\ INCOME\ HOUSEHOLDS)_i \\ + CONSTANT\end{aligned}, \quad (4.10)$$

$$\begin{aligned}(Ln(PRODUCTIONS))_i = \beta_1 * (NO.\ OF\ LOW\ INCOME\ HOUSEHOLDS)_i \\ + \beta_2 * (NO.\ OF\ HIGH\ INCOME\ HOUSEHOLDS)_i \\ + CONSTANT\end{aligned}, \quad (4.11)$$

$$\begin{aligned}(\sqrt{PRODUCTIONS})_i = \beta_1 * (NO.\ OF\ LOW\ INCOME\ HOUSEHOLDS)_i \\ + \beta_2 * (NO.\ OF\ HIGH\ INCOME\ HOUSEHOLDS)_i \\ + CONSTANT\end{aligned}, \quad (4.12)$$

It was found that the model estimated by using number of trips transformed by finding the square root was better than the model estimated using non-transformed number of trips (Adj-R^2 of 0.78 against 0.69). Also, the t-statistics are higher in the model estimated using number of trips transformed by finding the square root. Therefore, this model was chosen as the best fit for non-spatial trip production model. Table 4.9 shows a summary of the results of the selected non-spatial model of trip production. The final equation with numerical coefficients is shown in Equation (4.13). Mathematically, the resulting model can be written as follows:

$$\left(\sqrt{PRODUCTIONS}\right)_i = 1.118*\left(NO.\,OF\ LOW\ INCOME\ HOUSEHOLDS\right)_i + 0.562*\left(NO.\,OF\ HIGH\ INCOME\ HOUSEHOLDS\right)_i + 6783.664 \quad , \qquad (4.13)$$

Table 4.9. Coefficient estimates of the non-spatial trip production model

	Coef.	**Std. Err**	**t**	**P>|t|**	**[95% Conf. Interval]**	
lowinc	1.118	0.212	5.29	0.000	0.664	1.572
highinc	0.562	0.058	9.67	0.000	0.437	0.686
_cons	6783.664	647.887	10.47	0.000	5394.085	8173.242
Number of obs: 17						
F(2, 14): 29.50						
Prob > F: 0.0000						
R-squared: 0.8082						
Adj R-squared: 0.7808						

4.4.2. Spatial Models of Trip Production

Similar to the trip attraction model, the variables were tested for spatial autocorrelation to investigate the effect of location on number of trips produced from a particular district. The Moran's I, the Geary's C and the Getis & Ord's G statistics were estimated and are shown in Table 4.10. The results indicated that there is no significant spatial autocorrelation in the observations of the response variable. However, in order to confirm this finding, the spatial models were estimated and compared with the non-spatial model.

Both the spatial error and the spatial lag models were estimated using the number of trips transformed by finding the square root. Their final loglikelihoods were compared to identify the best fit model. Table 4.11a shows a summary of the spatial lag and the spatial error trip production model results estimated using the binary contiguity spatial relationship. The results indicate that the spatial variable is not significant (p-value is greater than 0.05) at the 5% level of significance for both the spatial error and the spatial lag model. In addition to the coefficients of the variables for the two models being close, the final loglikelihood value of the spatial error model is similar to that of the spatial lag model (loglikelihood of -150.849 against -150.85).

Table 4.10. Moran's I statistics for trip production model variables

MORAN-I					
Variables	**I**	**E(I)**	**sd(I)**	**z**	**p-value**
expproduc	-0.167	-0.063	0.127	-0.823	0.205
expproduclog	-0.109	-0.063	0.125	-0.371	0.355
expproducsqrt	-0.147	-0.063	0.128	-0.664	0.253
lowinc	0.078	-0.063	0.114	1.232	0.109
highinc	0.001	-0.063	0.122	0.518	0.302
GEARY C					
Variables	**c**	**E(c)**	**sd(c)**	**z**	**p-value**
expproduc	1.149	1.000	0.140	1.065	0.143
expproduclog	0.987	1.000	0.145	-0.089	0.464
expproducsqrt	1.092	1.000	0.138	0.666	0.253
lowinc	0.904	1.000	0.164	-0.588	0.278
highinc	0.992	1.000	0.150	-0.050	0.480
GETIS & ORD G					
Variables	**G**	**E(G)**	**sd(G)**	**z**	**p-value**
expproduc	0.321	0.309	0.025	0.466	0.321
expproduclog	0.312	0.309	0.002	1.089	0.138
expproducsqrt	0.318	0.309	0.012	0.806	0.210
lowinc	0.447	0.309	0.065	2.139	0.016
highinc	0.330	0.309	0.035	0.616	0.269

Mathematically, the resulting model (spatial lag) can be written as follows:

$$
\begin{aligned}
\left(\sqrt{PRODUCTIONS}\right)_i = & \ 0.86*\left(NO.\ OF\ LOW\ INCOME\ HOUSEHOLDS\right)_i \\
& +0.48*\left(NO.\ OF\ HIGH\ INCOME\ HOUSEHOLDS\right)_i \\
& +0.0012*\left(SPATIAL\right)_i \\
& +7455.67
\end{aligned}
\quad , \qquad (4.14)
$$

Table 4.11a. Spatial model of production using binary contiguity spatial relationship

	Spatial Lag Model				Spatial Error Model			
	Coef.	Std. Err.	z	P>|z|	Coef.	Std. Err.	z	P>|z|
lowinc	0.86	0.25	3.42	0.001	0.88	0.26	3.31	0.001
highinc	0.48	0.07	7.15	0.000	0.49	0.07	7.15	0.000
_cons	7455.67	1893.05	3.94	0.000	7660.45	1522.12	5.03	0.000
spatial	0.0012	0.03	0.04	0.965	-0.005	0.05	-0.10	0.923
	Likelihood ratio test of spatial=0: chi2(1) = 0.002 (0.965)				Likelihood ratio test of spatial=0: chi2(1) = 0.009 (0.923)			
	Number of obs = 17				Number of obs = 17			
	Log likelihood = -150.85303				Log likelihood = -150.84926			

Table 4.11b shows a summary of the spatial lag and the spatial error trip production model results estimated using the separation spatial relationship. The results indicate that the spatial variable is not significant (p-value is greater than 0.05) at the 5% level of significance for both the spatial error and the spatial lag model. While the coefficients of the variables for the two models are close, the final loglikelihood values indicate that the spatial error model is slightly better than the spatial lag model (loglikelihood of -146.15 against -146.34 respectively). Mathematically, the resulting model (spatial lag) can be written as follows:

$$
\begin{aligned}
\left(\sqrt{PRODUCTIONS}\right)_i = & \ 1.014*\left(NO.\ OF\ LOW\ INCOME\ HOUSEHOLDS\right)_i \\
& +0.582*\left(NO.\ OF\ HIGH\ INCOME\ HOUSEHOLDS\right)_i \\
& +0.073*\left(SPATIAL\right)_i \\
& +5393.86
\end{aligned}
\quad , \qquad (4.15)
$$

Table 4.11b. Spatial model of production using separation spatial relationship

	Spatial Lag Model				Spatial Error Model			
	Coef.	**Std. Err.**	**z**	**P>|z|**	**Coef.**	**Std. Err.**	**z**	**P>|z|**
lowinc	1.014	0.217	4.68	0.000	1.022	0.230	4.45	0.000
highinc	0.582	0.055	10.50	0.000	0.573	0.054	10.65	0.000
_cons	5393.864	1568.864	3.44	0.001	6456.935	796.109	8.11	0.000
spatial	0.073	0.077	0.95	0.341	0.111	0.152	0.73	0.467
	Likelihood ratio test of spatial=0: chi2(1) = 0.882 (0.348)				Likelihood ratio test of spatial=0: chi2(1) = 0.504 (0.478)			
	Number of obs = 17				Number of obs = 17			
	Log likelihood = -146.15213				Log likelihood = -146.3408			

Table 4.11c shows a summary of the spatial lag and the spatial error trip production model results estimated using the contiguity-separation spatial relationship. The results indicate that the spatial variable is not significant (p-value is greater than 0.05) at the 5% level of significance for both the spatial error and the spatial lag model.

Table 4.11c. Spatial model of production using contiguity-separation spatial relationship

	Spatial Lag Model				Spatial Error Model			
	Coef.	**Std. Err.**	**z**	**P>|z|**	**Coef.**	**Std. Err.**	**z**	**P>|z|**
lowinc	0.88	0.28	3.19	0.001	0.91	0.29	3.13	0.002
highinc	0.48	0.07	6.55	0.000	0.48	0.07	6.70	0.000
_cons	7705.03	1437.47	5.36	0.000	7380.13	900.62	8.19	0.000
spatial	-0.02	0.13	-0.14	0.887	-0.06	0.23	-0.28	0.778
	Likelihood ratio test of spatial=0: chi2(1) = 0.020 (0.887)				Likelihood ratio test of spatial=0: chi2(1) = 0.080 (0.777)			
	Number of obs = 17				Number of obs = 17			
	Log likelihood = -150.84389				Log likelihood = -150.81381			

While the coefficients of the variables for the two models are close, the final loglikelihood values indicate that the spatial error model is slightly better than the spatial lag model (loglikelihood of -150.81 against -150.84 respectively). Mathematically, the resulting model (spatial error) can be written as follows:

$$\begin{aligned}\left(\sqrt{PRODUCTIONS}\right)_i = & \; 0.88*\left(NO.\,OF\ LOW\ INCOME\ HOUSEHOLDS\right)_i \\ & + 0.48*\left(NO.\,OF\ HIGH\ INCOME\ HOUSEHOLDS\right)_i \\ & - 0.02*\left(SPATIAL\right)_i \\ & + 7705.03\end{aligned} \quad , \qquad (4.16)$$

Table 4.11d shows a summary of the spatial lag and the spatial error trip production model results estimated using the economic linkage (accessibility) spatial relationship. The results indicate that the spatial variable is not significant (p-value is greater than 0.05) at the 5% level of significance for both the spatial error and the spatial lag model.

Table 4.11d. Spatial model of production using economic linkage (accessibility) spatial relationship

	Spatial Lag Model					**Spatial Error Model**			
	Coef.	**Std. Err.**	**z**	**P>|z|**		**Coef.**	**Std. Err.**	**z**	**P>|z|**
lowinc	0.87	0.27	3.25	0.001		0.89	0.28	3.23	0.001
highinc	0.48	0.07	7.15	0.000		0.49	0.07	7.08	0.000
_cons	7563.62	1045.42	7.23	0.000		7649.5	967.4	7.91	0.000
spatial	-0.0003	0.01	-0.04	0.964		-0.0022	0.01	-0.19	0.847
	Likelihood ratio test of spatial=0: chi2(1) = 0.002 (0.964)					Likelihood ratio test of spatial=0: chi2(1) = 0.038 (0.846)			
	Number of obs = 17					Number of obs = 17			
	Log likelihood = -150.85298					Log likelihood = -150.83507			

While the coefficients of the variables for the two models are close, the final loglikelihood values indicate that the spatial error model is slightly better than the spatial lag model (loglikelihood of -150.84 against -150.85). However, the spatial variable is not significant in both models. Mathematically, the resulting model (spatial lag) can be written as follows:

$$\begin{aligned}\left(\sqrt{PRODUCTIONS}\right)_i = & \ 0.87*\left(NO.\ OF\ LOW\ INCOME\ HOUSEHOLDS\right)_i \\ & +0.48*\left(NO.\ OF\ HIGH\ INCOME\ HOUSEHOLDS\right)_i \\ & -0.0003*\left(SPATIAL\right)_i \\ & +7563.62\end{aligned} \quad , \quad (4.17)$$

4.4.3. Comparison of the Trip Production Models

The criteria for selecting the best model were based on the goodness-of-fit of the models. Table 4.12 presents the computed RSS, AIC, SIC and the final loglikelihood for the three selected spatial models, one for each alternative quantification of the spatial relationship. The non-spatial model was used as a base for comparison.

Table 4.12. Comparison of the models of trip production

	Non Spatial	Contiguity	Separation	Contiguity-Separation	Accessibility
RSS (10^7)	6.00	5.81	6.17	6.25	6.29
AIC (10^6)	5.03	5.47	5.81	5.89	5.92
SIC (10^6)	5.82	6.65	7.07	7.16	7.21
Final Loglikelihood	-	-150.85	-146.15	-150.81	-150.84

The results indicate that while the spatial models minimize RSS, they have higher AIC and SIC. This implies that the spatial variable in the spatial models of trip production is unnecessary as shown in Tables 4.11b, 4.11c, and 4.11d. In addition, the likelihood ratio test indicates that the coefficient associated with the spatial variable is not significantly different from zero (p-value greater than 0.05) as shown in these tables. Figure 4.4 shows the graphical comparison results. Appendix F shows the details of computations of RSS, AIC and SIC.

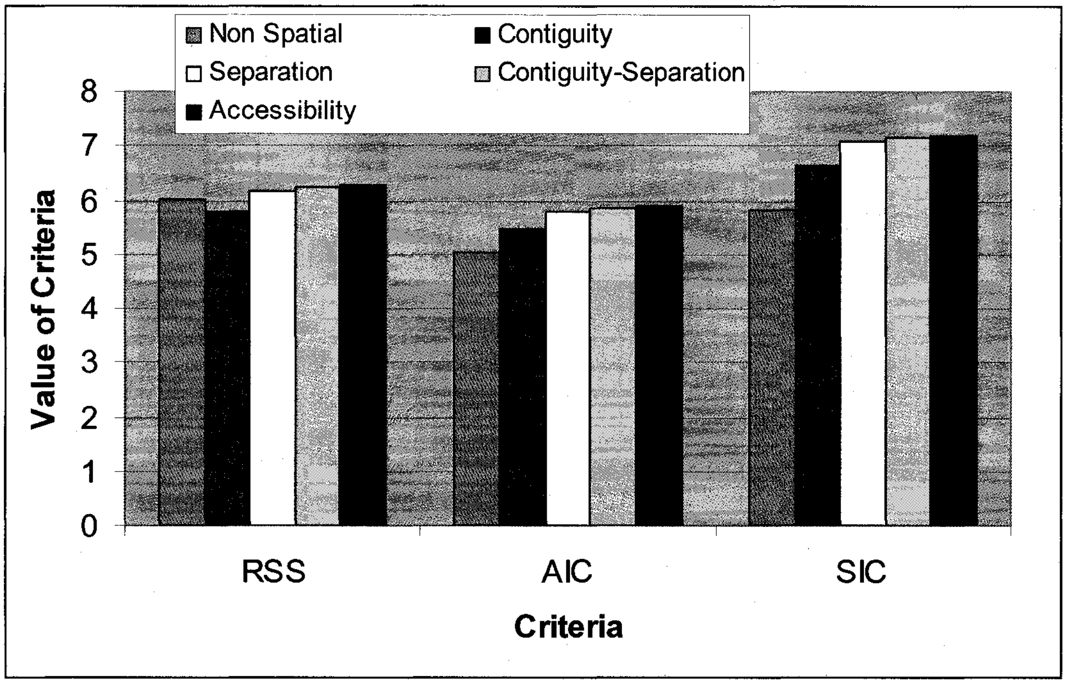


Figure 4.4. Comparison of the models of trip production

Furthermore, the average absolute percentage deviation of modeled trips from observed trips were compared for non-spatial model and spatial models estimated using the four alternatives of defining spatial relationship. Figure 4.5 shows that the models developed with spatial variable do not minimize percentage deviation when compared to the non-spatial model.

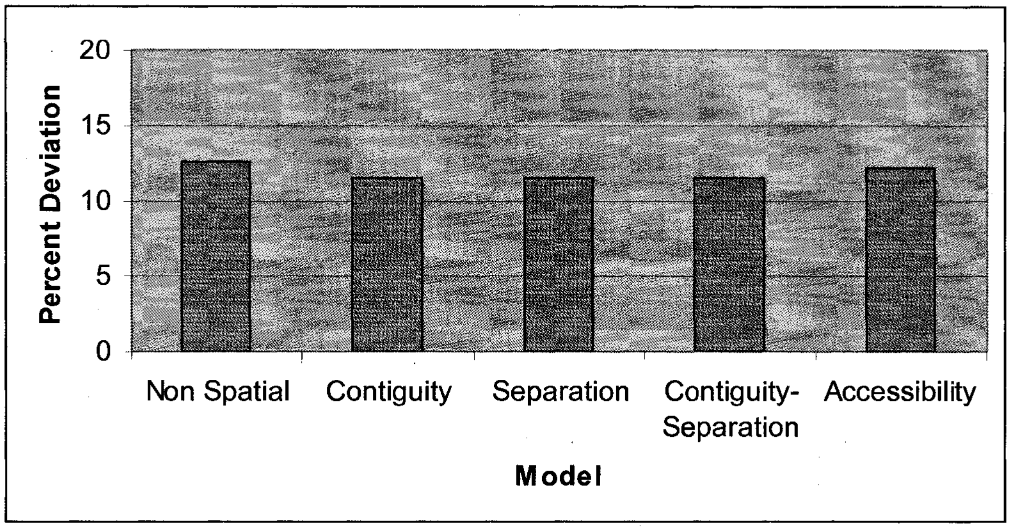


Figure 4.5. Comparison of the percent deviation for the models of trip production

4.5. General Findings on Trip Generation Models

The first finding was that for both trip attraction and trip production models, the non transformed observations do not satisfy the linear assumption of the models. This was resolved by transforming the response variable into other forms. The trip attraction observations were transformed by applying natural logarithm, while the trip production observations were transformed by finding the square root. The second finding was that the trip attraction observations are spatially autocorrelated significantly while the trip production observations are not spatially autocorrelated. This finding is consistent with intuition because for trip production, there is no direct linkage between one making a trip from one zone to someone else making a trip from an adjacent zone. However, for trip attraction, it is possible that whatever attracts someone (from somewhere) to a specific zone is related to what attracts someone else (from somewhere else) to an adjacent zone. If this case is true, the observations made at these two adjacent zones would be spatially correlated. The last key finding is that using the contiguity spatial relationship only does not fully explain the effect of spatial relationship in the observations. However, introduction of separation between the observations improves the explanation of the effect of spatial relationship.

B. DESTINATION CHOICE MODELS

4.6. Descriptive Statistics and Correlation Test

Table 4.13 shows the descriptive statistics of the variables used. The results indicate that the average age of people who participated in the survey was 46.53 years while the maximum age was 87.

Table 4.13. Descriptive statistics for selected variables

Variable	Obs	Mean	Std. Dev.	Min	Max
cbd	1113	1.43	2.24	0.02	42.99
hotelre	1113	41939.39	35868.50	2709.58	116928.30
otherjo	1113	2143.64	1306.28	172.34	4990.91
vehhhworker	1113	0.83	0.45	0	4
age	1112	46.58	13.47	16	87
income	1113	2.35	0.62	1	3

Table 4.14 presents the correlation coefficients for selected variables. The results indicate that the variables selected are less correlated. This suggests that all variables can be used together to estimate the destination choice model.

Table 4.14. Correlation coefficients of the variables

	cbd	*hotelre*	*otherjo*	*vehhhworker*	*age*	*income*
cbd	1.000					
hotelre	-0.076	1.000				
otherjo	-0.145	0.410	1.000			
vehhhworker	-0.026	-0.008	-0.040	1.000		
age	-0.010	0.022	0.091	0.104	1.000	
income	-0.171	-0.008	-0.042	0.226	-0.031	1.000

4.7. Destination Choice Model Estimation Results

The models with all variables were initially specified and estimated. Table 4.15 presents a summary of the results for the model with all variables while Appendix G presents the full results. District eight and fifteen were used as basis since they constitute the resort corridor and the downtown area. The two locations contain most of the hotel/casino jobs. The model results indicated that only hotel/casino and retail jobs (*hotelre*) and other jobs (*otherjo*) were the significant variables. The spatial variable was also significant in the spatial model. Since the alternative specific constant was also not significant, the final models were estimated with generic constant. It should be noted that a multinomial logit model with generic is the same as the one without a constant—the constant does not affect the utilities.

The non-spatial model and the spatial model using hotel/casino and retail, and other jobs only were finally estimated. The spatial model included a variable quantifying the separation between the origin and the destination.

Table 4.16 shows the results for the non-spatial model. The results indicate that all coefficients have the signs consistent with intuition. The coefficient for casino/hotel and retail jobs, which is the number of jobs in hotel/casino and retail available in a district have a positive sign indicating that the increase in casino/hotel and retail jobs results in increase in utility of the alternative being chosen as a destination. The same interpretation is applicable to the number of jobs other than hotel/casino and retail. Mathematically, the resulting model can be written as follows:

$$P_n(j) = \frac{e^{0.17*(HOTEL\,/\,CASINO\ JOBS)_j + 0.47*(OTHER\ JOBS)_j + 0.24}}{\sum_{p \in C_n} e^{0.17*(HOTEL\,/\,CASINO\ JOBS)_p + 0.47*(OTHER\ JOBS)_p + 0.24}}, \quad p \neq j \tag{4.18}$$

Table 4.15. Results for the model with all variables

	Socioeconomics			Destination Attributes				
Alt.	*Vehicles per drivers*	*age*	*Income*	*CBD*	*Jobs in hotel and retail*	*Other jobs*	*Spatial*	*Constant*
1	0.789 (1.82)	0.015 (0.87)	-0.350 (-0.73)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	-2.143 (-1.15)
2	-0.306 (-0.65)	-0.018 (-1.27)	0.084 (0.20)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	-0.360 (-0.47)
3	0.688 (1.96)	-0.008 (-0.64)	0.127 (0.32)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	-1.320 (-1.80)
4	0.682 (0.99)	-0.030 (-1.09)	0.965 (0.89)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	-1.057 (-0.60)
5	0.406 (0.83)	0.014 (0.85)	0.223 (0.43)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	-2.890 (-1.85)
6	0.644 (1.09)	-0.019 (-0.87)	-0.538 (-0.87)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	-1.734 (-1.47)
7	0.286 (0.50)	-0.037 (-1.94)	-0.902 (-1.77)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	-0.141 (-0.15)
8				0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	
9	0.074 (0.22)	-0.010 (-1.03)	0.014 (0.05)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	0.605 (1.06)
10	-0.404 (-1.01)	-0.006 (-0.47)	-0.177 (-0.52)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	0.813 (1.21)
11	0.487 (0.96)	-0.011 (-0.61)	-0.436 (-0.86)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	0.947 (0.85)
12	0.188 (0.50)	-0.024 (-1.95)	0.610 (1.50)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	0.830 (1.13)
13	0.021 (0.04)	-0.033 (-1.96)	1.175 (1.80)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	-1.007 (-1.00)
14	0.352 (1.29)	-0.009 (-1.04)	0.233 (0.87)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	0.478 (0.95)
15				0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	
16	0.489 (1.63)	0.001 (0.10)	0.109 (0.36)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	-0.652 (-1.13)
17	0.749 (1.40)	0.003 (0.14)	-0.389 (-0.66)	0.023 (1.54)	0.168 (20.21)	0.476 (17.56)	2.640 (3.15)	-1.926 (-1.60)

Table 4.16. Non-Spatial model results

Name	Value	Std err	t-test	p-value	Robust Std err	Robust t-test	p-value
B_hotelre	0.17	0.008	20.17	0.000	0.008	20.30	0.0000
B_otherjo	0.47	0.026	17.70	0.000	0.026	17.69	0.0000
C	0.24	0.104	2.33	0.020	0.104	2.33	0.0200
Number of individuals: 1,113 Init log-likelihood: -3,153.366 Final log-likelihood: -2785.457 Likelihood ratio test: 735.818 Rho-square: 0.117 Adjusted rho-square: 0.116							

Table 4.17a shows the results for spatial model using separation (distance) as the definition of spatial relationships between the origin and destinations. The results indicate that all coefficients have the signs consistent with intuition. The coefficient for casino/hotel and retail jobs, which is the number of jobs in hotel/casino and retail available in a district have a positive sign indicating that the increase in casino/hotel and retail jobs results in increase in utility of the alternative being chosen as a destination. The same interpretation is applicable to *otherjo*, which is the number of jobs other than hotel/casino and retail. The spatial coefficient is positive indicating that the higher the spatial relationship between the origin and the destination, the higher the utility of the alternative to the decision maker. It should be noted that the higher separation measure means shorter distance between origin and destination since the measure is an inverse of the distance. Mathematically, the resulting model for individual n to choose destination j can be written as follows:

$$P_n(j) = \frac{e^{0.17*(HOTEL\,/\,CASINO\ JOBS)_j + 0.47*(OTHER\ JOBS)_j + 2.64*(SPATIAL)_p + 0.158}}{\sum_{p \in C_n} e^{0.17*(HOTEL\,/\,CASINO\ JOBS)_p + 0.47*(OTHER\ JOBS)_p + 2.64*(SPATIAL)_p + 0.158}}, \quad p \neq j \qquad (4.19)$$

Table 4.17a. Spatial model results using separation spatial relationship

Name	Value	Std err	t-test	p-value	Robust Std err	Robust t-test	p-value
B_hotelre	0.17	0.008	20.18	0.0000	0.008	20.30	0.0000
B_otherjo	0.47	0.026	17.70	0.0000	0.026	17.69	0.0000
B_spatial	2.64	0.838	3.15	0.0004	0.834	3.17	0.0004
C	0.158	0.187	0.85	0.4000	0.191	0.83	0.41
Number of individuals: 1113 Init log-likelihood: -3153.366 Final log-likelihood: -2780.124 Likelihood ratio test: 744.485 Rho-square: 0.118 Adjusted rho-square: 0.117							

Table 4.17b shows the results for spatial model using economic linkage—accessibility measure as the definition of spatial relationships between the origin and destinations. The results indicate that all coefficients have the signs consistent with intuition. The coefficient for casino/hotel and retail jobs, which is the number of jobs in hotel/casino and retail available in a district have a positive sign indicating that the increase in casino/hotel and retail jobs results in increase in utility of the alternative being chosen as a destination. The same interpretation is applicable to *otherjo*, which is the number of jobs other than hotel/easino and retail. The spatial coefficient is positive indicating that the higher the spatial relationship between the origin and the destination, the higher the utility of the alternative to the decision maker. Also, it should be noted that the spatial relationship variable has relatively higher marginal impact on the utility of the alternatives. Mathematically, the resulting model for individual n to choose destination j can be written as follows:

$$P_n(j) = \frac{e^{0.17*(HOTEL/CASINO\ JOBS)_j + 0.47*(OTHER\ JOBS)_j + 4.69*(SPATIAL)_p - 0.07}}{\sum_{p \in C_n} e^{0.17*(HOTEL/CASINO\ JOBS)_p + 0.47*(OTHER\ JOBS)_p + 4.69*(SPATIAL)_p - 0.07}}, \quad p \neq j \tag{4.20}$$

Table 4.17b. Spatial model results using economic linkage (accessibility measure) spatial relationship

Name	Value	Std err	t-test	p-value	Robust Std err	Robus t t-test	p-value
B_hotelre	0.17	0.008	20.18	0.0000	0.008	20.30	0.0000
B_otherjo	0.47	0.026	17.70	0.0000	0.026	17.69	0.0000
B_spatial	4.69	2.110	2.22	0.0300	2.020	2.32	0.0200
C	-0.07	0.168	-0.41	0.6800	0.164	-0.42	0.6700
Number of individuals: 1113 Init log-likelihood: -3153.366 Final log-likelihood: -2782.962 Likelihood ratio test: 740.808 Rho-square: 0.117 Adjusted rho-square: 0.116							

4.8. Model Evaluation and Comparison

The final loglikelihoods and adjusted rho-squared values were compared to identify the best model as shown in Figure 4.6. The results indicated that while all models had the same initial loglikelihood, the non spatial model had the lowest final loglikelihood and adjusted rho-squared. The model developed using separation spatial relationship had the maximum final loglikelihood and adjusted rho-squared. This indicates that the spatial model developed using separation spatial relationship is the best fitted model. Intuitively, the model using accessibility as the measure of spatial interaction was expected to be the best fitted model. However, since the accessibility measure uses total employment which contains employments in hotel/casino and retail and other jobs, it creates dependence between the variables, and hence disqualifies it from being the best fitted model.

The model estimated with separation measure was further compared with the non-spatial model by conducting the likelihood ratio test with the null hypothesis that the coefficient of the separation measure variable was not different from zero. The test statistic obtained (10.67) was greater than the critical value (3.84), suggesting that the

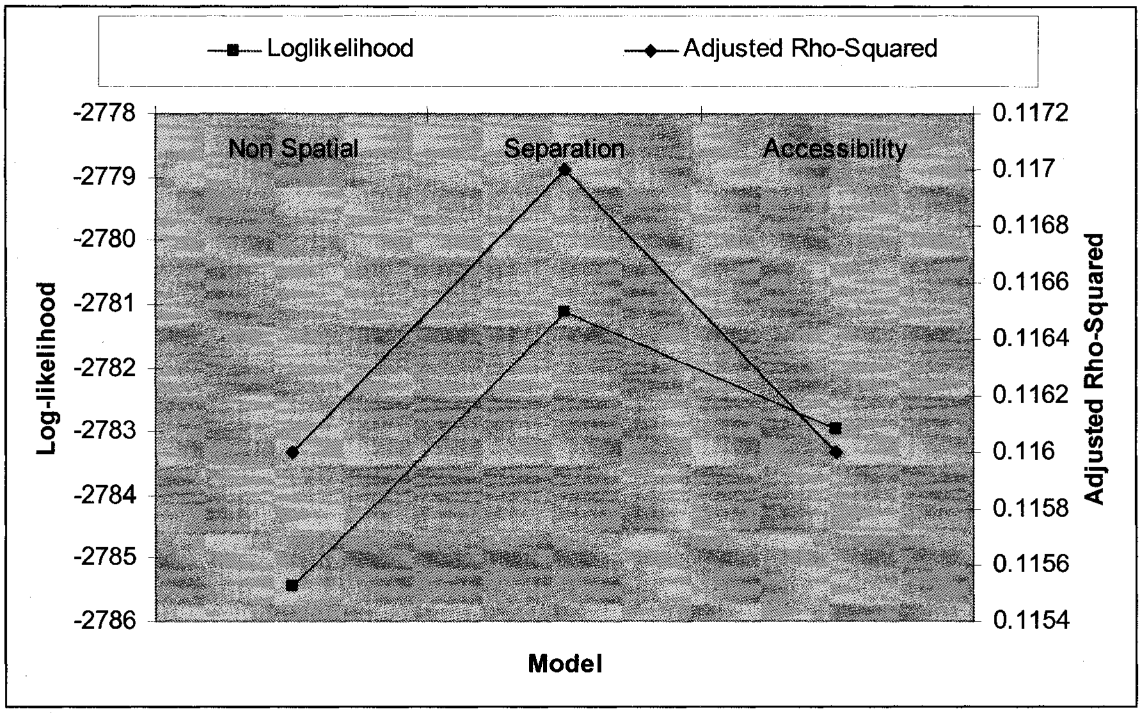


Figure 4.6. Comparison of the models of destination choice

coefficient associated with the spatial variable was significantly different from zero at 5% level of significance. This suggests that the model incorporating separation measure is the best specification for the destination choice model. In order to quantify the difference between the model without separation measure and the model with separation measure, sensitivity analysis as well as forecasting capability was analyzed. The following sections provide the details of sensitivity analysis. Finally, a demonstration of how to use the model results to create origin-destination (O-D) matrix in practice is given.

4.9. Sensitivity analysis

In order to perform sensitivity analysis, the Las Vegas valley was divided into two areas: (1) Inner Districts (7, 8, 9, 10, 14, 15, and 16) and (2) Outer Districts (1, 2, 3, 4, 5, 6, 11, 12, 13, and 17). The inner districts include all districts containing resort corridor

(Strip) and central business district (CBD) and other neighboring districts. The outer districts contain all other districts, as shown in Figure 4.7. The reason for dividing the valley into these parts was the assumption that jobs in hotel/casino and retails may have different growth rates for the inner and outer districts. Three scenarios were considered:

(a) Increase in casino/hotel and retail jobs in outer districts only,

(b) Increase in casino/hotel and retail jobs in inner districts only, and

(c) Equal increase in casino/hotel and retail jobs in outer and inner districts.

The increase in casino/hotel and retail jobs was assumed at 10%, 20% and combinations of 10%/5% and 20%/10%. The shares for each alternative destination were computed for each scenario by using the following equation:

$$P_{ij}^{N}(j) = \frac{\sum_{all\ N_{ij}} P_n(j)}{N_{ij}}, \tag{4.21}$$

in which $\sum_{all\ N_{ij}} P_n(j)$ is the sum of individuals, N_{ij}, who choose district j.

The difference between spatial model forecasted shares and non-spatial model forecasted shares was analyzed. The spatial model was used for sensitivity analysis of scenario (a) and (b) in which the percentage change was measured as the difference between the base year shares and the forecasted shares. Finally, the difference in shares computed using the non-spatial model and those computed using the spatial model at different percentage changes in casino/hotel and retail jobs was compared in scenario (c) in which uniform increase in casino/hotel and retail jobs in inner and outer districts was assumed. The difference was computed by subtracting the shares computed using the non-spatial model from those computed using the spatial model.

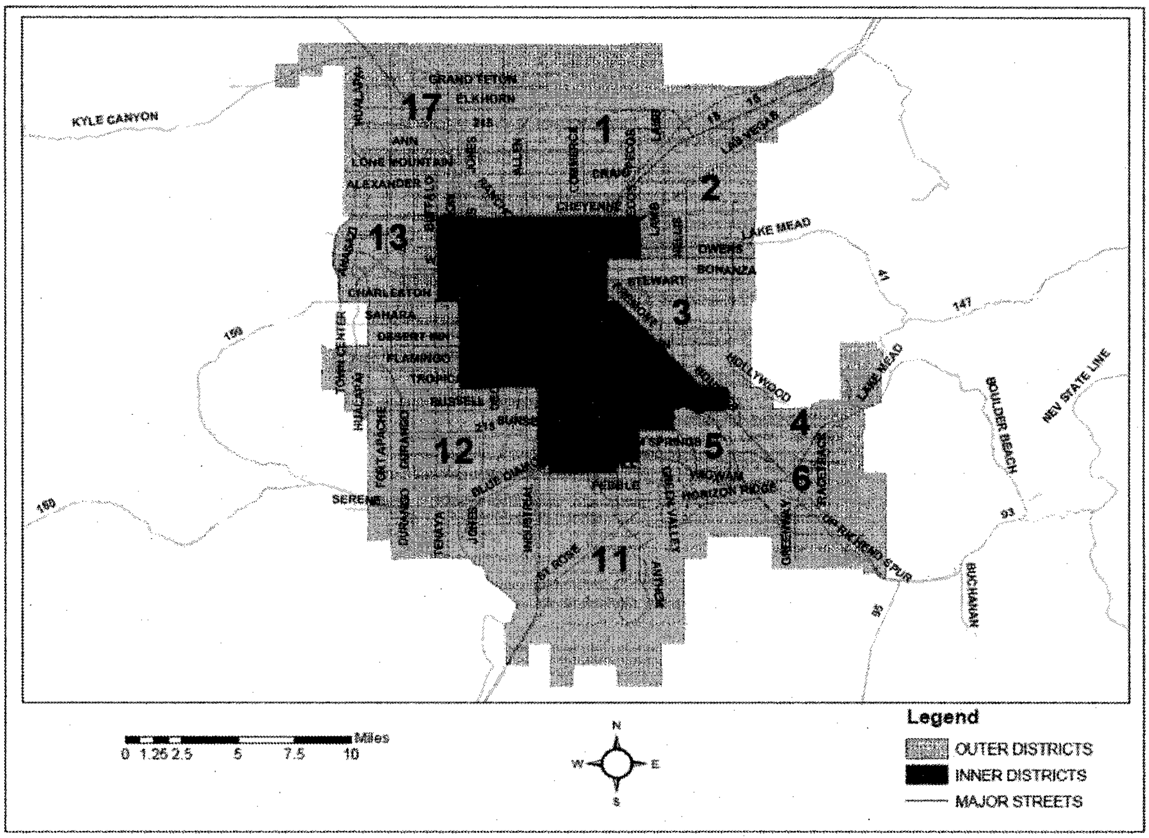


Figure 4.7. Inner and outer districts

4.9.1. Increase in casino/hotel and retail jobs in outer districts only

In this scenario, the number of jobs in hotel/casino and retail (*hotelre*) was increased by 10% and 20% in outer districts only. The spatial destination choice model was used for this analysis. The forecasted proportions (shares) of individuals choosing each alternative destination district at each percentage increase were computed. The change between base year shares and new shares was computed by subtracting the base share from the new share at a specified percentage increase in casino/hotel and retail jobs. Figure 4.8 shows the changes in probability for each percentage increase in casino/hotel and retail jobs. The results indicate that there is a small change in the shares for almost all districts. Also, the results indicate that the change is higher for higher percentage increase in casino/hotel and retail jobs. The result imply that an increase in hotel/casino and retail employment in outer districts which exclude the resort corridor and downtown areas is likely to change destination choice patterns. Individuals will prefer making work trips to outer districts than inner districts in response to the additional jobs in those districts.

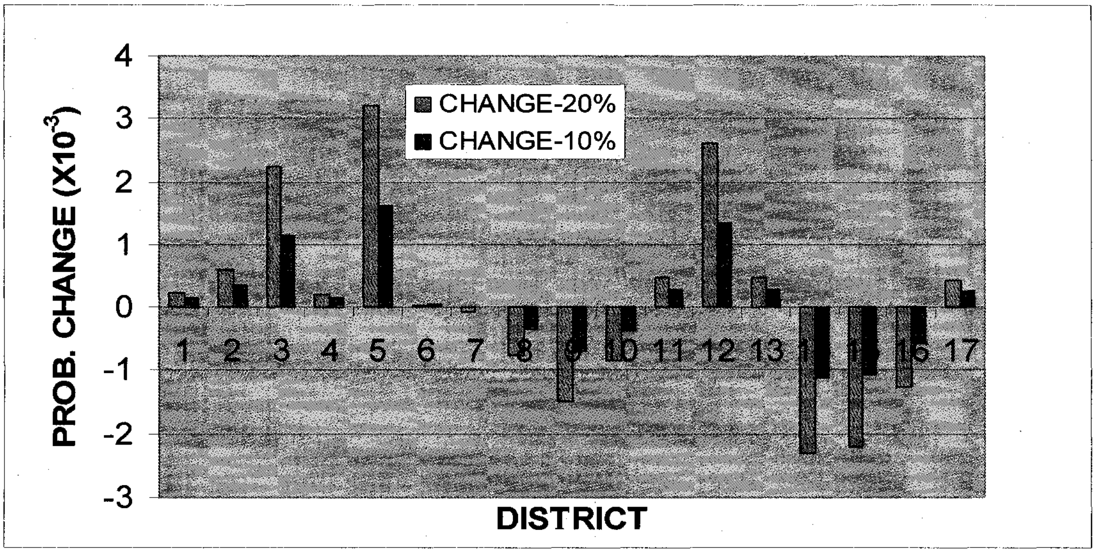


Figure 4.8. Increase in casino/hotel and retail jobs in outer districts only

4.9.2. Increase in casino/hotel and retail jobs in inner districts only

In this scenario, the number of jobs in hotel/casino and retail (*hotelre*) was increased by 10% and 20% in inner districts only. The spatial destination choice model was used for this analysis. The forecasted proportions (shares) of individuals choosing each alternative destination district at each percentage increase were computed. The change between base year shares and new shares was computed by subtracting the base share from the new share at a specified percentage increase in casino/hotel and retail jobs. Figure 4.9 shows the changes in probability for each percentage increase in casino/hotel and retail jobs.

The results indicate that there is a reduction in the shares for almost all districts except district 15 and 14, which are among the inner districts. However, different from the case when there is an increase in outer districts only (Figure 4.7), the highest increase of share is only for district 15 which contains the resort corridor. The result suggests that increase in hotel/casino and retail jobs in the resort corridor will attract more workers. Also, minor changes in shares in other districts suggests that the base year shares for inner districts are higher and hence adding new jobs only in these districts only may not result into new destination choice patterns.

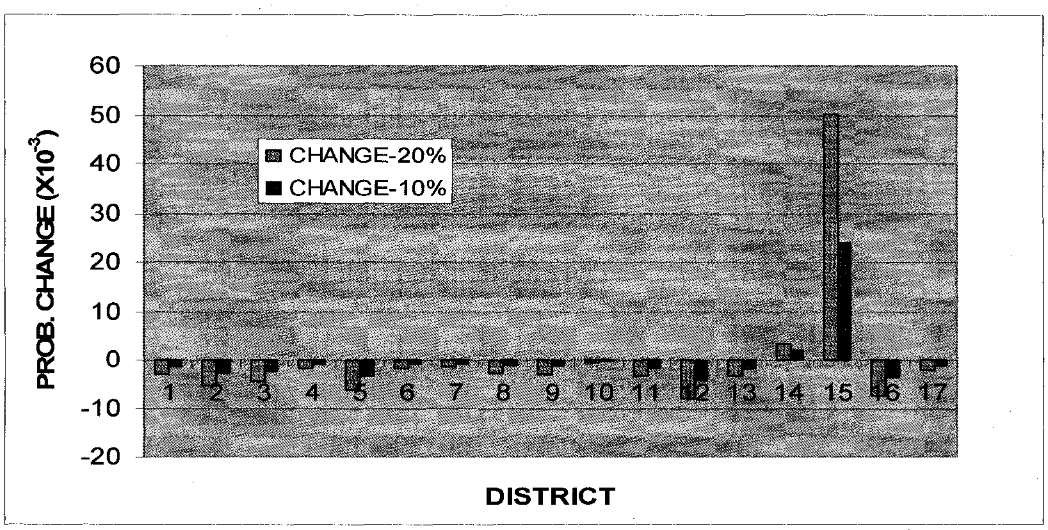


Figure 4.9. Increase in casino/hotel and retail jobs in inner districts only

4.9.3. Equal increase in casino/hotel and retail jobs in outer and inner districts

In order to compare the forecasts of the model with separation measure to those of the model without separation measure, the difference between forecasted destination shares at different percentage increase in hotel/casino and retail jobs were computed and plotted as shown in Figure 4.10. The difference was computed by subtracting the shares computed using the non-spatial model from those computed using the spatial model. The results indicate that the difference between the forecasted destination shares is negligible when there is a small percentage increase in casino/hotel and retail jobs. However, the difference between the forecasted shares is noticeable at higher percentage increase in casino/hotel and retail jobs. The implication of this finding is that the difference between spatial model and non-spatial model can be seen only when there is a higher percentage increase in hotel/casino jobs in the valley.

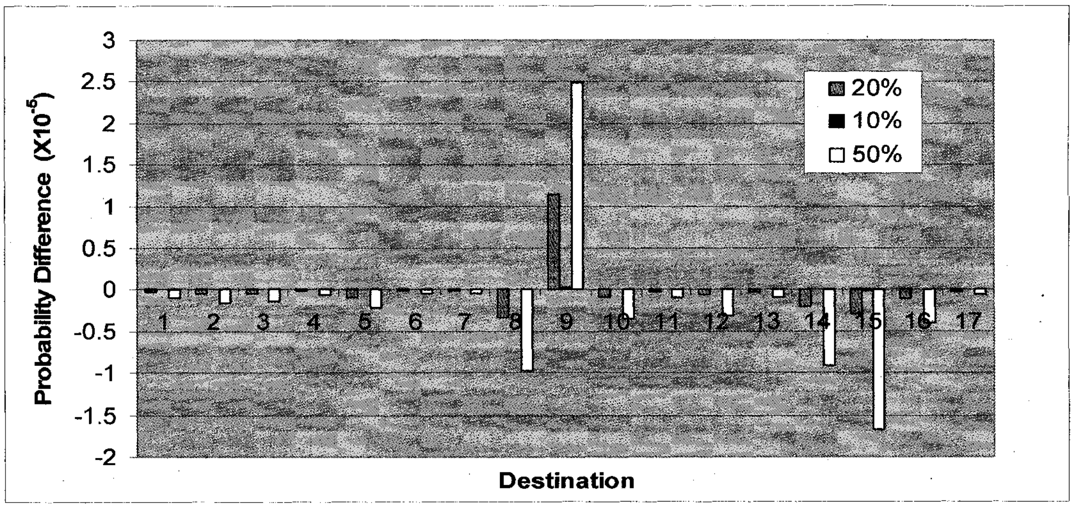


Figure 4.10. Difference in forecasted destination shares

4.9.4. Implication of sensitivity analysis results

The sensitivity analysis focused on three scenarios: increase in casino/hotel and retail jobs in outer districts only, increase in casino/hotel and retail jobs in inner districts only, and equal increase in casino/hotel and retail jobs in outer and inner districts. The results indicated that if there is an increase in casino/hotel and retail jobs in outer districts only, there is a change in the shares for almost all districts. Also, the results indicated that the change is higher for higher percentage increase in casino/hotel and retail jobs.

When there is an increase in casino/hotel and retail jobs in inner districts only, there is a change in the shares for almost all districts. However, different from the case when there is an increase in outer districts only, the changes are only higher for inner districts. This suggests that the base year shares for inner districts are originally higher and hence adding new jobs only in these districts may not result in new destination choice patterns.

Comparing the non-spatial and the spatial models of destination choice, indicated that there is a negligible difference between the forecasted shares when there is low percentage increase in casino/hotel and retail jobs, but the difference is noticeable when there is high percentage increase in casino/hotel and retail jobs. It should be noted that higher percentage increase in hotel/casino and retail jobs would be realized after a long time because, based on land use data, there was an increase of 5% in total employment between year 2000 and 2005. The models developed in this study did not incorporate housing location as a variable and hence they may not be accurate for long-term forecasting. Since the model with separation measure is comparable to the model without separation measure when there is less than 10% increase in hotel/casino and retail jobs, it can be concluded that the separation measure does not make a big difference in forecasting capabilities.

4.10. Example Origin-Destination (O-D) Matrix

The aggregate probabilities for a group of individuals from a specific origin choosing a specific destination were computed using Equation (3.14). Table 4.18 shows a portion of the aggregate probabilities for each group at base year while Appendix H shows the detailed computations. The number of observed trip productions (*Pi*) and trip attractions (*Aj*) at base year are also shown in Table 4.18. Multiplying the probabilities (shares) with trip productions provides distribution of trips among the available destinations.

Table 4.18. Aggregate O-D probabilities

		Attractions (at destination, *j*)							
		1	**2**	**3**	**4**	**…**	**16**	**17**	***Pi***
Productions (at origin, *i*)	**1**	0.0226	0.0420	0.0334	0.0135		0.0988	0.0163	17126
	2	0.0227	0.0420	0.0334	0.0135		0.0987	0.0163	16457
	3	0.0226	0.0409	0.0333	0.0134		0.0985	0.0163	41721
	4	0.0227	0.0413	0.0345	0.0109		0.0959	0.0150	9074
	5	0.0226	0.0409	0.0334	0.0109		0.0956	0.0150	9023
	6	0.0234	0.0405	0.0331	0.0127		0.0969	0.0158	2299
	7	0.0216	0.0397	0.0467	0.0013		0.0958	0.0152	7529
	8	0.0211	0.0377	0.0541	0.0020		0.0947	0.0146	11616
	9	0.0172	0.0308	0.1057	0.0069		0.0854	0.0104	40485
	10	0.0214	0.0383	0.0515	0.0017		0.0957	0.0149	10072
	11	0.0199	0.0352	0.0710	0.0036		0.0919	0.0133	21047
	12	0.0227	0.0409	0.0334	0.0067		0.0910	0.0129	23637
	13	0.0226	0.0411	0.0334	0.0078		0.0922	0.0134	19699
	14	0.0226	0.0409	0.0334	0.0082		0.0986	0.0163	48263
	15	0.0229	0.0384	0.0338	0.0136		0.0998	0.0165	9258
	16	0.0227	0.0321	0.0335	0.0135		0.0989	0.0164	21679
	17	0.0227	0.0410	0.0334	0.0135		0.0988	0.0163	16375
Aj		7052	12676	15101	2963		30975	4824	**325360**

Table 4.19 presents a portion of the estimated O-D matrix at base year. In order to evaluate the performance of the logit model, the sum of the trips in a row should be equal to the total number of trips (*Pi*) originating from that zone. Also, the sum of the trips in a

column should match the number of trips (A_j) attracted to that zone. It can be seen that the sum of the estimated number of trips (A_j) of Table 4.19 matches the sum of the observed number of trips (A_j) of Table 4.18 with an error of 1.33%. With the forecasted probabilities for each group of individuals, the trip patterns can therefore be forecasted.

Table 4.19. Estimated O-D trip distributions

		Attractions (at destination, *j*)							
		1	**2**	**3**	**4**	**…**	**16**	**17**	***Pi***
Productions (at origin, *i*)	**1**	387	719	573	231		1692	280	17126
	2	374	691	550	222		1625	269	16457
	3	943	1706	1391	561		4109	680	41721
	4	206	375	304	99		870	136	9074
	5	204	369	301	98		863	135	9023
	6	54	93	76	29		223	36	2299
	7	163	299	352	10		721	114	7529
	8	245	438	628	23		1100	170	11616
	9	696	1247	4279	279		3459	422	40485
	10	216	386	519	17		964	150	10072
	11	419	741	1495	75		1934	279	21047
	12	537	967	789	158		2150	304	23637
	13	445	810	657	154		1816	265	19699
	14	1091	1974	1611	396		4758	787	48263
	15	212	356	313	126		924	153	9258
	16	492	696	726	293		2143	354	21679
	17	372	671	548	221		1617	267	16375
Aj		7054	12537	15112	2989		30967	4801	**325360**

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

5.1. Spatial Trip Generation Models

Spatial models of trip attraction and trip production were estimated and the results were compared to those of non-spatial models. Only the home-based work (HBW) trips were considered in this analysis. Relevant explanatory variables for each model were tested for correlation prior to inclusion in the model in order to avoid multicolinearity effect. The dependent variable for the trip attraction model was the number of trips attracted to zone and the explanatory variables were employment opportunities in gaming industry (*Gaming*) and employment in other industries (*Non_Gaming*). It was hypothesized that while employment opportunities in other industries might have similar impact (positive) on trip generations, the impact of employment in gaming industry might be different.

The dependent variable for trip production model was the number of work trips produced by a zone, and the relevant predictors were the district population density, average household size, number of households with income less than $17.5K, number of households with income between $17.5K and $47.5K, and number of households with income greater than $47.5K. A variable quantifying the effect of spatial relationship of the observations was included in both the models of trip attraction and trip production.

The spatial variable was developed based on four alternative methods for quantifying spatial relationship: contiguity spatial relationship, separation between districts, a combination of contiguity and separation, and a combination of employment opportunities and separation (accessibility measure).

The models developed were compared to identify the best method for quantifying spatial relationship. Different model specifications for the non-spatial and spatial models were estimated using STATA® software. The Ordinary Least Square (OLS) method of estimation was used for the non-spatial model while the Maximum Likelihood (ML) estimation method was used for the spatial model. In order to demonstrate the effect of omission of spatial correlation in trip generation models, comparison of trip prediction of the non-spatial model and the best spatial model, was conducted.

5.1.1. Trip Attraction Model

For the trip attraction model, the observed zonal trips were transformed by applying natural logarithm in order to establish linear relationship among the variables. In order to estimate the spatial trip attraction model, variables were tested for spatial dependency. The Moran's I, the Geary's C and the Getis & Ord's G statistics were estimated for all variables as well as three forms of the response variable. All three test results indicated that there was significant spatial autocorrelation in the number of trips attracted to zones. This finding implied that it was important to develop trip attraction models accounting for spatial autocorrelation. Therefore, in addition to the non-spatial model, the models of trip attraction were developed using the four alternatives for quantifying spatial relationships. These models are shown in Equation (4.5) through Equation (4.9). The criteria for selecting the best method of quantifying spatial relationship were based on the goodness-of-fit of the models estimated using different spatial relationships and their accuracy in predicting the observed trips. The statistical tests (RSS, AIC, SIC and the final loglikelihood) for the four spatial models indicated that the model estimated using the spatial weight matrix defined with separation spatial relationship has the best fit.

The accuracy of this spatial model (Equation 4.7) in predicting the observed trips was compared with that of the non-spatial model (Equation 4.5) by computing the percentage deviation of modeled trips from observed trips for each district. The results indicated a smaller deviation for the spatial model, which suggests a much better prediction capability of the spatial model when compared to the non-spatial model. Specifically, while the non-spatial model had an average absolute deviation of 31%, the spatial model had an average absolute deviation of 17% only — an increase of 14% in prediction accuracy.

5.1.2. Trip Production Model

For the non-spatial trip production model, the dependent variable was transformed by finding square root. Similar to the spatial model for trip attraction, variables were tested for spatial dependency in order to estimate the spatial trip production model. The Moran's I, the Geary's C and the Getis & Ord's G statistics indicated that there was no significant spatial autocorrelation in the observations of the variables. This finding is consistent with intuition because for trip production, there is no direct linkage between someone making a trip from one zone to someone else making a trip from an adjacent zone. However, for trip attraction, it is possible that whatever attracts someone (from somewhere) to a specific zone is related to what attracts someone else (from somewhere else) to an adjacent zone. If this case is true, the observations made at these two adjacent zones would be spatially correlated. However, to confirm this finding, the spatial models were estimated and compared with the non-spatial model.

Similar to the trip attraction model, the criteria for selecting the best model were based on the goodness-of-fit of the models. These models are shown in Equation (4.13) through Equation (4.17). The results indicated that while the spatial models minimize RSS, they have higher AIC and SIC. This implies that the spatial variable in the spatial models of trip production is unnecessary. The result confirmed the initial finding of no spatial autocorrelation in the variables. Furthermore, the accuracy of the models developed was evaluated by comparing the average absolute percentage deviation of the modeled trips from the observed trips. The results indicated that the models estimated by incorporating the spatial variable do not improve prediction accuracy. Therefore, the non-spatial model was selected as the best model.

5.2. Destination Choice Models

Trip distribution is the second step of the widely used four-step procedure in which trip productions and trip attractions for each zonal pair are linked. Gravity models are commonly used in trip distribution and are functions of activity system attributes (indirectly through the generated productions and attractions) and network attributes (typically, inter-zonal travel times). Trip distribution is essentially destination choice and therefore discrete choice models can be used. The choice set for a destination choice model is comprised of Traffic Analysis Zones (TAZ). Since the TAZs are geographical units, their spatial relationships may have an impact on individuals' destination choices.

Home-based work trips are normally viewed as trips involving long–term decisions such as job location and residential location. Since most job locations are fixed, workers have no alternative choices and therefore do not really make a destination choice for such trips. However, this assumption may not be applicable to all urban areas, such as Las Vegas, in which numerous comparable job opportunities exist and are distributed over the area. Thus, spatial locations of the jobs and residence of the workers may be important factors in estimating long-term trip generation as well as destination choice models.

The efficiency of alternative methods for quantifying spatial relationship of the origins and destinations in a multinomial logit model for destination choice was evaluated. Two alternative methods were evaluated: separation (distance only) and economic linkage (accessibility). The spatial model developed using separation spatial relationship was found to be the best fitted model with maximum adjusted rho-squared. The model estimated using separation spatial relationship was further compared with the non-spatial model by conducting the likelihood ratio test with the null hypothesis that the coefficient of the spatial variable is not different from zero. The test statistic obtained (10.67) was greater than the critical value (3.84), suggesting that the coefficient associated with the spatial variable was significantly different from zero, and therefore that the spatial model is the best specification for destination choice model.

In order to evaluate the impact of omitting spatial variable from the destination choice model, the difference between forecasted destination shares at different percentage increase in hotel/casino and retail jobs were computed. The results indicated that for when there is low percentage increase in casino/hotel and retail jobs, the difference between the forecasted destination shares is small while when there is high percentage increase in casino/hotel and retail jobs, the difference between the forecasted probabilities increases. The implication of this finding is that incorporation of spatial factors does not affect the destination choice pattern when there is a small change in employment opportunities in hotel/casino and retail. However, it is worthy noting that higher percentage increase in hotel/casino and retail jobs would be realized after a long time because, based on Las Vegas land use data, there was an increase of 5% in total employment between year 2000 and 2005. In order to evaluate the impact of spatial factors in long-term destination choice patterns, additional variables should be considered. For example, the models developed in this study did not incorporate housing location as a variable and hence they may not be accurate for long-term forecasting. Since the alternative shares of the model with spatial factor are comparable to the model without spatial factor when there is less than 10% increase in hotel/casino and retail jobs, it can be concluded that the factor does not make a big difference in forecasting capabilities when there is a small change in employment opportunities.

To demonstrate the applicability of the spatial multinomial logit model for destination choice, the origin-destination (O-D) matrix for base year was created. The sum of the estimated number of trips attracted to a specific zone matched the sum of the observed number of trips for that zone with a maximum absolute error of less than 1.33%.

5.3. Recommendations for Future Work

The Modifiable Areal Unit Problem (MAUP) in spatial analysis, which consists of scale effect and zonal aggregation effect, needs to be addressed in future similar studies. Variation of spatial effect can occur when data from one scale of areal units is aggregated into larger or smaller areal units. Also, variability of analytical or statistical results derived from data for the same region, but aggregated or partitioned in different ways, with the number of areal units in different partitioning schemes being the same, can occur.

The models developed in this study were based on districts developed by the Regional Transportation Commission of Southern Nevada (RTC) by aggregating homogeneous Traffic Analysis Zones (TAZs). These districts may not be the optimal zones that minimize the MAUP. It is recommended that an attempt to design new analysis zones through a careful aggregation process be conducted. The basic unit to begin with could be the existing TAZs and aggregate them to a practical number that can be handled. Homogeneity of the zone attributes should be carefully controlled. Automated zone design procedure which implements statistical design rules can be used. Also, an analysis in which different number of zones or different configurations of the zones are utilized can be used to assess the best zonal structure by comparing the goodness-of-fit of the models and their forecasting capability.

There were no households interviewed from two districts. Two possible reasons for this were that the sampling design did not incorporate spatial distribution of the zones (districts), or individuals selected from these districts did not complete their diaries as required. The sample drawn for the survey was a stratified random sample of households based on vehicles available and household size—not location. It is important to have zones (districts) represented proportionally in the sample. The zones can be weighted by population.

While the results of the destination choice model suggested that spatial relationship of the origins and destinations have impact on work trips for the Las Vegas Valley, the conclusion may not be applicable to cities where there is no numerous comparative employment opportunities distributed over the area. However, for cities with numerous job opportunities distributed over the area similar to Las Vegas, the author recommends incorporation of spatial interaction of the origin and destination.

5.4. Research Contribution

This study contributed to the existing knowledge of the effects of spatial autocorrelation in travel demand models. First, it was found that the variables for trip attraction models are significantly correlated in space while the variables for trip production models are not significantly correlated. This finding called for the design and incorporation of specific variables to account for spatial autocorrelation in trip attraction models. The results from the developed models indicated an improvement in predicting the observed work trips when compared to models without specific variables to account for spatial autocorrelation. This improvement can lead to more accurate design traffic estimations and hence air quality emissions. The accurate estimates of emissions could lead to identifying the proper countermeasures and strategies for minimizing the emissions.

Second, this research compared the alternatives for defining spatial relationship of the variables. It was found that different methods for defining spatial relationship give different model results. Most importantly, the study found that the widely used contiguity method does not capture the spatial relationship well. This is an important contribution to the knowledge since most of spatial analysis studies use contiguity as the sole method for defining spatial relationship.

Finally, the developed models established the basis for conducting similar studies for other trip purposes such as shopping trips. Also, the methodology established herein can be used to develop models for other areas.

APPENDICIES

A. Alternative spatial weight matrices

A. 1. Contiguity spatial relationship

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
2	1	0	1	1	0	0	1	0	0	0	0	0	0	0	0	1	0
3	0	1	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0
4	0	1	1	0	1	1	0	0	1	0	1	0	0	0	0	0	0
5	0	0	0	1	0	1	0	0	1	1	1	0	0	0	0	0	0
6	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0
7	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0
8	0	0	1	0	0	0	1	0	1	0	0	0	0	0	1	1	0
9	0	0	1	1	1	0	0	1	0	1	0	0	0	0	1	0	0
10	0	0	0	0	1	0	0	0	1	0	1	1	0	1	1	0	0
11	0	0	0	1	1	1	0	0	0	1	0	1	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	0
13	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1
14	0	0	0	0	0	0	0	0	0	1	0	1	1	0	1	1	1
15	0	0	0	0	0	0	0	1	1	1	0	1	0	1	0	1	0
16	1	1	0	0	0	0	1	1	0	0	0	0	0	1	1	0	1
17	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0

A.2. Separation spatial relationship

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	0.00	0.10	0.07	0.03	0.05	0.04	0.15	0.10	0.06	0.06	0.04	0.05	0.06	0.07	0.08	0.12	0.10
2	0.10	0.00	0.16	0.04	0.06	0.06	0.18	0.09	0.08	0.05	0.04	0.04	0.05	0.07	0.07	0.10	0.06
3	0.07	0.16	0.00	0.06	0.09	0.09	0.17	0.19	0.16	0.09	0.05	0.05	0.07	0.09	0.09	0.13	0.05
4	0.03	0.04	0.06	0.00	0.10	0.17	0.05	0.05	0.06	0.05	0.07	0.04	0.03	0.04	0.04	0.04	0.03
5	0.05	0.06	0.09	0.10	0.00	0.20	0.08	0.08	0.15	0.14	0.12	0.06	0.04	0.05	0.09	0.07	0.04
6	0.04	0.06	0.09	0.17	0.20	0.00	0.06	0.06	0.09	0.07	0.10	0.05	0.04	0.04	0.05	0.05	0.03
7	0.15	0.18	0.17	0.05	0.08	0.06	0.00	0.29	0.12	0.08	0.05	0.06	0.08	0.12	0.13	0.29	0.06
8	0.10	0.09	0.19	0.05	0.08	0.06	0.29	0.00	0.13	0.09	0.06	0.07	0.09	0.15	0.16	0.32	0.07
9	0.06	0.08	0.16	0.06	0.15	0.09	0.12	0.13	0.00	0.18	0.10	0.08	0.05	0.11	0.22	0.09	0.05
10	0.06	0.05	0.09	0.05	0.14	0.07	0.08	0.09	0.18	0.00	0.13	0.13	0.05	0.09	0.19	0.08	0.04
11	0.04	0.04	0.05	0.07	0.12	0.10	0.05	0.06	0.10	0.13	0.00	0.08	0.04	0.06	0.08	0.06	0.03
12	0.05	0.04	0.05	0.04	0.06	0.05	0.06	0.07	0.08	0.13	0.08	0.00	0.06	0.12	0.10	0.07	0.06
13	0.06	0.05	0.07	0.03	0.04	0.04	0.08	0.09	0.05	0.05	0.04	0.06	0.00	0.13	0.07	0.12	0.09
14	0.07	0.07	0.09	0.04	0.05	0.04	0.12	0.15	0.11	0.09	0.06	0.12	0.13	0.00	0.19	0.18	0.09
15	0.08	0.07	0.09	0.04	0.09	0.05	0.13	0.16	0.22	0.19	0.08	0.10	0.07	0.19	0.00	0.15	0.06
16	0.12	0.10	0.13	0.04	0.07	0.05	0.29	0.32	0.09	0.08	0.06	0.07	0.12	0.18	0.15	0.00	0.07
17	0.10	0.06	0.05	0.03	0.04	0.03	0.06	0.07	0.05	0.04	0.03	0.06	0.09	0.09	0.06	0.07	0.00

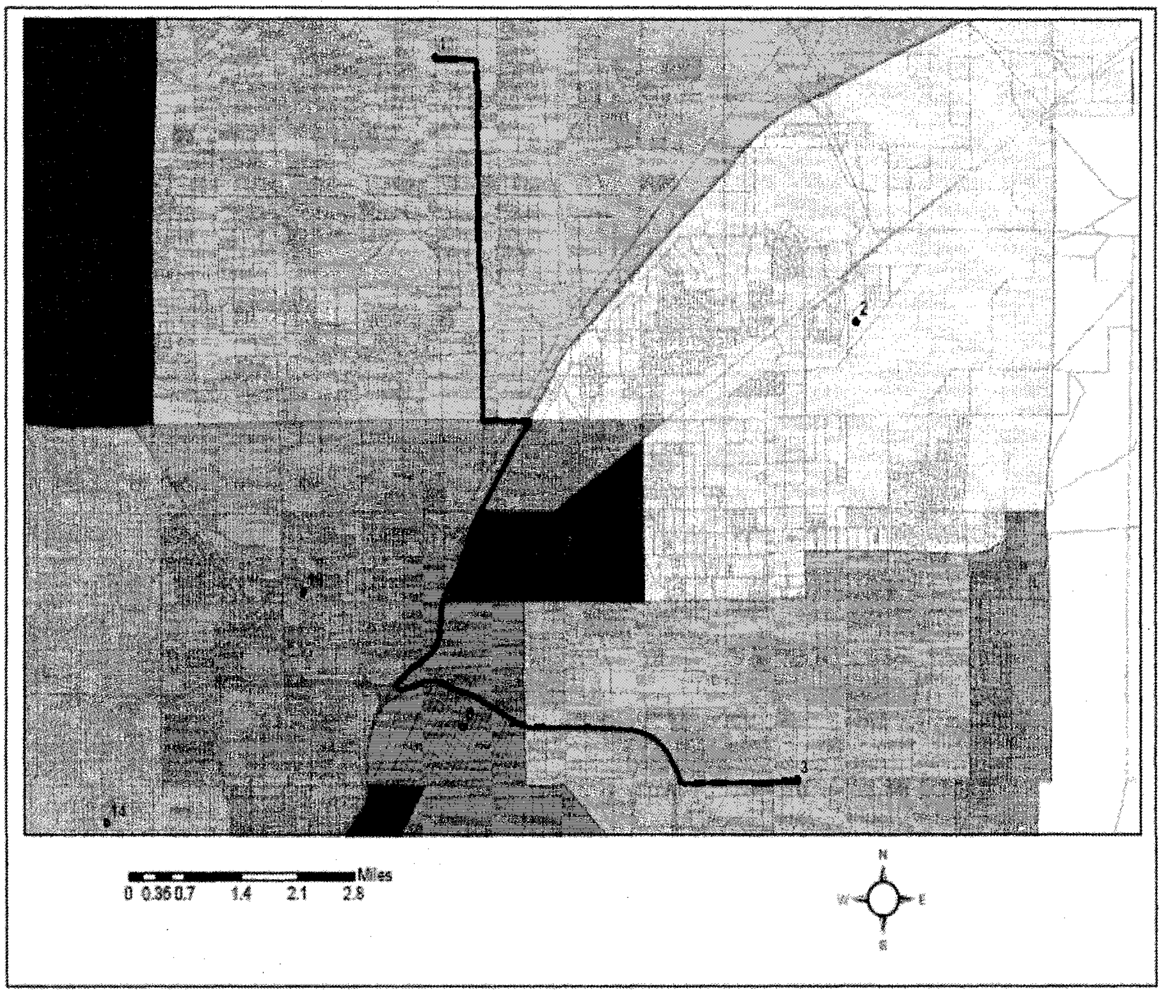
A.3. Contiguity-Separation spatial relationship

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	0	0.31	0	0	0	0	0	0	0	0	0	0	0	0	0	0.31	0.26
2	0.31	0	0.37	0.16	0	0	0.42	0	0	0	0	0	0	0	0	0.25	0
3	0	0.37	0	0.23	0	0	0.49	0.45	0.46	0	0	0	0	0	0	0	0
4	0	0.16	0.23	0	0.31	0.46	0	0	0.21	0	0.16	0	0	0	0	0	0
5	0	0	0	0.31	0	0.44	0	0	0.41	0.31	0.32	0	0	0	0	0	0
6	0	0	0	0.46	0.44	0	0	0	0	0	0.23	0	0	0	0	0	0
7	0	0.42	0.49	0	0	0	0	0.82	0	0	0	0	0	0	0	0.58	0
8	0	0	0.45	0	0	0	0.82	0	0.4	0	0	0	0	0	0.54	0.76	0
9	0	0	0.46	0.21	0.41	0	0	0.4	0	0.48	0	0	0	0	0.49	0	0
10	0	0	0	0	0.31	0	0	0	0.48	0	0.32	0.35	0	0.28	0.5	0	0
11	0	0	0	0.16	0.32	0.23	0	0	0	0.32	0	0.2	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0.35	0.2	0	0.19	0.3	0.31	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0.19	0	0.33	0	0	0.3
14	0	0	0	0	0	0	0	0	0	0.28	0	0.3	0.33	0	0.53	0.53	0.19
15	0	0	0	0	0	0	0	0.54	0.49	0.5	0	0.31	0	0.53	0	0.4	0
16	0.31	0.25	0	0	0	0	0.58	0.76	0	0	0	0	0	0.53	0.4	0	0.21
17	0.26	0	0	0	0	0	0	0	0	0	0	0	0.3	0.19	0	0.21	0

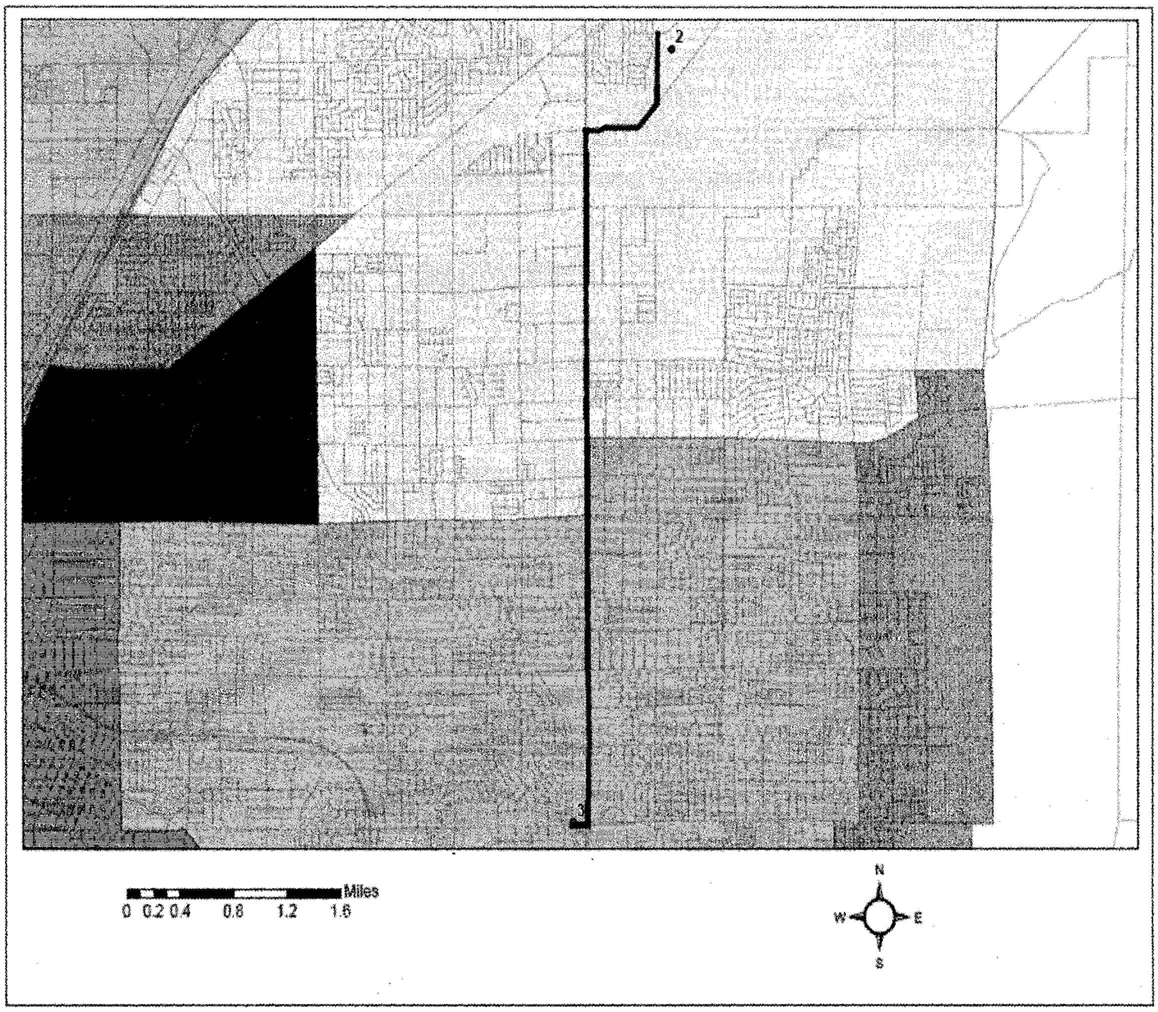
A.4. Economic Linkage (accessibility) spatial relationship

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	0	0.74	0.48	0.26	0.28	0.23	0.77	0.6	0.38	0.31	0.23	0.28	0.44	0.47	0.42	0.72	0.62
2	0.74	0	1.04	0.47	0.45	0.38	1.2	0.81	0.59	0.42	0.32	0.33	0.38	0.5	0.54	0.72	0.41
3	0.48	1.04	0	0.96	1.16	0.81	2.1	1.91	1.96	1.01	0.68	0.66	0.6	0.94	1.28	1.24	0.54
4	0.26	0.47	0.96	0	0.41	0.6	0.2	0.2	0.27	0.21	0.21	0.14	0.12	0.15	0.19	0.17	0.11
5	0.28	0.45	1.16	0.41	0	2.21	0.93	1.04	2.07	1.53	1.58	0.82	0.55	0.81	1.16	0.82	0.46
6	0.23	0.38	0.81	0.6	2.21	0	0.11	0.12	0.18	0.15	0.18	0.1	0.07	0.1	0.12	0.1	0.06
7	0.77	1.2	2.1	0.2	0.93	0.11	0	0.86	0.32	0.22	0.14	0.17	0.19	0.31	0.34	0.6	0.18
8	0.6	0.81	1.91	0.2	1.04	0.12	0.86	0	2.4	1.64	0.92	1.19	1.21	2.48	3.19	4.55	1
9	0.38	0.59	1.96	0.27	2.07	0.18	0.32	2.4	0	3.93	1.94	1.74	1.2	2.12	4.02	2.2	0.97
10	0.31	0.42	1.01	0.21	1.53	0.15	0.22	1.64	3.93	0	1.8	1.93	0.86	1.54	2.81	1.25	0.63
11	0.23	0.32	0.68	0.21	1.58	0.18	0.14	0.92	1.94	1.8	0	0.55	0.29	0.41	0.54	0.37	0.23
12	0.28	0.33	0.66	0.14	0.82	0.1	0.17	1.19	1.74	1.93	0.55	0	0.98	1.53	1.59	1.01	0.62
13	0.44	0.38	0.6	0.12	0.55	0.07	0.19	1.21	1.2	0.86	0.29	0.98	0	0.9	0.55	0.72	0.8
14	0.47	0.5	0.94	0.15	0.81	0.1	0.31	2.48	2.12	1.54	0.41	1.53	0.9	0	4.94	4.98	1.76
15	0.42	0.54	1.28	0.19	1.16	0.12	0.34	3.19	4.02	2.81	0.54	1.59	0.55	4.94	0	5.44	1.95
16	0.72	0.72	1.24	0.17	0.82	0.1	0.6	4.55	2.2	1.25	0.37	1.01	0.72	4.98	5.44	0	1.34
17	0.62	0.41	0.54	0.11	0.46	0.06	0.18	1	0.97	0.63	0.23	0.62	0.8	1.76	1.95	1.34	0

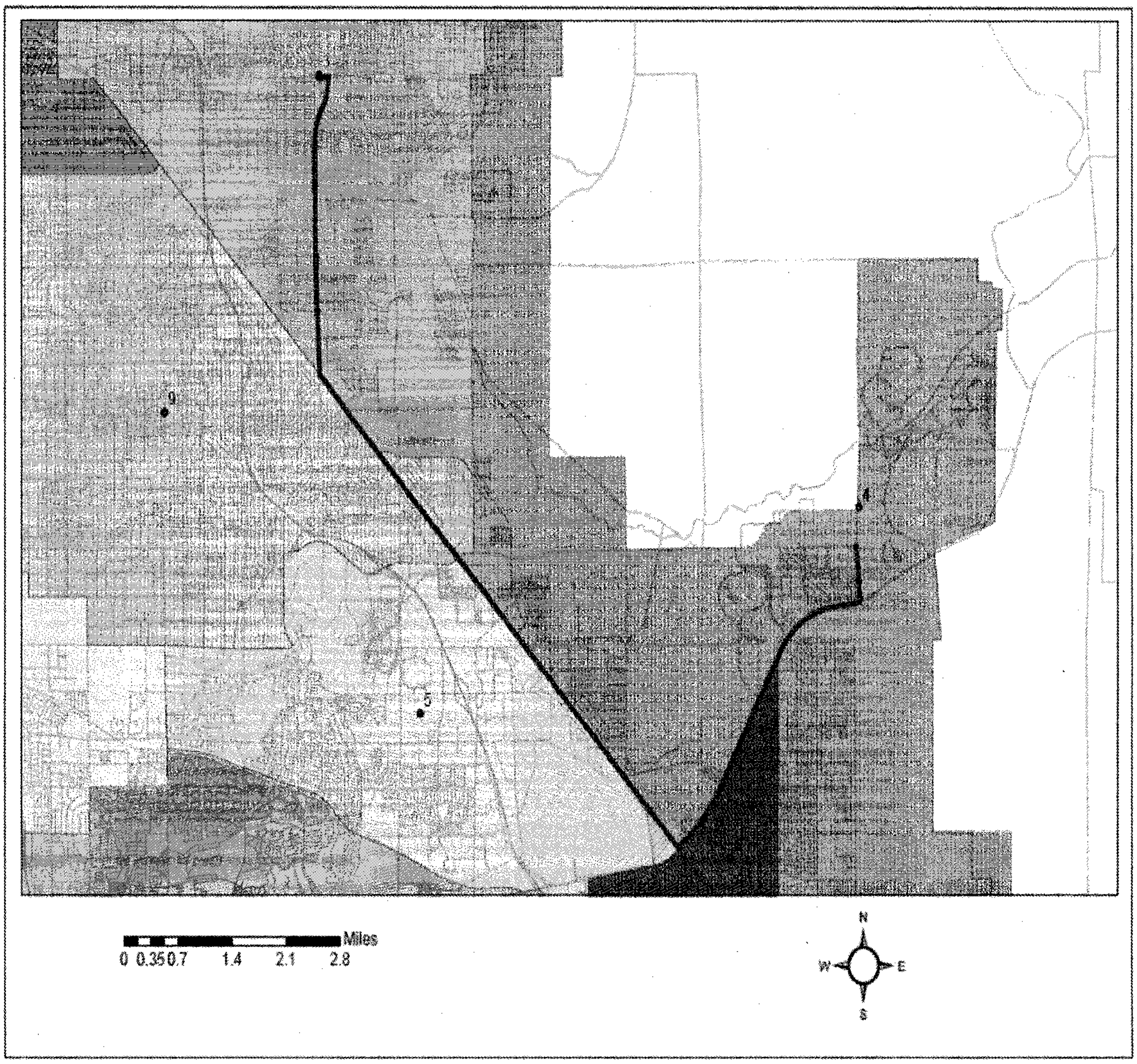
B. Network route distance from district 3 to other districts



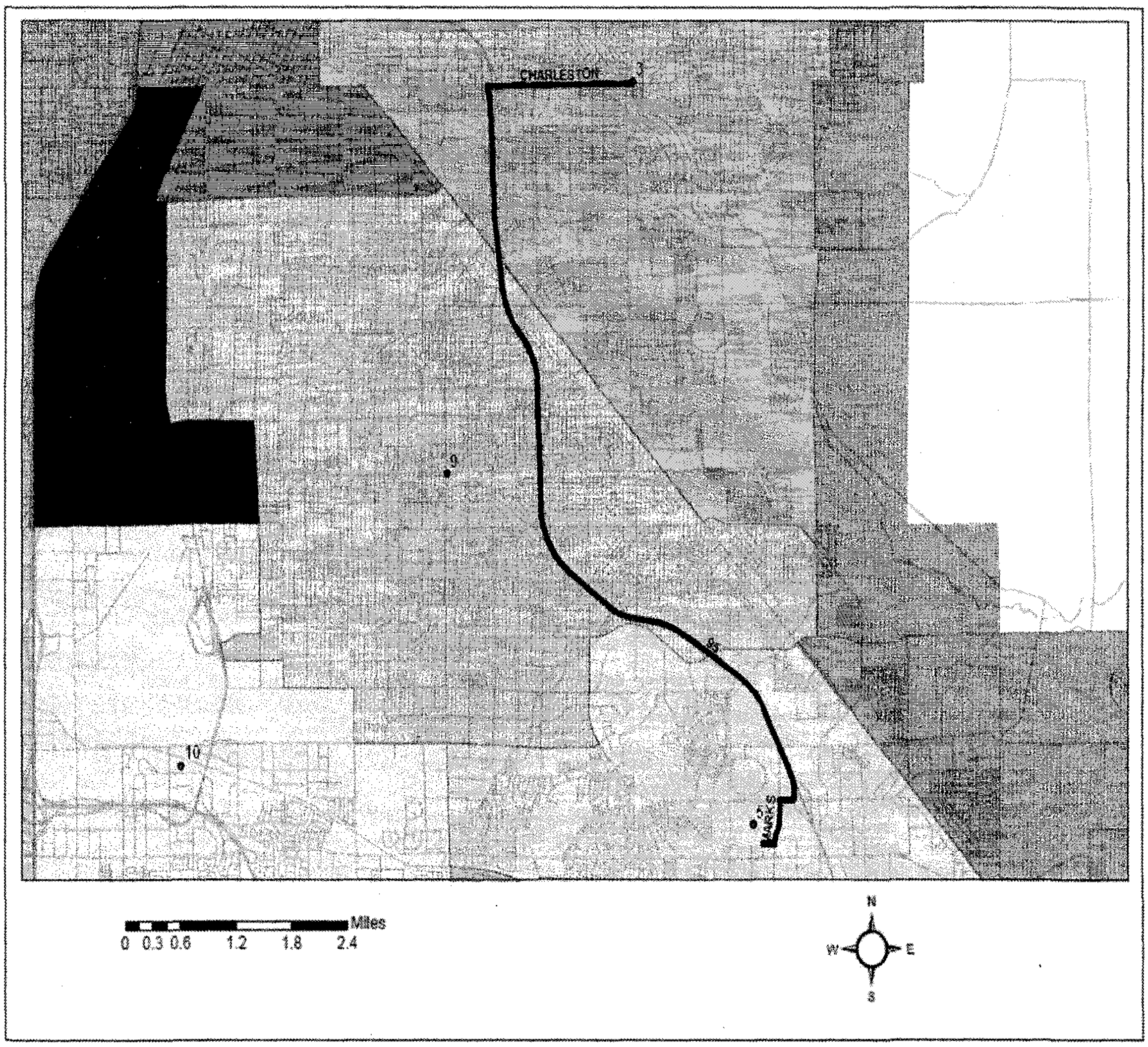
Appendix B.1. Network distance from District 3 to District 1



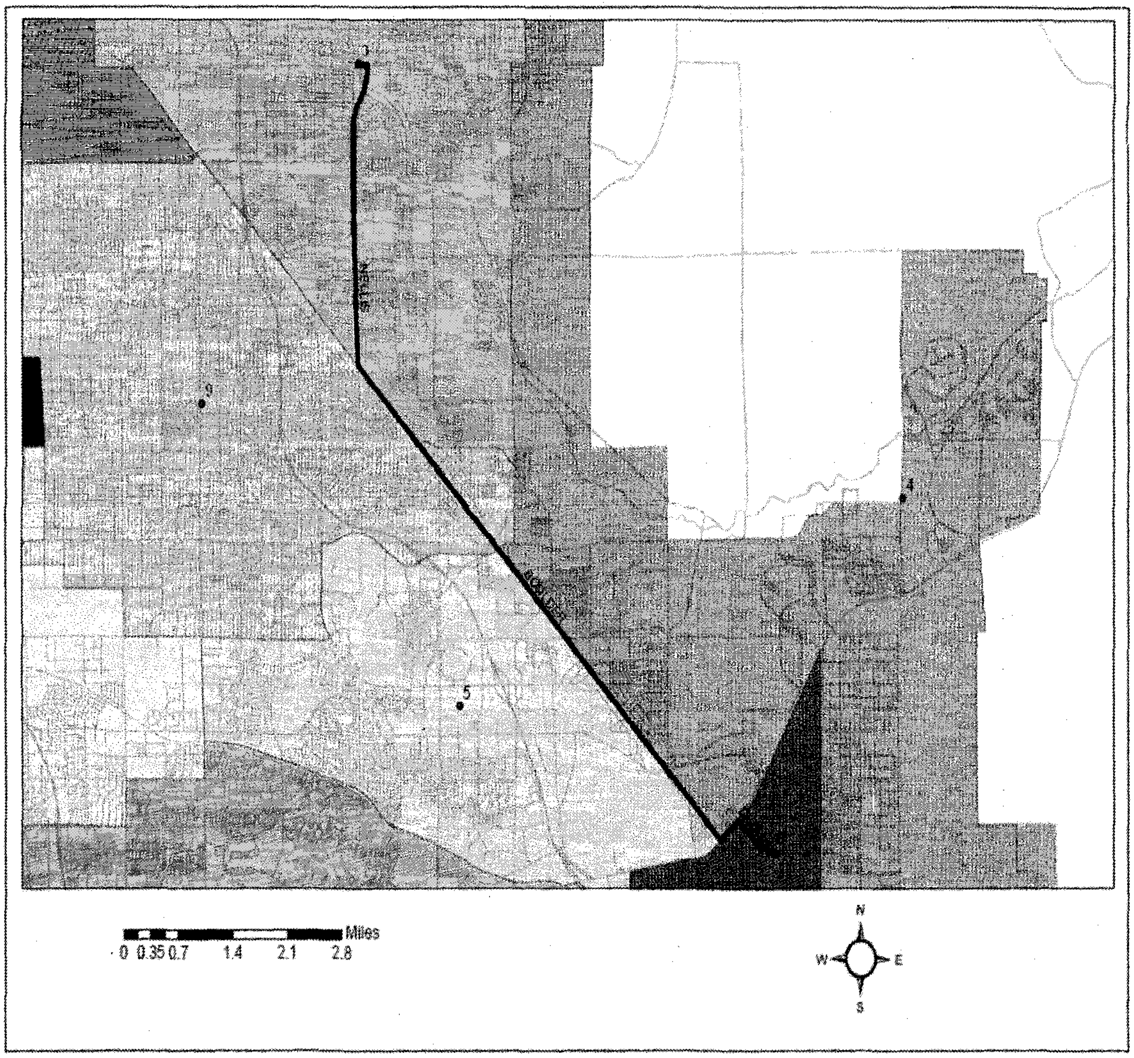
Appendix B.2. Network distance from District 3 to District 2



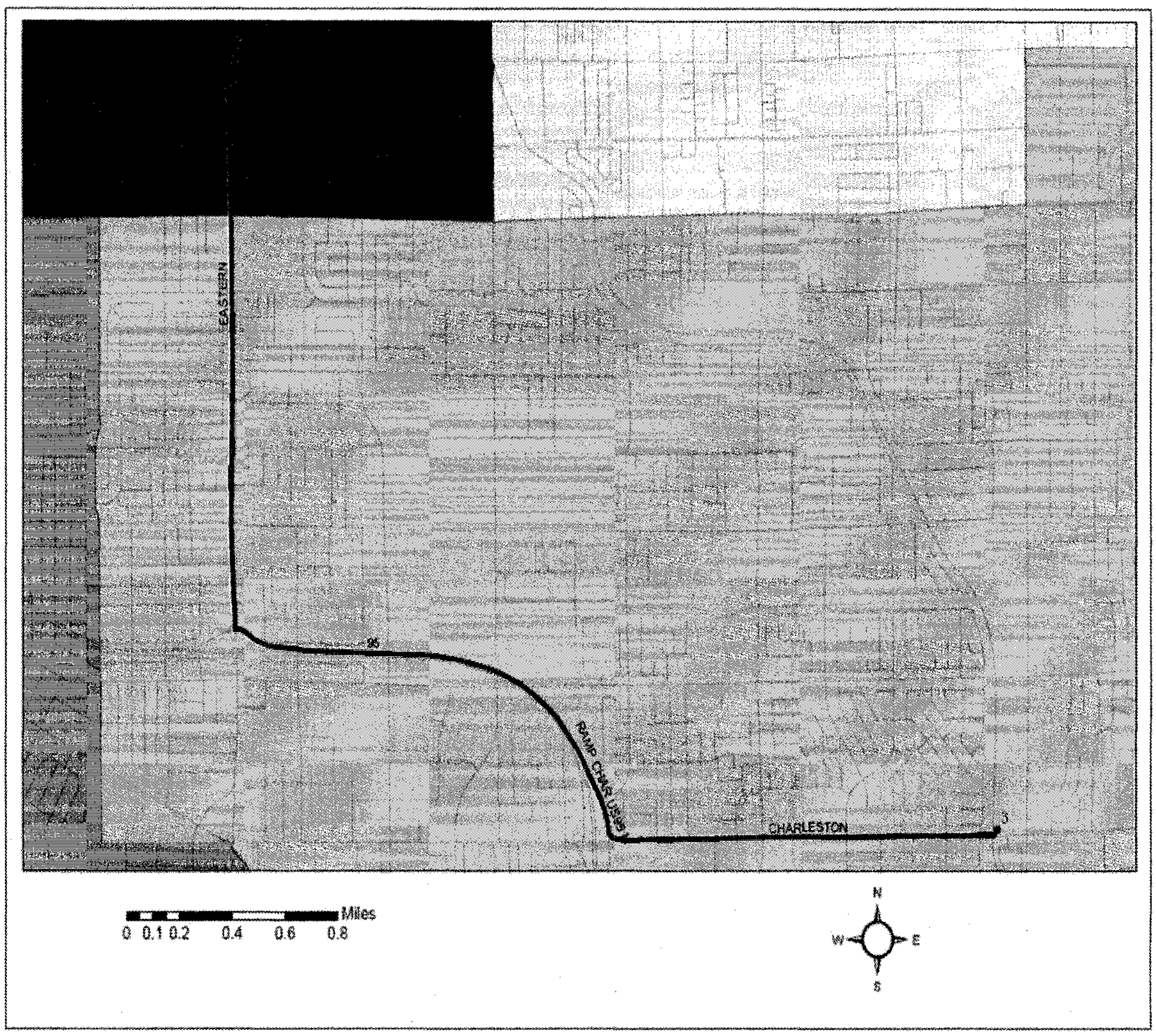
Appendix B.3. Network distance from District 3 to District 4



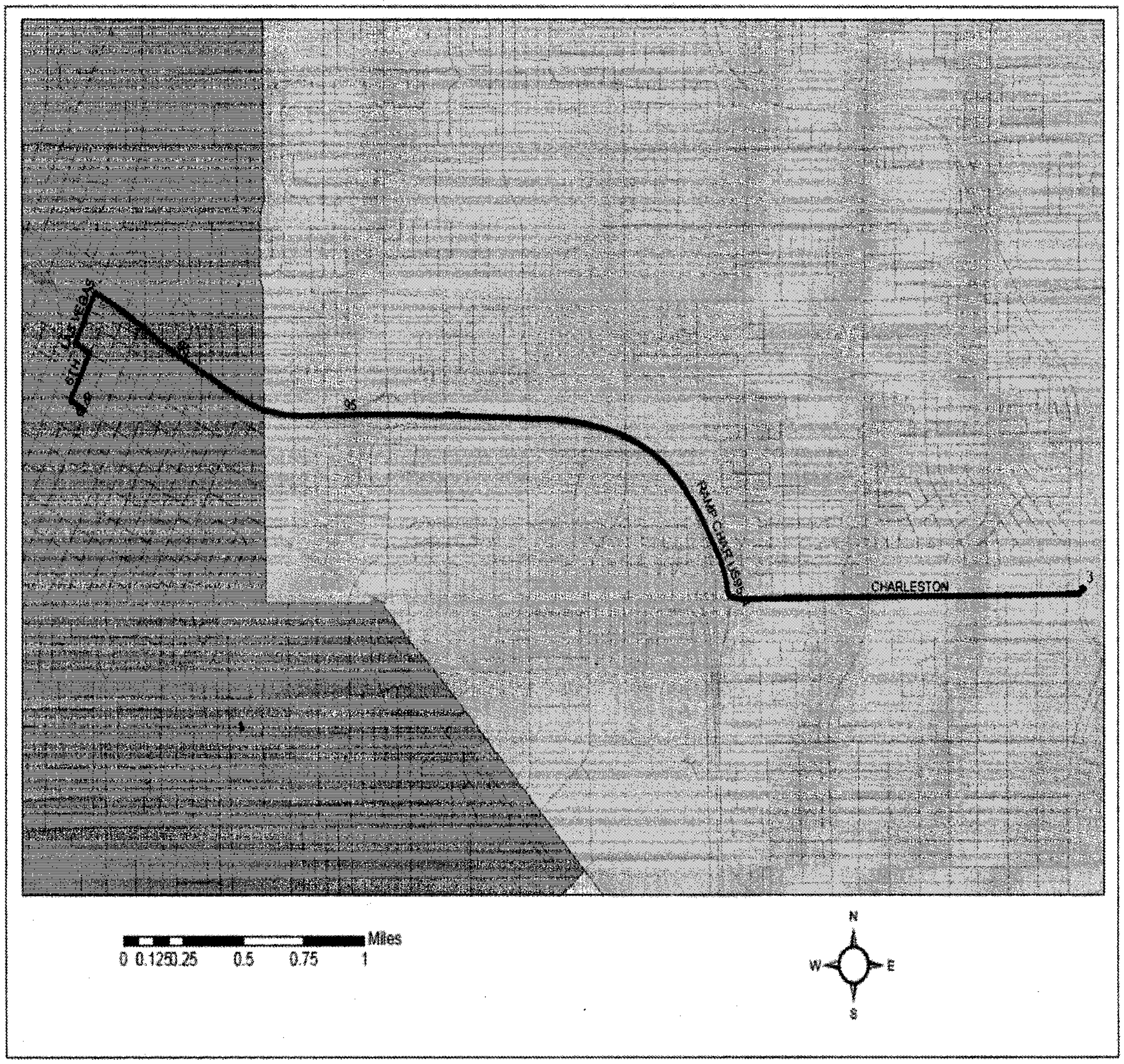
Appendix B.4. Network distance from District 3 to District 5



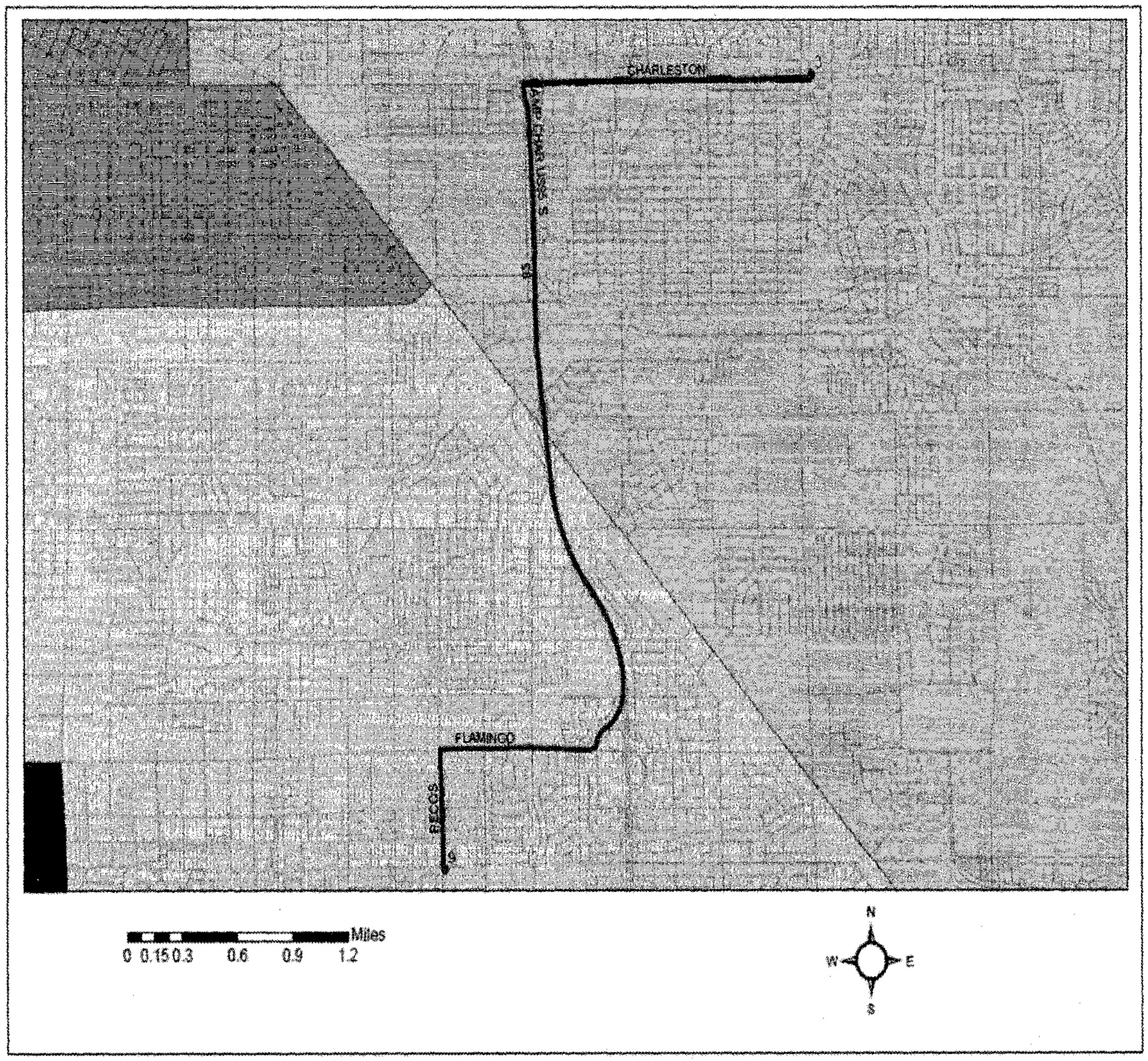
Appendix B.5. Network distance from District 3 to District 6



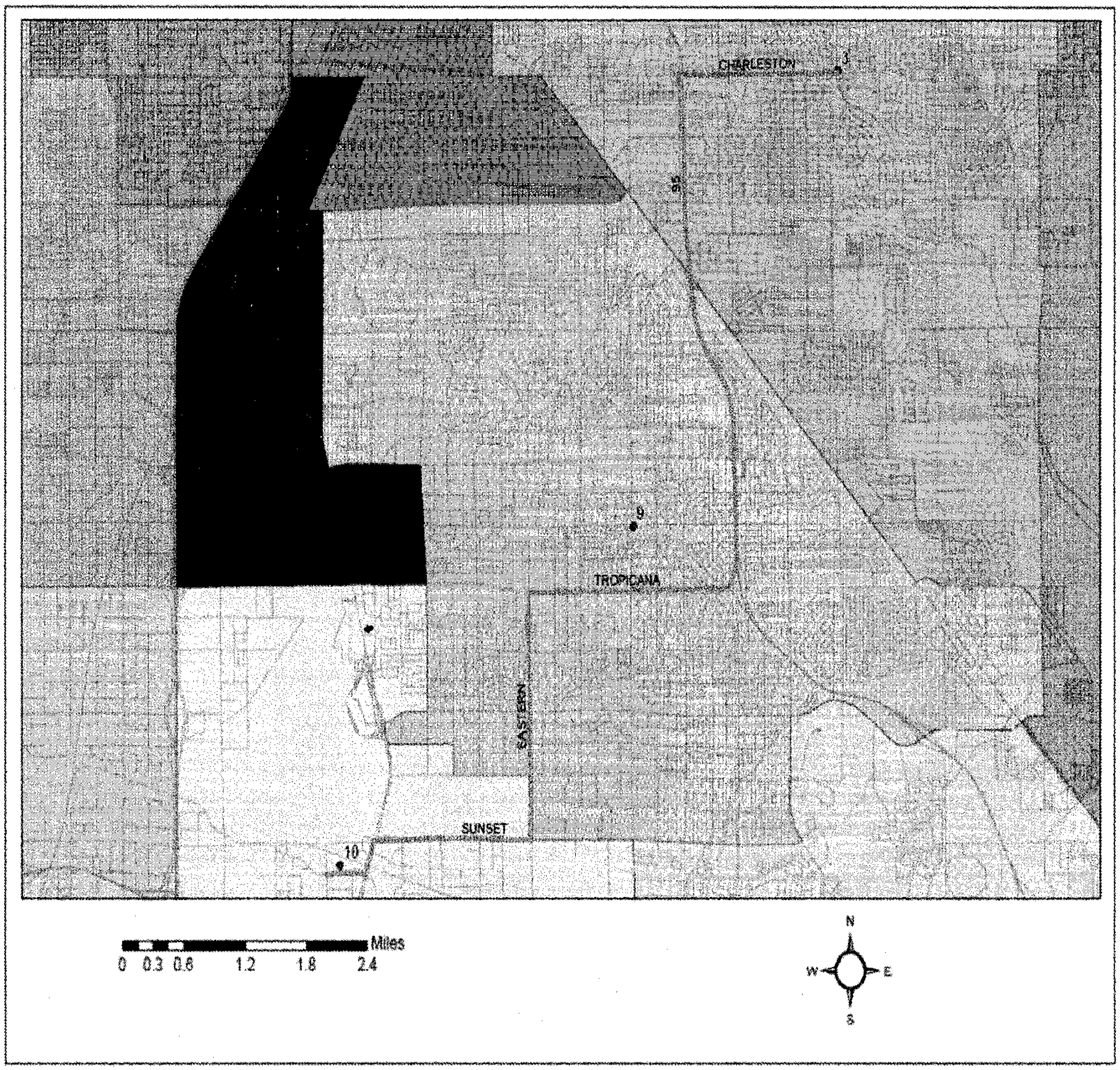
Appendix B.6. Network distance from District 3 to District 7



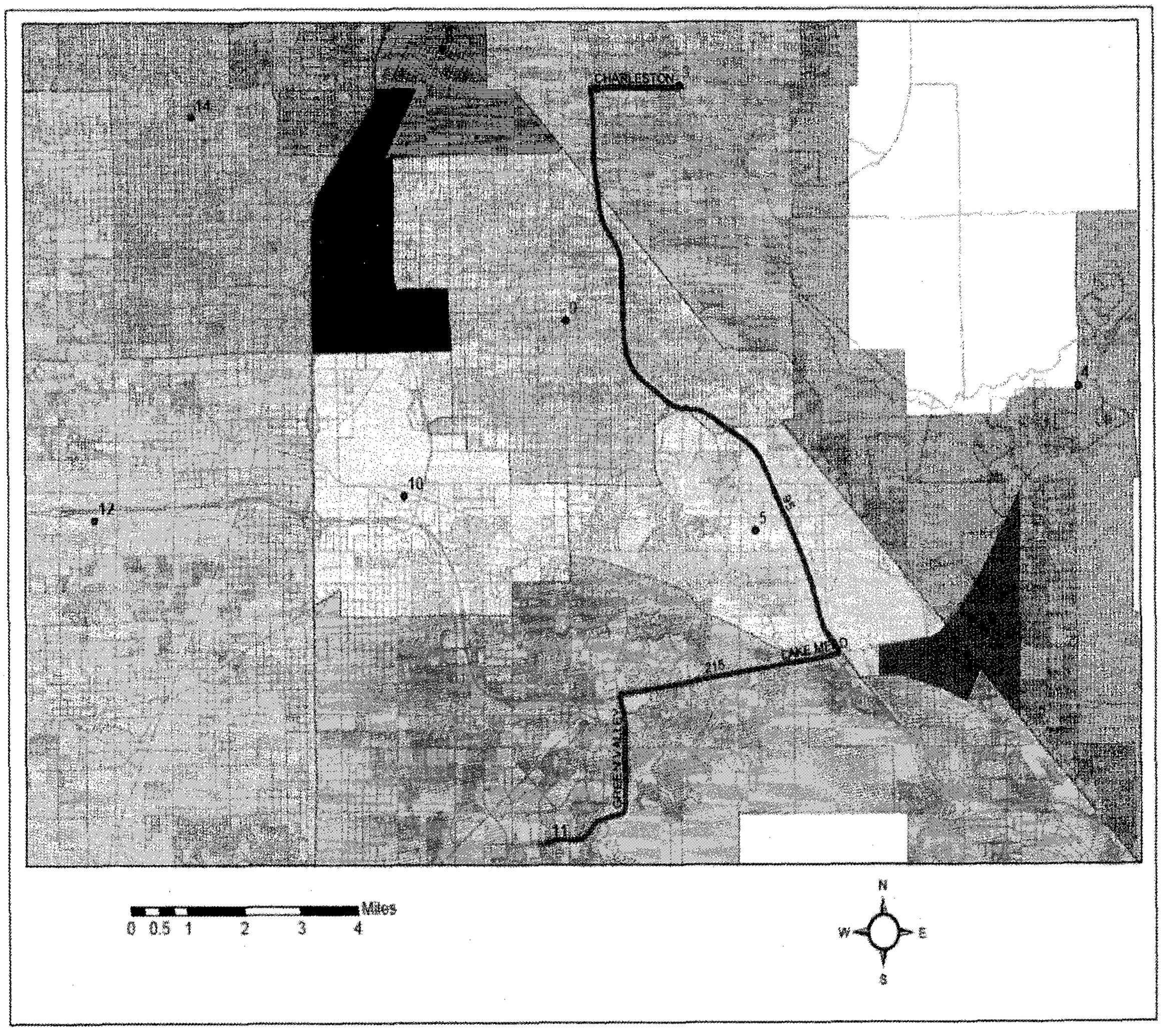
Appendix B.7. Network distance from District 3 to District 8



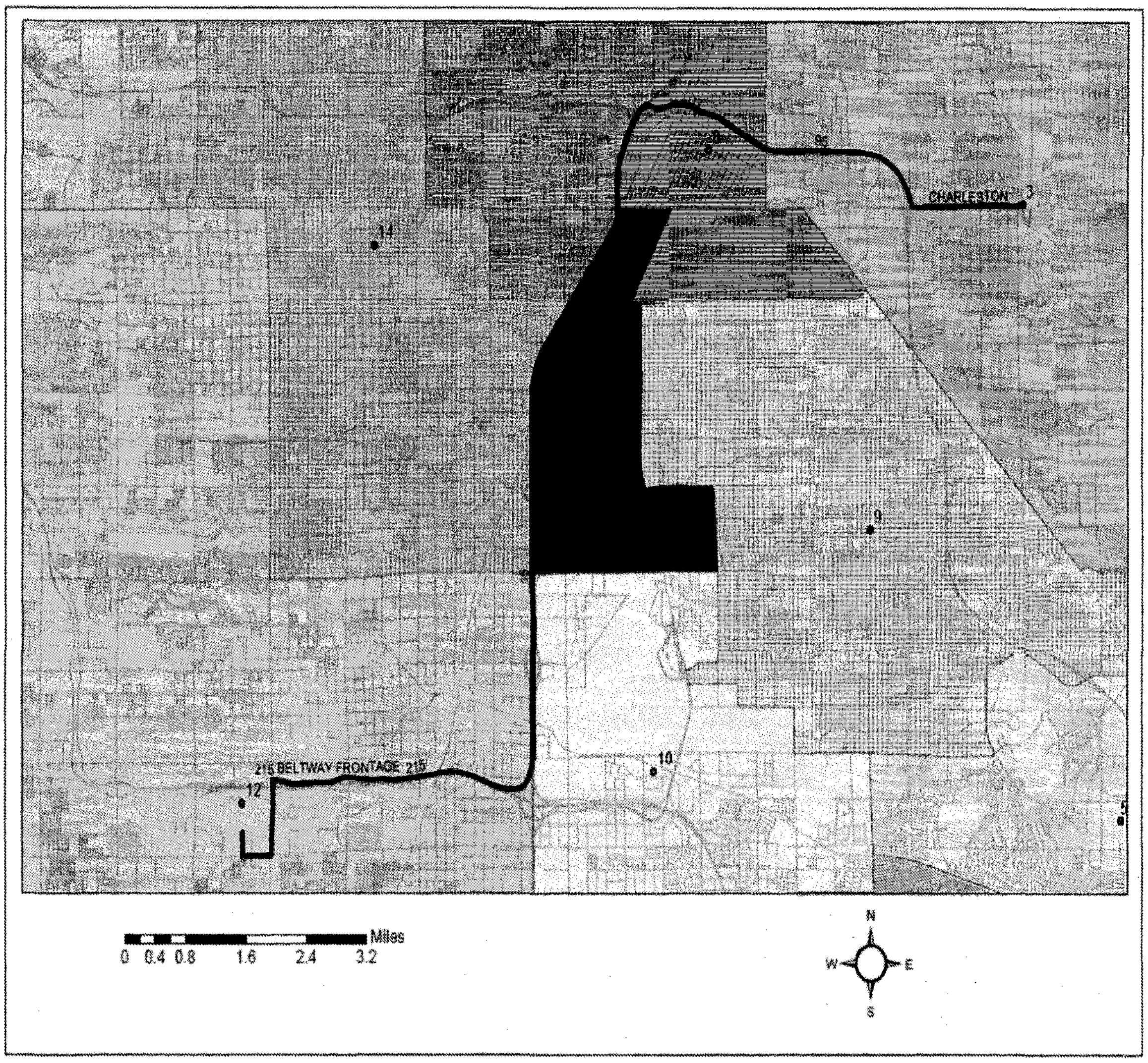
Appendix B.8. Network distance from District 3 to District 9



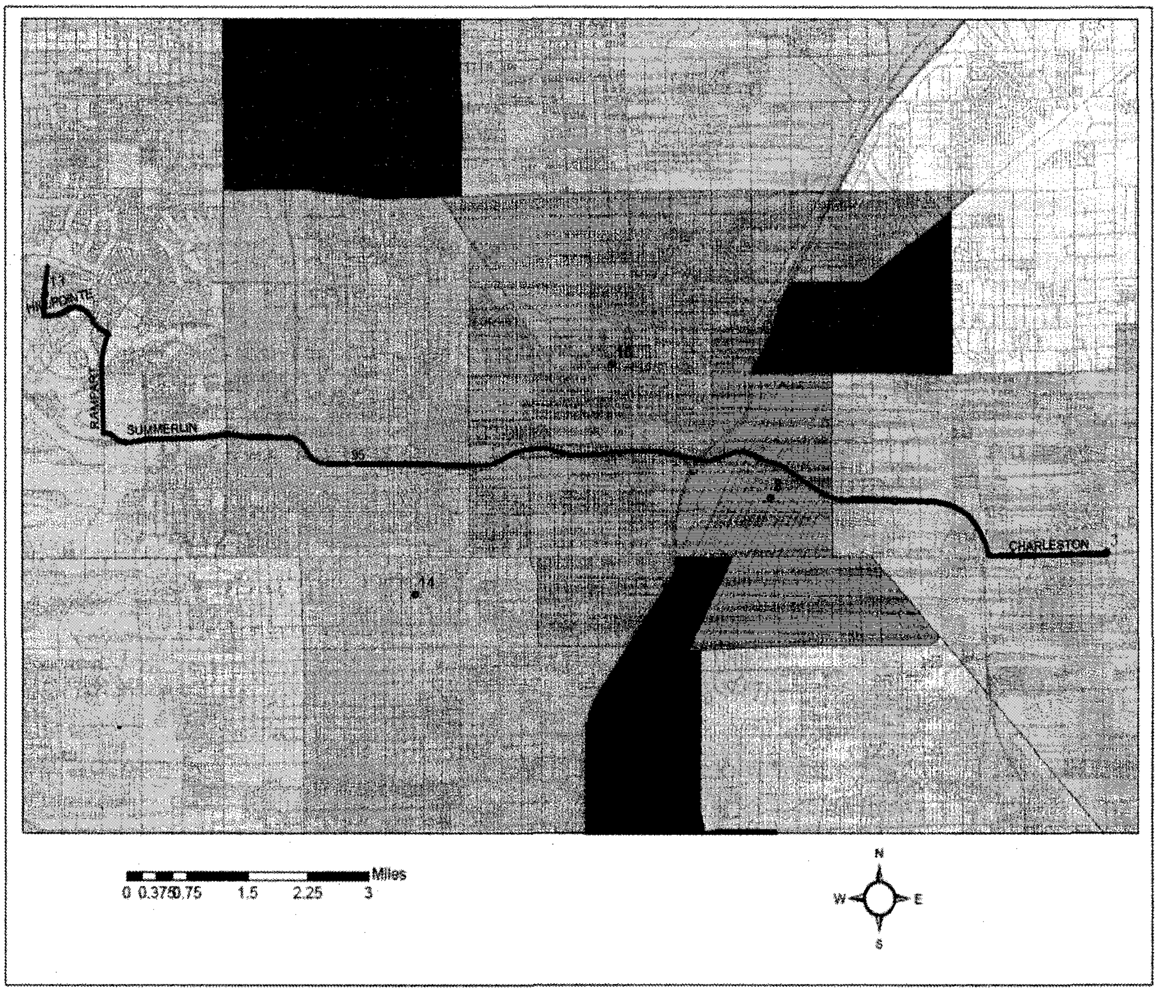
Appendix B.9. Network distance from District 3 to District 10



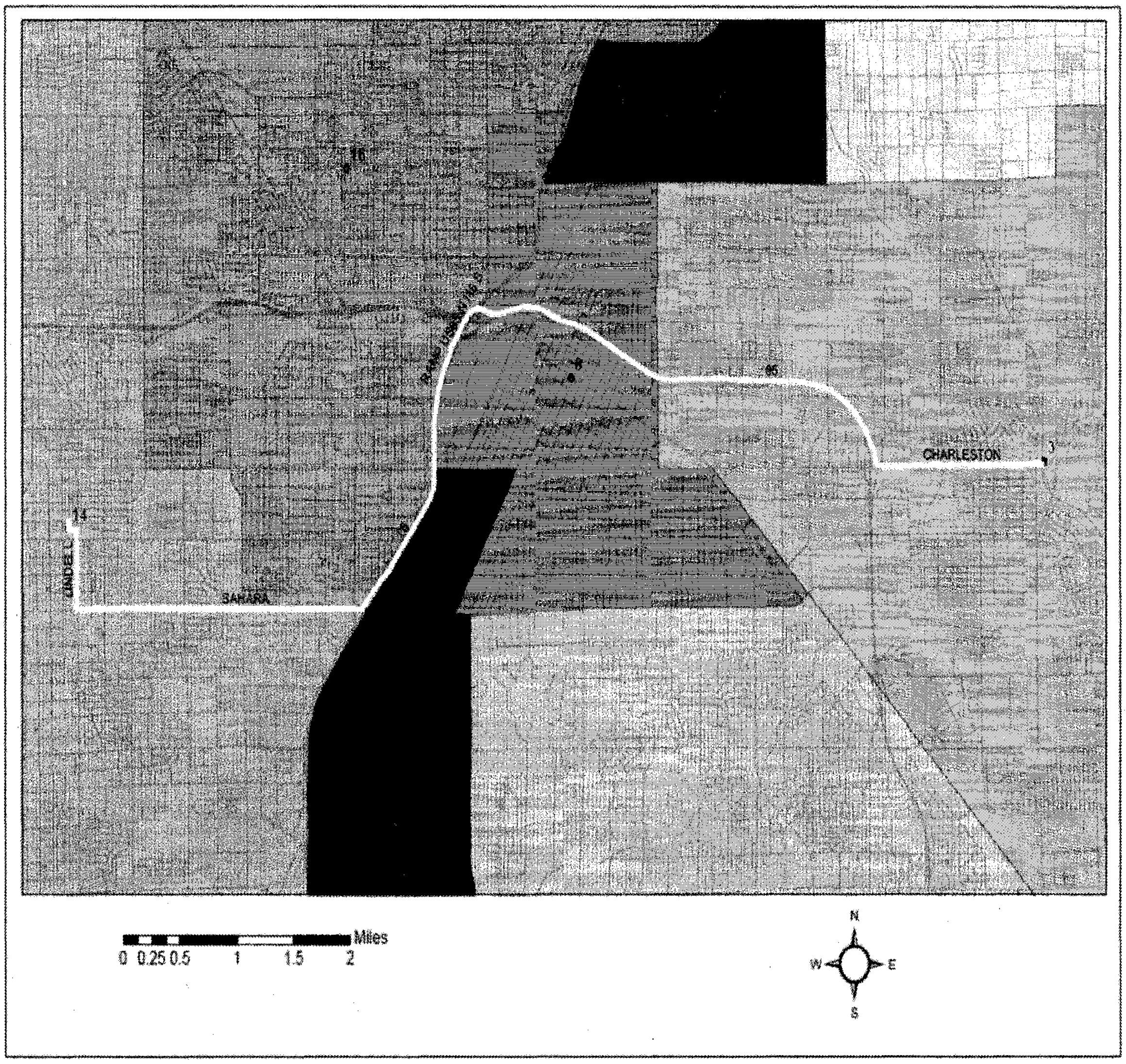
Appendix B.10. Network distance from District 3 to District 11



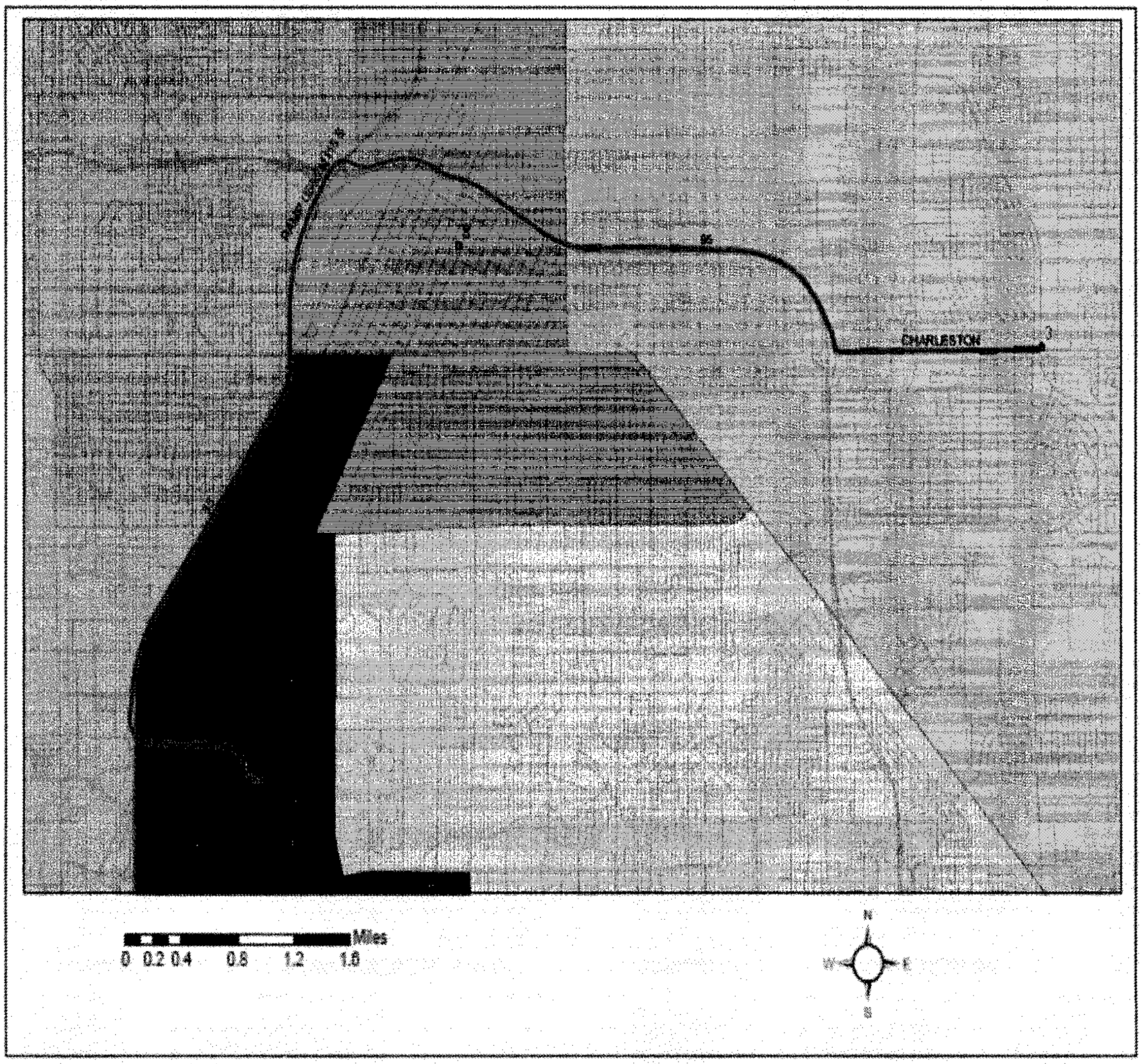
Appendix B.11. Network distance from District 3 to District 12



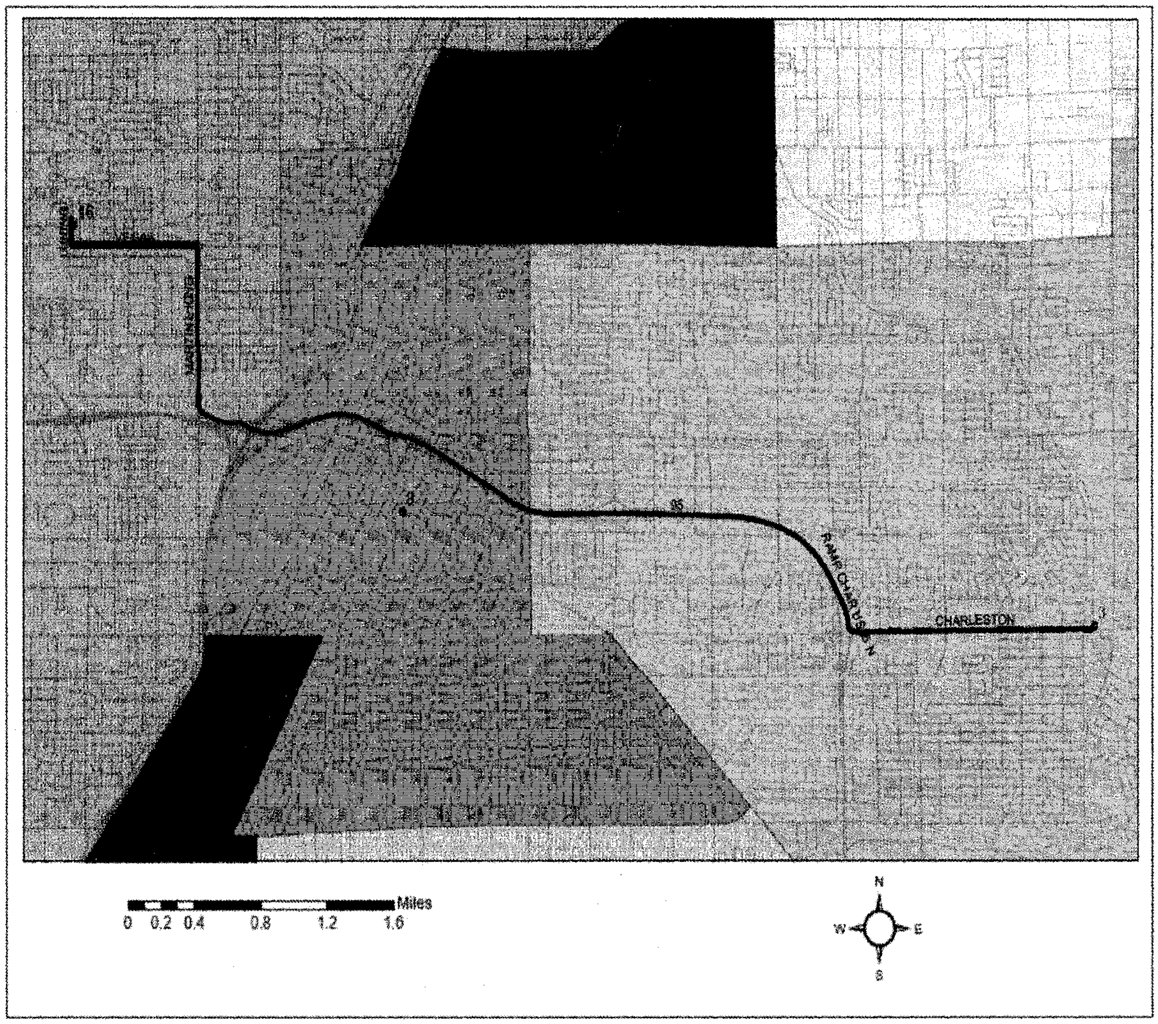
Appendix B.12. Network distance from District 3 to District 13



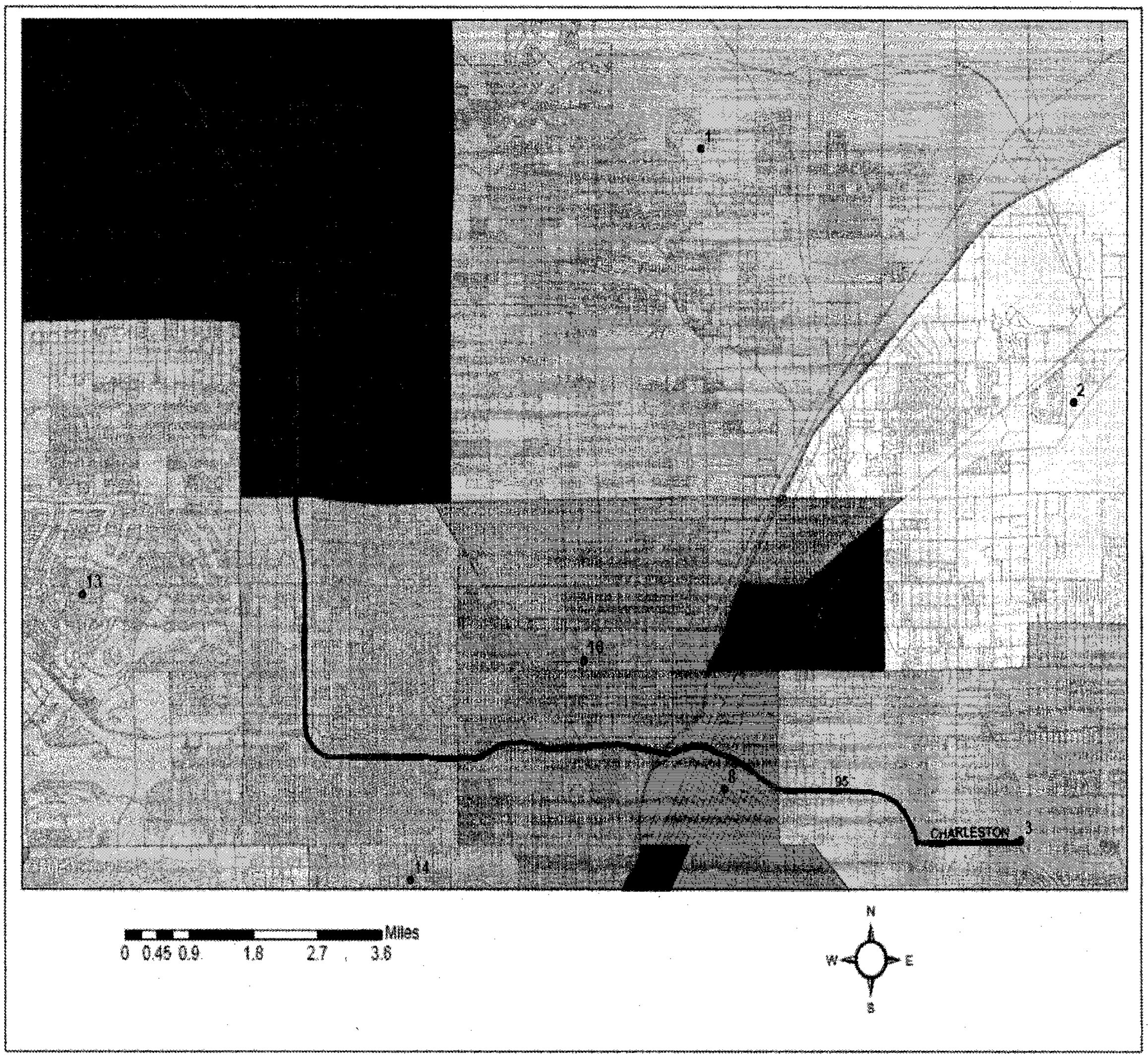
Appendix B.13. Network distance from District 3 to District 14



Appendix B.13. Network distance from District 3 to District 15



Appendix B.14. Network distance from District 3 to District 16



Appendix B.15. Network distance from District 3 to District 17

C. Parcel-based planning areas growth factors

	AREA	1990	1996	Factor	Average Factor
District 1					
	North Las Vegas	15070	31191	2.070	2.070
District 2					
	Sunrise Manor	37173	59356	1.597	1.597
District 3					
	Sunrise Manor	37173	59356	1.597	1.597
District 4					
	Whitney	3975	6574	1.654	1.654
District 5					
	Henderson	24241	58384	2.408	2.408
District 6					
	Henderson	24241	58384	2.408	2.408
District 7					
	North Las Vegas	15070	31191	2.070	2.070
District 8					
	Las Vegas	106689	167413	1.569	1.569
District 9					
	Paradise	61117	74102	1.212	
	Winchester	14679	15402	1.049	1.184
District 10					
	Paradise	61117	74102	1.212	
	Enterprise	2161	9952	4.605	1.614
District 11					
	Paradise	61117	74102	1.212	
	Henderson	24241	58384	2.408	
	Enterprise	2161	9952	4.605	1.940
District 12					
	Enterprise	2161	9952	4.605	
	Spring Valley	21832	42402	1.942	2.448
District13					
	Lone Mountain	1732	3048	1.760	
	Las Vegas	106689	167413	1.569	1.573
District 14					
	Winchester	14679	15402	1.049	
	Las Vegas	106689	167413	1.569	1.525
District 15					
	Paradise	61117	74102	1.212	
	Las Vegas	106689	167413	1.569	1.460
District 16					
	North Las Vegas	15070	31191	2.070	2.070
District 17					
	Lone Mountain	1732	3048	1.760	1.760
OVERALL (AVERAGE) VALLEY GROWTH					**1.820**

D. Trip attraction model results

D.1. Non-transformed non-spatial trip attraction model

Source	SS	df	MS
Model	2.8417e+09	2	1.4209e+09
Residual	459552464	14	32825176
Total	3.3013e+09	16	206331062

Number of obs =	17
F(2, 14) =	43.29
Prob > F =	0.0000
R-squared =	0.8608
Adj R-squared =	0.8409
Root MSE =	5729.3

expattrac	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Gaming	.3111028	.0570635	5.45	0.000	.1887137	.4334918
Non_Gaming	.5027336	.0714967	7.03	0.000	.3493885	.6560787
_cons	-3258.857	2760.691	-1.18	0.257	-9179.95	2662.235

D.2. Non-spatial trip attraction model transformed by inverse of the square

Source	SS	df	MS
Model	2113160.8	2	1056580.4
Residual	640131.899	14	45723.7071
Total	2753292.7	16	172080.794

Number of obs =	17
F(2, 14) =	23.11
Prob > F =	0.0000
R-squared =	0.7675
Adj R-squared =	0.7343
Root MSE =	213.83

expattracs~v	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Gaming	-.0066468	.0021297	-3.12	0.008	-.0112147	-.002079
Non_Gaming	-.0153216	.0026684	-5.74	0.000	-.0210448	-.0095985
_cons	1565.2	103.035	15.19	0.000	1344.212	1786.188

D.3. Spatial trip attraction models using contiguity spatial relationship

```
 Weights matrix
 Name: W
 Type: Imported (binary)
 Row-standardized: No

Spatial lag model                                   Number of obs    =         17
                                                    Variance ratio   =      0.859
                                                    Squared corr.    =      0.859
Log likelihood = -121.84752                         Sigma            =     313.74

------------------------------------------------------------------------------
expattracl~2 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
expattracl~2 |
      Gaming |   .0156511   .0036318     4.31   0.000     .0085329    .0227692
  Non_Gaming |   .0319323    .004916     6.50   0.000     .0222971    .0415676
       _cons |   8083.894   347.9537    23.23   0.000     7401.918    8765.871
-------------+----------------------------------------------------------------
         rho |   .0019825   .0092328     0.21   0.830    -.0161134    .0200785
------------------------------------------------------------------------------
Wald test of rho=0:                       chi2(1) =    0.046 (0.830)
Likelihood ratio test of rho=0:           chi2(1) =    0.046 (0.830)
Lagrange multiplier test of rho=0:        chi2(1) =    0.045 (0.832)

Acceptable range for rho: -1.885 < rho < 1.000
```

```
 Weights matrix
 Name: W
 Type: Imported (binary)
 Row-standardized: No

Spatial error model                                 Number of obs    =         17
                                                    Variance ratio   =      0.850
                                                    Squared corr.    =      0.859
Log likelihood = -121.86881                         Sigma            =     314.13

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
expattracl~2 |
      Gaming |   .0159568   .0034963     4.56   0.000     .0091041    .0228095
  Non_Gaming |   .0323962   .0049142     6.59   0.000     .0227645    .0420279
       _cons |   8137.349   279.8344    29.08   0.000     7588.883    8685.814
-------------+----------------------------------------------------------------
      lambda |   .0006461   .0109686     0.06   0.953    -.0208519    .0221441
------------------------------------------------------------------------------
Wald test of lambda=0:                    chi2(1) =    0.003 (0.953)
Likelihood ratio test of lambda=0:        chi2(1) =    0.003 (0.953)
Lagrange multiplier test of lambda=0:     chi2(1) =    1.037 (0.309)

Acceptable range for lambda: -1.885 < lambda < 1.000
```

D.4. Spatial trip attraction models using contiguity-separation spatial relationship

```
Weights matrix
 Name: W
 Type: Imported (non-binary)
 Row-standardized: No

Spatial lag model                                 Number of obs    =         17
                                                  Variance ratio   =      0.892
                                                  Squared corr.    =      0.892
Log likelihood = -119.58413                       Sigma            =     274.48

------------------------------------------------------------------------------
expattracl~2 |      Coef.   Std. Err.      z    P>|z|      [95% Conf. Interval]
-------------+----------------------------------------------------------------
expattracl~2 |
      Gaming |   .0123367   .0031775     3.88   0.000       .006109    .0185644
  Non_Gaming |   .0284224    .003874     7.34   0.000      .0208294    .0360154
       _cons |   7863.543   182.3276    43.13   0.000      7506.187    8220.898
-------------+----------------------------------------------------------------
         rho |   .0710566   .0310028     2.29   0.022      .0102922     .131821
------------------------------------------------------------------------------
Wald test of rho=0:                      chi2(1) =    5.253 (0.022)
Likelihood ratio test of rho=0:          chi2(1) =    4.573 (0.032)
Lagrange multiplier test of rho=0:       chi2(1) =    4.271 (0.039)

Acceptable range for rho: -1.896 < rho < 1.000


Weights matrix
 Name: W
 Type: Imported (non-binary)
 Row-standardized: No

Spatial error model                               Number of obs    =         17
                                                  Variance ratio   =      0.628
                                                  Squared corr.    =      0.858
Log likelihood = -119.72389                       Sigma            =     276.71

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|      [95% Conf. Interval]
-------------+----------------------------------------------------------------
expattracl~2 |
      Gaming |   .0128311   .0031296     4.10   0.000      .0066972    .0189649
  Non_Gaming |   .0285717    .003888     7.35   0.000      .0209513    .0361922
       _cons |   8543.684    230.395    37.08   0.000      8092.118     8995.25
-------------+----------------------------------------------------------------
      lambda |   .0794103   .0350824     2.26   0.024      .0106501    .1481705
------------------------------------------------------------------------------
Wald test of lambda=0:                   chi2(1) =    5.124 (0.024)
Likelihood ratio test of lambda=0:       chi2(1) =    4.293 (0.038)
Lagrange multiplier test of lambda=0:    chi2(1) =    2.745 (0.098)

Acceptable range for lambda: -1.896 < lambda < 1.000
```

D.5. Spatial trip attraction models using economic linkage (accessibility) spatial relationship

```
Weights matrix
 Name: W
 Type: Imported (non-binary)
 Row-standardized: No

Spatial lag model                                   Number of obs    =        17
                                                    Variance ratio   =     0.875
                                                    Squared corr.    =     0.875
Log likelihood = -120.84953                         Sigma            =    295.85

------------------------------------------------------------------------------
expattracl~2 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
expattracl~2 |
      Gaming |   .0104074   .0048321     2.15   0.031      .0009367    .0198781
  Non_Gaming |   .0236106   .0071156     3.32   0.001      .0096642     .037557
       _cons |   8015.681   169.6626    47.24   0.000      7683.149    8348.214
-------------+----------------------------------------------------------------
         rho |   .0046344   .0031462     1.47   0.141     -.0015321    .0108009
------------------------------------------------------------------------------
Wald test of rho=0:                     chi2(1) =   2.170 (0.141)
Likelihood ratio test of rho=0:         chi2(1) =   2.042 (0.153)
Lagrange multiplier test of rho=0:      chi2(1) =   1.912 (0.167)

Acceptable range for rho: -3.131 < rho < 1.000

 Weights matrix
 Name: W
 Type: Imported (non-binary)
 Row-standardized: No

Spatial error model                                 Number of obs    =        17
                                                    Variance ratio   =     0.427
                                                    Squared corr.    =     0.858
Log likelihood = -120.89914                         Sigma            =    296.72

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
expattracl~2 |
      Gaming |   .0105292   .0048119     2.19   0.029      .0010981    .0199603
  Non_Gaming |   .0235784   .0072358     3.26   0.001      .0093964    .0377604
       _cons |   8054.397   158.4667    50.83   0.000      7743.809    8364.986
-------------+----------------------------------------------------------------
      lambda |   .0052865   .0034448     1.53   0.125     -.0014652    .0120382
------------------------------------------------------------------------------
Wald test of lambda=0:                  chi2(1) =   2.355 (0.125)
Likelihood ratio test of lambda=0:      chi2(1) =   1.943 (0.163)
Lagrange multiplier test of lambda=0:   chi2(1) =   0.052 (0.820)

Acceptable range for lambda: -3.131 < lambda < 1.000
```

E. Trip production model results

E.1. Non-transformed non-spatial trip production model

```
      Source |       SS       df       MS              Number of obs =      17
-------------+------------------------------           F(  2,    14) =   34.82
       Model |  1.5741e+09     2   787037001           Prob > F      =  0.0000
    Residual |   316427179    14  22601941.3           R-squared     =  0.8326
-------------+------------------------------           Adj R-squared =  0.8087
       Total |  1.8905e+09    16   118156324           Root MSE      =  4754.1

------------------------------------------------------------------------------
  expproduc2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lowinc |   2.495531   .6784971     3.68   0.002       1.0403    3.950763
     highinc |   1.290447   .1863043     6.93   0.000     .8908641     1.69003
       _cons |   3628.106   2078.161     1.75   0.103    -829.1059    8085.317
------------------------------------------------------------------------------
```

E.2. Non-spatial trip production model transformed by applying natural logarithm

```
      Source |       SS       df       MS              Number of obs =      17
-------------+------------------------------           F(  2,    14) =   19.16
       Model |  5378225.21     2   2689112.6           Prob > F      =  0.0001
    Residual |  1964585.24    14  140327.517           R-squared     =  0.7324
-------------+------------------------------           Adj R-squared =  0.6942
       Total |  7342810.45    16  458925.653           Root MSE      =   374.6

------------------------------------------------------------------------------
expproducl~2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lowinc |   .1279699   .0534622     2.39   0.031      .013305    .2426349
     highinc |   .0783078   .0146798     5.33   0.000     .0468226    .1097929
       _cons |   8768.306   163.7486    53.55   0.000       8417.1    9119.512
------------------------------------------------------------------------------
```

E.3. Spatial trip production models using contiguity spatial relationship

```
Weights matrix
 Name: W
 Type: Imported (binary)
 Row-standardized: No

Spatial lag model                                 Number of obs    =        17
                                                  Variance ratio   =     0.808
                                                  Squared corr.    =     0.808
Log likelihood = -150.85303                       Sigma            =   1728.08

------------------------------------------------------------------------------
exppproducs~2 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
exppproducs~2 |
      lowinc |   .8643563   .2527388     3.42   0.001     .3689974    1.359715
     highinc |   .4847316   .0677979     7.15   0.000       .35185    .6176131
       _cons |   7455.674   1893.048     3.94   0.000     3745.368    11165.98
-------------+----------------------------------------------------------------
         rho |   .0012061   .0276853     0.04   0.965    -.053056     .0554682
------------------------------------------------------------------------------
Wald test of rho=0:                    chi2(1) =   0.002 (0.965)
Likelihood ratio test of rho=0:        chi2(1) =   0.002 (0.965)
Lagrange multiplier test of rho=0:     chi2(1) =   0.002 (0.966)

Acceptable range for rho: -1.885 < rho < 1.000
```

```
Weights matrix
 Name: W
 Type: Imported (binary)
 Row-standardized: No

Spatial error model                               Number of obs    =        17
                                                  Variance ratio   =     0.813
                                                  Squared corr.    =     0.808
Log likelihood = -150.84926                       Sigma            =   1727.70

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
exppproducs~2 |
      lowinc |   .8763154    .264552     3.31   0.001      .357803    1.394828
     highinc |   .4850585   .0678064     7.15   0.000     .3521605    .6179566
       _cons |    7660.45   1522.117     5.03   0.000     4677.155    10643.74
-------------+----------------------------------------------------------------
      lambda |  -.0049651   .0512499    -0.10   0.923    -.1054132    .0954829
------------------------------------------------------------------------------
Wald test of lambda=0:                 chi2(1) =   0.009 (0.923)
Likelihood ratio test of lambda=0:     chi2(1) =   0.009 (0.923)
Lagrange multiplier test of lambda=0:  chi2(1) =   0.652 (0.419)

Acceptable range for lambda: -1.885 < lambda < 1.000
```

E.4. Spatial trip production models using contiguity-separation spatial relationship

```
 Weights matrix
 Name: W
 Type: Imported (non-binary)
 Row-standardized: No

Spatial lag model                                   Number of obs   =        17
                                                    Variance ratio  =     0.808
                                                    Squared corr.   =     0.808
Log likelihood = -150.84389                         Sigma           =   1727.09

------------------------------------------------------------------------------
expproducs~2 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
expproducs~2 |
      lowinc |   .8848052     .277295     3.19   0.001      .341317    1.428293
     highinc |   .4805628    .0733799     6.55   0.000     .3367408    .6243848
       _cons |   7705.026    1437.465     5.36   0.000     4887.647    10522.41
-------------+----------------------------------------------------------------
         rho |  -.0188477    .1327059    -0.14   0.887    -.2789465    .2412512
------------------------------------------------------------------------------
Wald test of rho=0:                     chi2(1) =   0.020 (0.887)
Likelihood ratio test of rho=0:         chi2(1) =   0.020 (0.887)
Lagrange multiplier test of rho=0:      chi2(1) =   0.022 (0.883)

Acceptable range for rho: -1.896 < rho < 1.000
```

```
 Weights matrix
 Name: W
 Type: Imported (non-binary)
 Row-standardized: No

Spatial error model                                 Number of obs   =        17
                                                    Variance ratio  =     0.810
                                                    Squared corr.   =     0.808
Log likelihood = -150.81381                         Sigma           =   1723.37

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
expproducs~2 |
      lowinc |    .912093    .2916457     3.13   0.002     .3404778    1.483708
     highinc |   .4778554    .0713462     6.70   0.000     .3380195    .6176913
       _cons |   7380.128    900.6243     8.19   0.000     5614.937     9145.32
-------------+----------------------------------------------------------------
      lambda |  -.0635266    .2255099    -0.28   0.778    -.5055178    .3784646
------------------------------------------------------------------------------
Wald test of lambda=0:                  chi2(1) =   0.079 (0.778)
Likelihood ratio test of lambda=0:      chi2(1) =   0.080 (0.777)
Lagrange multiplier test of lambda=0:   chi2(1) =   0.377 (0.539)

Acceptable range for lambda: -1.896 < lambda < 1.000
```

E.5. Spatial trip production models using economic linkage (accessibility) spatial relationship

```
 Weights matrix
 Name: W
 Type: Imported (non-binary)
 Row-standardized: No

Spatial lag model                                   Number of obs    =        17
                                                    Variance ratio   =     0.808
                                                    Squared corr.    =     0.808
Log likelihood = -150.85298                         Sigma            =   1728.08

------------------------------------------------------------------------------
expproducs~2 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
expproducs~2 |
      lowinc |   .8714846   .2682629     3.25   0.001       .345699    1.39727
     highinc |   .4846677   .0677419     7.15   0.000      .3518959   .6174395
       _cons |   7563.622   1045.423     7.23   0.000      5514.631   9612.613
-------------+----------------------------------------------------------------
         rho |  -.0002681   .0059922    -0.04   0.964     -.0120126   .0114765
------------------------------------------------------------------------------
Wald test of rho=0:                      chi2(1) =    0.002 (0.964)
Likelihood ratio test of rho=0:          chi2(1) =    0.002 (0.964)
Lagrange multiplier test of rho=0:       chi2(1) =    0.002 (0.964)

Acceptable range for rho: -3.131 < rho < 1.000
```

```
 Weights matrix
 Name: W
 Type: Imported (non-binary)
 Row-standardized: No

Spatial error model                                 Number of obs    =        17
                                                    Variance ratio   =     0.824
                                                    Squared corr.    =     0.808
Log likelihood = -150.83507                         Sigma            =   1726.26

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
expproducs~2 |
      lowinc |    .890813   .2760239     3.23   0.001       .349816    1.43181
     highinc |   .4865869   .0687131     7.08   0.000      .3519118    .621262
       _cons |   7649.465   967.3805     7.91   0.000      5753.434   9545.496
-------------+----------------------------------------------------------------
      lambda |  -.0022578   .0117161    -0.19   0.847      -.025221   .0207053
------------------------------------------------------------------------------
Wald test of lambda=0:                   chi2(1) =    0.037 (0.847)
Likelihood ratio test of lambda=0:       chi2(1) =    0.038 (0.846)
Lagrange multiplier test of lambda=0:    chi2(1) =    0.685 (0.408)

Acceptable range for lambda: -3.131 < lambda < 1.000
```

F. Trip generation models comparison tests

F.1. RSS, AIC, and SIC calculations for trip attraction models

OBSERVED	PREDICTED					RESIDUAL SQUARED				
expattraclog	Non Spa	Contiguity	Separation	Conti-Sep	Access	Non Spatial	Contiguity	Separation	Contiguity-Separation	Accessibility
8988.132	8864.543	8815.842	8809.448	8700.649	8784.258	15274.16	29684.01	31928.11	82646.26	41564.78
9315.109	9004.387	9036.789	8992.975	8980.886	8966.919	96548.37	77462.06	103770.5	111705.3	121235.9
9490.159	9345.66	9371.63	9465.813	9384.564	9560.886	20880.07	14049.12	592.745	11150.24	5002.296
7896.67	8429.076	8527.449	8248.639	8413.597	8361.35	283456.2	397882.2	123882.5	267213.7	215927.8
9222.603	9638.367	9648.737	9587.687	9608.145	9700.703	172859.7	181590.2	133286.7	148642.8	228579.4
8401.871	8313.426	8268.042	8336.801	8281.405	8328.197	7822.459	17910.17	4234.158	14512.04	5427.895
8752.216	8366.298	8388.312	8778.317	8680.568	8520.939	148932.8	132426.2	681.2794	5133.399	53488.85
10206.36	9777.981	9802.237	10085.03	10000.15	10185.5	183508.5	163315.8	14721.56	42522.52	435.2103
10120.81	10726.9	10767.43	10746.35	10734.49	10932.18	367342.4	418123.6	391301	376598.3	658328.4
9735.833	9525.081	9603.176	9541.569	9623.339	9808.385	44416.52	17597.97	37738.62	12654.94	5263.737
8937.212	8982.813	9012.647	8876.203	8902.165	9023.36	2079.411	5690.426	3722.111	1228.284	7421.459
10092.69	9763.073	9793.018	9600.197	9620.622	9695.821	108647.6	89803.42	242549.7	222848.1	157504.8
8789.588	8908.257	8861.282	8773.134	8713.214	8941.078	14082.39	5140.098	270.7325	5833.049	22949.08
10801.35	11013.4	11045.47	10935.47	10900.65	11004.73	44966.91	59595.09	17988.13	9859.907	41362.39
10691.17	11058.97	11067.05	10833.68	10959.5	11077.43	135278.5	141287	20308.68	71998.66	149197.8
10468	10120.16	10229.07	10411.98	10417.41	10580.42	120989.9	57088.35	3138.69	2559.01	12638.95
8446.189	8667.617	8677.577	8488.501	8549.522	8891.827	49030.34	53540.43	1790.273	10677.78	198593.6
						1816116	1862186	1131906	1397784	1924922
	RSS					18.16	18.62	11.32	13.98	19.25
	RSS (x106)					1.82	1.86	1.13	1.40	1.92
	AIC					1.52	1.75	1.07	1.32	1.81
	SIC					1.76	2.13	1.30	1.60	2.21
	Final Loglikelihood					-	-121.52	-118.91	-119.42	-119.60

F.2. RSS, AIC, and SIC calculations for trip production models

OBSERVED	PREDICTED					RESIDUAL SQUARED				
expproducsqrt	Nonspat	Contiguity	Separation	Conti-Sep	Access	Non Spatial	Contiguity	Separation	Contiguity-Separation	Accessibility
13950.74	11292.1	10769.77	11137.04	10964.23	10981.86	7068367	10118555	7916923	8919263	8814277
12011.01	12281.74	11826.19	11986.81	12032.07	11881.44	73294.73	34157.83	585.4035	443.3196	16788.81
19120.94	20940.3	19630.28	20550.56	20521.48	20470.73	3310071	259430.3	2043815	1961504	1821946
9080.975	9618.593	9739.737	9038.825	9260.51	9124.828	289033.1	433967.6	1776.59	32232.68	1923.063
10923.83	10286.01	9964.905	10201.41	10190.75	10393.93	406814.4	919537.2	521889.5	537403.6	280797.8
5519.053	7948.927	7042.902	7636.325	7418.046	7484.449	5904288	2322116	4482840	3606175	3862780
9252.893	8282.636	7891.77	8765.369	8572.074	8045.367	941398.6	1852656	237679.6	463514.1	1458119
10000.92	10607.72	10031.36	10875.54	10842.57	11054.79	368206.2	926.2965	764952.5	708373.2	1110645
16217.47	17956	17613.79	18142.48	18148.19	18367.24	3022487	1949704	3705681	3727688	4621506
9446.418	11121.4	11469.4	11277.98	11526.6	11575.18	2805565	4092461	3354635	4327153	4531639
14971.22	14464.01	14222.38	14338.87	14338.93	14326.44	257262	560768.1	399866.2	399784.8	415738.8
17820.08	13291.63	13291.42	13096.97	13223.78	13305.63	20506859	20508803	22307735	21125993	20380244
13042.1	15131.99	14929.86	15093.95	15053.84	14991.26	4367640	3563648	4210094	4047096	3799211
20095.06	22967.72	23061.28	23319.32	23458.09	23294.04	8252175	8798448	10395821	11309982	10233456
8612.159	7948.927	8232.243	8274.643	8508.034	9015.764	439876.7	144335.9	113917.1	10841.96	162897.2
15692.48	14942.17	14551.64	15018.91	15190.46	15339.92	562965.1	1301526	453700.8	252026.9	124295.1
12579.18	13792.76	13681.63	13478.24	13622.08	13710.39	1472776	1215404	808317.7	1087640	1279628
						60049079	58076446	61720231	62517114	62915894
				RSS		60.05	58.08	61.72	62.52	62.92
				RSS (x107)		6.00	5.81	6.17	6.25	6.29
				AIC		5.03	5.47	5.81	5.89	5.92
				SIC		5.82	6.65	7.07	7.16	7.21
				Final Loglikelihood		-	-150.85	-146.15	-150.81	-150.84

G. Destination Choice Model Full Results

```
Iteration 0.   log likelihood = -2761.3481
Iteration 1:   log likelihood = -2721.4657
Iteration 2:   log likelihood = -2693.7093
Iteration 3:   log likelihood = -2685.5188
Iteration 4:   log likelihood =   -2683.727
Iteration 5:   log likelihood = -2683.5952
Iteration 6:   log likelihood = -2683.5935
Iteration 7:   log likelihood = -2683.5935

Multinomial logistic regression                   Number of obs   =       1113
                                                  LR chi2(64)     =     155.51
                                                  Prob > chi2     =     0.0000
Log likelihood = -2683.5935                       Pseudo R2       =     0.0282
```

choice	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
1						
vehhhworker	.7894907	.4335827	1.82	0.069	-.0603157	1.639297
age	.0145235	.0167386	0.87	0.386	-.0182836	.0473306
income	-.3498629	.4786136	-0.73	0.465	-1.287928	.5882026
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	-2.142848	.9959254	-2.15	0.031	-4.094826	-.1908704
2						
vehhhworker	-.3056246	.4693065	-0.65	0.515	-1.225448	.6141992
age	-.0176456	.013927	-1.27	0.205	-.0449421	.0096508
income	.0836478	.4109282	0.20	0.839	-.7217566	.8890523
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	-.3598748	.7592224	-0.47	0.635	-1.847923	1.128174
3						
vehhhworker	.688124	.3505605	1.96	0.050	.001038	1.37521
age	-.0082278	.0127916	-0.64	0.520	-.0332989	.0168434
income	.1272641	.3935141	0.32	0.746	-.6440095	.8985376
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	-1.319949	.7329005	-1.80	0.072	-2.756408	.1165094
4						
vehhhworker	.6824683	.6888575	0.99	0.322	-.6676676	2.032604
age	-.0302502	.0278	-1.09	0.277	-.0847372	.0242368
income	.9646376	1.081831	0.89	0.373	-1.155711	3.084987
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	-1.057457	1.757841	-0.60	0.547	-4.502762	2.387847

Full Destination Choice Model Results (continued)

5						
vehhhworker	.4058932	.4867115	0.83	0.404	-.5480439	1.35983
age	.014231	.0166587	0.85	0.393	-.0184195	.0468815
income	.2230513	.5206919	0.43	0.668	-.7974861	1.243589
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	-2.8899	1.012828	-2.85	0.004	-4.875006	-.9047937
6						
vehhhworker	.6438898	.5920078	1.09	0.277	-.5164242	1.804204
age	-.0192185	.0220994	-0.87	0.384	-.0625326	.0240956
income	-.5376516	.6197367	-0.87	0.386	-1.752313	.6770099
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	-1.733936	1.178485	-1.47	0.141	-4.043723	.5758511
7						
vehhhworker	.2855315	.5714962	0.50	0.617	-.8345803	1.405643
age	-.0367047	.0188864	-1.94	0.052	-.0737214	.0003119
income	-.9021742	.5092281	-1.77	0.076	-1.900243	.0958944
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	-.1408222	.9430957	-0.15	0.881	-1.989256	1.707611
8						
vehhhworker	-.0466236	.3255469	-0.14	0.886	-.6846839	.5914366
age	.0092031	.0099094	0.93	0.353	-.010219	.0286252
income	.0216671	.2968356	0.07	0.942	-.5601199	.6034541
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	-.7023613	.575027	-1.22	0.222	-1.829394	.4246709
10						
vehhhworker	-.4040309	.3982756	-1.01	0.310	-1.184637	.376575
age	-.0055197	.0118625	-0.47	0.642	-.0287698	.0177304
income	-.1772244	.3391512	-0.52	0.601	-.8419485	.4874996
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	.812844	.6741554	1.21	0.228	-.5084762	2.134164
11						
vehhhworker	.4865976	.505949	0.96	0.336	-.5050442	1.478239
age	-.0112252	.0183531	-0.61	0.541	-.0471965	.0247462
income	-.4362422	.5074886	-0.86	0.390	-1.430902	.5584172
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	.9467044	1.108772	0.85	0.393	-1.226448	3.119857

Full Destination Choice Model Results (continued)

12						
vehhhworker	.1878075	.3770611	0.50	0.618	-.5512186	.9268336
age	-.0242498	.0124322	-1.95	0.051	-.0486165	.0001169
income	.6101774	.4055434	1.50	0.132	-.184673	1.405028
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	.8297037	.7342285	1.13	0.258	-.6093578	2.268765
13						
vehhhworker	.0206892	.5290322	0.04	0.969	-1.016195	1.057573
age	-.033429	.0170805	-1.96	0.050	-.0669062	.0000482
income	1.174792	.6527348	1.80	0.072	-.1045444	2.454129
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	-1.007179	1.007789	-1.00	0.318	-2.982409	.9680507
14						
vehhhworker	.3524887	.273767	1.29	0.198	-.1840848	.8890622
age	-.0091213	.0087748	-1.04	0.299	-.0263196	.008077
income	.2331835	.2675491	0.87	0.383	-.291203	.7575701
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	.4780546	.5025591	0.95	0.341	-.5069431	1.463052
15						
vehhhworker	.2849998	.2765424	1.03	0.303	-.2570133	.827013
age	-.0036832	.0087631	-0.42	0.674	-.0208586	.0134922
income	.0987135	.2643001	0.37	0.709	-.4193051	.6167321
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	.1987228	.5018729	0.40	0.692	-.7849299	1.182376
16						
vehhhworker	.4889187	.2994338	1.63	0.103	-.0979607	1.075798
age	.0010443	.009949	0.10	0.916	-.0184554	.0205441
income	.1094127	.3028389	0.36	0.718	-.4841406	.702966
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	-.6515511	.5771901	-1.13	0.259	-1.782823	.4797208
17						
vehhhworker	.7486465	.534291	1.40	0.161	-.2985447	1.795838
age	.0028652	.0206548	0.14	0.890	-.0376174	.0433479
income	-.3886135	.5852728	-0.66	0.507	-1.535727	.7585
spatial	2.6402154	0.838163	3.15	0.004	.454265	3.2959015
CBD	0.0226812	0.014724	1.54	0.131	-.0958253	.0915704
hotelre	0.1684324	0.008334	20.21	0.000	.0847163	.2896424
otherjo	0.4763584	0.027127	17.56	0.000	.2838354	.5753672
_cons	-1.926138	1.200937	-1.60	0.109	-4.279932	.4276562

H. Creation of Origin-Destination (O-D) Matrix using logit model results

H.1. Average Probabilities

		Attractions (at destination, j)																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Productions (at origin, i)	1	0.0227	0.0410	0.0334	0.0135	0.0478	0.0119	0.0123	0.0623	0.1154	0.0690	0.0247	0.0616	0.0242	0.1763	0.1687	0.0988	0.0163	17126
	2	0.0227	0.0410	0.0334	0.0135	0.0478	0.0119	0.0123	0.0622	0.1159	0.0689	0.0247	0.0616	0.0242	0.1761	0.1686	0.0987	0.0163	16457
	3	0.0226	0.0409	0.0333	0.0134	0.0477	0.0119	0.0123	0.0621	0.1180	0.0688	0.0246	0.0615	0.0241	0.1757	0.1682	0.0985	0.0163	41721
	4	0.0227	0.0410	0.0335	0.0109	0.0300	0.0112	0.0164	0.0798	0.1162	0.0604	0.0224	0.0564	0.0236	0.1799	0.1689	0.0959	0.0150	9074
	5	0.0226	0.0409	0.0334	0.0109	0.0300	0.0111	0.0163	0.0795	0.1187	0.0602	0.0224	0.0562	0.0235	0.1794	0.1684	0.0956	0.0150	9023
	6	0.0224	0.0405	0.0331	0.0127	0.0428	0.0116	0.0132	0.0660	0.1258	0.0660	0.0238	0.0596	0.0237	0.1751	0.1668	0.0969	0.0158	2299
	7	0.0216	0.0387	0.0467	0.0013	0.0328	0.0112	0.0156	0.0765	0.1208	0.0615	0.0227	0.0569	0.0235	0.1784	0.1680	0.0958	0.0152	7529
	8	0.0211	0.0377	0.0541	0.0020	0.0248	0.0109	0.0175	0.0845	0.1195	0.0577	0.0217	0.0547	0.0233	0.1804	0.1684	0.0947	0.0146	11616
	9	0.0172	0.0298	0.1057	0.0069	0.0000	0.0084	0.0304	0.1402	0.1216	0.0305	0.0144	0.0379	0.0213	0.1921	0.1691	0.0854	0.0104	40485
	10	0.0214	0.0383	0.0515	0.0017	0.0281	0.0111	0.0169	0.0818	0.1147	0.0596	0.0222	0.0559	0.0235	0.1807	0.1692	0.0957	0.0149	10072
	11	0.0199	0.0352	0.0710	0.0036	0.0064	0.0101	0.0217	0.1028	0.1182	0.0490	0.0194	0.0493	0.0227	0.1846	0.1690	0.0919	0.0133	21047
	12	0.0227	0.0409	0.0334	0.0067	0.0013	0.0099	0.0229	0.1078	0.1192	0.0465	0.0187	0.0478	0.0225	0.1855	0.1689	0.0910	0.0129	23637
	13	0.0226	0.0409	0.0334	0.0078	0.0090	0.0102	0.0211	0.1001	0.1194	0.0502	0.0197	0.0500	0.0227	0.1838	0.1687	0.0922	0.0134	19699
	14	0.0226	0.0409	0.0334	0.0082	0.0000	0.0077	0.0123	0.1552	0.1172	0.0688	0.0247	0.0615	0.0241	0.1759	0.1684	0.0986	0.0163	48263
	15	0.0229	0.0414	0.0338	0.0136	0.0483	0.0121	0.0125	0.0629	0.1062	0.0697	0.0250	0.0623	0.0244	0.1781	0.1705	0.0998	0.0165	9258
	16	0.0227	0.0410	0.0335	0.0135	0.0479	0.0120	0.0124	0.0623	0.1146	0.0690	0.0247	0.0617	0.0242	0.1764	0.1689	0.0989	0.0164	21679
	17	0.0227	0.0410	0.0334	0.0135	0.0478	0.0119	0.0123	0.0623	0.1155	0.0690	0.0247	0.0616	0.0242	0.1762	0.1687	0.0988	0.0163	16375
		7052	12676	15101	2963	7664	3373	5570	31458	38269	19103	7151	17996	7609	58679	54897	30975	4824	325360

H.2. O-D Matrix

		Attractions (at destination, j)																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Productions (at origin, i)	1	389	702	573	231	819	204	211	1067	1976	1181	423	1056	414	3019	2890	1692	280	17126
	2	373	674	550	222	787	196	203	1024	1908	1134	406	1014	398	2899	2775	1625	269	16457
	3	944	1705	1391	561	1989	497	514	2591	4922	2869	1028	2564	1006	7332	7019	4109	680	41721
	4	206	372	304	99	272	101	149	724	1054	548	203	512	214	1633	1533	870	136	8929
	5	204	369	301	98	270	100	147	717	1071	544	202	507	212	1619	1520	863	135	8880
	6	52	93	76	29	98	27	30	152	289	152	55	137	55	403	383	223	36	2289
	7	162	291	352	10	247	85	118	576	909	463	171	429	177	1343	1265	721	114	7432
	8	245	438	628	23	288	127	203	981	1388	671	252	635	270	2095	1956	1100	170	11469
	9	697	1207	4279	279	0	340	1230	5676	4922	1234	583	1536	861	7779	6844	3459	422	41348
	10	215	386	519	17	283	112	170	824	1155	600	224	563	237	1820	1705	964	150	9943
	11	418	741	1495	75	135	213	457	2164	2488	1031	408	1039	477	3885	3557	1934	279	20795
	12	535	967	789	158	30	233	541	2547	2818	1099	442	1130	531	4384	3992	2150	304	22651
	13	446	806	657	154	177	201	416	1972	2352	988	388	986	448	3620	3323	1816	265	19014
	14	1093	1974	1611	396	0	373	595	7490	5657	3321	1190	2969	1165	8489	8127	4758	787	49994
	15	212	383	313	126	447	112	115	583	983	645	231	577	226	1649	1578	924	153	9258
	16	492	889	726	293	1038	259	268	1351	2484	1496	536	1338	525	3825	3661	2143	354	21679
	17	371	671	548	221	783	196	202	1020	1891	1129	405	1009	396	2886	2763	1617	267	16375
		7055	12669	15112	2989	7664	3374	5570	31459	38267	19104	7147	17999	7613	58679	54891	30967	4801	**325360**

REFERENCES

Applied Management and Planning Group, and Parsons Brinckerhoff, Q & D, Inc., *"1996 Las Vegas Valley Household Travel Survey"*, Final Report, prepared for Regional Transportation Commission of Southern Nevada (RTC), 1998.

Ashtakala, B., (1987) *Generalized Power Model for Trip Distribution* Transportation Research, Part B, Volume 21B, Issue No. 1, pp. 59 – 67

Bartfay, E. and Donner, A. (2000) *The Effect of Collapsing Multinomial Data when Assessing Agreement* International Journal of Epidemiology, Volume 29, pp. 1070 – 1075

Ben-Akiva, M. and Lerman, S. R. (1985) Discrete Choice Analysis: Theory and Application to Travel Demand, The MIT Press, Cambridge, Massachusetts, London, England.

Bento, A. M., Maureen L. C., Ahmed M. M., and Katja V. (2005). The effects of urban spatial structure on travel demand in the United States. *The Review of Economics and Statistics,* 87, 3, pp. 466-478.

Bierlaire, M. (2003). BIOGEME: a free package for the estimation of discrete choice models. Conference paper, 3rd Swiss Transport Research Conference (STRC), Monte Verita, Ascona.

Bowman, J. L. and Ben-Akiva, M. E., (2000), *Activity-based disaggregate travel demand model system with activity schedules,* Transportation Research Part A: Policy and Practice, Volume 35, Issue 1, pp. 1-28.

Butler, C. J. (1972). Gravity Models as planning tools: a review of theoretical and operational problems, Geografiska Annaler. Series B, Human Geography, Vol. 54, Issue No. 1, pp. 68-78.

Casey, H. J., (1955) *Applications to traffic Engineering of the Law of Retail Gravitation* Traffic Quarterly, IX (1), pp 23 – 35 cited by Ortùzar and Willumsen 2001

Clark County, (1996) *July 1996 Population Estimate for Clark County Incorporated Cities and Unincorporated Towns,* prepared by the Clark County Department of Comprehensive Planning

Cramer, J. S. (1991), The Logit Model: An Introduction for Economists, Edward Arnold, London

D'Juran, C. A., (1995) A Combined Fratar-Gravity Model for Trip Distribution, Dissertation submitted to the Department of Civil Engineering, University of Toronto, available at http://wwwlib.umi.com/dissertations

Daganzo, C. F., (1979) Multinomial Probit: The Theory and Its Application to Demand Forecasting. Academic Press, Inc. New York.

DeDonnea, F. X., (1971) The Determinants of Transport Mode Choice in Dutch Cities, Rotterdam University Press, in Kanafani (1983).

Doubleday, C., (1977) *Some Studies of the Temporal Stability of Person Trip Generation Models* Transportation Research, Volume 11, pp 255 – 263

Duffus, L. N., A. S. Alfa., and A. H. Soliman, (1987) The reliability of using gravity model for forecasting trip distribution, Transportation, Vol. 14. pp. 175-192

ESRI, (2007). *ArcGIS—The Complete Geographic Information System*. Retrieved February 23, 2007, from ESRI Web site: http://www.esri.com/software/arcgis/

Ghaeli, R. and Bruce G. H. (1998). Spatial variations in travel behavior within greater Toronto area. *Journal of Transportation Engineering,* 124, 2, pp.179-187.

Greene, W. H., (1997) Econometric Analysis, 3rd Edition, Prentice-Hall, Inc., Upper Saddle River, New Jersey

Gujarati, D. N., Basic Econometrics, 4th Edition, McGraw Hill, Singapore, 2003.

Haining, R., (2003) Spatial Data Analysis, Theory and Practice, Cambridge University Press, United Kingdom.

Hamilton, C. L. (2004). Statistics with STATA, updated for Version 8, Thomson Learning™, Belmont, California, USA.

Hammadou, H., Thomas, I., Tindemans, H., Hofstraeten, D., and Verhestel, A., *How to incorporate the spatial dimension in destination choice models? The case of Antwerpen,* Draft version, April 2004. Available at http://www.ersa.org/ersaconfs/ersa04/PDF/70.pdf. Accessed October, 2006.

Handy, S. (1993). Regional versus local accessibility: implications for nonwork travel. *Transportation Research Record,* 1400, pp. 58-66.

Hensher D. A, Johnson, L. W (1981) Applied Discrete-Choice Modelling. Croom Helm, London

Hensher, D. A., and Button, K. J., Handbook of Transport Modelling, 1st Edition, Elsevier Science Ltd, Kidlington, Oxford, UK.

ITE, (1992). Transportation Planning Handbook, Edwards, J. D. (Jr.), editor, PTR Prentice Hall, Inc., Englewood Cliffs, New Jersey, USA.

Jacobson, J., (1982) *Alternative Specifications of the Random Disturbances for Trip Frequency Models,* Transportation Research, Part A, Volume 16A, No 3, pp 219 – 223

Johnson, N. L. & Kotz, S. (1970) Continuous Univariate Distributions-1, Houghton Mifflin Company, New York

Kanafani, A. (1983) Transport Demand Analysis, McGraw-Hill, United States of America.

Khatib, Z., Kang-tsung C. and Yanmei O. (2001). Impacts of zone structures on modeled statewide traffic. *Journal of Transportation Engineering,* 127, 1, pp.31-38.

Kitamura, R., (1984) *Incorporating Trip Chaining into Analysis of Destination Choice* Transportation Research, Part B, Volume 18B, Issue No. 1, pp. 67 – 81.

Kockelman, K. M. (1997). Travel behavior as a function of accessibility, land use mixing, and land use balance: evidence from San Francisco Bay Area. *Transportation Research Record,* 1607, pp. 116-125.

Koenig, J. G., Indicators of Urban Accessibility: Theory and Application. *Transportation,* Vol. 9, 1980, pp. 145 - 172.

Kwan, M.-P and Joe W. (2007). Scale and accessibility: Implications for the analysis of land use-travel interaction. *Applied Geography,* doi: 10.1016/j.apgeog.2007.07.02.

Limanond, T. (2001) Effects of Neighborhood Setting and Intra-Neighborhood Location on Shopping Travel Behavior of Residents in Traditional Neighborhoods, Dissertation submitted to the Department of Civil and Environmental Engineering, University of California, Davis, available at http://wwwlib.umi.com/dissertations.

Limanond T. and Debbie A. N. (2004). Effect of land use on decisions of shopping tour generation: A cas study of three traditional neighborhoods in WA. *Transportation.* 31, 2, pp.153-181

Maddala, G. S. (1983) Limited Dependent and Qualitative Variables in Econometrics, Cambridge University Press, Cambridge

Manski, C. F., (1973) The Stochastic Utility Model of Choice. Unpublished Ph.D. Thesis. MIT, Department of Economics, Cambridge, MA.

Matula, D. W. and Sokal, R. R., Properties of Gabriel graphs relevant to Geographic variation and the clustering of points in the plane. Geographic Analysis, 12, 205-222.

McNally, M. G., (2000) *The Four Step Model,* available at http://repositories.cdlib.org/itsirvine/casa/UCI-ITS-AS-WP-96-1 , Accessed September, 2006.

Meurs, H., (1990) *Dynamic Analysis of Trip Generation,* Transportation Research, Part A, Volume 24 A, Issue No. 6, pp 427 – 442

Meyer, M. D. and Miller, E. J., (2001) Urban Transportation Planning: A decision-Oriented Approach, 2nd Edition, McGraw-Hill Companies, Inc., New York.

Naess, P. (2006). *Urban Structure Matters: Residential location, car dependence and travel behaviour,* Routledge, New York, USA.

Ortűzar, J., and Willumsen, L. G., (2001) Modelling Transport, John Wiley & Sons, West Sussex, England.

Pindyck, R. S., and Rubinfeld, D. L., (1998) Econometric Models and Economic Forecasting, 4th Edition, McGraw-Hill Companies, Inc., United States of America.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review,* 15, 3, pp. 351-357.

Rodrigue, J. P., Comtois, C., and Brian Slack, B., (2006) The geography of Transport System, New York: Routledge. Available at http://people.hofstra.edu/geotrans/index.html, Accessed September 2006.

Ruiter, E. R. and M. Ben-Akiva (1978). Disaggregate travel demand models for the San Francisco bay area, Transportation Research Record, Volume 673, pp. 121-128.

Said, G. M. and Young, D. H., (1990) *A General Linear Model Framework for Estimating Work Trip Rates for Households in Kuwait* Transportation Research, Part A, Volume 24A, No 3, pp 187 – 200

Southworth, F., (1981) *Calibration of Multinomial Logit Models of Mode and Destination Choice* Transportation Research, Volume 15A, pp. 315 – 325.

Srinivasan, S. (2001). Quantifying Spatial Characteristics for Travel Behavior Models. *Transportation Research Record,* 1777, pp. 1-15.

Stewart, J. Q., and W. Warntz. Physics of Population Distribution. *Journal of Regional Science*, Vol. 1, 1958, pp 99-123.

Stopher, P. R. and McDonald, K. G., (1983) *Trip generation by Cross-Classification; An Alternative Methodology* Transportation Research Record, Volume 944, pp 84 – 91.

Train, K. (1986) Qualitative Choice Analysis, Theory, Econometrics, and an Application to Automobile Demand, MIT Press, Cambridge (Mass.) and London.

Unwin, D. J. (1996). GIS, spatial analysis and spatial statistics. *Progress in Human Geography,* 20, 4, pp. 540-551.

Wansbeek, T. J., (1977) *Least-Squares Estimation of Trip Distribution Parameters: A Note* Transportation Research, Volume 11, pp. 429 – 431

Weiner, E. (1999). Urban Transportation Planning in the United States: an historical overview, Revised and Expanded Edition, Praeger Publishers, Westport, CT, USA.

White, M. T., (1976), *An Examination of Residual Distributions in Ordinary Least Squares (OLS) Household-Based Trip Generation Models,* Transportation Research, Volume 10, pp 249 – 254.

Wong, D. (1996). Aggregation effects in geo-referenced data, *Practical Handbook of Spatial Statistics,* Arlinghaus, S. L. (Editor). CRC Press, Inc., Boca Raton, FL, USA.

VITA

Valerian Kwigizile

Local Address:
1600 E Rochelle Ave, Apt 128
Las Vegas, NV 89119

Home Address:
c/o Mr. H. T. Kwigizile
P.O. Box 77476
Dar Es Salaam, Tanzania

Degrees:
Bachelor of Science in Civil Engineering, 2001
University of Dar Es Salaam, Tanzania

Masters of Science in Civil Engineering, 2004
Florida State University, USA

Publications:

1. Teng, H., **Kwigizile, V**., Karakouzian, M., James, D. E., and Etyemezian, V. (2007) Evaluation of the AP-42 sampling method, *Journal of the Air & Waste Management Association,* **In print**.

2. Teng, H., **Kwigizile, V**., James, D. E., and Merle, R. (2006) "Identifying Influencing Factors On Paved Roads Silt Loading", *Journal of the Air & Waste Management Association,* Vol. 57, No. 7, pp.778-784.

3. **Kwigizile, V.,** Mussa, R., and Selekwa, M. Connectionist Approach to Improving Highway Vehicle Classification Schemes – The Florida Case. Transportation Research Record, Volume 1917, December 2005, pp. 182-189.

4. Mussa, R., **Kwigizile, V**., and Selekwa, M.F. Probabilistic Neural Networks Application for Vehicle Classification. ASCE Journal of Transportation Engineering, Volume 132, Number 4, April 2006, pp. 293-302.

5. Selekwa, M.F., **Kwigizile, V.**, and Mussa, R. Setting Up a Probabilistic Neural Network for Classification of Highway Vehicles. International Journal of Computation Intelligence and Applications, Volume 5, Number 4, December 2005, pp. 411-423.

Dissertation Title: Incorporating Spatial Characteristics in Travel Demand Models

Chair, Dr. Hualiang (Harry) Teng, Ph. D.
Committee Member, Prof. Edward Neumann, Ph. D., P.E.
Committee Member, Dr. Mohamed Kaseko, Ph. D.
Committee Member, Dr. David James, Ph. D., P. E.
Graduate Faculty Representative, Dr. Sandra Catlin, Ph. D.