EXPLORING SPATIAL VARIATIONS IN THE RELATIONSHIP BETWEEN NATIONAL PARK VISITATION AND ASSOCIATED FACTORS IN TEXAS COUNTIES

A Dissertation

by

KYUNGHEE LEE

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Michael A. Schuett
Committee Members,	Scott Shafer
	Wm. Alex McIntosh
	Douglas Wunneburger
Head of Department,	Gary Ellis

December 2013

Major Subject: Recreation, Park, and Tourism Sciences

ABSTRACT

Recreation demand such as national park visitation is influenced by various social, demographic, and economic factors. These key variables are important indicators in predicting future trends and provide beneficial information about potential park visitors for managers and planners. As parks and protected areas become impacted by socio-economic changes, it is important to understand the relationship between specific factors of recreation participation and national park visitation. From a practitioner perspective, recreation agencies require multi-scale levels of information in order to address visitor and facility needs. While site-based research or using disaggregated models are helpful to satisfy specific purposes for a park, they often do not provide this information in spatially distributed data on a statewide or regional level. Recreation planners and managers need recreation demand forecasts at levels of spatial aggregations.

This study tried to identify the spatial relationships between national park visitation and its associated factors using large aggregated data. Guided by the idea of opportunity theory and Pigram's conceptual framework, this study empirically investigated what and how factors associated with national park visitation influence demand within the Texas boundary. Specifically, this study developed a spatial regression model of national park visitation demand in Texas using Geographically Weighted Regression (GWR). This model estimated the strength of the relationship between visitation and selected demographic, socioeconomic and situational factors. Methodologically, traditional regression models (e.g., OLS) yield only a single estimate

in a relationship. In comparison, GWR allows an estimate of the spatial variation of the relationship within the study area. Several private and public data sources were used in the model to create reliably aggregated data. Several explanatory variables, e.g., poverty rate, family structures, recreation-related spending patterns and level of education, were hypothesized to influence the level of national park visitation for spatially varying relationships across the study area. From a methodological perspective, this study found interesting methodological implications (e.g., rethinking the traditional regression model for recreation demand estimation) and the potential associated with the use of spatial statistics to analyze the relationships between recreation participation and societal factors. This research demonstrated the importance of including spatial variables as part of recreation demand analysis. Relatively little work has used spatial models in the field of recreation. The results of this study demonstrate the usefulness of spatial analysis for detecting various relationships within the state over traditional statistical analysis.

ACKNOWLEDGEMENTS

Over the past five years I have received support and encouragement from a great number of people. Dr. Michael Schuett, my advisor, is the first person I have to thank. This dissertation would not be possible without his patient guidance and consistent support. He is always there when I need help. He showed me what is the real meaning of "a cold head and a warm heart". I am forever grateful for his mentorship.

I also want to express my gratitude to my committee members, Dr. Shafer, Dr. McIntosh, and Dr. Wunneburger for their thoughts and comments throughout the research and writing process.

I would like to give my deepest thanks to my parents, Jung-Soo Han and Eui-gon Lee, for always standing by me and supporting my graduate studies. Finally, I want to thank my brother, Taehee Lee, for his help in my study.

NOMENCLATURE

ACS	American Community Survey
AIC	Akaike Information Criterion
ESDA	Exploratory Spatial Data Analysis
ESRI	Environmental Systems Research Institute
GIS	Geographical Information Systems
GWR	Geographically Weighted Regression
LISA	Local Indicators of Spatial Association
MAUP	Modifiable Areal Unit Problem
MPI	Market Potential Index
NPS	National Park Service
OLS	Ordinary Least Squares
RUM	Random Utility Model
VIF	Variance Inflation Factor
WUI	Wildland-Urban Interface

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

The USDA Forest Service recently published a national study titled “Outdoor Recreation Trends and Futures” (Cordell, 2012), which shows that Americans' current choices for outdoor recreation differ remarkably from those made by previous generations of Americans. Participation in "traditional" activities such as hunting and fishing is in decline, while participation in activities which involve bird watching and nature viewing is growing. The study compared, by activity, the percentage of participants in each demographic strata to the percentages of participants in each the regional strata for the different activity groups in the United States. This study reveals that there are variations in recreation participation among the different regions and among the demographic groups. The comparisons show that, visitation of recreation and historic sites was highest in the North Region and lowest in the South. Participation in hunting and fishing activities was highest in the South and Rocky Mountain Regions and lowest in the North and Pacific Coast Regions. These differences in recreation preferences demonstrate potential differences in recreation that are present based on regional and demographic factors.

In order to respond to the varied demands for recreation across the country, it is important that recreation planners understand recreation demand and its associated factors (Wisconsin State, 2006). There is, however, no consensus in the field of recreation planning as to the most appropriate methods for measuring current and future demand for outdoor recreation resources and facilities (Outdoor Recreation in Florida, 2008). Specifically, no commonly accepted methods exist for determining, on a statewide or broad regional basis, the amount of outdoor recreation in which a person would participate under certain social conditions. Several factors, relating to demographic, societal, and economic trends, affect recreation participation (Bowker et al., 2012; Zarnoch, et al., 2010). Because of the varied approaches to understanding recreation participation and demand, it is difficult to determine the most effective way to plan for recreation needs. To better understand the demand for recreation, the aim of this dissertation is to suggest an approach to better understand (estimate) the relationship between specific recreation demand (e.g., national park visitation) and its related social factors. This alternative method considers a larger scale perspective by using spatial analysis and statistical techniques through Geographical Information Systems (GIS).

Previous recreation demand models

Many studies in outdoor recreation have tried to identify the relationships between social variables and recreation demand. The first and most straight forward form of research into the social aspects of outdoor recreation is the measurement of recreation activity and its demand[1] (Manning, 2011, p.23). Researchers have adopted a variety of methods for describing, analyzing and predicting recreation demand including

[1] Recreation demand defined as the estimated number of people who are projected to participate in a particular recreation opportunity at some predetermined future time and location (Haas et al., 2007).

the demand for national park visitation. Recreation demand models can be classified as disaggregated models or aggregated (zonal) models from a spatial context (Cui, 2010).

Phaneuf and Smith (2005) summarized the history of recreation demand estimation in three stages. The first stage of models can be divided into two types: 1) travel cost demand models estimated with zonal data, i.e., aggregate visit rates from population zones at varying distances from recreation sites, and 2) activity participation models that are best interpreted as reduced form models. The first stage of recreation demand studies focused on the difficulties caused by using aggregate data, without specific socioeconomic information about the recreationists involved. Specifically, Clawson (1959) mentioned that

> "The methodological problem is admittedly difficult, but a more basic difficulty is the nature of the demand for outdoor recreation. This is dynamic and changing; the future may be very different from the past, and we know relatively little about the past. No methodology can yield wholly satisfactory answers when the problem is so difficult and the data poor" (p.141).

Clawson (1959) notes the difficulty of combining disparate data and obtaining sufficient socioeconomic and demographic information about the participants in outdoor recreation and related regions. The second stages of recreation demand studies during the 1970's focused primarily on disaggregate demand estimations. Burt and Brewer (1971) introduced the first application of the travel cost method to micro data, estimating a system of demand equations for water-based recreation. Their study, which began in the second stage of research, shifted attention to the opportunity cost of travel time, role of substitute sites, trip length, and site attributes in recreation demand. The random utility model (RUM) has been used to estimate recreation demand since the third and contemporary stage. The RUM describes the decision process associated with individuals selecting, from a number of alternatives, which recreation site they will visit. This model was developed by McFadden (1974) and was first applied to recreation valuation by Hanemann (1978, 1984, 1999). The RUM has been used for valuing aspects of recreation by several investigators, and has been regarded as a theoretically consistent method for resolving the mixed discrete/continuous choice problem.

Limitations of disaggregated demand estimation and its implementation

The mainstream of recreation demand studies have focused on disaggregated demand estimation. This method is used to valuate natural resources as non-market products related to specific recreation sites from the benefit-cost implication in the second stage (Cui, 2010). Despite the value of the disaggregated model, many studies have shown that this type of model has various limitations (Feather et al., 1995; Garber-Yonts, 2005, Marvasti, 2012). First, the disaggregated model requires information from users of recreation sites through on-site surveys. Second, it is difficult to get high quality data (reliable and validated data) at specific sites and times. Surveying must be conducted during all seasons and at different time periods (e.g., week vs. weekend or day time vs. night time). In addition, surveys need to cover all combinations of uses for the site. Third, most individual-level models based on cross sectional data are one-shot pictures of the variables and cannot provide information over time (Légaré & Haider, 2008, Schuett et al., 2008). Fourth, disaggregated data based on individual level surveys is not compatible with secondary data which can be only used in an aggregate form for reflecting social trends. For example, census data is reliable secondary data that includes demographic and socioeconomic information. Census data is zone-based aggregated data (e.g., county, census block, census track, etc.) so it is very difficult to combine with non-spatial individual-level data.

Further, recreation agencies require multi-scale levels of information for conducting recreation planning. When recreation planners or managers have applied estimated recreation demand in the real world, there have been challenges and limitations. Many recreation-related plans are devoid of recreation demand information or the demand information is too narrow to be useful (Haas, 2007, p.1). Previous literature focused heavily on local or site-level recreation demand estimation so it is hard to make regional level or state level planning alternatives (Hall & Page, 2006; Smith, 2005). Garber-Yonts (2005) mentions that the reliability of extending demand estimation from individual sites to larger spatial regions is problematic:

> "That is, no single measurement of demand can address the full range of issues confronting recreation resource use and policy. However, a more integrated method of measuring demand at the scale of concern to national forest planners and managers would help them anticipate the response of potential users to management changes and assist in better coordination between forests at regional levels" (pp. 6-7).

The problem described by Garber-Yonts suggests that broad regional or statewide recreation demand estimation is needed. The regional recreation demand model should cover all regional boundaries but also provide enough information about recreation demand within the region. In order to effectively gather and synthesize regional information to estimate demand, the use of GIS mapping and analysis will be beneficial in illustrating this type of model.

Advantage of GIS in the regional planning process

A spatial context is important in understanding variations of recreation demand within regional boundaries. The definition of spatial analysis is "to describe mathematical methods that use locational information to better understand processes generating observed attribute values" (Fotheringham & Rogerson, 2009, p.1). Common software tools for spatial analysis are geographic information systems (GIS). GIS refers to computer-based databases used to collect, store, analyze, integrate, model in layers and visualize spatial data, therefore they are connected to discrete locations (Grimshaw, 2000; Longley et al., 2008). In this sense, GIS methods have attempted to link some of the technical and policy gaps at "the source of the business–social–environmental divide by providing a platform to combine information and knowledge measured from different sources and pertaining to social, economic and ecological domains"(Tremblay, 2005, p.163). GIS is a useful tool for describing whole areas and the variation of sub-areas within specific boundaries by combining disparate data.

GIS has been widely applied by researchers and practitioners in recreation management and planning (Lee et al. 2003; Leung et al., 2002; Morse et al., 2009; Prato & Fagre, 2012; Tomczyk, 2011). The following diagram (See Figure 1) shows the role of GIS in the planning process:

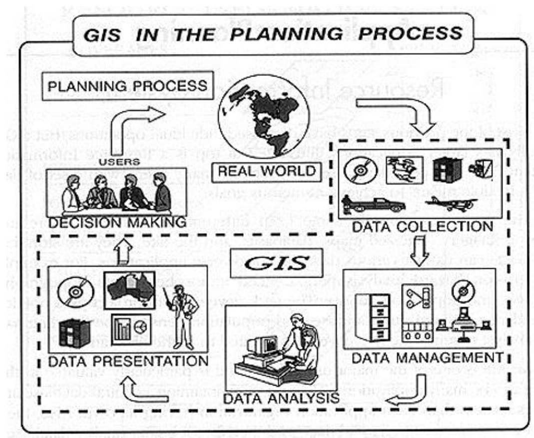


Figure 1 GIS: a visual approach

Source: Davis (2001)

A powerful function of GIS is demonstrating spatial relations between disparate data, such as private marketing data, census data, and socio-economic data. GIS maps can simultaneously and holistically show the distribution of recreation resources and recreation participation activity patterns along with other information. This multi-dimensional perspective allows researchers and practitioners to test and describe complex interactions and relationships. The wide implementation of GIS as a primary decision-making tool has been adopted among recreation resource planners and managers (Nicholls & Shafer, 2001). Along with computer technology development, spatial statistics provide sophisticated and detailed results of analysis with controlling spatial interactions. Spatial statistics were designed specifically for use with spatial data. These statistical techniques use area, length, proximity, direction, orientation, or some notion of how the features in a dataset interact with each other.

Spatial statistics have been used in past research because traditional statistics cannot control spatial relationships and interactions. When a traditional regression model such as ordinary least squares (OLS) is applied in the analysis of spatial data, it sometimes violates the assumptions of regression models (e.g., residuals are uncorrelated with each other).

Social and physical phenomena are often highly clustered within the region (e.g., regional voting patterns, racial segregation, and the poverty belt). However, spatial relationships often are ignored. The following section explains the limitation when a non-spatial regression model (traditional statistic method) is used in regional spatial analysis.

Limitations of using OLS (global) regression model in regional spatial analysis

Recreation demand should be determined from a regional perspective. The earliest recreation demand models relied on ordinary least squares (OLS) with aggregate zonal data (Phaneuf & Smith, 2005). Normally, the effects investigated with OLS regression are fixed effects and pertain to the relationships between the dependent and independent variables based on the whole analytic area. However, if this relationship is different in one part of the study area than it is in another, we have a form of *spatial heterogeneity* where the effects are geographically conditional.

Spatial heterogeneity offers another way to disaggregate regression results so they are formed to be more revealing in differing regions. On the other hand, it resists standard regression assumptions of constant variance across the domain of analysis.

Spatial dependency should also be considered within a regression framework. Spatial dependency is "the propensity for nearby locations to influence each other and to possess similar attributes" (Goodchild, 1992, p.33). One of the major effects of spatial dependency is that regression residuals in the model exhibit spatial autocorrelation (Anselin, 1993; Cliff & Ord, 1981; Odland, 1988). The problem is that previous recreation demand models usually depend on ordinary least squares (OLS) to analyze aggregate zonal data. However, the violation of regression assumptions has been overlooked from a spatial context. Important regression assumptions include (1) constant variance, (2) independence, and (3) normality (Griffith & Layne, 2000). A number of issues may arise if these assumptions are not met. For example, heteroskedasticity in the residuals causes estimates of the regression coefficients to be less precise. Spatial autocorrelation in the residuals results in an underestimation of the standard error of the estimates of the regression coefficients. It also results in a bias towards rejecting the null hypothesis that the value of the coefficient is zero. Non-normality of the residuals compromises interpretability of significance tests of the regression coefficients. Finally, multicollinearity results in overestimates of the variances of the regression coefficients.

Without controlling limitations of traditional statistical methods, current recreation demand models do not provide enough detailed information to the practitioners (e.g., varying spatial relationship between national park visitation and its societal factors). To solve this problem, we need to expand the traditional regression model to a spatial regression model with the help of geographically weighted regression.

Advantage of using GWR in regional spatial analysis

Geographically Weighted Regression (GWR) is useful for conducting spatial analysis. Known as GWR, geographically weighted regression is a tool for dealing with spatial heterogeneity. Exploratory Spatial Data Analysis (ESDA) can provide estimates of regression coefficients for each geographical location based on a weighting of other observations near that location. This approach to spatial analysis was developed by Brunsdon, Fotheringham and Charlton (1996, 1997, 1999)[2]. Recently, GWR has been frequently used to detect the relationships between the spatial distributions of crime and its related factors to find the cause of crime spatial distributions. For example, Lu and Tang (2011) found that many social factors have some influence on the spatial distribution of theft crime rates. The GWR model is indeed able to analyze the varying relationship between crime rates and its related factors from the spatial context (e.g., a negative correlation between population density and theft crime rate; negative and positive correlation between theft crime rate and average land price depending on the areas). In another use of the GWR model, Troy (2012) and his colleagues suggested that there is slight geographic variability in the relationships between crime and trees. They found that that a few isolated areas showed a positive relationship between crime and trees in the Baltimore region.

[2] GWR is more fully explained in Fotheringham, Charlton and Brunsdon (2002).

Although the advantages of GWR have been demonstrated in a broad regional analysis, few recreation studies have used GWR to detect the relationship between recreation demand and its associated factors on a regional or state scale. Specifically, national park visitation and demand has not been studied at the regional level. In this context, this study will use the visitation demand of the national parks in each of the 254 Texas counties as a dependent variable. Spatial analysis and statistics will then be used to show varying relationships between national parks visitation and social factors within the state of Texas.

National park visitation and societal changes

National parks are primarily established to protect the natural and cultural environment; they are also expected to provide all citizens with equal opportunities to experience these resources. Some researchers suggest that additional wilderness and public lands should be designated to meet increasing demand (e.g., Gartner & Lime 2000). An interesting question is whose demand is actually met when decisions about the provision of recreation opportunities are made? The awareness of recreation areas and the opportunities and inclination to use them vary among different segments of society. According to Eagles (2004), “Park use and park management are reflections of societies’ ideas and culture” (p.18) Because American society is so diverse, it can be expected that park use will vary across the population. Demographics in the United States are changing rapidly. The Census Bureau reported that the median age of Americans is now 37.2, with seven states recording a median age of 40 or older. The report also shows that the male population grew 9.9 percent between 2000 and 2010, while the female population grew 9.5 percent. Of the total 2010 Census population, 157 million people were female (50.8 percent) and 151.8 million were male (49.2 percent). The Hispanic population increased by 15.2 million between 2000 and 2010 and accounted for more than half of the total U.S. population increase of 27.3 million. Between 2000 and 2010, the Hispanic population grew by 43 percent, or four times the nation's 9.7 percent growth rate (Census, 2010).

Dramatic racial transformations of the past decade have accelerated economic polarization, leading to the increased isolation of minorities in central cities, growing minority unemployment, and other forms of economic dislocation (Eitzen et al., 2011, p.208). Demographics, economics, technology and changing lifestyles all play important roles in shaping the future of outdoor recreation and park management (Gartner & Lime, 2000). For example, a recently-published article found that the current economic recession has led to lower visitation at national parks (Poudyal et al., 2012). Although, this research will not identify geographical variations national park visitation, the study will explore how socioeconomic variables are important indicators in predicting future trends of national park visitation and outdoor recreation.

Visitation has fluctuated over the last 15 years, with the highest reported visitation occurring in 1999 at 287 million (NPS Public Use Statistics, 2012). National parks and protected areas exist within a dynamic social and political setting that is sometimes difficult to understand and challenging to predict (Gartner & Lime, 2000).

Understanding what and how societal factors impact national park visitation is a key issue for determining future national park use and predicting demand.

Nature of problem

Previous approaches to outdoor recreation planning have been site-based focusing on the balance between conservation and visitor experiences. However, recreation resource planning studies have not addressed how a specific recreation setting might contribute to regional recreation participation and demand or how recreational opportunities can be linked to larger systems (The Society of Outdoor Recreation Professionals, 2012). In the past, national and state based population surveys have simply tested the relationship between national park visitation and societal factors. As Manning (2011) points out, "the percentage of respondents who had visited or more national parks rose progressively with increased socioeconomic status" (pp.36-37). However, these kinds of comprehensive survey results do not provide realistic alternatives for recreation planning nor do they provide enough detail as to why visitation fluctuations may be occurring (e.g., which area or clusters shows lower national park visitation; why this kind of spatial patterns exist; what factors affect those kinds of patterns).

Local communities contain recreation resources such as national parks, state parks, forests, watersheds, and open spaces from a regional approach (Allmendinger, 2009; McLoughlin, 1969). These recreational resources are part of a larger system for local community recreational use. However, few studies have examined different approaches to recreation demand estimation using county level data. From the practitioner perspective, recreation agencies require multi-scale levels of information for the purpose of siting facilities. While site based research or disaggregated models are helpful to satisfy a specific purpose at a park, it does not provide all the available information needed which is spatially distributed over several regions or an entire state.

Previous research (Bowker et al., 1999; Bowker et al., 2006; Leeworthy et al., 2005) has established that factors including race, ethnicity, gender, age, income, and supply or proximity to settings, affect outdoor recreation participation as well as participation intensity and consumption. Reliable information about these factors is often available from external sources, such as the U.S. Census or other research efforts aimed at modeling and simulating exogenous variables into the future. Such information is thus available long before recreation survey results can be obtained.

In this context, recreation planners and managers need recreation demand forecasts at different levels of spatial aggregations within a regional system. Recreation trends over time are important indicators of what may happen with outdoor recreation in the near future (Hall et al., 2009). However, simple descriptive statistics or trends do not formally explore underlying factors and associations which may be driving these trends.

Purpose of the study

The purpose of this study is to develop a spatial regression model to improve the ability to estimate the relationship between recreation demand and its associated factors at the county level in Texas. Among various recreation activities, this dissertation will use *national park visitation* to show how to build the estimate model of recreation demand. Specifically, key variables that affect national park visitation will be tested by OLS regression model (cf. population shifts do not always translate into changes in outdoor recreation participation). In addition, this dissertation will test regional variations of the relationship between national park visitation and factors associated with visitation using GWR. In order to explore this line of research, the following research questions are proposed:

Research questions

1) Will national park visitation vary spatially with visitation rates at the county level? Do spatial patterns of national park visitation show clustered patterns within the state?
2) What and how do socio-demographic and economic factors (e.g., age, gender, income) influence the level of national park visitation?
3) What and how do situational factors (e.g., urbanization, population growth) influence the level of national park visitation?
4) Does the GWR model improve prediction when compared with the OLS regression model?
5) Is there spatial heterogeneity based on the relationship between the aforementioned factors and national park visitation among Texas counties? How does GWR reflect this kind of spatial relationship?

CHAPTER II

LITERATURE REVIEW

Concept of recreation demand

Several theories have been proposed to help explain participation in outdoor recreation. Socioeconomic demographic theory and Opportunity theory both emerged in the 1970s and have been used to explain specific causes of participation or nonparticipation.

Socioeconomic demographic theory suggests that individuals with similar socioeconomic status are more likely to participate in similar outdoor recreation activities (Karlis & Dawson, 1995; McDonald & Hutchison, 1987).

Opportunity theory also has been frequently tested in the field of outdoor recreation. Opportunity theory postulates that “all things being equal, individuals from different segments of society have the propensity to participate in outdoor recreation activities” (Romsa & Hoffman, 1980, p.322). Specifically, opportunity theory attributes ethnic variations in recreation participation to differences in the cost and proximity to outdoor recreation facilities (McDonald & Hutchison, 1987; Hung, 2003). In this sense, expensive transportation costs, entrance fees, and far distance to the recreation resources have been regarded as a barrier to lower income groups within opportunity theory. For example, Lee et al. (2001) insisted that opportunity theory “tends to focus on locational distribution of recreation resources rather than on income itself as factor affecting outdoor recreation participation” (p. 429).

In addition to demographic and socioeconomic factors, individual preference also influence recreation participation, and may explain differences in participation across and within groups.

Recreation demand is generally defined as based on an individual's preference or desires, whether or not the individual has the economic, physical and other resources necessary for their satisfaction (Driver & Brown, 1978). In spite of the importance of motivation, it is apparent that no single theory or clear consensus exists in relation to recreation. Instead, "in theories of motivation need is seen as a force within the individual to gain satisfactions and completeness" (Hall & Page, 2006, p.50). According to Torkildsen (1992), "There appear to be many levels and types of need, including the important needs of self-actualization and psychological growth" (p.86). Understanding of the needs and desires implicit in studies of recreational motivation may offer a range of insights into why people engage in recreational activities. Not only is it necessary to understand why people engage in recreation, but also what factors or barriers may inhibit them from participating. Torkildsen (1992) summarized the influences on leisure participation into three categories: 1) personal; 2) social and circumstantial; 3) opportunity factors. These influences are also beneficial in understanding some of the constraints on recreation participation. The ambiguous "nature of leisure and the capricious characteristics of recreation decisions make generalization and prediction" about leisure choices and demand more challenging (Pigram & Jenkins, 2006, p.18). Further, national parks and protected areas exist within a dynamic social situation that is difficult to understand and challenging to predict (Eagles, 2004).

Pigram (1983) simplified the representative factors which affect the decision to participate in recreation activities (Figure 2). The variables affecting recreation demand can be categorized into demographic, socioeconomic and situational characteristics which generate a tendency to recreate, and the external factors which facilitate or constrain recreation participation. It is important to understand the relationship between specific factors of recreation participation and park visitation. An exploration into Pigram's conceptual framework will help explain this relationship more completely.

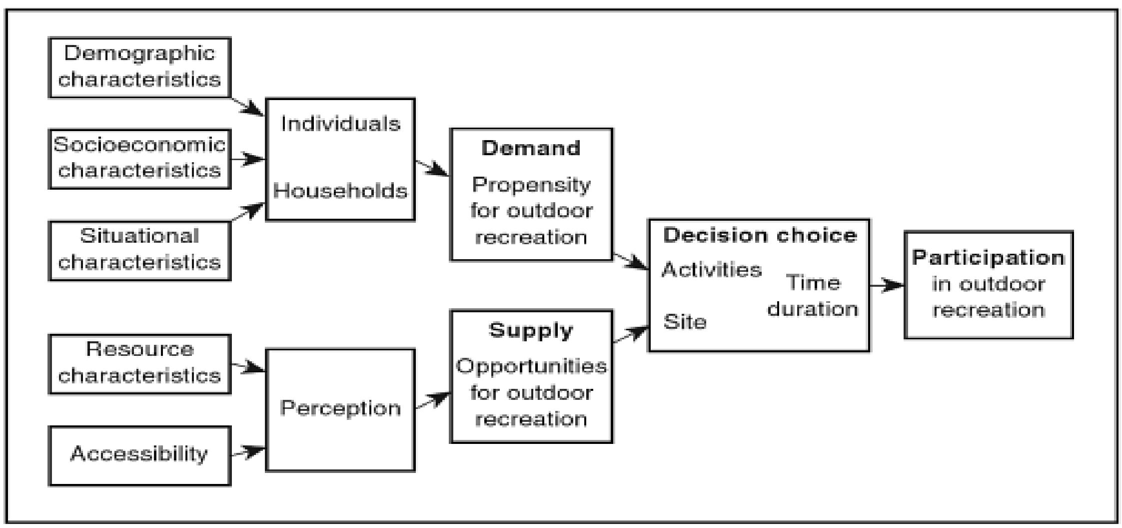


Figure 2 The decision process in outdoor recreation

Source: Pigram (1983)

Demographic factors affecting recreation participation and park visitation

Gender

According to the recent American Time Use survey (2011), 20% of men and 17% of women were likely to participate in sports, exercise, or recreation on any given day. Research suggests that men report higher levels of physical activity than do women, and are more likely to participate in mobile activities such as sports and walking, either alone or with friends. Women were more likely than men to see the forest environment as threatening. Additionally, women expressed a preference for spending time in developed settings with park manager present, whereas men preferred to spend time in remote, natural settings (Virden & Walker, 1999).

Early outdoor recreation studies suggest that "women participate in fewer recreation activities than men and are less likely to participate in gender inappropriate activities (i.e., hunting) but are equally as likely to participate in other outdoor recreation activities" (Manning, 2011, p. 53). In particular, clear gender differences exist in terms of amount and type of leisure constraints, with women experiencing more and different constraints than men (Shaw & Henderson, 2005). Additionally, women are primarily constrained by time, stress, and lack of time for self (Shaw & Henderson, 2005).

Race/ ethnicity

In 2000, the Hispanic population became the largest minority group[3] in the U.S., surpassing the Black population. The median household income and educational attainment of Hispanics and of Blacks is much lower than that of Whites. Though minority populations are growing, people of color are significantly underrepresented in outdoor recreation (McKinney, 1999). According to Jensen and Guthrie (2006) "Outdoor recreation has been for those with more money and a better educational background. Some race/ethnic minority groups (Hispanic and Black) make substantially less money, and are less likely to finish high school" (p.66). Although, many studies have suggested that the lower participation rate of minorities in outdoor recreation is the result of lower levels of income and educational attainment, other contributing factors should be considered as well.

Recently, Bustam and his colleagues (2011) found that demographic differences within a race/ethnic group were related to recreation constraint differences. The Black cohort showed significant differences in perceived constraints across income groups. This finding supports past research which found that constraints to recreation are prominent among racial and ethnic minorities with low socio-economic status (Shores et al., 2007). Hispanics revealed significant differences in perceived constraints across age and income. Specifically, those with the lowest annual income perceived greater intrapersonal constraints compared to other income groups (Bustam et al., 2011).

[3] A minority group is one who, because of its race or ethnicity, experiences a wide range of discriminatory treatment and is assigned to a low status position in the broader society (Yetman, 1999).

Further, Gómez (2008) suggests a direction for additional research regarding racial/ethnic differences in outdoor recreation. He insists that geographically scattered research needs to be done in different geographic areas and on a broader scale to determine the relationship between diverse racial/ethnic groups and recreation participation:

> "The focus has traditionally been on national parks, regional recreation areas, or areas predominantly west of the Appalachians. Most African American studies have been conducted in the Great Lakes Region or in the U.S. Southeast; most studies on Latinos in the U.S. Southwest; Very little has been done in the larger metropolitan areas, in general (with Chicago and Los Angeles being the exception), …. Additionally, the Latino/Hispanic population grew faster from 1990 through 2000 in much of the U.S. South than in other areas of the United States (Kochlar et al. 2005), and these new Latinos have distinctive characteristics (e.g., they tend to be male, unmarried, born abroad, and young), different from their predecessors in traditional settlement areas" (p.78).

Related to park visitation, current and estimated future changes in the United States population suggest that racial and ethnic minority groups will continue to grow rapidly. Minority groups, however, continue to be underrepresented in visitation to parks (Whiting et al., 2012). Research published in 1994 indicates that racial/ethnic minorities were largely absent among visitors to national parks (Goldsmith, 1994), and that 90% of national park visitor groups were Whites of European descent (Floyd, 1999).

Contemporary research still stresses the issue of disparity in national park visitation. Non-Hispanic Whites comprised 78% of park visitors in 2008–2009; by comparison, Hispanics accounted for 9% of visitors (The National Park System Comprehensive Survey, 2011).

Furthermore, Le (2012) found that Hispanics and women visited parks in larger groups, came from areas closer to the national park and spent more time at the parks than did Whites. These results support differences in meanings and preferences among different ethnicity/race groups. Ethnic variations in park preferences and use have implications for park and recreation planners and managers. Planners and managers need to consider these kinds of variations in their services based upon the ethnic/racial characteristics of potential users.

Age

While limited research has investigated the influence of age on national park visitation, appreciable studies have been conducted to measure the effects of age on leisure participation. Even though there are variations according to one's, "…income level, personality, interest, health condition, ability level, transportation, education level and a number of social characteristics" (Hayslip & Panek 1989, p.425), results usually insist that leisure participation is directly and negatively correlated to aging. For example, while time constraints, often related to family obligations are constraining to the young, poor health and lack of transportation constrain older individuals (Jackson & Henderson, 1995; Scott & Munson, 1994; Shaw et al., 1991). Kelly (1980) insisted a negative correlation between recreation participation and age. Other research asserts that as people age there is some evidence that they participate more frequently in appreciative and learning activities and less frequently in active expressive activities (Foot, 1990). In summary, age has an important influence on recreation participation but its effect will, "vary depending on the person, the opportunities and the type of activity" (Torkildsen, 2005, p.109).

Family structure

Family is a major context in which leisure can be studied (Shaw, 1994). The basic assumption has shown that family structure and parental influence significantly affect children's leisure participation. However, relatively little work has been done to examine family structure as it relates to recreational choices, patterns, and barriers (Blanco, 2009).

In recent years, a dramatic shift has taken place in the composition of the American family (McKinney, 1999). American society increasingly has become diversified in family structure. Today, 30% of families are headed by a single-parent and 42% are dual-worker families, replacing both the traditional two-parent, male-worker model (King et al., 2010). The changing structure of the American family has provided outdoor recreation and tourism managers and planners with challenges. One recently published article examined the relationships between different family-structures and level of park use (See Fan et al., 2012). In this research, family-structure differences in the variables (traditional family structure vs. non-traditional) were examined using multivariate regression analyses. Results show that , "working single parents and dual-worker parents have lower levels of park use than parents in two-parent, single-worker families" (p.520).

Overall, family structure has influenced demand for recreation participation and park visitation in several ways; however, few studies have examined this relationship. Examining different family structures can help explain and predict trends in recreation participation and its demand.

Socioeconomic factors

Income, poverty, and recreation-related spending

Many studies emphasized lack of money as a barrier to engaging in recreational activities (Argyle, 1996). Scott and Munson (1994) found income was the single best predictor of perceived constraints to park visitation among the socio-demographic variables tested. Sessoms (1993) insisted that individuals with lower incomes have fewer recreation opportunities due to relatively expensive cost of leisure participation. This kind of argument was empirically tested by previous researches. For example, Chubb and Chubb found that "the poor do not have the recreation rooms, landscaped backyards, automobiles, recreation environments of those with higher incomes" (1981, p.94). Further, the previous literatures insisted that low income limits the "expression of tastes" (Howard & Crompton, 1980). Recently, McNiel (2011) indicated that median household income, education levels, and percent of poverty are the best predictors of park visitation through OLS regression analysis. The above tenet frequently assumes that the poor and some minority groups largely lack the monetary ability to travel to national parks (Bultena & Field, 1978, 1980).

Entrance fees on public lands and protected areas have been an argumentative topic (Buckley, 2003; Burns & Grafe, 2006; Crompton, 2002; Kyle et al., 2002; Ostergren et al., 2005; Schwartz & Lin, 2006) and affect the participation of some groups in recreation. Another related factor which affects participation is recreation related spending. Socio-economic information obtained from outdoor recreation participants are an increasingly important component of recreation resource management and its planning. People are motivated to participate in consumptive outdoor recreation for a variety of reasons. An experience-based recreation management model proposes that motivation to participate in a recreational activity is a function of two expectations: (1) that effort (e.g., purchasing equipment and licenses, driving to the site) will lead to participation; and (2) that notion that participation will lead to positive psychological outcomes (Haas, 2001; Driver & Rosenthal 1983; Manfredo et al., 1983). In this sense, people who conduct research with the intent to purchase recreation related equipment would lead to higher rates of participation in recreation activities.

Education

Several research studies have shown that education is heavily related to recreation participation. Specifically, previous literatures found that there is a positive relationship between recreation participation and higher education level (Lucas, 1990; Lee et al., 2001, Kelly, 1996; Scott & Munson, 1994). Some of these study findings indicated that individuals who have higher educational attainment more likely to participate in outdoor recreation activities because, higher education level and income are positively related with each other (Lee et al., 2001; Scott and Munson, 1994). Along with this, Kelly (1983) regarded education attainment as the most significant predictor of recreation participation rather than occupation or income.

Recently, Li and colleagues (2010) noted that "individuals with higher educational levels know more about how to access information regarding leisure activities and better understand the associated benefits; that knowledge may further increase the probability of participating in leisure activities" (p.72). These studies generally insisted that higher education level contributes to participation in outdoor recreation activities. However, a few studies in outdoor recreation fields have focused on education as a significant variable affecting park visitation constrains (Zanon et al., 2013).

Occupation

Reid (1981) defined social class as "a grouping of people into categories on the basis of occupation." (p.6). Because of the interrelationship between social status and income, education level, it is usually regarded that "social class, as determined by occupation, is the most influential factor in determining recreational participation" (Torkildsen, 2005, p.109). In England, the General Household Surveys (2003) found that "Generally, it was professional workers who tended to have the highest participation rates in leisure activities and unskilled workers who had the lowest rates. Moreover, the surveys conclude that the middle classes are not only more active culturally, socially and intellectually, but they also play more sport and travel more widely" (as cited in Torkildsen, 2005, p.110).

Several studies report that occupational demand predetermined workers' choice of recreational activities. The studies also argue that the advancement in technology throughout the world has made many workers sedentary and devoid of physical fitness. Additionally, workers who engage in more physical labor tend not to engage in recreation, but instead opt to sleep in their spare time (Alla, 2001). Burton and Turrell (2000) found that those in blue-collar occupations were roughly 50% more likely than white collar occupations to be classified as insufficiently active outside of work hours.

Further, Warnick and his colleagues (2010) found that U.S national park visitation rates were not equally distributed among occupational groups of adults. In this article, the professional occupation group inversely showed a declining national park visitation rate (negative 2.1 percent per year from 2000 to 2008). In sum, occupation is not the only factor that influences recreation participation, rather it is associated more closely with other socioeconomic factors.

Situational factors

Urbanization

Many regions in North America have experienced dispersed land development patterns and suburban housing growth, typically referred as urban sprawl (Rodrigue, 2006). Urban sprawl has widespread ramifications for habitat conservation and human safety. Sprawl expands land development toward suburban and rural territories, expanding the wildland-urban interface (WUI) (Radeloff et al., 2005; Stewart et al., 2007) and enhancing human/wildlife conflict (Johnson, 2001; Hussain et al., 2007). In addition, urban sprawl affects hunting and other outdoor recreation opportunities. Urban sprawl changes the socioeconomic, demographic and cultural characteristics of rural communities (Katz, 2002) and introduces modern indoor recreational opportunities that can eventually replace traditional outdoor activities (Brown et al., 2000). Of course, relocation and urbanization may not always lead to decreasing recreation, as participants might move to areas with increased opportunities for recreation, perhaps due to more conducive climate or more sites for preferred activities, or the novelty of new activities might attract participants to recreate in a manner new to them. While most research findings indicate that rural residency is a primary indicator of nature based recreation participation (e.g., hunting), Stedman and Heberlein (2001) found that a rural upbringing is only one of several significant variables that contribute to outdoor recreation participation. Marcouiller and his colleagues (2009), found that "wide variation exists in the type of outdoor recreation provided throughout the state and results suggest that recreation preference is strongly influenced by key spatially distinct differences in population" such as a function of a local region's level of urbanization (p.104). Although recreation researchers generally have regarded urbanization as a barrier to recreation participation, it is necessary to consider when examining recreation participation from a regional perspective.

Distance (accessibility)

The broad notion of recreation demand is supply-independent (Phaneuf & Smith, 2005). It assumes no constraints on recreation opportunities or access to them. However, actual outdoor recreation participation is greatly influenced by a function of the supply of the opportunities (e.g., the ease of access, residential area of visitors, capacity of facilities, etc.). For example, participants in outdoor recreation more focus on "locational distribution of recreation resources rather than on income itself as factor affecting outdoor recreation participation" (Lee et al., p.429). Further some of studies concluded that the distance between residential area and location of recreation facilities affect outdoor recreation participation (Craig, 1972; Hung, 2003).

Tourism researchers also have examined that integrating various forms of spatial factors (e.g., proximity) into destination choice models can improve the predictability of a model (Fesenmaier, 1988; Lin & Morais, 2008; Murphy & Keller, 1990; Prideaux, 2000).

The distance decay effect in "the gravity model used to explore the aggregate choice behavior was referred to as the frictional effect on forecasting tourist demands or spatial interactions" (Lin & Morais, 2008, p.385). As showed in the figure 3 below, the distance decay effect curve shows that there is a specific distance beyond which visitors would be "unwilling to take a pleasure vacation in order to avoid the escalation of vacation costs" (Lin & Morais, 2008, p.385; Cook & McCleary, 1983).

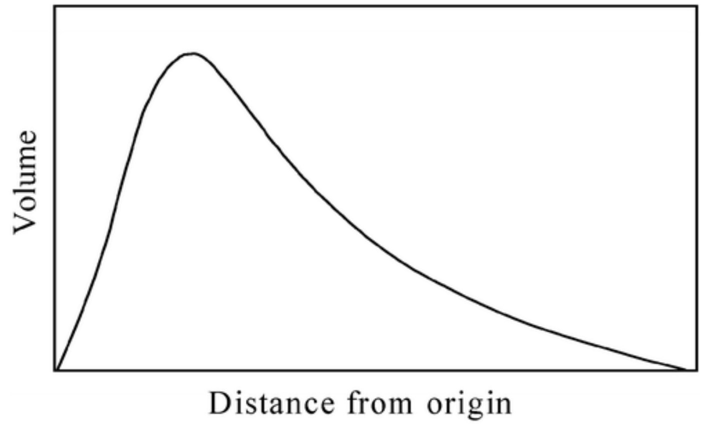


Figure 3 Theoretical distance decay curve

Source: McKercher & Lew (2003)

Furthermore, Kim and Fesenmaier (1990) examined the effects of the spatial structure of recreation opportunity on recreation travel. They insisted that recreation activity increases as a function of the number of nearby facilities located within 25 miles. They also insisted that "recreation facilities exhibit both competitive and agglomerative relationships with other nearby recreation opportunities" (p.1).

In sum, the literature reviewed makes note that people with higher socioeconomic status are more likely to participate in recreation activities and visit national parks rather than people with lower socioeconomic status. As you see the above, the previous studies insisted that various socioeconomic factors, lower educational attainment, limited financial resource and a lack of accessibility, have led people with lower socioeconomic status to less visit national parks. Although there might be overlap between approaches, it is argued that no single approach to assessment is sufficient, and combining multiple theories is recommended for a comprehensive assessment.

This study will use various data in the analysis to identify the relationship between national park visitation and its associated variables from the literature. Global regression models (e.g., OLS) have some limitations in regard to regional data therefore the following section illustrates these limitations when applied to regional analysis using aggregated data.

Limitations of regional analysis using aggregated data: Simpson's paradox & modifiable areal unit problem (MAUP)

From the methodological perspective, studies in the social sciences generally use global statistics to estimate/predict socioeconomic trends by using aggregated data. However, this approach has some problems when used in data analysis on a regional scale. According to Simpson' Paradox, the extension of global estimates of relationships could show biased interpretations of local relationships. Simpson's Paradox indicates that the reversal of results when groups of data are analyzed separately and then combined (Simpson, 1951). The following figures show the relationship between the participation rate and the poverty rate of the area visitors came from (See Figure 4 & 5). There were different interpretations between using spatially aggregated data and disaggregated data. These data are separated by location and in both locations the relationship between participation rate and poverty rate is a negative one. For both individual locations, there is a negative relationship between poverty rate and participation rate, but aggregated date from the two locations shows the positive relationship. Simpson's Paradox stresses the risk of analyzing aggregate data. While it is generally demonstrated in non-spatial data sets “where the aggregation is over population subgroups, the paradox applies equally to spatial data where the aggregation is over locations” (Fotheringham et al., 2002, p.8).

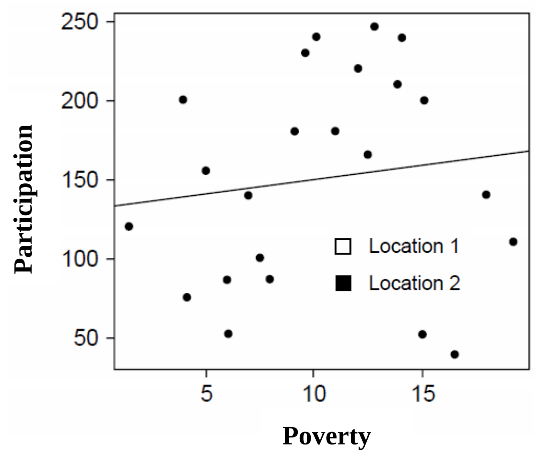


Figure 4 Spatially aggregated data

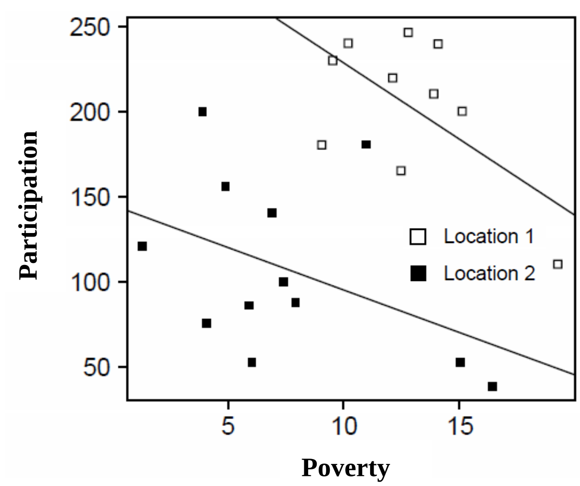


Figure 5 Spatially disaggregated data

Modified in Fotheringham et al (2002)

Extended to Simpson's Paradox, researchers also consider when the estimates at one level of aggregation are different from the estimates obtained at a different level of aggregation (Openshaw & Taylor, 1979). For instance, geographers have long been intrigued with the "modifiable areal unit problem" (MAUP), a problem that is isomorphic to the statistical inference problem. The different scale related problems have been identified in the analysis of spatially aggregated data (Fotheringham et al., 2002, p.144). The MAUP "arises from the fact that areal units are usually arbitrarily determined and modifiable, in the sense that they can be aggregated to form units of different sizes or spatial arrangements" (Johnston and Semple, 1983; Jelinski & Wu, 1996, p.130). Thus the MAUP has two related but distinctive components: the scale problem and the zoning (or aggregation) problem (Openshaw, 1984). Openshaw and Taylor (1979) explained the scale problem as "the variation in results that may be obtained when the same areal data are combined into sets of increasingly larger areal units of analysis," and the zoning problem as "variations in results due to alternative units of analysis where n, the number of units, is constant" (p.128). Specifically, Tita and Radil (2010, p.473) suggested the potential solutions as "increasing the aggregation of units by increasing the area covered by the units decreases the variance in the data between the units" for the scale problem, and "rezoning the areas contained by each unit while holding the total number of units the same can impact both the mean and variance of any measured data" for the zoning problem. The scale and zoning effects (problems) are related to the changing definition of areal units (Openshaw, 1984; Wong & Lee, 2005).

To improve accuracy and prediction of the model, it is necessary to control these problems by reflecting spatial variations. The advantage of using GWR is that it can be used to look for localized exceptions or deviations from regional trends. This method is useful to reduce Simpson's paradox and MAUP.

A large-scale of recreation demand estimation usually has been based on general population surveys. For example, data from National Survey on Recreation and the Environment (NSRE) has been used to predict demand for outdoor recreation (Bowker et al., 2006, 2007). These predictions continuously show positive correlations between recreation demand and income, education, and gender and negative correlations with age, residence (urban), race (black), and ethnicity (Hispanic). Though these estimations address regional, state, or national trends of recreation demands, they do not provide detailed information at sub-areas (e.g., county level).

As noted earlier, a majority of recreation demand studies have been developed at a site-level or unit-level. If recreation managers or planners need more sophisticated information (e.g., varying relationship between social factors and recreation demand at county level), current recreation demand models cannot satisfy this need. From a methodological perspective, the OLS model has been utilized to estimate the relationship between recreation demand and its related social factors in a large spatial boundary.

When relationships are consistent across a study area, the OLS regression equation models the relationships well. However, when the relationships are shown differently in part of the whole study area, the regression equation produces an average model of present relationships. When the relationships represent two extremes, the global average will not model either relationship very well. Because of this limitation, OLS cannot provide enough information for outdoor recreation planning.

The following chapter introduces data that was used for the study and its source. The specific methods and model specifications will then be described.

CHAPTER III

METHOD

Social scientists, including outdoor recreation researchers, have tried to find a universal theory or trend to determine the factors which dictate recreation participation. However they have long been faced with a difficult question and a potential dilemma: Is there any universal theory that governs social processes, and if there is not, does a quantitative approach have any validity? (Fotheringham et al., 2002, p.9). For example, some researchers have found that the recent economic recession caused lower visitation at national parks. However, spatial variation always exists beyond the trend (e.g., Yosemite National Park visitation has steadily increased since 2006). A traditional regression technique such as ordinary least squares (OLS) has been a popular statistical method in social science because it simplifies complex social situations and helps find relationships between cause and effect. Though the nature of social science is to find a general theory or trend to explain some phenomenon, spatial outliers or exceptions always exist. OLS can hide important spatial variations in the model parameters and are not able to deal with spatial autocorrelations existing in the variables. In this study, a recently developed technique, geographically weighted regression (GWR), was used to examine the varying relationships between national park visitation and factors related to the visitation at a county level within the state of Texas.

Description of the study area

From 2000 to 2010, Texas experienced a population increase of nearly 4.3 million people. This was the largest population growth in any decade in Texas history and the largest numerical increase of any state in the nation. Additionally, the population in Texas is becoming increasingly diverse ethnically, racially, and in age (U.S. Census, 2010). According to the 2010 Census, since 2000 the U.S. Hispanic population has increased 43% and now makes up 16.3% of the total U.S population. The White population increased by 464,032, but decreased from 52.4% of the total population in 2000 to 45.3% of the total population in 2010. The Black population was about 11.5 % of the population in 2010 but accounted for only 12.2% of the growth from 2000 to 2010; the Asian population was 3.8% of the population but accounted for 9.2% of the 2000-2010 growth. The combination of all other population groups accounted for 1.8% of the population but accounted for only 2.8% of the total population increase from 2000 to 2010.

Notably, Hispanic (of any race) populations now make up 38% of the Texas population. These 9.5 million Hispanic residents represent 19 percent of the total Hispanic population in United States (See Figure 6 below). According to Texas State Data Center (2008), by 2040 more than 53% of the Texas population will be Hispanic. Given this growth, it is important to understand the outdoor recreation trends of the Hispanic population.

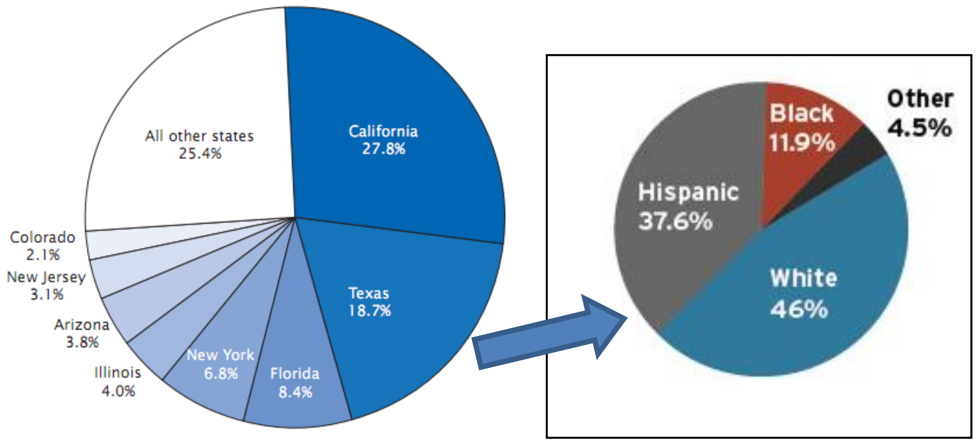


Figure 6 Percent distribution of the Hispanic population by state (Census, 2010)

Along with diverse population shifts, there may be increased demand for a variety of recreation opportunities that reflect these changes.

Texas is divided into 254 counties. This is the highest number of counties in any state. Although the growth in Texas was extensive it was not uniformly distributed across the state as 79 of the state's 254 counties lost population from 2000 to 2010. The counties which experienced a population decrease were concentrated in West Texas and the Panhandle. Based on these population shifts, there may be a spatial variation in recreation demands due to population changes at a county level.

Variable selection

Dependent variable

In any statistical analysis, a critical step is variable selection. The market potential index (MPI) from the Environmental Systems Research Institute (ESRI) was chosen as a dependent variable to represent the national park visitation ratio at a county level. The MPI describes the expected number of consumers in a given area compared with the U.S. national average. The information used to derive the MPI index is usually collected through market segmentation cluster analytic techniques. These methods integrate information such as census data, purchasing and consumptive behaviors to describe the lifestyle preferences, spending habits, and favorite brands and products of neighborhood residents. Although the specific manner in which the classifications are constructed is proprietary, the methodological approach is similar and is offered by several companies, such as ESRI, Claritas, and Experian. These systems enable businesses to exploit on possible markets that might be missed when solely relying upon census statistics (See Kwate et al., 2012). This information can be useful in identifying locations for new stores, selecting merchandise that matches consumer preferences, and targeting advertising with the right message.

ESRI's retail demand is calculated as (ESRI, 2009):

$$\text{Expected number of consumers} = \sum_{n=1}^{n=65}(\text{countn} \times \text{consumption rate})$$

Where n is a count of community tapestry segments (descriptions of the lifestyle preferences and spending habits of neighborhood residents) and the consumption rate is for the unique segments that characterize a neighborhood:

$$\text{Local consumption} = \text{expected consumers} / \text{base count}$$

$$\text{MPI} = (\text{local consumption} / \text{US consumption}) \times 100$$

In this formulation, 100 indicates the U.S. national average; values above or below 100 indicate retail demand that is higher or lower than the national average, respectively. Thus, a value of 70 would indicate that the retail demand was 30% lower than the national average, and a value of 200 would indicate that the demand was two times as high as the national average.

MPI consists of seven categories with different types of market potential data. Sports and leisure market potential data was used to build the model. Sports and leisure market potential measures the relative likelihood of adults or households in a specified area to participate in activities. These data are then compared to the U.S. average. Although MPI is useful to compare a national visitation averages to local averages, this study used the percentage of national park visitors (who visited national parks while on a domestic vacation within the past12 months) in each county to detect the spatial pattern of the dependent variable and the relationship with its associated societal factors within a Texas region.

Independent variables

Various public and private levels of secondary data as independent variables were transformed into spatial data while preserving their attributes. Disparate secondary data were manipulated in an appropriated form to estimate and predict national park visitation. There are many sources of demographic and socioeconomic data that can provide information for this research including U.S Census Bureau, Texas State Data Center, ESRI, and Community health status indicators. Decennial census data (2010) was used to show the general demographic information of each county. The American Community Survey (ACS, 2011) was used to obtain median household income, family structure, age distribution, and education attainment at the county level. The ACS is an ongoing statistical survey conducted by the U.S. Census Bureau, and provides communities with current information useful for planning investments and services. More specifically, the survey helps determine how more than $400 billion in federal and states funds will be distributed each year. Conducted every year and sent to approximately three million addresses a year, the survey asks for various demographic, economic, social, and housing information such as age, sex, race/ethnic background, place of birth, family status, income, benefits, health insurance, education, veteran status, and disabilities. Aside from the decennial census, the ACS is the second largest survey conducted by the U.S. Census (ACS, 2011). In this research, ESRI consumer spending data was used to detect spending patterns within specific counties including average recreation expenditure. It was necessary to match the above data sources with MPI data within each county boundary.

Table 1 shows the independent variables and sources used in this study:

Table 1 Independent variable & its source

Independent variables		
Demographic Factors		
Total population	The number of population	Decennial Census, 2010
Age	Percentage of the population over age 65	American Community Survey, 2006-2010.
Gender	Percentage of the male population in 2010	Decennial Census, 2010
Race	Racial Segregation Index	Census Scope, 2010
	Proportion of Hispanic or Latino	Decennial Census, 2010
	Proportion of White	Decennial Census, 2010
	Proportion of Black or African American	Decennial Census, 2010
Family Structure	Householder who lives with parent(s)	Decennial Census, 2010
	Householder who lives with children	Decennial Census, 2010
	Householder who lives with grandchildren	Decennial Census, 2010
	Householder who living alone (nonfamily household)	Decennial Census, 2010
	Average household size	Decennial Census, 2010
Socioeconomic Factors		
Income	Median Family Income	American Community Survey, 2006-2010.
	Median house Income	American Community Survey, 2006-2010.
Poverty rate	Percentage of individuals living in poverty	Community health status indicators, 2009
Occupation	Unemployment rate	American Community Survey, 2006-2010.

Table 1 Continued

Socioeconomic Factors		
Education Attainment	Proportion of the population over 25 years old with less than a high school education	American Community Survey, 2006-2010.
	Proportion of the population over 25 years old with a high school education or more	American Community Survey, 2006-2010.
	Proportion of the population over 25 years old with less than Bachelor's degree	American Community Survey, 2006-2010.
	Proportion of the population over 25 years old with Bachelor's degree or higher	American Community Survey, 2006-2010.
Spending Pattern	Amount spent on high end sports/recreation equipment in the past12months: $250 +	Consumer Spending, 2011
	Recreation expenditure: average spending to purchase recreation equipment	Consumer Spending, 2011
	Average annual travel expenditure	Consumer Spending, 2011
Situational Factors		
Urbanization		
Population density	Population per square mile	Us Census Bureau, 2010
Urban population	The percentage of urban population	Us Census Bureau, 2010
Urban area	The percentage of urban area	Texas A&M GIS center
Locational attribute		
Distance to national park	Euclidean distance from the centroid of the counties to the centroid of the nearest national park or national forest	ESRI, 2009

Analysis procedure and description

This study used three stages of analysis: the first stage was spatial cluster analysis to detect which areas show clustered patterns based on the dependent variable (national park visitation). The same analysis was employed to determine whether regression residuals are spatially autocorrelated. The second stage adopted the OLS model to test what and how factors affect national park visitation with the variance inflation factor (VIF). In the third stage of analysis, GWR was used to analyze the variations of relationships between national park visitation and its associated factors from a regional scale. Various spatial and non-spatial statistical software programs were employed to manipulate data, construct models and visualize results: SPSS 21 was used for non-spatial analyses (e.g., correlation analysis) and combining various secondary datasets based on geocodes of Texas counties. Spatial autocorrelation (global and local) tests were performed in Geoda 1.4.1. Spatial relationships and outputs of GWR were analyzed and mapped in ArcGIS 10.0 and GWR 4.0.

Spatial autocorrelation and spatial cluster analysis

In 1970 the famous geographer Tobler identified the first law of geography: “Everything is related to everything else, but near things are more related to each other” (Tobler, 1970, p.236). Spatial cluster analysis, based on the above tenet, is the reason why researchers use spatial data when they analyze data using spatial autocorrelation. Specifically, the structure of spatial data consists of integrating map data and attribute data. Map data contains the location and shape of geographical features. Attribute data is the descriptive data that geographic information systems linked to map features. Although inferential tests are standard practice in much of science, they are problematic for geographic or spatial data. First, the samples in standard statistical inference are obtained independently, while geographic datasets have a fixed location in the population. Further, researchers regard datasets as a sample of a larger area; however, the concept of sampling, which is the basis for statistical inference, does not transfer easily to the spatial context (Lee & Wong, 2005). Along this line, detecting spatial autocorrelation, which is the correlation of a variable with itself through space, means that the spatial distribution of a variable shows existing systematic patterns between one area and its neighboring areas (Anselin, 1995). If nearby or neighboring areas are more alike, this is a positive spatial autocorrelation (spatial dependence) while a negative autocorrelation describes patterns in which neighboring areas are unlike (spatial heterogeneity). Thus, the objective of spatial cluster analysis is to detect the non-randomness of spatial patterns or the existence of spatial autocorrelation (Anselin, 1994).

There are several indicators that exist to measure spatial autocorrelation. The previous studies (Ding & Fotheringham, 1992; Bao et al. 1995; Barkley et al 1995) have been conducted based on the standard global and local spatial statistics such as the Moran I, Geary C (Cliff & Ord 1973, 1981), G statistics (Getis, 1992) and LISA (Anselin, 1995). In this study, Moran's I was used as an indicator to identify which kind of spatial autocorrelation exists among adjacent counties around Texas national park units. This indicator has been used extensively in spatial analysis: "Moran's I is one of the oldest indicators of spatial autocorrelation and has remained a de facto standard" (Lai et al., 2009, p.58). The Moran's I value ranges from -1 (indicating perfect dispersion) to 1 (perfect correlation). A zero value indicates a random spatial pattern (Figure 7).

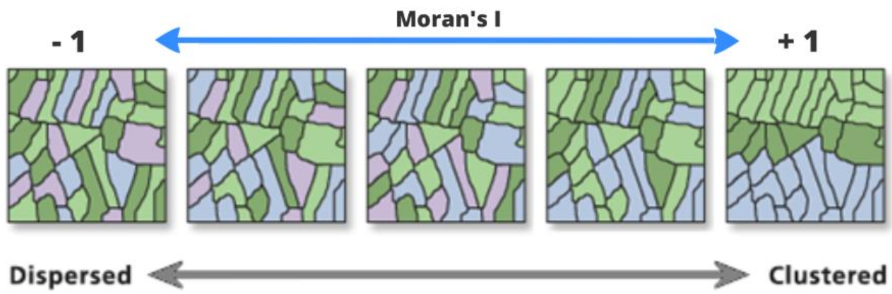


Figure 7 Spatial dependence vs. spatial heterogeneity

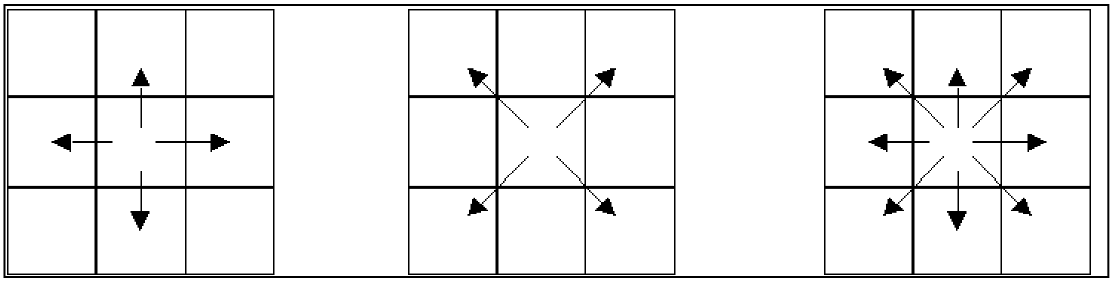
Source: www.esri.com

There are two different types of Moran's I estimation: 1) global Moran's I and 2) local Moran's I. Global Moran's I is used to measure the degree of overall clustering tendency over the whole study area. Local Moran's I called, the Local Indicators of Spatial Association (LISA) assesses significant local spatial clustering around an individual location (Anselin, 1995). Indicators based on the exclusive measurement of an entire data set (global view) are usually accurate judgments for the existence of clustering through whole data distribution. However, these indicators provide little information about where clusters exist. Indicators can show local patterns and measure local instabilities (local view) to identify specific clusters existing in data sets. The generalized form for the local Moran's I can be defined as follows (Anselin, 1995):

$$I_i(d) = Z_i \sum_{j \neq i}^{n} w_{ij} Z_j$$

where the observations Z_i and Z_j are in standardized form (with a mean of zero and variance of one). The spatial weight W_{ij} is in row-standardized form. So, Ii is a product of Z_i and the average of the observations in the surrounding locations. In a conditional permutation approach, a test can be conducted on the null hypothesis that all values are randomly distributed over space. A pseudo-significance level of the Ii may be obtained by a "conditional" randomization or permutation approach (Anselin, 1995). The observed value of Z_i at location i is held fixed and the remaining values are randomly permuted over all the locations in equal probability.

In actual computation, each resample data set can be selected from the population randomly without replacement. The significance level p-value can be obtained by calculating the proportion of data permutations in the data sets that have emulated Ii greater than (or less than) or equal to the actual Ii. Since Z_i is fixed in each emulated Ii, the computation of the significance level p-value can be simplified by calculating the proportion of data permutations in the data sets that have an emulated average greater than, less than, or equal to the actual average of the observations surrounding location i. The experimental p-value provides the basis for a test on the null hypothesis that the average of the observations surrounding location i is in no extremes (i.e., all values are randomly distributed over the space). In spatial autocorrelation analysis some measure of contiguity is required (a standard of defining neighborhoods). Most analyses in spatial autocorrelation adhere to a common definition of neighborhood relations. Explicitly, neighborhood relations are defined as either rooks case, bishops case or queens case (Figure 8).



Rook Bishops Queen

Figure 8 Defining neighborhoods

Source: Dubin (2008)

This study adopted queen contiguity as a standard for defining neighborhood because counties show spatial interactions within the vertex area.

OLS regression model with spatial data (global regression model)

Regression encompasses a wide range of methods for modeling the relationship between a dependent variable and a set of one or more independent variables. A regression model is expressed as an equation:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad \text{for } i=1 \ldots n$$

In this equation y_i is the response variable, here measured at some location i, x_i is the independent variable, e_i is the error term, and β_0 and β_1 are coefficient which are to be estimated such that the $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ ninimized over the n observations in the data set. The term $(y_i - \hat{y}_i)$ is known as the residual for the *i*th observation, and the residuals should be both independent and drawn identically from a normal distribution with a mean of zero. Such a model is usually fitted using a procedure known as Ordinary Least Squares (OLS).

There are a number of assumptions underlying the OLS regression model, one of which is that the observations should be independent of one another. This is not always the case with data for spatial units and Tobler's first law of geography. Not only might the variables in the model exhibit spatial dependence (that is, nearby locations will have similar values) but also the model's residuals might exhibit spatial dependence (Fox, 1997). The latter characteristic can be observed if the residuals from the basic regression are plotted on a map where the residuals in neighboring spatial units commonly have a similar magnitude and sign. These characteristics of spatial data have implications for the estimates of the parameters in the OLS model. If there is spatial structure in the residuals from the model, this will lead to inefficient estimates of the parameters, which in turn means that the standard errors of the parameters will be too large (Faraway, 2002). This has implications for inference where potentially significant parameter estimates may not exist. Spatial structure in the data means that the value of the dependent variable in one spatial unit is affected by the independent variables in nearby units. This leads to parameter estimates which are both biased and inefficient. A biased estimate is one that is either too high or too low to be an estimate of the unknown true value. For example, if we want to analyze the relationship between income and recreation demand within Texas by using OLS regression, we would start by using all 254 counties. To demonstrate this, imagine the state of Texas in which y (recreation demand) is regressed on x (income) and the resulting OLS regression model is

$$(1)\ y_i = 3.8 + 4.5x_i$$

However, suppose that one county in the state shows the following relationship

$$(2)\ y_i = 3.8 + 5.2x_i$$

and in another county of the state this relationship is

$$(3)\ y_i = 3.8 + 3.8x_i$$

Then applying Equation (1) to that county of the state in which Equation (2) holds will lead to underestimates of the yi values in that county and positive residuals, assuming xi is positive. Conversely, applying Equation (1) to that county of the region in which Equation (3) holds, will lead to overestimates of the yi values and negative residuals. Hence, there will be a strong positive autocorrelation of the error terms resulting from the inability of the global model to deal with the spatial nonstationarity of the relationships being measured.

For this reason, spatial cluster analysis is utilized to test whether or not residuals from the model have spatial dependency before the GWR analysis. The spatial independency of residuals will be evaluated by the spatial autocorrelation coefficient (Moran's I). The diagnoses of an OLS model will be determined by assessing the multicollinearity of the independent variables. The multicollinearity will be assessed through variance inflation factor (VIF) values, and if "VIFs are greater than 7.5", multicollinearity is indicated (Terrón et al., 2011, p.756).

Geographically Weighted Regression (local regression model)

In OLS, error terms are generally assumed to be independent, normally distributed, random variables with zero means and constant variance. This model, in its unconstrained form, is not implementable for investigating spatial processes because the number of parameters increases with the number of observations. A technique is needed for estimating a parameter "drift" (Leung et al., 2000). Brunsdon et al. (1996, 1997) and Fotheringham et al. (1997a, 1997b) suggest a geographically weighted regression (GWR) technique in which the parameters are estimated by a weighted least squares procedure. GWR allows local rather than global parameters to be estimated (See Table 2).

Table 2 Comparison chart between OLS and GWR

Global (e.g., OLS)	**Local (e.g., GWR)**
Summarize data for whole region	Local disaggregations of global statistics
Single-valued statistics	Multi-valued statistics
Non-mappable	Mappable
GIS-unfrendly	GIS-friendly
Aspatial or spatially limited	Spatial
Emphasizes similarities across space	Emphasizes differences across space
Searches for regularities or laws	Searches for exceptions or local hot-spots

Cited in Fotheringham (2002)

The GWR model captures the heterogenic nature of national park visitation by allowing the equation to "alter over space to reflect the structure within the data" (Brunsdon, Fotheringham, & Charlton 1996, p.281). The typical output from a GWR model is a set of parameters that can be mapped in the geographic space to represent spatial heterogeneity (non-stationarity). This is the essence of GWR in comparison to traditional regression models in a regional analysis. GWR allows the relationships to vary over space (i.e., βs do not need to be the same everywhere). A GWR model is expressed as an equation:

$$y_i = \beta_{i0} + \beta_{i1}x_{1i} + \beta_{i2}x_{2i} + \ldots\ldots + \beta_{in}x_{ni} + \varepsilon_i$$

In the GWR model, βs (coefficient) vary in terms of location (i) instead of remaining the same everywhere. Separate regression is run for each county (observation), using a spatial kernel that centers on a given point and weights observations subject to a distance decay function. Researchers can use fixed size kernel or adaptive kernel to determine the number of local points that will be included in each local regression (See. Figures 9 & 10).

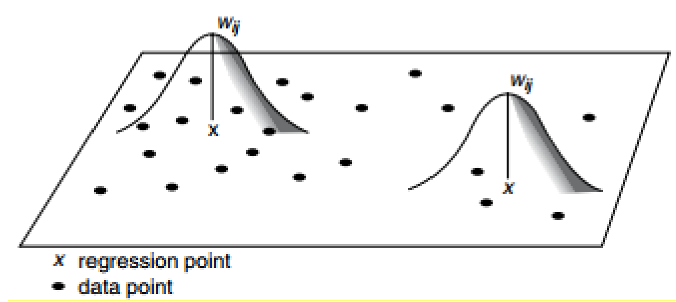


Figure 9 GWR with fixed spatial kernels

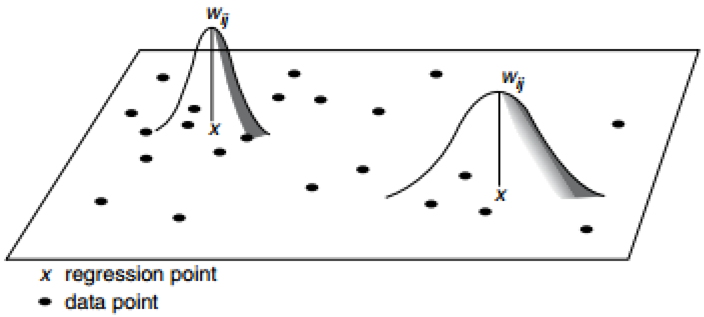


Figure 10 GWR with adaptive spatial kernels

Cited in Fotheringham (2002)

Adaptive kernels are usually used when data are not evenly distributed. Points are weighted based on distance from center of kernel e.g., Gaussian kernel where weighting is given by: $w_i(\mathbf{g}) = \exp[-1/2(d_{ij}/b)^2$ where b is bandwidth.

Bandwidth size

There are various automatic techniques for determining bandwidth (Foody, 2003; Wang et al., 2005). However, an adaptive kernel has been popularly regarded as an appropriated technique to find the optimal bandwidth that minimized the corrected AIC when there were no specific criteria or solid evidences for determining bandwidth (Fotheringham et al., 2008). In this study an AIC approach was adopted, as it offers more flexibility when GWR is used in an exploratory context. Adaptive kernel width is determined through minimization of the AIC. In particular, a bias-corrected AIC statistic is employed, which can be estimated using the following formula:

$$\mathrm{AIC}_c = 2n\log_e(\hat{\sigma}) + n\log_e(2\pi) + n\left\{\frac{n+\mathrm{tr}(\mathbf{S})}{n-2-\mathrm{tr}(\mathbf{S})}\right\}$$

Where σˆ is the maximum likelihood estimate of the standard deviation of the error term. "The rationale for the corrected AIC approach is to provide a bandwidth that avoids the tendency to undersmooth, which is common with (uncorrected) AIC. It is assumed that any bandwidth selected is reasonable (i.e., not too small), even though any spatial autocorrelation in the data is not accounted for in the GWR model defined" (Harris et al., p.289). The smaller the bandwidth, the more variance there is. On the other hand, the lower the bias, the larger the bandwidth, but more bias also reduces the variance. This is because we assume there are many regression coefficients over space and the more it is like a global regression, the more biased it is. AIC minimization provides a way of choosing bandwidth that makes optimal tradeoff s between bias and variance (See Bowman & Azzalini, 1997; Brunsdon et al., 1999).

Prediction in GWR

GWR allows researchers to make predictions. The potential recreation demand model includes the national park visitation ratio of each of the 254 Texas counties as dependent variables and various societal factors as independent variables. The best model for using OLS can be identified by using a multicollinearity test. GWR calibrates all of the local regression equations, one for each feature using national park visitation and its related variables.

The second part of the GWR tool allows people to make predictions based on the calibrated equations (See Figure 11). The dependent variable does not change, and there is an opportunity to provide modified explanatory variables. For instance, the same explanatory variables have been used, but instead of using current demographic information, projected demographic information is used to obtain an estimate of the impact of projected demographic changes. Population projections are useful to outdoor recreation agencies as they try to understand the potential impacts of population growth and demographic change on recreational resources. Population projections help natural resource managers identify future stakeholders; understand preferences for various management actions and people's resource-related attitudes, beliefs, and expectations for management, and plan organizational changes within an agency (Decker et al., 1996).

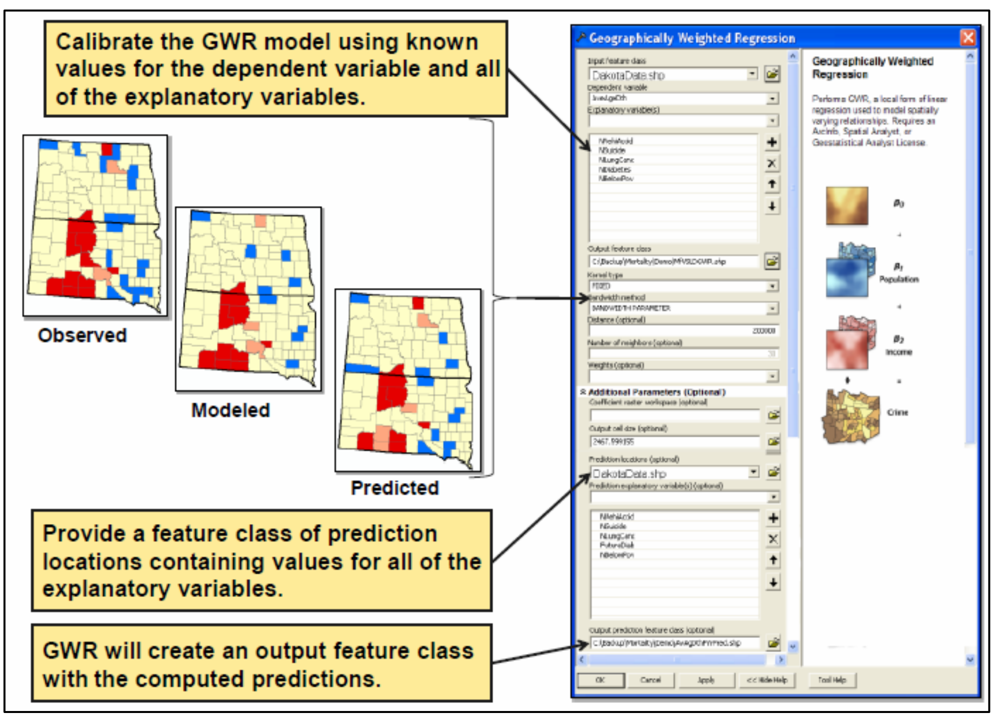


Figure 11 GWR prediction using projected data

Source: ESRI Resource Center, 2012

In summary, this study expected that the GWR regression model would make noticeable improvements of model performance over the OLS regression model, which was proven by the comparison of the models' R-squared and corrected Akaike Information Criterion (AIC). The GWR model would also improve the reliability of the relationships by reducing spatial autocorrelation in three ways: (1) increasing the local predictive power of the regression; (2) reducing autocorrelation in the residuals, and (3) relaxing the assumption of stationary regression coefficients.

CHAPTER IV

RESULTS

The main purpose of this study is to develop a spatial regression model to improve the ability to estimate the relationship between national park visitation and its associated factors (socio-demographic, economic, and situational factors) at the county level in Texas. This dissertation suggests various spatial methodological approaches to examine the effects associated factors have on national park visitation ratios (the visitation per capita) of a specified cause. To accomplish this, an Exploratory Spatial Data Analysis (ESDA) was first employed. ESDA is a set of geographic information system (GIS) spatial statistical techniques that are useful in describing and visualizing the spatial distribution, detecting patterns of hotspots, and suggesting spatial regimes or other forms of spatial heterogeneity (Anselin, 1998).

In the first stage of analysis, a spatial cluster analysis (Moran's I statistic) was used to detect the spatial cluster patterns of national park visitation ratios. There are two different types of spatial autocorrelation (Boots et. al., 1988). Spatial cluster shows a positive spatial autocorrelation when similar values are spatially clustered together. On the opposite side is the distribution of similar values separated or dispersed from each other, which is called negative spatial autocorrelation. Spatial distribution may indicate patterns of an underlying process. Spatial cluster analysis could reveal information about the underlying geographical processes that generate the spatial pattern, which can further aid in the understanding of the underlying geographical processes and their relationships with the social phenomenon under investigation. In the next step, an OLS regression model was employed to explain what kind of associated variables that affect generation of certain spatial cluster patterns for national park visitation ratios.

To select appropriate variables from large spatial databases which influence visitation and represent Pigram's conceptual framework, a stepwise regression and VIF test were used. A non-spatial correlation test and local bivariate correlation test were employed to show how each selected independent variable was related to national park visitation. Specifically bivariate local indicator of spatial association (LISA) maps could visually show the spatial relationships between a dependent variable and selected independent variables. It may be helpful to comprehensively determine how independent variables interact with a dependent variable. After this analysis, a GWR model based on the final OLS model showed how selected variables explained (spatial non-stationary) national park visitation within the state through spatial visualization and statistical tests.

First-stage analyses: Spatial cluster analysis

ESDA provides measures for both global and local spatial autocorrelation, which are fundamental in outlining the spatial inequity. As mentioned in the previous chapter, spatial distributions with values at certain locations showing relationships with values at neighboring locations are called spatial autocorrelation. A global autocorrelation detects this correlation from the general perspective by incorporating all samples. In contrast, a local autocorrelation only focuses on the specific relationships in a particular location over space. The most frequently used measurement of global and local spatial autocorrelation is Moran's I statistic (Cliff & Ord, 1981). To provide a better interpretation of the results, the following figure is presented (Figure 12). Moran's I scatterplot shows the relationship between a variable and the average value of its neighbors for the same variable.

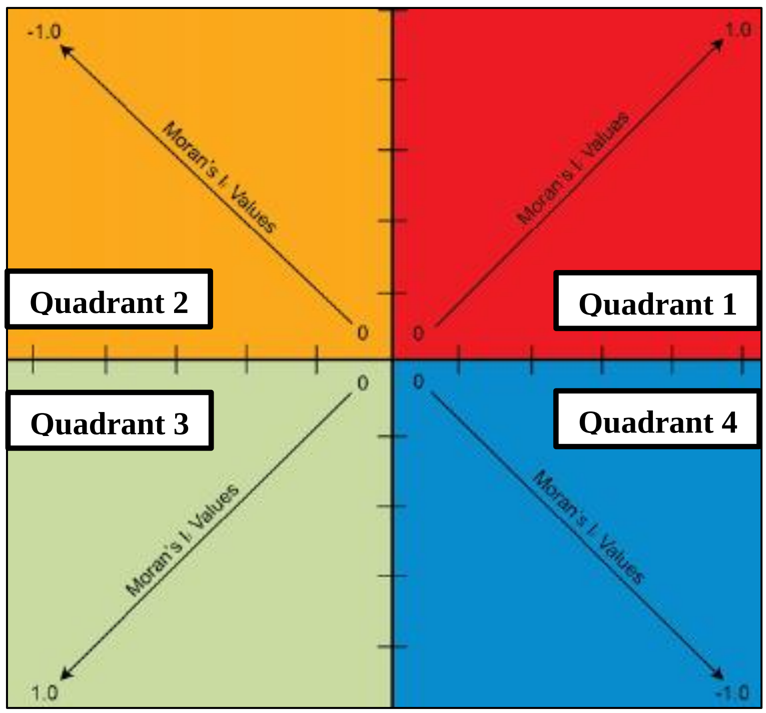


Figure 12 The interpretation of Moran`s I scatter plot

Quadrant 1 (HH) and Quadrant 3 (LL) refer to positive spatial autocorrelation while quadrants LH and HL refer to negative spatial autocorrelation. Hence, through Moran's I scatterplots, the relationships between each county and its neighbors can be identified. Quadrant I (HH) shows the regions in the data set which have variable values above the mean, and the average values of the neighboring variables are also above the mean (called hotspot). Quadrant II (LH) shows the regions which have variable values below the mean and the average of the neighboring variable values are above the mean. Quadrant III (LL) shows the regions which have variable values below the mean and the average of the neighboring variable values are below the mean. Quadrant IV (HL) shows the regions which have variable values above the mean and the average of the neighboring variable values are below the mean.

Global/Local Moran's I test of national park visitation

The first spatial measure used, the Moran's I, is a global measure of spatial autocorrelation that quantifies the degree to which areas are clustered or uniformly distributed overall. The global Moran's I test statistic was used to test for global spatial autocorrelation of national park visitation ratios (adult per capita visitation) across the counties in Texas. The spatial cluster analysis of the national park visitation for each county showed a statically significant value. The global Moran's I test demonstrated that national park visitation ratios were spatially clustered (Figure 13). The global Moran's I value for national park visitation was 0.439 and showed a positive spatial autocorrelation across the state: Moran's I can be interpreted as the correlation between dependent variable (national park visitation ratio) on the horizontal axis and the spatial lag of the variable (average national park visitation ratio in the county's neighbors) on the vertical axis. Both variables are standardized and the graph is divided into four quadrants. The slope of the regression line is Moran's I.

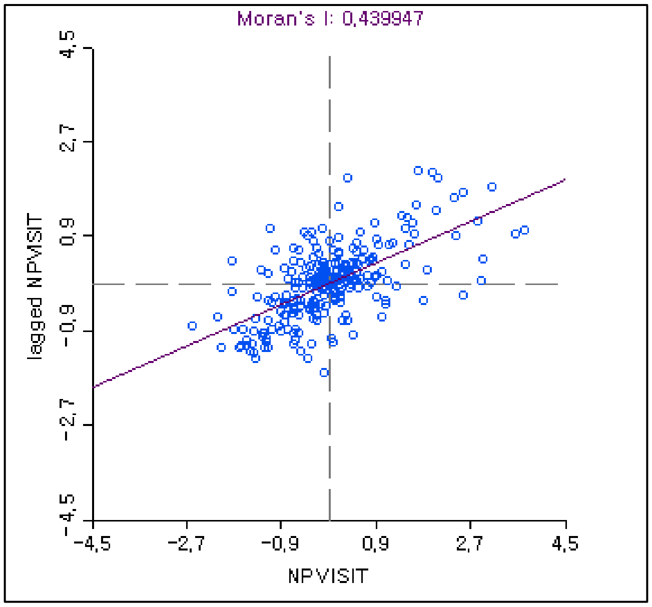


Figure 13 Global Moran's I scatter plot of national park visitation ratio

Furthermore, the mapping of the local Moran's I test visually indicated which counties were clustered and the cluster pattern (See Figure 14).

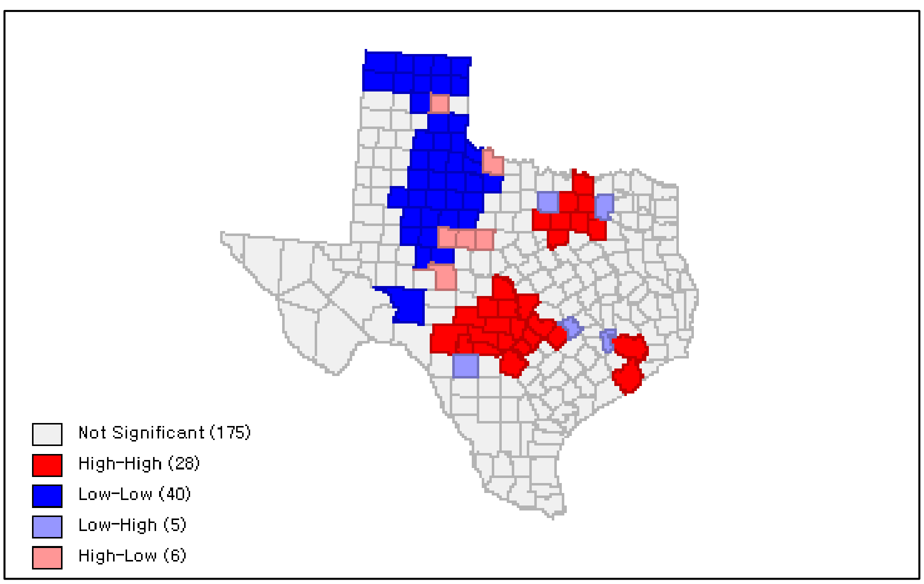


Figure 14 Local Moran's I cluster map of national park visitation ratio

Twenty-eight counties located in the mid and some northeastern Texas regions showed HH cluster patterns. HH signifies that one specific county and its neighboring counties have a higher national park visitation ratio than the mean of the Texas visitation ratio. Conversely, 40 counties located in the northern and northwestern regions showed LL cluster patterns. LL indicates that a single county and its neighboring counties have a lower national park visitation ratio than other areas in Texas. The local Moran's I summary table provides detailed information about the cluster maps (See Appendix A). The results of the local indicator of spatial association (LISA) were statistically significant within a 95% confidence level (See Figure 15).

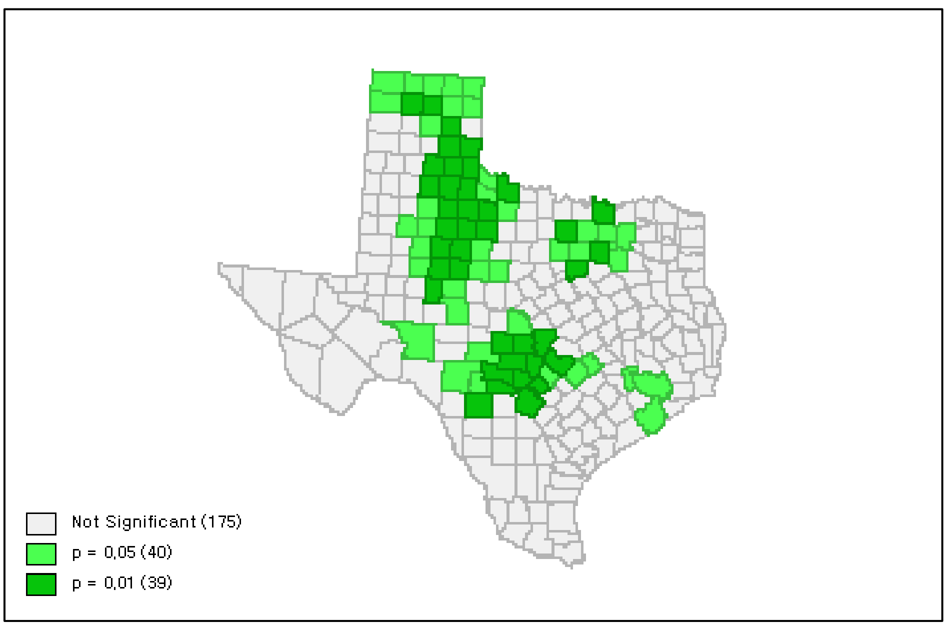


Figure 15 Local Moran's I significance map of national park visitation ratio

The next step examined why these spatial patterns exist and what factors affect these spatial distributions through the use of an OLS regression model.

Second-stage analyses: Explanatory variables selection and OLS analysis

In the OLS regression procedure, variables were carefully selected by controlling for multicollinearity problems among potential independent variables. Fifty-two models were tested based on 27 potential independent variables to get an appropriate model to detect the relationships between national park visitation and its related factors. In the initial step in the analysis, a non-spatial scatter plot matrix was used to identify the correlation among independent variables. To determine the income level, poverty rate, average incomes, and median incomes were considered as representative variables. However, median house income (VIF: 9.995) and median family income (VIF: 10.295) were highly correlated with each other. Also, race/ethnic variables were highly correlated with each other: the proportion of the Hispanic population (VIF: 158.035), the proportion of the white population (VIF: 135.631), and the proportion of the black population (VIF: 16.908093). Because a proportion of three different race/ethnic populations in each Texas county were used, there was a multicollinearity problem. To control this limitation, extended models based on different race/ethnic groups were needed. Conceptually, race/ethnicity is an important variable in understanding national park visitation because they have been shown to influence visitation levels across various groups (Floyd, 1994; Henderson, 1996).

A stepwise regression was chosen because this investigation was exploratory (Stebbins, 2001). Stepwise regression is a semi-automated process of building a model by successively adding or removing variables based on the t-statistics of their estimated coefficients. An OLS stepwise regression was applied to test for significant factors. Result showed that the distance from each county's center to the centroid of the closest national park was not a statistically significant predictor for national park visitation at a 95% confidence level. The population density and log transformed total population were also not statistically significant in several of the models. The previous analysis steps detected global the multicollinearity problem through a VIF test. However, finding the local multicollinearity is more complex and problematic (ESRI, 2012). In this sense, a thematic map for each of the explanatory variables was created and areas with little or no variation in values were used for model construction to reduce the problems when there were local multicollinearity problems in the models. After a stepwise regression, six explanatory variables were selected for the final OLS model (Table 3). These variables should satisfy two criteria. First, the variables represented associated socio-demographic or economic information with national park visitation. Second, these variables showed minimized multicollinearity problems and the best linear unbiased estimator (BLUE). "Best" means giving the lowest variance of the estimate. The OLS estimates β are obtained by minimizing the sum of squared prediction errors, hence, least squares. In order to obtain the BLUE property and make statistical inferences about the population regression coefficients from the estimated b, certain assumptions about the random error of the regression equation need to be made.

As mentioned in the previous chapter, OLS regression assumptions may not always be satisfied in practice[4]. When a value observed in one location depends on the values observed at neighboring locations, there is a spatial dependence. Also spatial data may show spatial dependence in the variables and error terms. With spatial error in OLS regression, the assumption of uncorrelated error terms is violated (Moon & Farmer, 2013). As a result, the estimates are inefficient. Spatial error is indicative of omitted (spatially correlated) covariates that if left unattended would affect inference. In this context, a spatial autocorrelation test of standard residuals was employed.

Table 3 Description of selected explanatory variables

Selected explanatory variables	Description	Expected global relationship
Proportion of high recreation spending	Over $250 spent on high end sports/recreation equipment for 12 months	Positive
Proportion of low educational attainment	Proportion of the population over a certain age with less than a high school education	Negative
Proportion of elderly	Percentage of the population over age 65	Negative
Proportion of those living with parent(s)	Person in a household who lives with parents (at least 2 or more family members)	Positive
Proportion of those living alone	Person in a household who lives alone	Negative
Poverty rate	Percentage of individuals living in poverty	Negative

[4] "Spatial dependence is idea that the spatial dimension of social and economic variables may truly be important aspects of a modeling. Based on the premise that location and distance are important forces at work, regional science theory relies on notions of spatial interaction and diffusion effects, hierarchies of place and spatial spillovers" (Retrieved from http://www.s4.brown.edu/s4/courses/SO261-John/lab9.pdf).

The equation of the final OLS regression model was specified as:

National park visitation (%) = β_1 recreation related spending + β_2 poverty rate + β_3 less a high school education + β_4 65 age over + β_5 Living with parents + β_6 Living alone + e

High recreation related spending, the proportion of elderly, and the proportion of those living with parents positively influenced visitation, while the variables low educational attainment, poverty rate, and the proportion of those living alone negatively affected visitation (See Table 4).

Table 4 Summary of OLS results

Variable	Coefficient	StdError	t-value	Prob	Robust_SE	Robust_t	Robust_Pr	VIF
Intercept	2.5481	0.19703	12.9327	0.0000*	0.2016	12.6389	0.0000*	
High Recreation Spending	0.7058	0.03659	19.2859	0.0000*	0.0426	16.5407	0.0000*	1.2671
Low Education	-0.0152	0.0042	-3.58953	0.0004*	0.0041	-3.6544	0.0003*	2.4437
% of Elderly	0.04229	0.0075	5.57968	0.0000*	0.0073	5.7669	0.0000*	2.4943
Living with parents	0.4945	0.1144	4.3230	0.0000*	0.1148	4.3048	0.0000*	1.4190
Living Alone	-0.0578	0.0167	-3.45589	0.0006*	0.0156	-3.7013	0.0002*	2.8818
Poverty rate	-0.0216	0.0057	-3.78094	0.0002*	0.0052	-4.0856	0.0000*	2.1117

Table 5 Summary of OLS diagnostics

Number of Observations:	254
Number of Variables:	7
Degrees of Freedom:	247
Akaike's Information Criterion (AIC) :	206.396
Multiple R-Squared :	0.732
Adjusted R-Squared :	0.726
Joint F-Statistic : 112.758	Prob(>F), (6,247) degrees of freedom: 0.000*
Joint Wald Statistic : 718.087	Prob(>chi-squared), (6) degrees of freedom: 0.000*
Koenker (BP) Statistic*: 15.445	Prob(>chi-squared), (6) degrees of freedom: 0.017*
Jarque-Bera Statistic**: 0.909	Prob(>chi-squared), (2) degrees of freedom: 0.634

* Significant p-value indicates biased standard errors; used robust estimates.
**Significant p-value indicates residuals deviate from a normal distribution.

The regression analysis included an assessment of the OLS regression diagnostic tests of the Joint Wald Statistic, the Koenker (BP) Statistic for heteroscedasticity, and the Jarque-Bera Statistic for skewness and kurtosis (Rosenshein et al., 2011).

Both the Joint F-Statistic and Joint Wald Statistic are measures of overall model statistical significance. The Joint F-Statistic is trustworthy only when the Koenker (BP) statistic (See Table 5) is not statistically significant. If the Koenker (BP) statistic is significant, the Joint Wald Statistic should be checked to determine overall model significance. The null hypothesis for both of these tests is that the explanatory variables in the model are not effective. A p-value (probability) that is smaller than 0.05 indicates a statistically significant model at a 95% confidence level.

The Koenker (BP) Statistic (Koenker's studentized Bruesch-Pagan statistic) is used to determine if the explanatory variables in the model have a consistent relationship with the dependent variable (what this study is trying to predict/understand) both in geographic space and in data space. When the model is consistent in geographic space, the spatial processes represented by the explanatory variables behave the same everywhere in the study area (the processes are stationary). When the model is consistent in data space, the variation in the relationship between predicted values and each explanatory variable does not change with changes in explanatory variable magnitudes (there is no heteroscedasticity in the model). For example, to predict the relationship between national park visitation demand and the independent variable poverty rate, the model would have problematic heteroscedasticity if the predictions were more accurate for locations with lower poverty rates than they were for locations with higher poverty rates. The null hypothesis for this test is that the model is stationary. A p-value that is smaller than 0.05 indicates statistically significant heteroscedasticity and/or non-stationary processes at a 95% confidence level. When results from this test are statistically significant, the robust coefficient standard errors and probabilities should be checked to assess the effectiveness of each explanatory variable. Regression models with statistically significant non-stationary processes are especially good for GWR analysis. Based on the final OLS model, three additional extended models were constructed to show how race and ethnicity affect visitation. These three extended models are based on different race/ethnicity groups and show a stationary condition through the models (See the following Table). This means that applying a GWR to these models is not appropriate in later analysis. However, the results of these OLS models support previous literature (See Table 6). Higher national park visitation rates were found where there was a greater proportion of a white population (OLS coefficient: 0.0046). The rate of visitation decreases as the proportion of the Hispanic population increases (OLS coefficient: -0.005). However, the proportion of the black population positively affects national park visitation. This result contradicts conclusions of the previous literature (Bowker et al., 2006; Dwyer & Hutchison, 1990, Floyd, 1993; Henderson, 1996; Washburne, 1978).

Table 6 The final OLS model and 3 extended OLS models based on race/ethnic groups

	The final OLS model	Extended OLS model 1 (Hispanic population)	Extended OLS model 2 (White population)	Extended OLS model 3 (Black population)
Intercept	2.548*	2.467*	2.073*	2.441*
High Recreation Spending	0.705*	0.742*	0.730*	0.727*
Low Education	-0.015*	-0.005	-0.008	-0.011*
% of Elderly	0.042*	0.039*	0.034*	0.049*
Living with parents	0.494*	0.471*	0.519*	0.413*
Living Alone	-0.057*	-0.062*	-0.056*	-0.068*
Poverty rate	-0.021*	-0.02*	-0.018*	-0.026*
Race/Ethnicity	N/A	-0.005*	0.005*	0.011*
R – squared	0.732	0.742	0.737	0.740
Adjusted R – squared	0.726	0.735	0.730	0.733
AIC	206.396	199.047	203.934	200.811
Koenker statistics	15.445*	11.377	13.467	11.251
Jarque-Bera statistics	0.909	0.783	1.003	0.807

The Jarque-Bera statistic indicates whether or not the residuals (the observed/known dependent variable values minus the predicted/estimated values) are normally distributed. The null hypothesis for this test is that the residuals are normally distributed and if a histogram of those residuals were constructed, they would resemble the classic bell curve, or Gaussian distribution. When the p-value (probability) for this test is small (e.g., the p-value is smaller than 0.05 at a 95% confidence level), the residuals are not normally distributed, indicating model misspecification (a key variable is missing from the model). Results from a misspecified OLS model are not trustworthy. However, the final models did not show statistically significant values within this test.

Spatial autocorrelation (Moran's I) was again applied to the OLS regression residuals, in order to determine whether the errors in the estimates indicate the presence or absence of spatial clustering of like valued attributes. The OLS regression residuals were examined for spatial autocorrelation regression residuals showing strong spatial autocorrelation (See Figure 16 & Table 7).

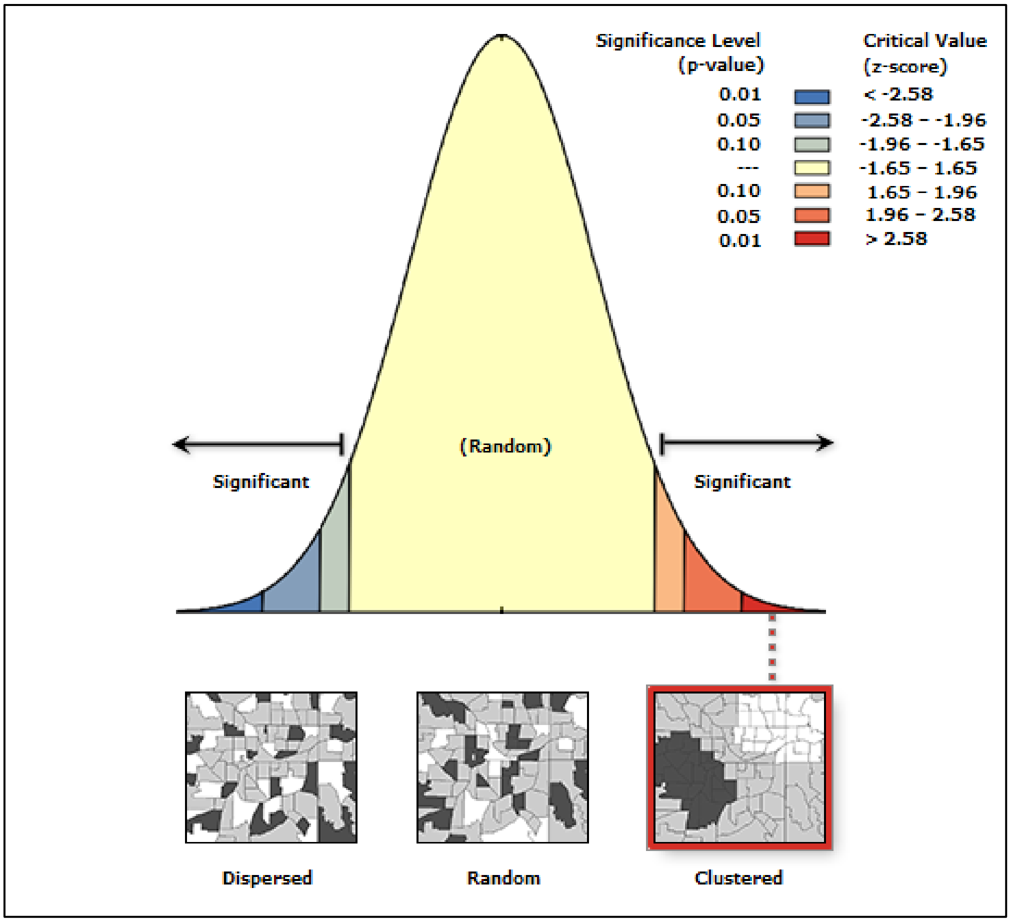


Figure 16 Spatial autocorrelation test of OLS residuals

Table 7 Summary of global Moran’s I test for OLS residuals

Moran's Index:	0.146
Expected Index:	-0.003
Variance:	0.001
z-score:	3.813
p-value:	0.000137

Given the z-score of 3.81, there was less than 1% likelihood that this clustered pattern could be the result of random chance. The results showed that linear regression models may result in flawed statistical inference as there is evidence that the residuals of linear models are spatially correlated across the state. Finally, the Koenker (BP) Statistic also showed statistically significant values and residuals show strong spatial autocorrelation. To address these problems and describe the non-stationary relationship, the OLS regression needed to be expanded through the use of the GWR model.

Also, selected independent variables in the OLS models were tested by non-spatial correlation and local bivariate correlation analysis. Using “NPVISIT” (adult per capita visitation) as the key variable (dependent variable), the following figures show the results of correlation and local bivariate correlation analysis. Spatial correlation with a scatter plot refers to the relationships between selected independent variables and the dependent variable. The local bivariate correlation analysis result indicated autocorrelation between a specified independent variables and the dependent variable.

The local bivariate correlation analysis

1. **Poverty rate and national park visitation**

Figure 17 indicates that a higher poverty rate is negatively related to national park visitation (slope = -0.284). Also, the local bivariate correlation analysis result showed that there was a negative spatial relationship between poverty rate and national park visitation with a Moran’s I of -0.11 (See Figure.18). Specifically, seven counties showed an HH relationship concentrated in the mid region of Texas. This means that one county has a higher poverty rate and its neighbor counties showed a higher average national park visitation ratio rather than the mean of the Texas visitation ratio. Twenty-four counties show an LL relationship focused on the midwestern and northwestern regions of Texas. This finding means that one county has a lower poverty rate and its neighbor counties showed a lower average national park visitation ratio rather than the mean of the Texas visitation ratio. Interestingly, most of the higher national park visitation clustered areas showed an LH relationship (See Figure 19). This means that one county had a lower poverty rate and its neighboring counties showed a higher average national park visitation ratio rather than the mean of the Texas visitation ratio. This result supports the contention that a better economic situation results in higher national park visitation. Results of the analysis were statistically significant within a 95% confidence level (See Figure 20).

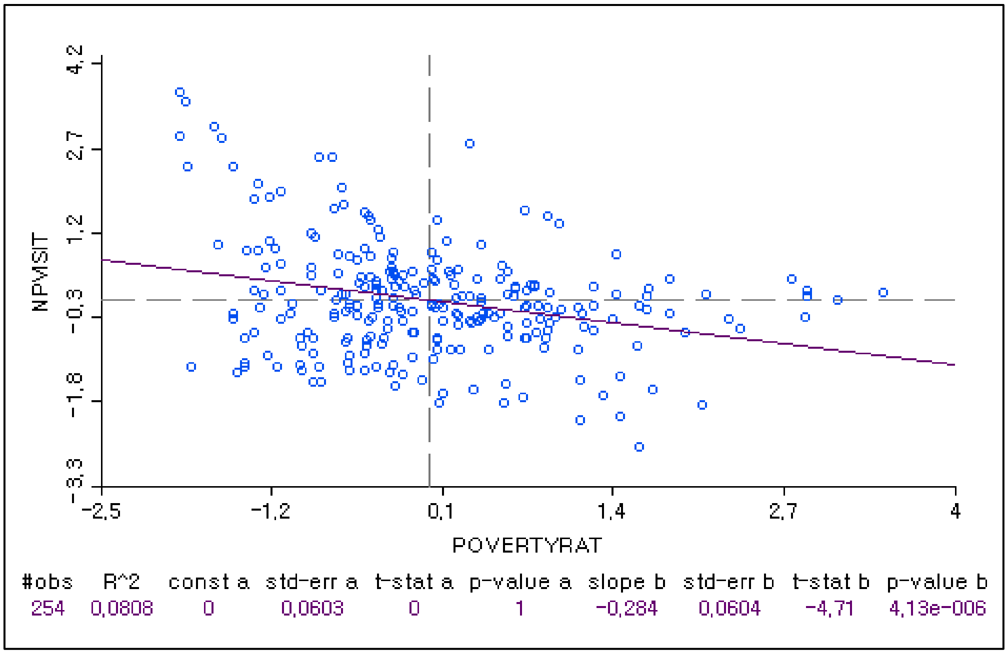


Figure 17 Non-spatial correlation between poverty rate and park visitation

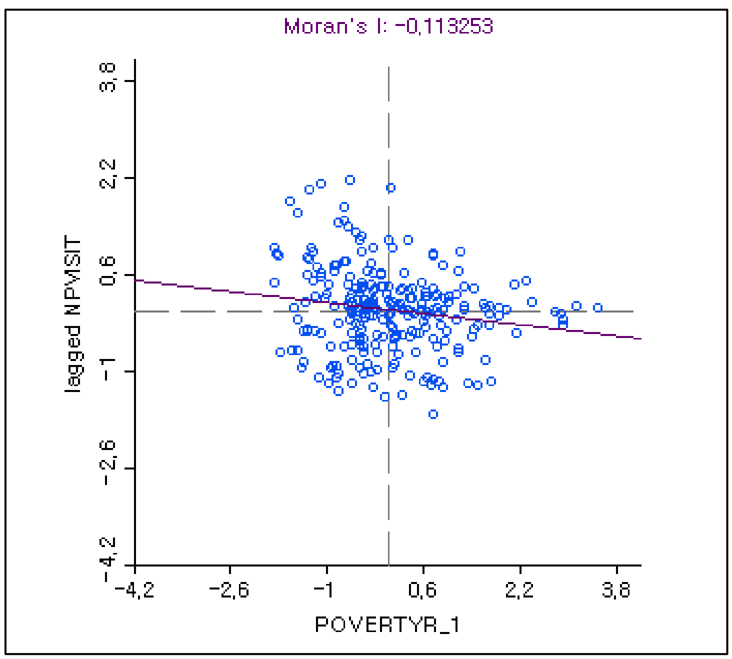


Figure 18 The global bivariate correlation between poverty rate and spatial lag of park visitation

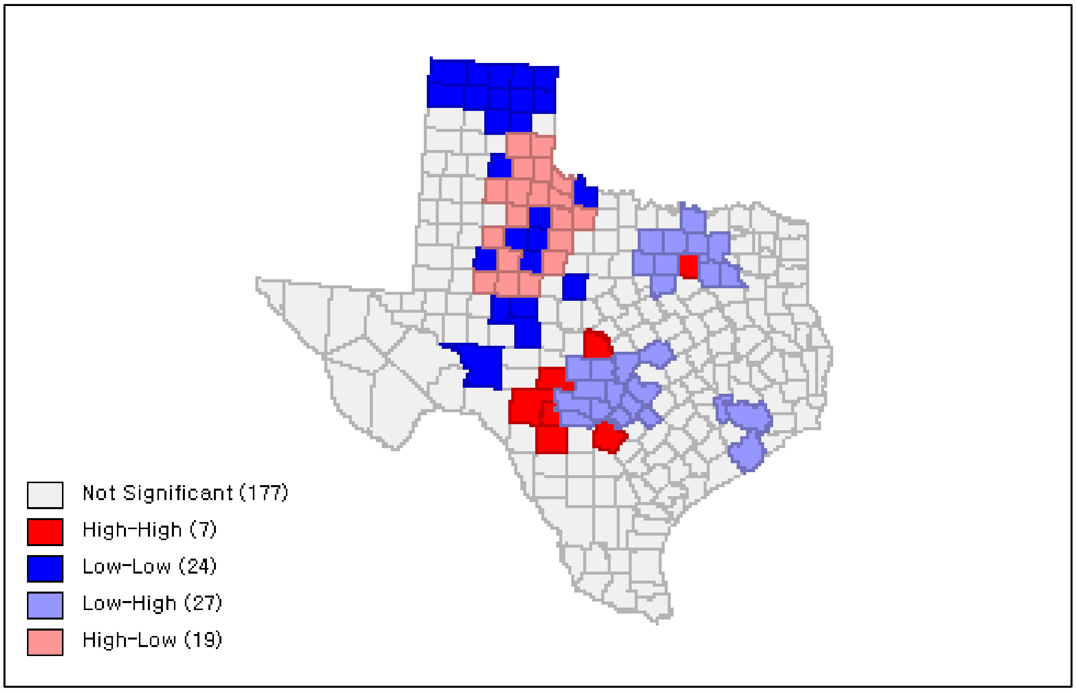


Figure 19 BiLisa Cluster map between poverty rate and spatial lag of park visitation

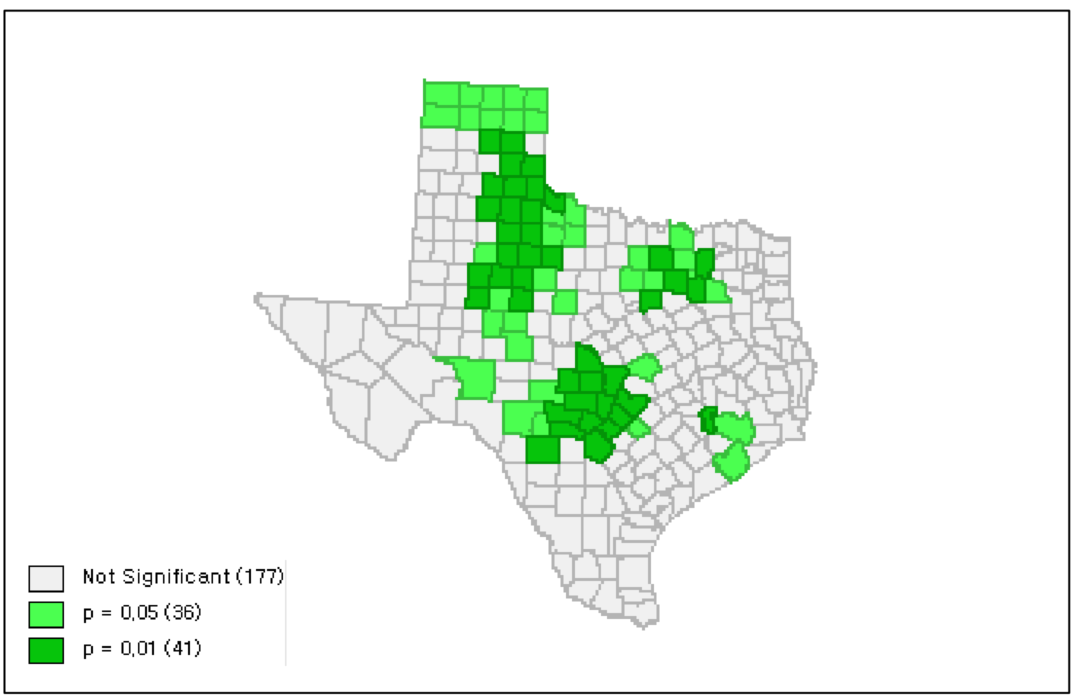


Figure 20 BiLisa significance map between poverty rate and spatial lag of park visitation

2. Recreation spending and visitation

Figure 21 illustrates that higher recreation spending is positively related to national park visitation (slope = 0.754). Also, the local bivariate correlation analysis result showed that there was a positive spatial relationship between higher recreation spending and national park visitation with a Moran's I of 0.356 (See Figure.22). Specifically, twenty-eight counties showed an HH relationship and 38 counties showed an LL relationship, the same as a LISA map of the dependent variable (See Figure 23). This means that recreation spending could be regarded as the most critical variable affecting clustered patterns in national park visitation in the first OLS model. Results of the analysis were statistically significant within a 95% confidence level (See Figure 24).

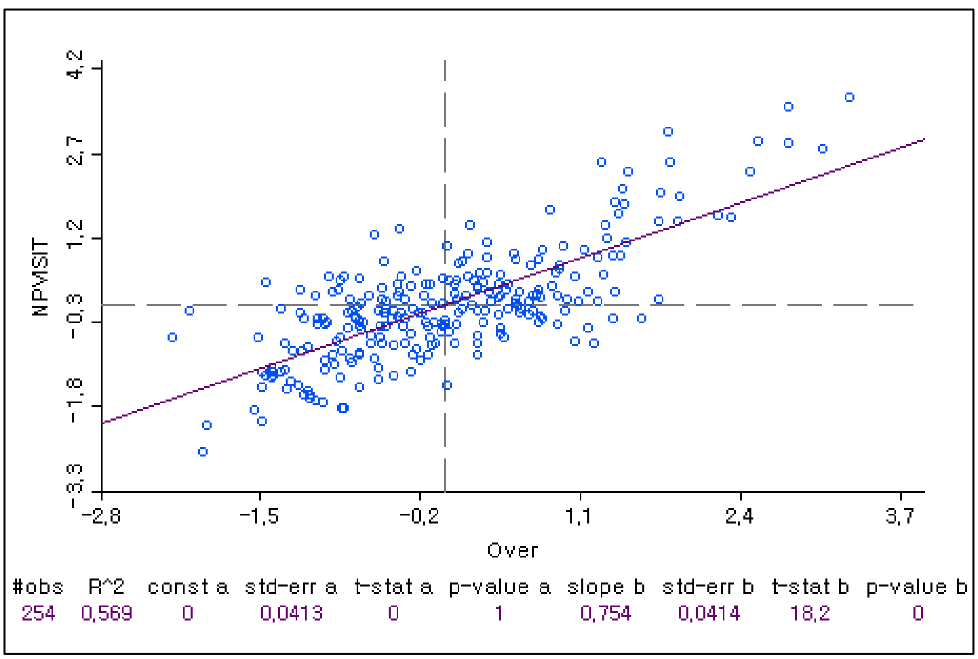


Figure 21 Non-spatial correlation between recreation spending and park visitation

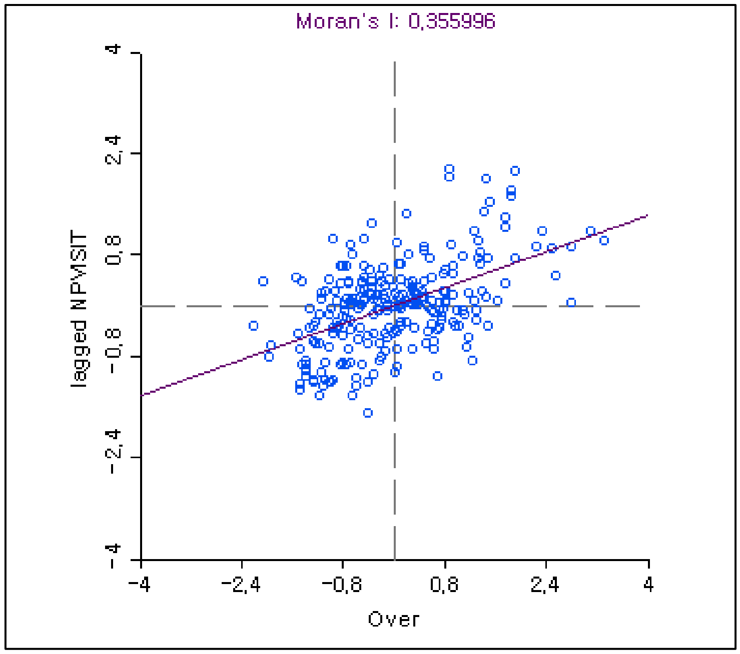


Figure 22 The global bivariate correlation between recreation spending and spatial lag of park visitation

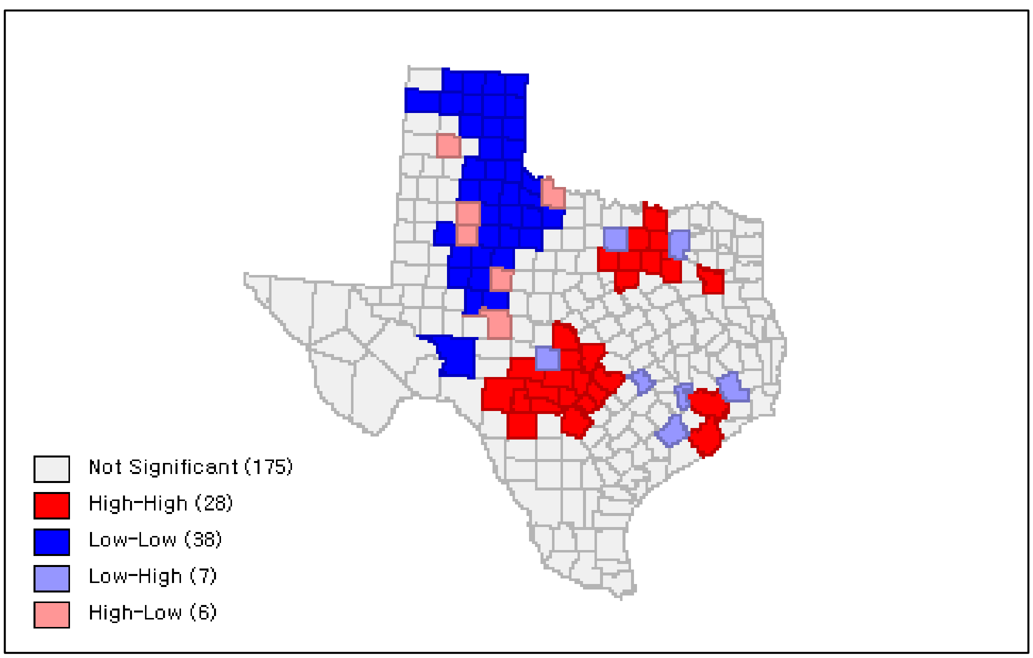


Figure 23 BiLisa Cluster map between recreation spending and spatial lag of park visitation

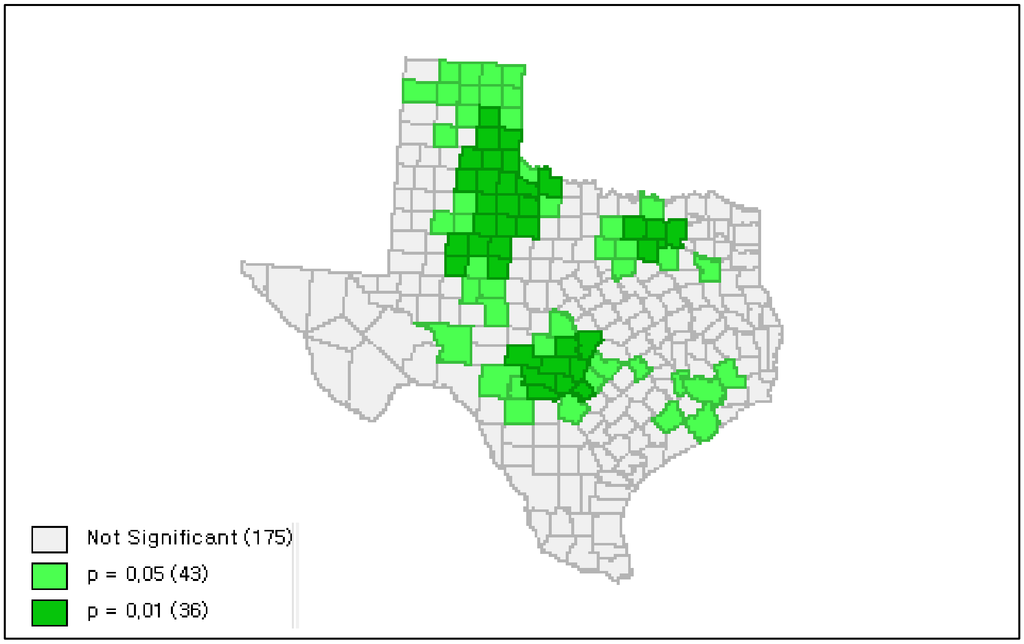


Figure 24 BiLisa significance map between recreation spending and spatial lag of park visitation

In the next analysis, two different types of family structure were used to test the relationship between family structure and national park visitation. The first variable is individuals who live with their parents in each county, which represented a family household of two or more; the second variable is the proportion of residents who live alone in each county, which represented a nonfamily household.

3. Family structure (living with parents) and national park visitation

Figure 25 shows that a higher proportion of residents who live with parents is positively related to national park visitation (slope = 0.387). Also, the local bivariate correlation analysis result showed that there was a positive spatial relationship between the family structure and national park visitation with a Moran's I of 0.267 (See. Figure

26). Specifically, twenty-four counties showed an HH relationship and forty-one counties showed an LL relationship (See Figure 27). Results of the analysis were statistically significant within a 95% confidence level (See Figure 28).

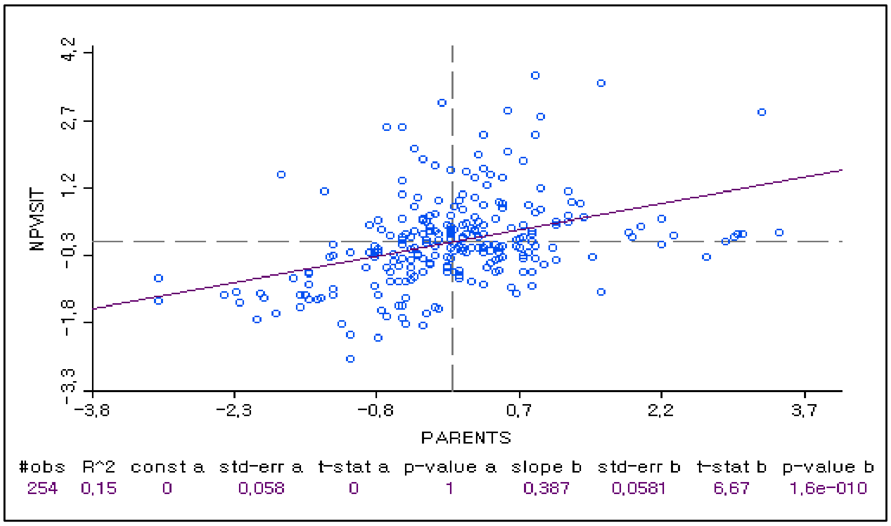


Figure 25 Non-spatial correlation between the proportion of residents living with parents and park visitation

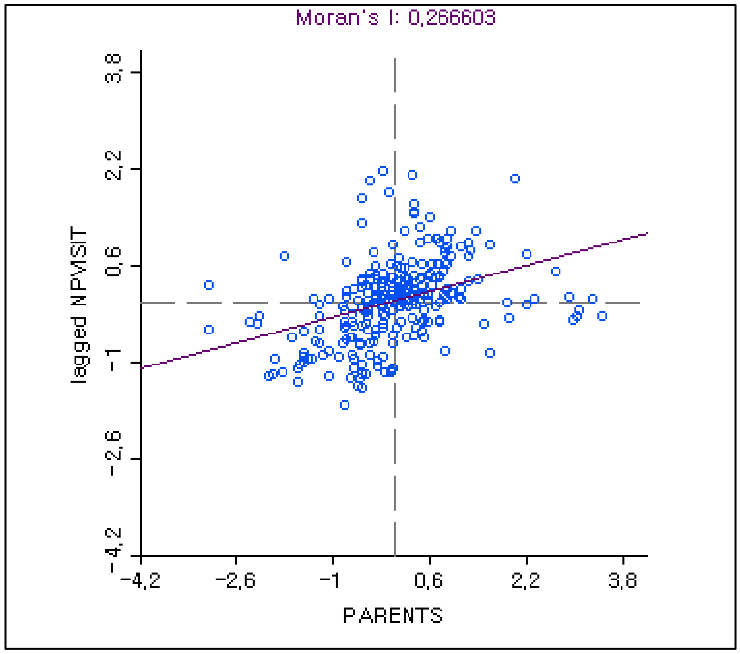


Figure 26 The global bivariate correlation between the proportion of residents living with parents and spatial lag of park visitation

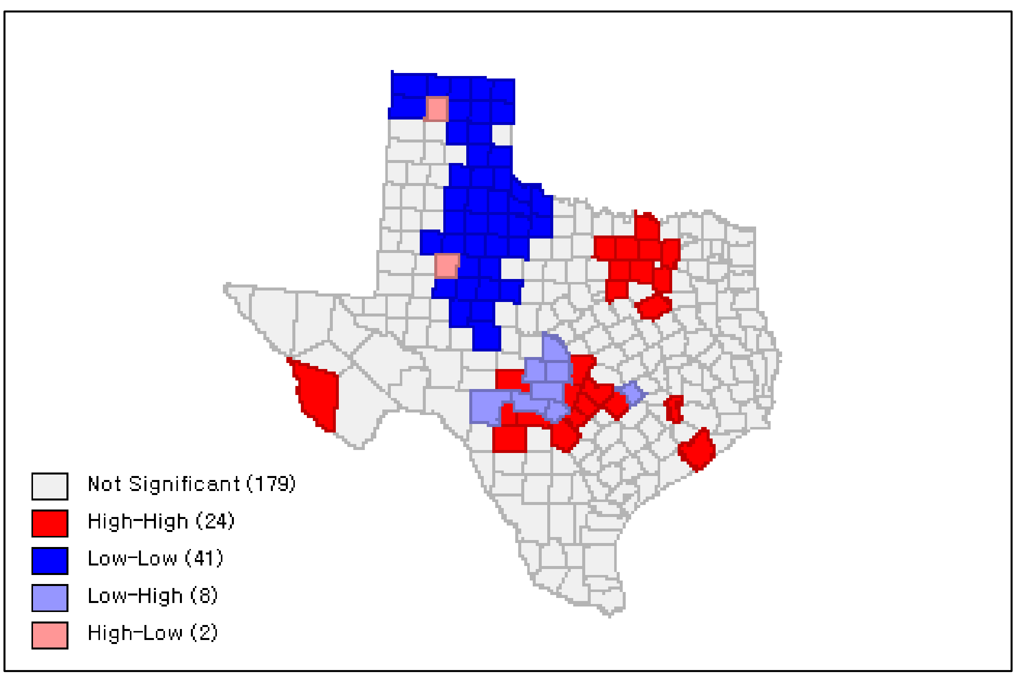


Figure 27 BiLisa Cluster map between the proportion of residents living with parents and spatial lag of park visitation

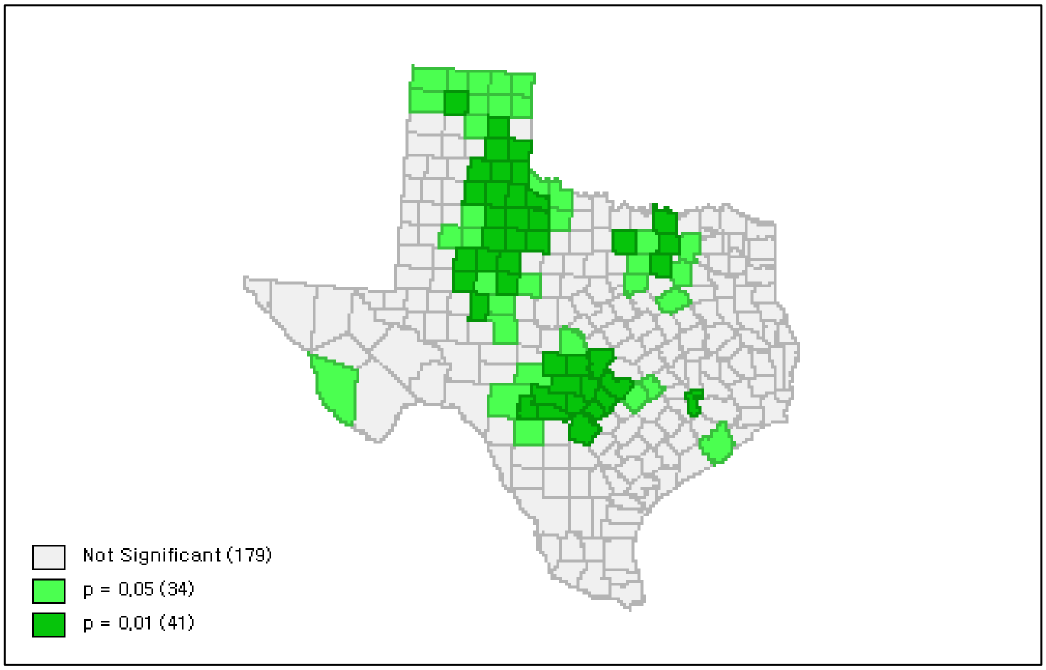


Figure 28 BiLisa significance map between the proportion of residents living with parents and spatial lag of park visitation

4. Family structure (living alone) and national park visitation

Figure 29 illustrates that a higher proportion of residents who live alone are negatively related to national park visitation (slope = -0.0275). Also, the local bivariate correlation analysis result showed that there was a negative spatial relationship between the family structure and national park visitation with a Moran's I of -0.075 (See Figure 30). Specifically, fourteen counties showed an HH relationship and nineteen counties showed an LL relationship (See Figure 31). Results of the analysis were statistically significant within a 95% confidence level (See Figure 32).

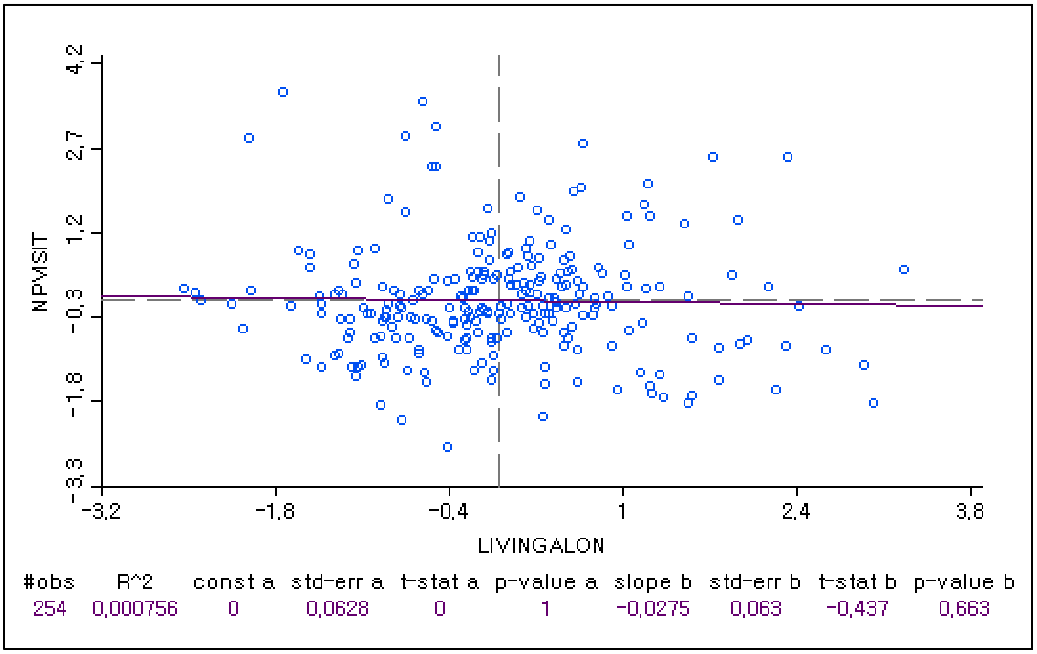


Figure 29 Non-spatial correlation between the proportion of residents who live alone and park visitation

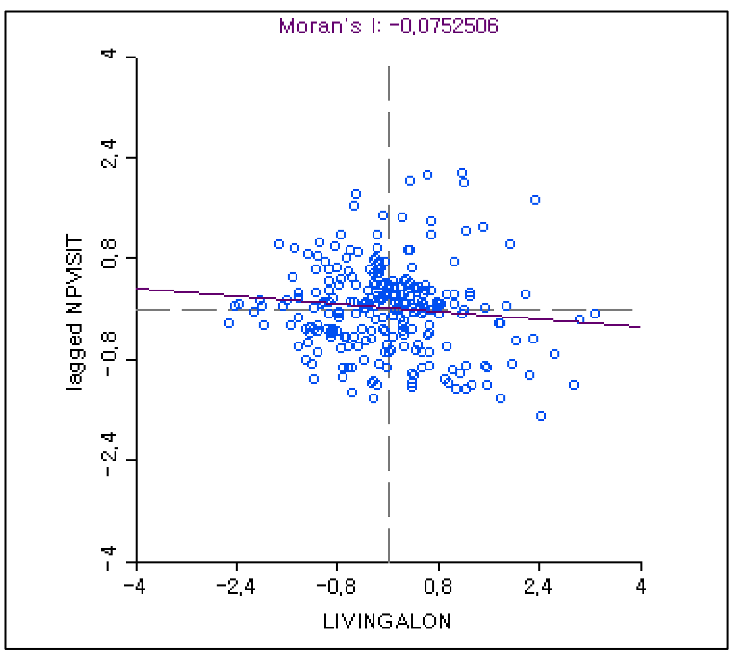


Figure 30 The global bivariate correlation between proportion of residents who living alone and spatial lag of park visitation

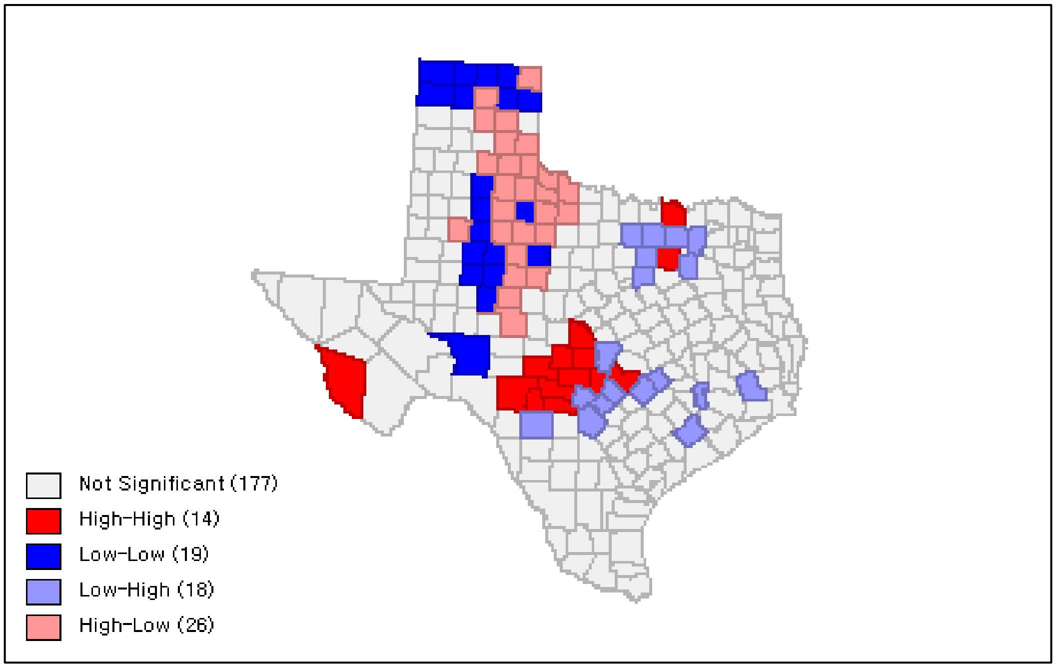


Figure 31 BiLisa Cluster map between the proportion of residents who live alone and spatial lag of park visitation

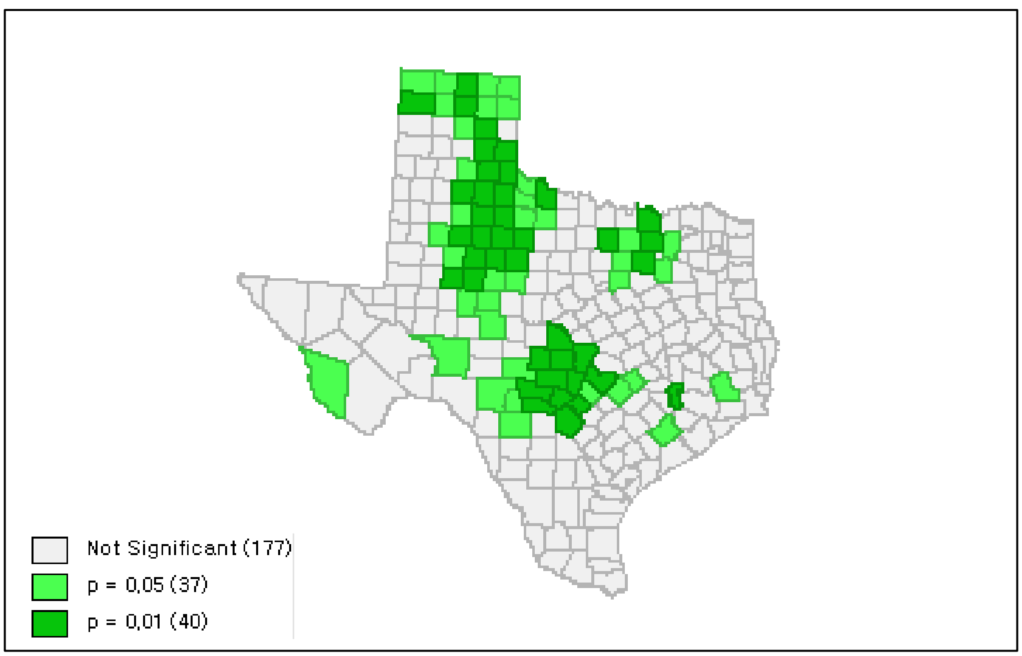


Figure 32 BiLisa significance map between the proportion of residents who live alone and spatial lag of park visitation

5. Lower education level and national park visitation

Figure 33 indicates that a lower education level (proportion of fewer high school graduations) is negatively related to national park visitation (slope = -0.306). Also, the local bivariate correlation analysis result showed that there was a negative spatial relationship between a lower education attainment and national park visitation with a Moran's I of -0.12 (See. Figure.34). Specifically, five counties showed an HH relationship concentrated in the midwestern region of Texas. This means that one county had a higher proportion of lower education attainment levels and its neighboring counties showed a higher national park visitation ratio rather than the mean of the Texas visitation ratio. Twenty-five counties show an LL relationship focused on the northwestern region of Texas. This means that one county has a lower proportion of lower education attainment levels and its neighbor counties show a lower national park visitation ratio rather than mean of Texas visitation ratio (See Figure 35). Results of the analysis were statistically significant within a 95% confidence level (See Figure 36).

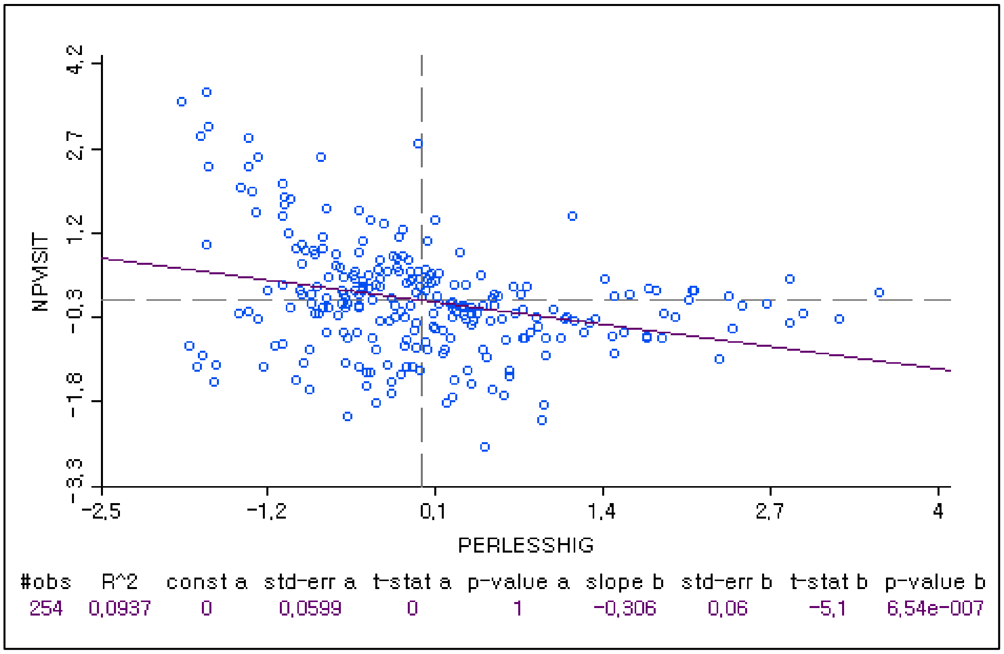


Figure 33 Non-spatial Correlation between low education attainment and park visitation

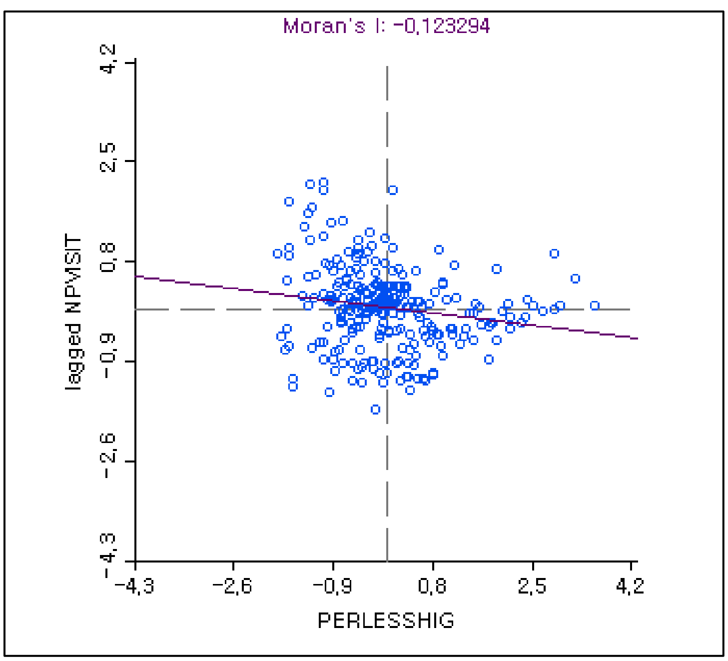


Figure 34 The local bivariate correlation between low education attainment and spatial lag of park visitation

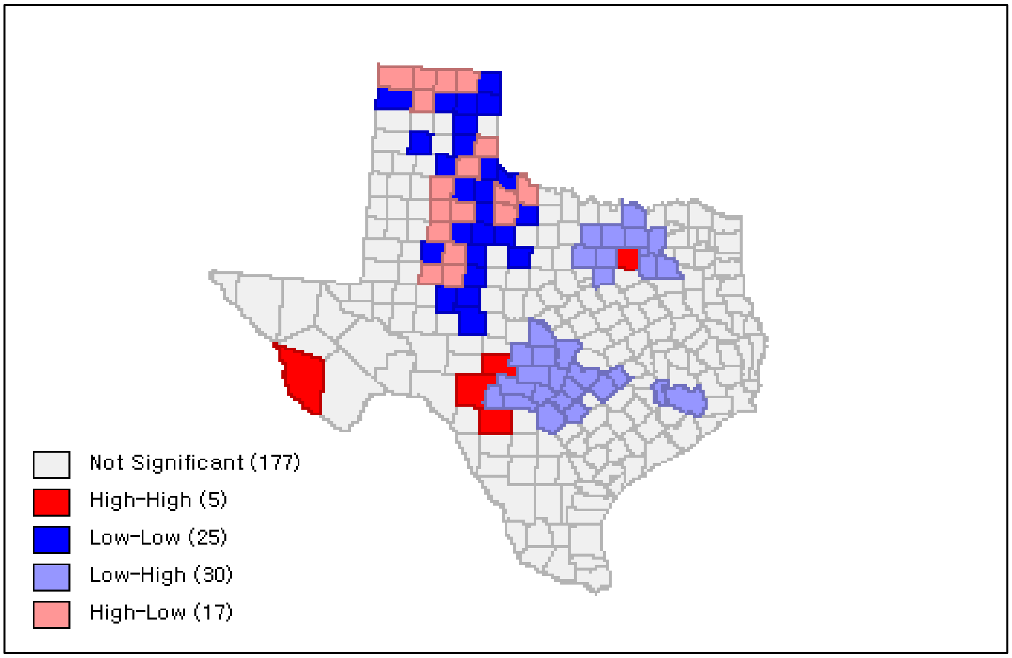


Figure 35 BiLisa Cluster map between low education attainment and spatial lag of park visitation

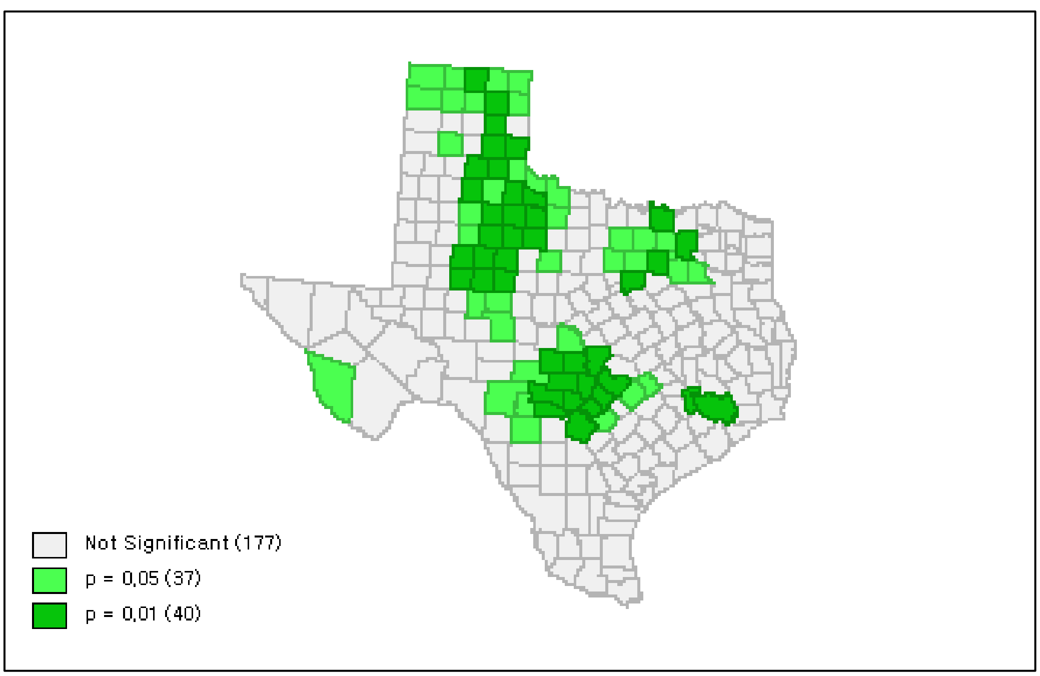


Figure 36 BiLisa significance map between low education attainment and spatial lag of park visitation

6. Proportion of elderly and national park visitation

Figure 37 shows that a higher proportion of elderly (age 65 or over) is negatively related to national park visitation (slope = 0.002). Also, the local bivariate correlation analysis result showed that there was a negative spatial relationship between the proportion of elderly and national park visitation with a Moran's I of -0.045 (See Figure 38). Specifically, fourteen counties showed an HH relationship concentrated in the mid region of Texas. This means that one county had a higher proportion of elderly and its neighboring counties showed a higher national park visitation ratio rather than the mean of the Texas visitation ratio. Twenty-one counties showed an LL relationship focused on the midwestern and northwestern regions of Texas. This means that one county had a lower proportion of elderly and its neighboring counties showed a lower national park visitation ratio rather than the mean of the Texas visitation ratio. Interestingly, most of the lower national park visitation clustered areas (coldspot areas) showed an HL relationship (See Figure 39). This finding that one county had a higher proportion of elderly and its neighboring counties showed a lower national park visitation ratio rather than the mean of the Texas visitation ratio. Otherwise, about half of the higher national park visitation clustered areas (hotspot areas) showed an LH relationship (See Figure 39). This means that one county has a lower proportion of elderly and its neighboring counties show a higher national park visitation ratio rather than mean of Texas visitation ratio. Results of the analysis were statistically significant within a 95% confidence level (See Figure 40).

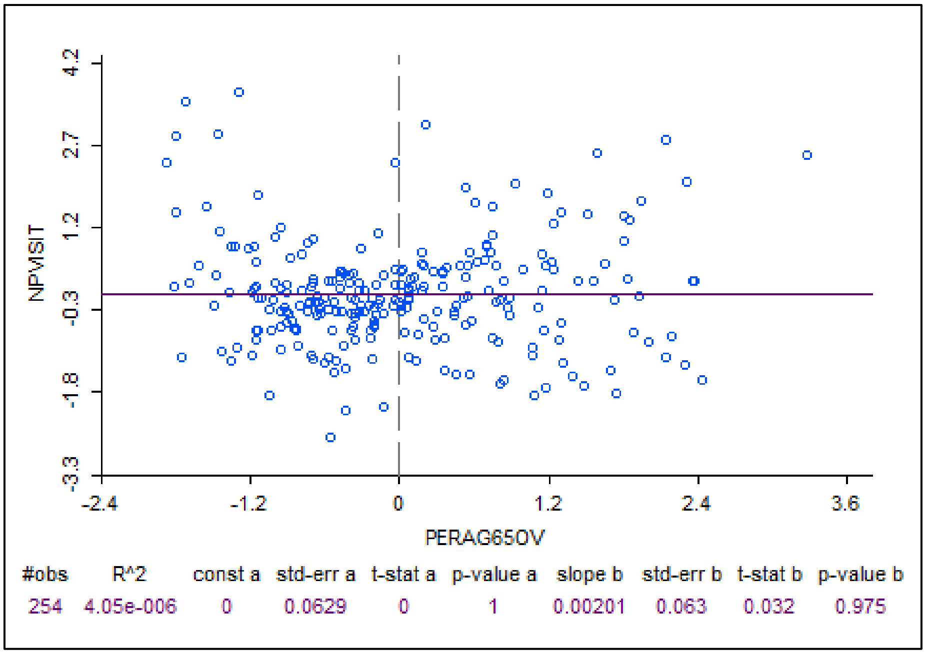


Figure 37 Non-spatial correlation between the proportion of elderly and spatial lag of park visitation

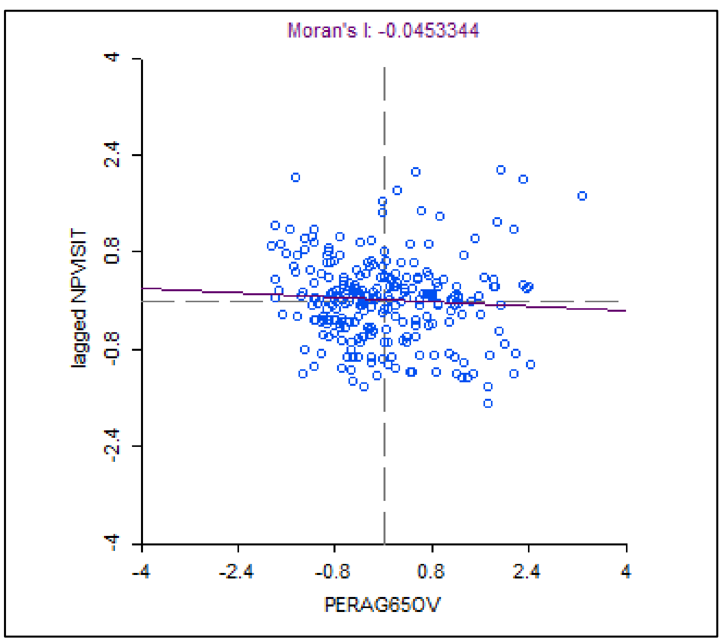


Figure 38 The global bivariate correlation between the proportion of elderly and spatial lag of park visitation

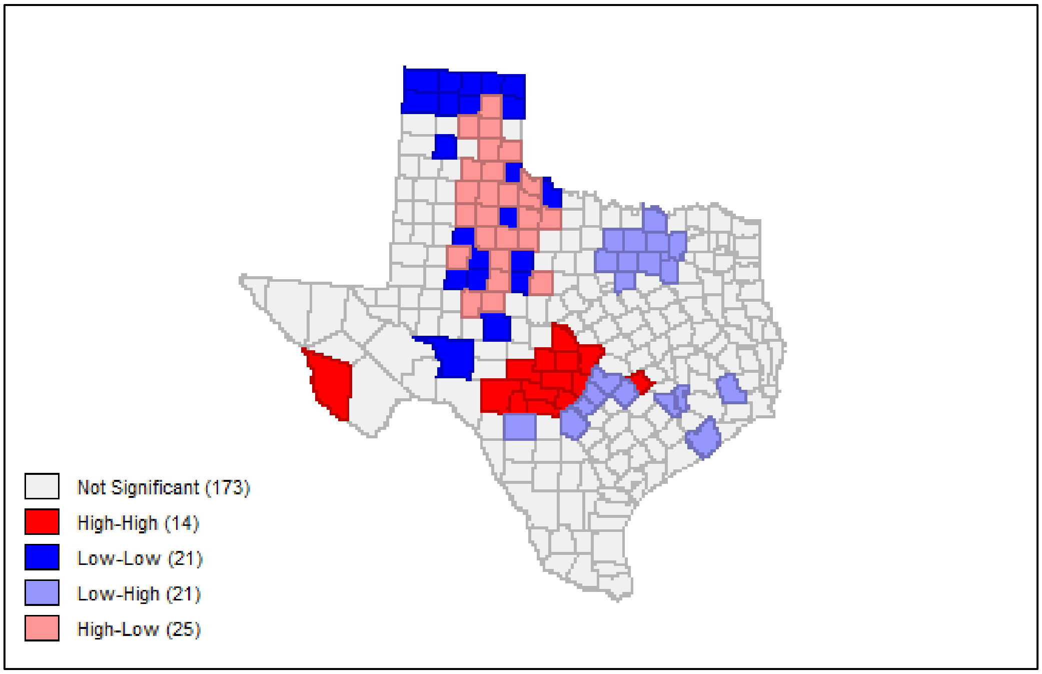


Figure 39 BiLisa Cluster map between the proportion of elderly and spatial lag of park visitation

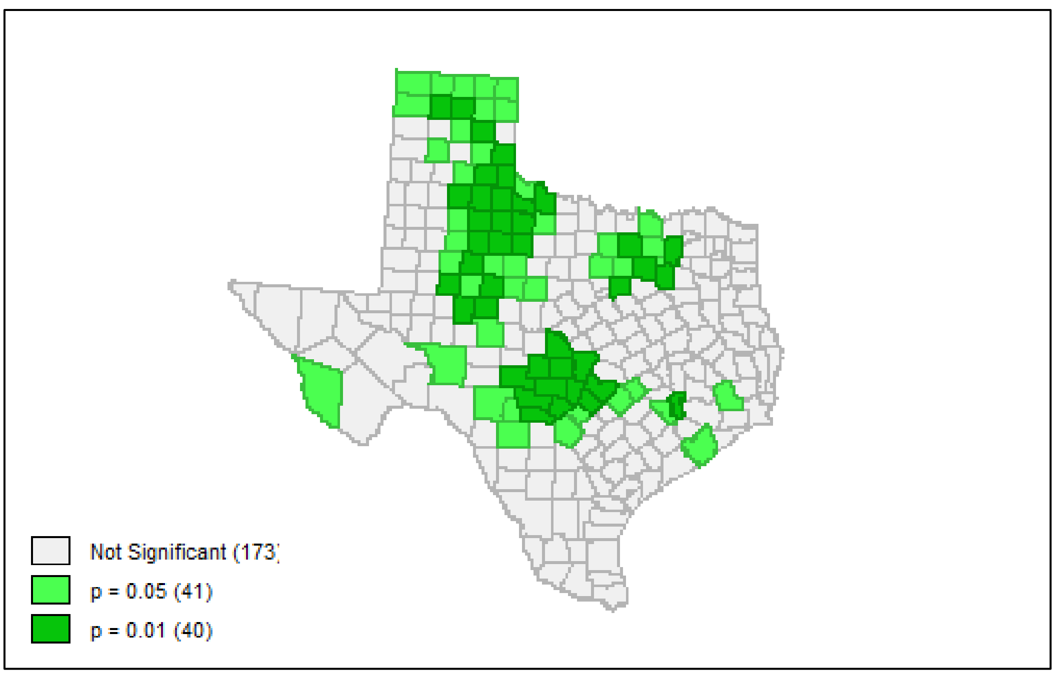


Figure 40 BiLisa significance map between the proportion of elderly and spatial lag of park visitation

Lifestyle segmentation and clustered areas

To create a more detailed interpretation, this study adopted an "ESRI Tapestry Segmentation." Tapestry classifies U.S neighborhoods into 65 market segments based on socioeconomic and demographic factors, and then consolidates them into life mode and urbanization groups (http://www.esri.com/data/esri_data). Previous literature has used ESRI tapestry segmentation data to measure the level of urbanization and lifestyle of residents from marketing perspective (See Wen el al., 2013). For this dissertation, the dominant tapestry in each county was combined with national park visitation cluster maps using ArcGIS.

Figure 41 shows the dominant tapestry number for each county. HH clustered areas consisted of 4, 7, 12, 19, 31, 46, 59 classifications (See Figure 41). The major description of these tapestry types supported the findings of this study. These neighborhoods showed higher income levels and educational attainment compared to the U.S. average. Their lifestyles are active and they are more likely to spend vacations with family members participating in outdoor recreation activities. Interestingly, tapestry number 59 showed a higher diversity index than other tapestry types within HH clustered areas (called hotspots). This neighborhood consisted of an 83% Hispanic population. Furthermore, tapestry code 37 within LL clustered areas (coldspots) showed little race/ ethnic diversity; 9 in 10 of these residents were white. These descriptions countered to the results of past research and extended OLS models in this study. Although many studies insisted that the Hispanic population was less likely to visit to a park and the white population was more likely to visit, however, this result would not be true for all areas within the state. The description of preference in segmentation code 59 showed that more information channels are needed to attract different ethnic groups: “They listen to Hispanic and contemporary hits on the radio. TV and radio are the best media to reach them instead of newspapers or magazines” (See Appendix D).

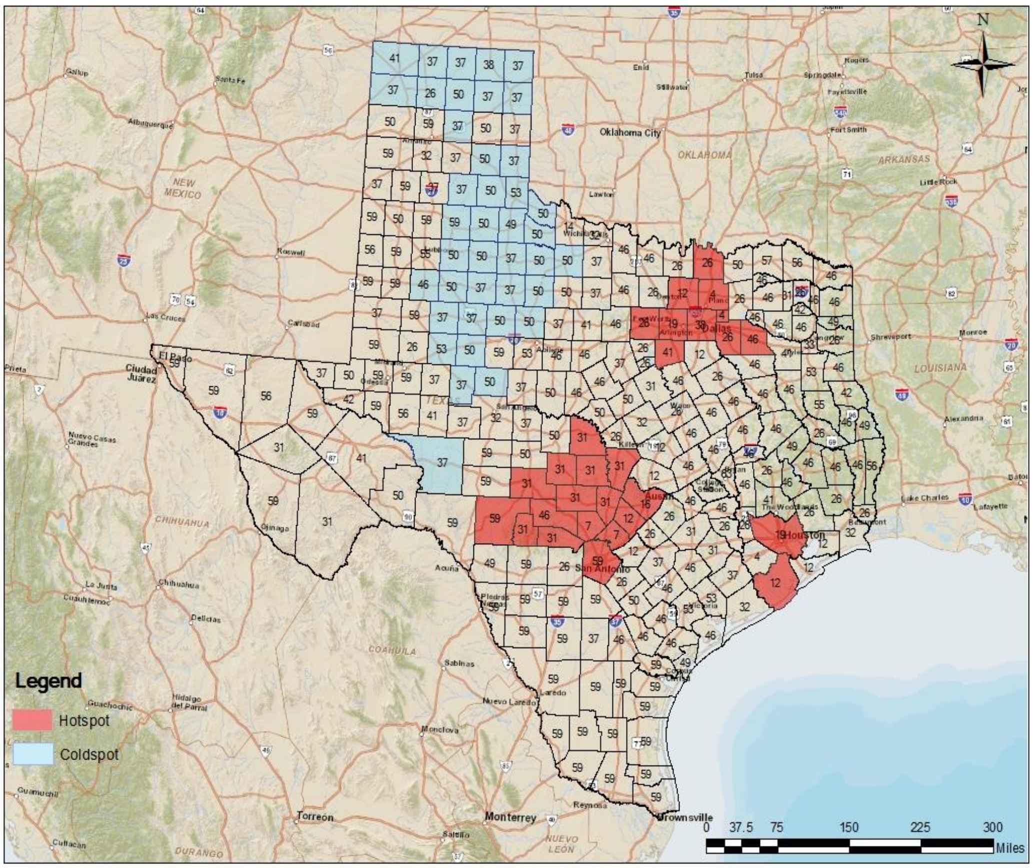


Figure 41 Dominant tapestry segmentation with cluster areas of park visitation

Previous research has rarely documented how potential visitors obtain information about local park opportunities (Byrne, 2012). For example, whether potential users circulate unevenly and through different channel: Tierney et al. (2001) suggest they do, noting that Latinos rely more on social networks to get access about urban protected areas in Los Angeles than do whites, and may be constrained as a result (also see Spotts and Stynes, 1985). Further, Byrne (2012, p.608) insisted, “One has to seriously question the motivations of park managers when their websites contain information about the park predominantly in English, when park images convey representations of ‘White’ nature (e.g. landscapes emptied of people representing ‘pristine’ or ‘pure’ nature)”. In this context, the result of the above analysis indicated that park planners and managers should investigate the information networks of potential visitors and target park use information to specific ethnic/race groups through the networks when park staffs are dominantly White, and park signs written only in English.

The public policy most often stated reflects a position of equal opportunity for all. However, this approach neglects the fact that certain groups of individuals are less able to take advantage of the opportunities. Spatially dominant residents of each Texas county show different socio-demographic and economic characteristics. The results of additional analysis supported the idea that these different characteristics could explain variations in national park visitation among the counties throughout the state. If target marketing is needed for future regional recreation planning, these findings could be useful. Target marketing is the process that an agency or other recreation delivery organization goes through to identify the characteristics of those it desires to serve. The results of this study allow recreation managers or planners to design programs and services to fit different segments of the populations. Locational information such as where national park visitors frequently or rarely come from, their dominant socio-demographic or economic circumstances, and how those circumstances relate to visitation, would allow for development of public information necessary to overcome barriers such as lack of knowledge of where to participate.

Third-stage analyses: GWR analysis

Contrary to the ordinary regression model, a GWR produces a series of local regression results based on weighted least squares for the location of each county. The adaptive kernel function was used to define the bandwidth of local relationships. As I mentioned in the method section, adaptive kernel width is determined through the minimization of the Akaike Information Criterion. AIC minimization provides a way of choosing bandwidth that makes an optimal tradeoff between bias and variance.

Table 8 Summary of GWR results

Neighbors	254
AICc	187.8027
R-squared	0.7636
Adjusted R-squared	0.7501

The GWR regression model made improvements in model performance over the OLS regression model. The adjusted R-squared value increased from 0.726 to 0.75 and the AIC value decreased from 206.39 to 187.80. With a lesser value and a difference of more than 3 in the AIC values, the GWR provided more specific and reliable information than the OLS (See Table 8). To test for spatial autocorrelation of the GWR results, the global Moran t I test of the residuals showed a random distribution (I = 0.059; z-score = 1.606; p value = 0.108) indicating the absence of misspecification in the model. Given the z-score of 1.606, the pattern did not appear to be significantly different than the random distribution. This means that there is little evidence of any spatial autocorrelation in the residuals. Any spatial dependencies which might have been present in the residuals for the OLS model have been removed with geographical weighting in the local model (See Figure 42 & Table 9).

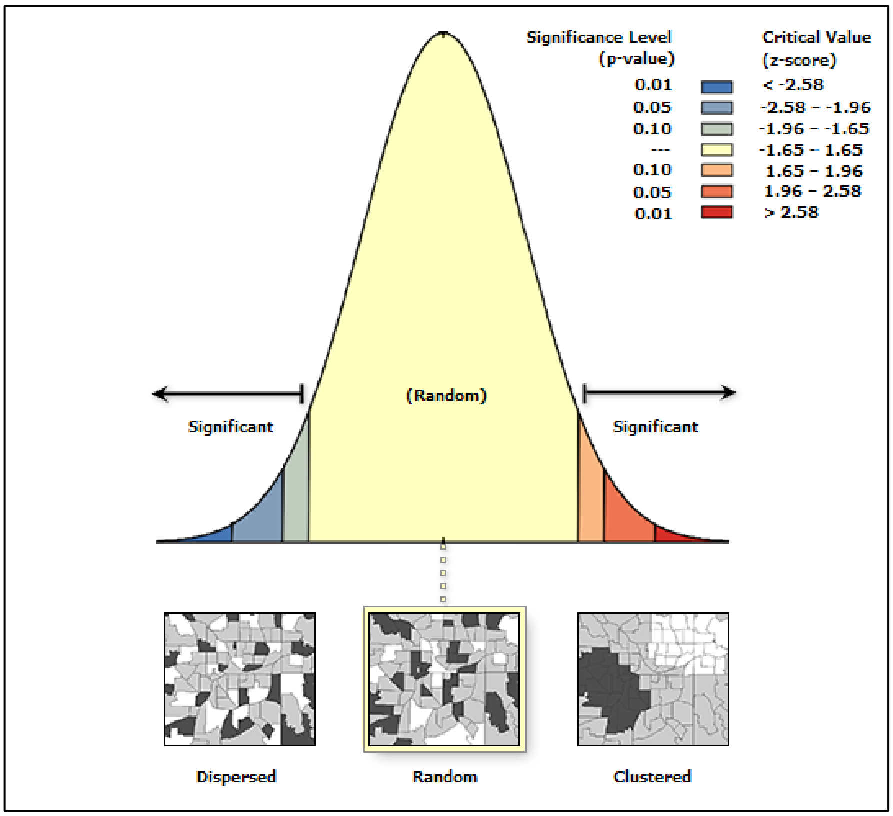


Figure 42 Spatial autocorrelation test of GWR residuals

Table 9 Summary of Global Moran's I test for GWR residuals

Moran's Index:	0.059587
Expected Index:	-0.003953
Variance:	0.001564
z-score:	1.606918
p-value:	0.108072

Local R-squared value

The local R-squared value ranges from 0 to 1. This statistic shows how well the local regression model fits the observed dependent value. If the local R-squared value is 1, the dependent variable is perfectly explained (predicted) by the model. The mapping of the local r-square shows different levels of model performance depending on the areas. The explanatory power of the OLS for explaining the relationship remained relatively good (R-squared value = 0.732). This was increased when the GWR was applied. More than 96% of Texas counties showed a higher local R-squared value than with the OLS. This indicates that 96% of Texas counties' national park visitation ratios were better explained by the GWR model rather than the OLS model and 4% of Texas counties still require more explanatory variables to better explain the visitation ratios.

The distributions of the local R-squared statistics ranging from 73.1% to almost 78.3% of variability were explained by the GWR model. The model predicted more than 75% of the variability in the midwestern region and 74% or less of the variability in counties in the eastern and northern regions. These variations indicated that the global regression model could not explain the non-stationary relationship across the state (See APPENDIX A-1). The condition number is the square root of the largest eigenvalue divided by the smallest eigenvalue (Lin & Wen, 2011). The condition number is a diagnostic that evaluates local collinearity (Siordia et al., 2012). If the condition numbers are greater than 30, multicollinearity would be a very serious concern (not reliable). The GWR model shows the condition number as less than 30 (from 23.9 to 29.5) in this study (See APPENDIX A-2). This means that there are no serious local multicollinearity problems.

The local coefficient estimates

The GWR tests and obtains local coefficients for each location and each variable; in this model 254 census counties and 6 variables. The map for the local coefficients revealed that the influence of the variables in the model varies considerably throughout Texas. Higher coefficient values (positive or negative) indicate a greater influence of the particular predictor variable on the dependent variable, whereas lower values mean lesser influence for the particular variable.

1. **Relationships between poverty rate and national park visitation**

The global estimate of the poverty rate variable was significantly negative as were all the spatial estimates (OLS coefficient: -0.021). This means that residents in poor economic situations negatively influenced national park visitation across the state. As shown in Figure 46, local coefficients of the poverty rate were negative and exhibited significant spatial variation. The range was from -0.034 to -0.021. Among residents who have the same poverty status, residents in the eastern region are more likely to visit a national park than residents in the western region of Texas (See APPENDIX A-3).

2. Relationships between recreation spending patterns and national park visitation

There was a variation in the coefficient estimates for the high recreation spending variable (See APPENDIX A-4). The estimated value for the global model was 0.751, with a standard error of 0.043. The map for the local coefficients revealed that the influence of this variable in the model varied considerably across Texas, with a strong western direction. The range of the local coefficient was from 0.666 in the eastern counties to 0.7704 in the western counties. Among the residents who spent 250$ more to purchase recreation equipment in the past 12 months, residents in the western region where more likely to visit a national park than residents in the eastern area.

3. Relationships between the proportion of elderly and national park visitation

The global estimate of the proportion of the elderly variable was significantly positive as are all the spatial estimates (OLS coefficient: 0.042). The map for the local coefficients revealed that the influence of this variable in the model varied considerably across Texas, with a strong southeastern direction. The range of the local coefficient was from 0.032 in the northeastern counties to 0.047 in the southwestern counties (See APPENDIX A-5).

4. **Relationships between proportion of low education attainment and national park visitation**

The global estimate of the proportion of the low educational attainment variable was significantly negative as were all the spatial estimates (OLS coefficient: -0.015). The range of the local coefficient was from -0.014 in the eastern counties to -0.0079 in the midwestern counties. As previous literature has shown (Floyd et al., 2006; Scott & Shafer, 2001) a higher education levels facilitate recreation participation otherwise lower education levels lead to lower national park visitation across the state; however, there are regional variations within the state (See APPENDIX A-6).

5. **Relationships between the proportion of different family structures and national park visitation**

In this study, two different types of family structure were used to test this relationship with national park visitation. The first variable was individuals who live with their parent(s), which represents a family of two or more. The global estimate of the proportion of individuals who live with their parent(s) was significantly positive as are all the spatial estimates (OLS coefficient: 0.494). The range of the local coefficient was from 0.227 in the southwestern counties to 0.677 in the northeastern counties (See APPENDIX A-7). On the other hand, the proportion of residents who live alone (nonfamily household) in each county showed a negative relationship in the global estimation (OLS coefficient: -0.057). The range of the local coefficient was from - 0.065 in the northwestern counties to -0.045 in the northeastern counties. These results demonstrate that family groups are more likely to visit a national park while those who live alone are less likely to visit a national park (See APPENDIX A-8).

The results noted by the GWR model are a reflection of the variability of the relationship between the response and predictor variables across different parts of the study area. The GWR was able to account for both spatial autocorrelation and spatial non-stationary processes, thereby providing a better foundation for prediction and explanation than the corresponding OLS model (See Table 10).

Table 10 Comparison result between OLS and GWR

	OLS	GWR			
		Mean	Minimum	Maximum	Standard Deviation
Intercept	2.548*	2.577693	2.453038	2.698183	0.076461
High Recreation Spending	0.705*	0.715951	0.666366	0.770422	0.032093
Low Education	-0.015*	-0.012188	-0.014533	-0.007936	0.001488
% of Elderly	0.042*	0.041184	0.031891	0.047117	0.005212
Living with parents	0.494*	0.463054	0.226753	0.677341	0.116156
Living Alone	-0.057*	-0.058849	-0.065724	-0.045814	0.006179
Poverty rate	-0.021*	-0.026169	-0.034075	-0.021211	0.012011
Local R – squared		0.728673	0.696087	0.762898	0.003423
R – squared	0.732	0.7636			
Adjusted R – squared	0.726	0.7501			
AIC	206.396	187.802			
Moran'I of Standard residual	0.146	0.059			
Koenker Statistics	15.445*	Neighbors: 254			
Jarque-Bera Statistics	0.909	Bandwidth methods: AICc			
		Kennel type: Adaptive			

This result was synchronized with the results of other spatial econometrics studies (e.g., Buyantuyev & Wu, 2010; Anselin et al., 2004; Yu, 2006; Newburg, 2011) that found significant variation in socio-economic or demographic parameters, and better prediction of the GWR model over alternative regression procedures such as OLS. According to the adjusted R-squared value, the GWR showed a higher value than OLS. Also, the AIC value of the GWR was 18.594 lower than the OLS.

The AICc is a relative measure, not an absolute measure (in comparing model performance the lower the better). The OLS model's value was lower by three or more, so there was "reasonable evidence that the local model was a better fit to the data, given the different model structures" (Fotheringham et al., 2008, p.232). This means that GWR models have improved reliabilities of the identified relationships by reducing spatial autocorrelations and by accounting for local variations and spatial non-stationarity between dependent and independent variables. The range for the local coefficients revealed that the influence of the variables in the model varied considerably throughout Texas. As seen in table 10 and in the mapping of the GWR results, the GWR model effectively showed the variations of OLS regression coefficients based on different sub-areas.

Summary results for research questions

1) ***Will national park visitation vary spatially with visitation rates at the county level? Do spatial patterns of national park visitation show clustered patterns within the state?***

According to the global Moran's I test, the dependent variable national park visitation ratio was spatially clustered (Moran's Index: 0.439). The local Moran's I test revealed that twenty-eight Texas counties showed HH cluster types concentrated in the mid and northwestern Texas regions. Forty counties located in the northwestern region shows an LL type cluster pattern. Interestingly, hotspot areas of national park visitation were concentrated in the areas of Dallas, San Antonio, Austin, and Houston. Even though urban population (population density) did not show statistically significant, there would be an impact of urbanization around the Texas triangle cities affecting national park visitation within the state (See Appendix E).

2) What and how do socio-demographic and economic factors (e.g., age, education attainment, economic condition) influence the level of national park visitation?

In the final model, this study selected six independent variables: high recreation related spending, the proportion of elderly, and family structure of two or more positively influenced visitation, while the low education attainment, poverty rate, and living alone variables negatively affected visitation. Gender was not a statistically significant variable in the previous tested OLS models. The diversity index, (the proportion of different races or ethnicities in a population), median house income, median family income, and unemployment rate showed statistical significance in some models. However, they were filtered (excluded) by a stepwise regression test and BP test.

3) What and how do situational factors (e.g., urbanization, proximity to the park) influence the level of national park visitation?

Population density, proportion of urban population, and distance from each county's center to the closest national park did not show statistical significance in the tested OLS models. These results might be due to original data not having specific destination information. Also, these results could be interpreted to mean that socio-demographic and economic variables affect visitation more than proximity and urbanized residential circumstances within the state of Texas.

4) Does the GWR model improve prediction when compared with the OLS regression model?

The OLS regression residuals were examined for spatial autocorrelation and regression residuals showed strong spatial autocorrelation. This means that the assumption or assumptions of the OLS regression model was not satisfied. Also the Koenker (BP) Statistic reveals that the OLS model has problematic heteroscedasticity. In this sense, the GWR (spatial statics) would be more appropriate to analyze the relationships with spatial data.

The GWR showed a higher R-squared value and a lower corrected AIC value compared to the OLS model. This means that the GWR regression model made noticeable improvements of model performance over the OLS regression model. Also, more than 96% of Texas counties showed higher local R-squared values than with the OLS.

5) Is there spatial heterogeneity based on the relationships between the aforementioned factors and national park visitation among Texas counties? How does the GWR reflect this kind of spatial relationship?

The GWR is a powerful tool for exploring spatial heterogeneity. Spatial heterogeneity exists “when the structure of the process being modeled varies across the study area” (Charlton & Fotheringham 2009, p.1). The map for the local coefficients revealed that the influence of the variables in the model varied considerably across Texas. Varying local relationships were tested and represented visually by mapping the distribution of local coefficients. For example, the coefficient estimates for the high recreation spending variable were evidence which points to heterogeneity in the model structure within Texas: its range of the local coefficient was from 0.6663 in the eastern counties to 0.7704 in the western counties. It was also demonstrated through GWR that the influence of each selected independent variable in the final OLS model varied considerably across Texas.

CHAPTER V

SUMMARY AND DISCUSSION

This study tried to identify the spatial relationships between national park visitation and its associated factors using large aggregated data. Guided by the idea of opportunity theory and Pigram's conceptual framework, this study empirically investigated what and how factors associated with national park visitation influence demand within the Texas boundary. Some of these study findings were in line with previous literature and some were not. Specifically, comparing the results of the final OLS regression model and extended OLS model, with the prediction of general outdoor recreation demand conducted with NSRE (Bowker et al., 2006, 2007), economic status, education attainment, and ethnicity (Hispanic) were in line with their prediction. However, race (black), age, and residence (urban) were not consistent with past research findings. Large-scale recreation demand estimation usually has been based on general population surveys such as the NSRE. Though these estimations address regional, state, or national trends of general recreation demands, they do not provide detailed information for sub-areas within the study area (e.g., county level). As mentioned at the beginning of this study, a recent published article found that the economic recession resulted in lower national park visitation (Poudyal et al., 2012). However, the results of that study did not show regional variations. Although recreation demand studies have used ordinary least squares (OLS) for estimating (predicting) recreation demand, one general relationship could not effectively explain the complex relationships within the study area. If sub-areas show variations or inverse relationships compared with OLS results, it would lead to an inappropriate recreation demand prediction (especially for the regional or national scale).

To overcome these statistical and spatial limitations, this dissertation tested regional variations of the relationship between national park visitation and factors associated with visitation using a combination of traditional OLS and spatial analysis: 1) spatial cluster analysis showed the location of hotspot areas in national park visitation; 2) local bivariate correlation tests and maps separately showed six selected explanatory variables related to national park visitation and how they contributed to the distribution of cluster areas within the state; 3) the OLS results showed the global relationships over the state; and 4) the GWR model described how varying local relationships exist based on the boundaries of all Texas counties within the state. Specifically, visualization of those varying relationships indicated that one specific model could not be applied to the complete study area. As was seen in the previous chapter, there were variations and different strengths of relationships within the state.

To sum up, this study discovered that the economic status of households greatly affected national park visitation, which varied by region, and what specific economic factors influenced the visit. The strongest positive influence on national park visitation was found for those who spent $250 or more to purchase recreation-related equipment. This variable led to the assumption that the economic ability to participate in recreation activities highly influenced the actual participation. Along the same lines, a lower economic status household negatively influenced national park visitation; lower poverty rates led to lower national park visitation throughout the state.

In each Texas county, the proportion of people over 65 years old positively influenced on the visitation rates across the state with different strength of local coefficients. As previous literature mentioned, those over 65 years old participate more frequently in appreciative and learning activities and less frequently in active expressive activities (Foot, 1990; Martinez, et al., 2009; McGuire et al., 2009). This indicates the possibility that those over 65 are more likely to regard national park visitation as a passive activity for their learning opportunities. These results are contrary to the findings in previous literature from tests examining relationships between aging and leisure behavior. Conclusions of previous studies usually indicated that leisure participation is directly and negatively correlated with aging (Kelly, 1980; Jackson & Henderson, 1995; Scott & Jackson, 1996; Shaw et al., 1991). However, limited research has examined the effects of specific age categories on national park visitation and these results would be helpful in understanding the impact of specific age groups on visitation.

Educational attainment also influences national park visitation. Different levels of education have been regarded as one of the most critical predictors of recreation participation (Kelly, 1983; Lucas, 1990; Lee et al., 2001). As previous literature has found, higher education levels facilitate recreation participation (Floyd et al., 2006). On the other hand, lower educational attainment has been regarded as a barrier for participation in recreation. Scott and Shafer (2001) insisted that participation in recreation activities would be more problematic for individuals with a lower level of income and education. Similarly, the proportion of those with lower rates of high school graduation showed a negative relationship across the state in this dissertation.

Relatively little work has been done to examine family structure as it relates to recreational choices, patterns, and barriers (Blanco, 2009). This study found that different family structure types affected visitation. Large family structure households with individuals that live with their parent(s) positively influenced visitation, while individuals who live alone negatively influenced visitation. However, it should be noted that the variables “average size of households” and “the proportion of household living with children” were excluded in the final model due to multicollinearity problems. These variables also positively influenced visitation in the other tested OLS models. In this sense, families with a larger family structure are more likely to visit the park rather than individuals living alone. This finding would support to the conclusion of previous study (McKinney, 1999): “park visitation tends to be a predominantly family-oriented activity” (p.104).

Mapping local variations in parameter estimates may also facilitate the identification of potential causes of missing variables or interaction terms and low estimation efficiencies (Jetz et al., 2005, Shi et al., 2006), which could provide positive implications towards future recreation data collection and research activities (Zhang & Shi, 2004). The GWR has the potential to play a leading role in this important new research area (Foody, 2003, 2004; Hanham & Spiker, 2004). For example, 4 % of Texas counties showed local R-squared values compared to OLS results. There would be more associated factors to explain the visitation. Additionally, all local regression coefficients of explanatory variables showed various inclinations within the state. Further research could investigate why those kind of various inclinations exist by using other factors.

Theoretical implications

The results of this study have contributed several significant theoretical revelations that will help guide future research on recreation demand and spatial econometrics. The following section will summarize the theoretical contributions of this research in detail.

The first law of geography

The geographic information was organized and stored, and common in the representations was that the information was always associated with geography. That is, geographic information always contains positional information and, hence, can be used to answer the question "Where is it?" With attribute information, researchers can answer the question "What is where?" More importantly, geographic information is always dictated by the "First Law of Geography," where "everything is related with everything else, but near things are more related than distant things". This first law immediately implies two unique characteristics for geographic information, commonly known as the spatial effects, spatial autocorrelation and spatial heterogeneity.

The existence of both autocorrelation and heterogeneity in geographic information is not accidental, but inherent. It is the existence of such unique characteristics in geographic information that makes analyzing spatial data rather different from other types of data.

Spatial autocorrelation and heterogeneity have been tested in various academic fields (e.g., criminology, medical geography, and real estate study) and these test results could be regarded as fundamental but critical clues for preparing for expected future social problems or creating alternatives. Unfortunately, there is a lack of studies in the recreation and parks field using geographic information based on the first law of geography. In this study, Moran's I statistics were used to test the first law of geography for detecting the clustered pattern of the dependent variable and the varying spatial relationships between dependent and independent variables by GWR. The results of spatial cluster analysis of national park visitation ratios supported the idea that neighboring counties showed similar patterns of national park visitation. This result confirmed half of the first law of geography, spatial autocorrelation. The remaining half of the law, spatial heterogeneity, was supported by the results of the GWR analysis. Locational information has been highlighted from many studies. Although researchers have found great theory or law without locational information, there could be problems applying the findings in the real world. Spatial context may reduce the gap between academics and practitioners. For example, if research tested a theory by using statistical analysis, the results could not effectively determine where specific social phenomena in the most intensely influential without locational information. Also, global statistics such as the OLS regression model should cover all research areas, but it could not detect the variations within the study areas.

MAUP (Modifiable Areal Unit Problem) theory

Modifiable Area Unit Problem (MAUP) is "the classical theory to solve the effects of spatial scale" (Yang et al., 2007, p.1). This theory presents many different ways to divide the study area into non-overlapping spatial zones (e.g. the different scale of administrative areas) for spatial analysis.

As a specific instance of the modifiable areal unit problem discussed in the field of geography, measures of variation in multilevel models are dependent on the arbitrary size and shape of the areas (Steel & Holt, 1996; Gotway & Young, 2002). As noted by Lee & Wong, (2005), relationships between variables of larger scale are stronger than the relationships between the same variables of smaller scale. This means that the "outcome of statistical analysis of data from different scales or spatial resolution levels do not yield the same results" (Lee & Wong, 2005, p. 9). The problems posed for statistical inference from MAUP have led some to conclude that all methods with results depending on areal units should be discarded and techniques independent of areal units should be used (Openshaw & Taylor, 1981; Openshaw, 1983; Fotheringham, 1997; Fotheringham & Wong, 1991; Fotheringham & Brunsdon, 1999).

The use of census enumeration units for analysis is common among spatial studies. Census data provide a basis for understanding the demographic constructs and socioeconomic characteristics of the target population. Using the U.S. example, population density data can be aggregated at the level of the entire United States, or they can be aggregated to the county subdivision level, as is the case with these map data. The following figure represents all of the U.S Census Bureau enumeration units.

As shown, the geographical areas covered by enumeration units are hierarchal in nature, going from larger regions to smaller regions (See Appendix G). For example, the nation can contain counties; counties can contain census tracts, etc. Lee and Wong (2005) compared the analysis results of income level based on different census levels, block groups vs. tract. The census tract data tended to show more areas with higher income (pp. 85-91). This comparison would be a good example of MAUP stemming from different scales (scale effect).

In this study, the results of the GWR model could reduce the MAUP scale problem to show the various local relationships within the state. The GWR might be a particularly useful statistical technique to examine and, in particular, compare spatial relationships across the study area and it can be complementary to global statistical analysis. This does not mean that the GWR removed the MAUP issue totally because the GWR also must define its appropriate bandwidth size. The MAUP as a challenging issue for the GIS spatial analysis has provided an analytical framework for understanding the spatial distribution involved. Although there are numerous methods, none provides a comprehensive solution that is capable of effectively and accurately quantifying the effects of the MAUP (Dark & Bram, 2007; Johnston et al., 2006). In general, the examination of multi-scale issues is still largely experimental. However many researchers have indicated that locally varying models such as the GWR may not be influenced by MAUP as are traditional global models (Fotheringham et al., 2002, p.144). The visualization of the GWR results provides more specific information: how associated variables differently influenced national park visitation through the local regression coefficients of each independent variable. This information covers all the state including its sub-areas (counties). This kind of multi-scale approach could reduce the MAUP to show local spatial relationships integrated with global regression model results.

Opportunity theory

In the field of outdoor recreation, opportunity theory has been frequently tested from many perspectives. Opportunity theory postulates that “all things being equal, individuals from different segments of society have the propensity to participate in outdoor recreation activities” (Romsa & Hoffman, 1980, p.322). However if barriers exist, some people would participate less and in fewer recreation activities than desired. This study also adopted the opportunity theory as the main background theory and borrowed Pigram’s conceptual frame work to choose and test appropriate factors related to the level of national park visitation. Actually, the previous studies have empirically verified relationships between participation rates (non-participation rate) and selected factors. These studies have evaluated the impact of a variety of variables representing space, time, past behavior, and socioeconomic demographic characteristics (Byrne & Wolch, 2009; Jackson, 1988; Lue et al., 1993; Hudson et al., 2010; Romsa & Hoffman, 1980). Their conclusions verified that the associated factors statistically affect participation or non-participation. For example, Doucouliagos,and Hall (2010) conducted meta-analysis based on five variables (income, education, age, gender, and race) associated with park use constraints:

> "Educated people are less likely to cite knowledge as a constraint to park visitation and, since education is linked to income, they are also less likely to cite cost as a constraint. Again not surprisingly, income has its greatest effect on easing cost as a constraint, while time is more of a constraint. For age, health is the more important factor limiting usage, while time and knowledge are the greatest factors easing constraints to park visitation" (p. 22).

Their findings, excluding age, are in line with the results of this dissertation. However, there were few studies that tested opportunity theory by using spatial data including geocoded secondary data and spatial analysis. From a data perspective, most of the statistical inference tests (specifically for predicting regional or broader area visitation trends) have weak points when the original dataset was collected from a limited time, site, and visitors. Although these efforts provided valuable contributions to creating a site-based planning agenda; however, regional or broader scale predictions should include all possibly associated information of sub-areas and the proportion of visitors and non-visitors in each sub-area. Use of secondary data may reduce this weak point.

A spatial context also can provide more specific information and detect locally different situations to answer how these factors differently affect the participation rate based on locational information. The OLS regression model supported the previous literature related to opportunity theory: Age, gender, and family composition are recognized as affecting recreation participation (Doucouliagos & Hall, 2010; Hatry & Dunn, 1971; Jackson & Henderson, 1995; Scott & Jackson, 1996). This dissertation confirmed that these demographic factors affect national park visitation from a spatial context. In the final model, gender was not a statistically significant variable affecting the level of national park visitation. Otherwise, proportions of those 65 and older in each county would positively affected visitation. Also different family structures influenced the level of visitation in various ways. Individuals who live alone negatively affect visitation while a family structure with children who live with their parents positively influenced the level of visitation. These findings are in line with Goldbloom's statement (1991): "in Texas, the individual least likely to visit a park would be single or a single parent".

This dissertation also tested what/how socioeconomic factors affect visitation. First, education attainment has been regarded as a critical socioeconomic factor. Earlier research (Wolfe, 1964) to more recent work (Sugiyam et al., 2009) insisted that highly educated people were likely to be more recreationally active; this may reflect a further correlation with a higher status occupation and reinforces already present income and class differences. Wellman and Propst (2004) found that the highly educated groups participate more in outdoor recreation because this group wanted to fulfill their needs in Maslow's hierarchy. In this dissertation, the proportion of fewer high school graduations showed a negative relationship. This finding also supports the finding that less educated people participate less in outdoor recreation activity.

Poverty rate showed a negative relationship in all four OLS models. Extended OLS model 3, which included a proportion of the black population, showed the highest regression coefficient of the poverty rate. This result supports Washburne's statement

(1978). He provided recreation researchers with the conceptual definitions for ethnicity and marginality as explanations of underutilization of recreation resources by ethnic/racial group members (specifically blacks). He said, “The marginality perspective suggests that Blacks do not participate because of poverty and various consequences of socioeconomic discrimination…” (1978, p. 176). Additionally, studies of recreation differences based on ethnicity/race reflect the longstanding societal goal of providing equal opportunity for participation in American culture (Martinez et al., 2007; Rose & Paisley, 2010; Washburn & Wall, 1978). Parks have “historically functioned as spaces of social control disciplining working class and racialized bodies, and redirecting ethno-racial and class tensions” (Byrne & Wolch, 2009, p. 755). Explanations of park visitation from leisure studies lack historical specificity and do not account for the spatiality of potential visitors (See Byrne & Wolch, 2009). Further studies of outdoor recreation should also concentrate more on regional spatial inequity of recreation participation beyond the ethnicity/race differences. Although leisure researchers have postulated about reasons for ethno-racial differences in park use, they also suggest that other factors may be responsible for observed differences in national park visitation and use. For example, Bryne (2013, p.381) found that the “three ethno-racial groups showed no statistically significant differences in the frequency of visitation, mode of travel to the park or in attitudes towards the park” located in the Los Angeles areas. His results are contrary to the ethnicity and assimilation/acculturation hypotheses, wherein people of color would be expected to vary significantly from white visitors across the variables.

The results of the GWR model provided more detailed information of spatial inequity of national park visitation and it would be helpful if future studies could investigate additional data about why spatial variations or inequity exist in any state or region.

Management implications

There are several managerial implications that may be derived from this study. The result of this study provides managers and planners with more specific information by using different types of aggregating data. This study suggests a way to develop demand estimation models from the spatial context. The total number of national park visits in Texas is third in the nation. However, the percentage of participation based on the current adult population is 5.18%, which is 6% below the U.S average. Many previous state or regional level reports only described the number of participants of specific outdoor recreation activities. These results are sometime not enough to provide specific information to prepare for future trends in outdoor recreation. The trend in outdoor recreation could be affected by various societal, demographic, and external factors. First, the results of spatial cluster analysis found the hotspot areas are based on national park visitation ratios. This result would be important to determine the circumstances of potential visitors. From a marketing perspective, spatial change in market concentration or dispersion has been used as an important indicator to prepare for the future market demand. This result showed spatially where visitors came from and questions what kinds of circumstances and characteristics of residents would increase national park visitation. As previous literature argued, socio-demographic or economic factors affect visitation from a spatial perspective. If so, are there outliers or variations within the state? To answer these questions, this study initially tested these relationships by using a global regression model as previous studies did throughout the state. The results of global regression would be helpful in finding general trends in the state. However, the global regression model showed positive relationships in some of the associated factors and national park visitation over the state. It does not mean that all sub-areas in the state showed the same positive relationships. There should be variations and different strengths of relationships (sometimes an inverse relationship could exist) between associated factors and visitation.

To reduce these limitations, the local bivariate correlation analysis and the GWR were employed. The local bivariate correlation results showed spatially how each explanatory variable in the final OLS model related with national park visitation. These results, including BiLisa cluster maps, identified what and how each factor influenced constructing cluster areas of higher (or lower) national park visitation. For example, higher recreation spending patterns and national park visitation showed similar cluster patterns. This means that higher recreation spending to purchase outdoor recreation equipment is closely related to national park visitation. On the contrary, the poverty rate and the proportion of lower educational attainment showed a negative spatial autocorrelation with visitation in the cluster area. Most of the hotspot areas in national park visitation showed LH cluster patterns. This means that most counties located in hotspot areas (higher national park visitation ratio than the mean of the Texas visitation ratio) that have lower poverty rates and lower rates of high school graduation led to higher national park visitation. Texas ranks 6th in terms of people living in poverty among U.S states (CNNMoney, 2011). About 18.4% of Texans were impoverished in 2010, up from 17.3% a year earlier, according to Census Bureau data released in 2011. The national average is 15.1%. For residents living in poverty, "the state doesn't offer many services or even make federally-funded benefits easily accessible" (CNNMoney, 2011).

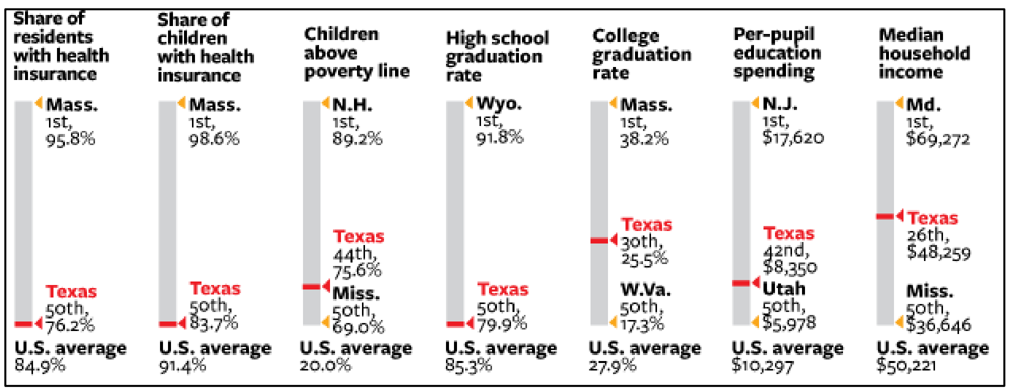


Figure 43 Comparison Texas education, poverty, and health insurance for residents with National average

Sources: National Journal, 2012; American Community Survey, 2009

The state has one of the lowest rates of spending on its citizens per capita and the highest share of those lacking health insurance (www.texaspolicy.com). Among people older than 25, only 79.9 percent of Texans have a high school degree. This is the lowest rate in the country (See Figure 43). Texas has "lagged far behind most other state in rankings on education, poverty, and health insurance for residents" (National Journal.com).

This study found that national park visitation was negatively affected by poverty status and lower educational attainment in the state. If national park managers or planners should consider these situations as problematic social issues and try to make more equivalent opportunities over the state. The findings of this study would be helpful in supporting their decisions and provide more specific information about where these situations emerge in the state. Although the local bivariate correlation analysis could show specific spatial relationships between selected explanatory variables and national park visitation, these separate relationships are not useful to explain (predict) national park visitation. In this context, the GWR was employed to show how associated factors affect visitation. When six selected explanatory variables were used, the same as used in the OLS, the GWR results showed locally varying local relationships among Texas counties throughout the state. This is helpful in applying different alternatives to the different areas.

In general, the results of this study showed general trend predictions in the state by using the OLS model and specifically provide detailed information about the relationships among sub-areas controlling for spatial autocorrelation, including local coefficients of each associated variable. All six selected explanatory variables showed different strengths based on the different counties. There were spatially varied impacts of explanatory variables on visitation. These results indicate that spatial analysis results could provide more detailed information for preparing future national park management and planning. Fragmented results of traditional regression models would be helpful to understand the situation of potential visitors and provide locational information about where managers or planners pay more attention in spending their effort and budget. If family structures change in the future, practitioners could find how and where these changes differently affect the specific outdoor recreation participation. Also the older population might grow far more rapidly (and recreation or leisure services would need to expand far more rapidly) if medical advances are made that enable people to live longer. Along with the trend toward a declining birth rate, it is possible that by the year 2050, elders will comprise 40% to 50% of the U.S. population, not 20% to 25% as is commonly projected (See. Leitner, 2004, p.4).

Spatial analysis and its visualization will be helpful as recreation managers and planners prepare for future societal changes. To make better predictions, park managers or planners could also use projection data of socio-demographic or economic variables based on the GWR results.

Limitations and future research

The GWR, while a powerful tool, should be used with caution. The GWR model may present some limitations because it is an exploratory technique, as statistical inferences from spatial modeling are still being debated (Carlos et al., 2012; Anselin, 2005; Voss et al., 2006). Different outcomes could result depending on the selection of the bandwidth. According to Moon and Farmer (2013), “spatial heteroskedasticity in the data can create a problem for the basic GWR mode” (p.616). There are various automatic techniques for determining bandwidth (Foody, 2003; Wang et al., 2005). However, an adaptive kernel has been popularly regarded as an appropriated technique to find the optimal bandwidth that minimized the AICc when there were no specific criteria or solid evidences for determining bandwidth (Fotheringham et al., 2008). Also, the residuals for the GWR model did not indicate heteroskedasticity in this study.

Another concern is the interpretation of GWR outputs. The interpretation of spatially-varying model coefficients is more complex in the case of the GWR than it is for other global models (Fotheringham et al., 2002; Foody, 2003). Although, this study used ESRI tapestry segmentation data integrated with clustered maps to make a more realistic interpretation, future research should investigate more to explain complex spatial heterogeneity across Texas areas with other underlying spatial information.

Additionally, the coefficients of the GWR model were examined and compared with their paired OLS model, and the results show that there was no difference in the sign of coefficients (e.g., combination of negative and positive local relationships for the same independent variable) between the OLS and GWR models. However, when the GWR model described the strength of local coefficients by visualization, there were some differences in the absolute values of the coefficients. These outputs revealed that the influence of independent variables in the models varies considerably across Texas.

Geographically, the GWR could not show local relationships beyond the study areas at which a model has been developed because of the potential for complex spatial-size interrelationships (Wang et al., 2005; Zhang & Gove, 2005). Beyond these statistical matters, there are some limitations present in the current dataset.

First, the national park visitation ratios of each Texas county (dependent variable) were calculated based on all U.S national park visitations. Although, this study used the Texas boundary as the study area, this study could not detect what percentage of national park visitors of each Texas county visited the Texas national parks. This limitation might affect the result that proximity to the closest parks was not statistically significant. Otherwise this result might support the idea that the tyranny of distance has been reduced by cheaper and faster transport. As David Harvey (1989) insisted, "time–space compression" often occurs as a result of technological innovations that condense spatial and temporal distances, including travels affected by the development of a transportation system. In this context, this result could be interpreted that national park visitors would be less affected by barriers of far distances in order to visit any U.S national park.

The best effort was made to use appropriate socio-demographic and economic variables based on previous literature and conceptual frameworks and provide realistic interpretations; however, a different set of variables could result in different conclusions. For example, educational attainments could be divided into seven categories based on the American community survey criteria: 1) Less than high school; 2) High school graduate (includes equivalency); 3) Some college; 4) Bachelor's degree; 5) Master's degree; 6) Professional school degree; and 7) Doctorate degree. Past literature has tested the relationship between educational attainment and outdoor recreation participation. However, they have not used the same criteria to define higher or lower educational attainment (See, Jinhee et al., 2009). In that sense, there were some difficulties in choosing appropriate variables.

This study also tried to explore the possible underlying factors affecting national park visitation because there were not solid theoretical standards or frameworks. Although the GWR itself has been demonstrated as a powerful exploratory model, and is useful to find exploratory factors, some variables are not confirmatory factors conducted by previous studies. As large spatial databases become available, people from research as well as application domains are exclusively facing a dilemma of being data rich while theory poor (Openshaw, 1991). However, Exploratory Spatial Data Analysis (ESDA) is effective and important in an era of exploding data to assist discovering patterns so that possible relationships can be identified and reasonable research questions or hypotheses established (Anselin, 1996).

This research has also raised several theoretical questions that deserve further exploration. From the extended OLS models, the proportion of the black population showed a positive relationship counter to the findings in previous literature (Bowker et al., 2006; Dwyer, 1994; Floyd, 1994; Henderson, 1996). Although, previous findings helped us understand the role of race/ethnicity in outdoor recreation, there are critical limitations. The previous studies collected data from the sampled populations that were mainly urban respondents (Johnson et al., 1992; Kaplan & Talbot 1988; Van Velsor, 2004). This dissertation used census data to reflect all populations in the state beyond an urban-rural distinction so it may have impacted the results. This finding could also be interpreted to mean that socio-economic status could have a stronger effect on park visitation, thus future studies should investigate why this inverse result occurred.

GWR models account for spatial heterogeneity and generate output that allows mapping and examination of the spatial distribution of coefficients. GWR could help decide which omitted variables should be included and also address the question of how to design and implement regional recreation policy across the state. Local estimates are more stable within the state, but there is still evidence that the magnitude of the impacts vary within the state. Expanded models are needed based on various recreation activities (e.g., hunting, fishing, etc.) and their associated factors. Future research should investigate factors related to the different strengths of relationships specific to northwestern and eastern Texas (e.g., proportion of residents living with parents in each county).

Additionally, park service providers should consider developing different strategies to reflect dynamic social changes in the planning process. Although, past researches in outdoor recreation more focused on ethnic/race minority groups however there has been the other types of minority groups. For example, the single parent family structure has increased in the U.S. and how it affects recreation participation, including park visitation would be examined. If park planners prepare for this kind of social trend or phenomenon and want to apply a different strategy in different areas, spatial analysis could satisfy their needs. They could avoid applying the same policy over a whole area which consumes more time and money.

From a methodological perspective, this study found interesting methodological implications (e.g., rethinking the traditional regression model for recreation demand estimation) and the potential associated with the use of spatial statistics to analyze the relationships between recreation participation and societal factors. This research demonstrated the importance of including spatial variables as part of recreation demand analysis. Relatively little work has used spatial models, which are commonly used in crime, public health, and environmental studies, in the field of recreation. The results of this study demonstrate the usefulness of spatial analysis for detecting various relationships within the state over traditional statistical analysis. However, spatial databases have limited application of these analyses in our field. Origin-destination data and spatial network analysis (e.g., how a transportation system affect visitation) should be employed to identify complex relationships and provide more detailed information for recreation planners or managers.

REFERENCES

American Community Survey (2011). About the American Community Survey (2006-2010). Retrieved from: http://www.census.gov/acs/

Alla, J.B. (2001). Employee recreation. A case study of international brewery limited (IBL) Ilesa. Ilorin Journal of Health, *Physical Education and Recreation*, 3(1), 20-26.

Allmendinger, P. (2009). *Planning theory*. Hampshire: Palgrave MacMillan UK.

Anselin, L. (1993). "SpaceStat: A Program for the Statistical Analysis of Spatial Data," Technical Report 93106–4060, Department of Geography, University of California at Santa Barbara: NCGIA.

Anselin, L. (1994). Exploratory spatial data analysis and geographic information systems. In Painho, M. (ed.), *New tools for spatial analysis.* Luxembourg: Eurostate, 45–54.

Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, *27*(2), 93-115.

Argyle, M. (1996). *The social psychology of leisure*. London: Penguin UK.

Bao, X., Dhliwayo, J., Heron, N., Webb, D. J., & Jackson, D. A. (1995). Experimental and theoretical studies on a distributed temperature sensor based on brillouin scattering. *Lightwave Technology, Journal of, 13*(7), 1340-1348.

Bella, L. (1992). *The Christmas imperative: leisure, family, and women's work*. Halifax, Nova Scotia: Fernwood Publishers.

Bialeschki, M. D., Michener, S., & Henderson, K. A. (1994). Re-entering leisure: transition within the role of motherhood. *Journal of Leisure Research*, *26*(1), 57-74.

Blanco, J. (2009). *Park Use Characteristics, Constraints, and Desired Strategies to Reduce Constraints: A Comparison of Single Parent and Dual Parent Leisure Behavior* (Doctoral dissertation, The Pennsylvania State University).

Bowker, J. M., English, D. B.K. & Cordell, H. K. (1999). "Projections of outdoor recreation participation to 2050". In *(Principal Investigator) Outdoor*

Recreation in American life: A National Assessment of Demand and Supply Trends, Edited by: Cordell, H. K. Thousand Oaks, CA: Sage.

Bowker, J. M., Murphy, D., Cordell, H. K., English, D. B., Bergstrom, J. C., Starbuck, C. M., ... & Green, G. T. (2006). Wilderness and primitive area recreation participation and consumption: an examination of demographic and spatial factors. *Journal of Agricultural and Applied Economics*, 38(2), 317.

Bowker, J. M., Murphy, D., Cordell, H. K., English, D. B. K., Bergstrom, J. C., Starbuck, C. M., ... & Reed, P. (2007). Wilderness Recreation Participation: Projections for the Next Half Century. 2007. *USDA Forest Service Proceedings RMRS-P-49.*

Bowker,J.M., Askew, Ashley E., Cordell,H. Ken, Betz, CarterJ., Zarnoch, Stanley J., & Seymour, Lynne. (2012). *Outdoor recreation participation in the United States – projections to 2060*. USDA Forest ServiceGeneral Technical Report.

Buckley, R. (2003). Adventure tourism and the clothing, fashion and entertainment industries. *Journal of Ecotourism*, 2(2), 126-134.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity.*Geographical Analysis*, *28*(4), 281-298.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1997). Geographical instability in linear regression modelling—a preliminary investigation. *New Techniques and Technologies for Statistics II*, 149-158.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1999). Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Science*, 39(3), 497-524.

Bultena, G. L., & Field, D. R. (1978). Visitors to national parks: A test of the elitism argument. *Leisure Sciences*, 1(4), 395-409.

Bultena, G. L., & Field, D. R. (1980). Structural effects in national park going. *Leisure Sciences*, 3(3), 221-240.

Burt, O. R., & Brewer, D. (2012). Estimation of Net Social Benefits from Outdoor Recreation. *Econometrica*, *39*(5), 813-27.

Burton, N. W., & Turrell, G. (2000). Occupation, hours worked, and leisure-time physical activity. *Preventive Medicine*, 31(6), 673.

Bureau of labor statistics, (2011). American Time Use Survey Retrieved from http://www.bls.gov/tus/charts/leisure.htm

Burns, R. C., & Graefe, A. R. (2006). Toward understanding recreation fees: Impacts on people with extremely low income levels. *Journal of Park and Recreation Administration*, 24(2), 1-20.

Bustam, T., Young, A., & Todd, S. (2003, April). Environmental sensitivity and youthful participation in outdoor recreation. In *Proceedings of the 2003 Northeastern Recreation Research Symposium* (pp. 270-275).

Brown, T. L., Decker, D. J., Riley, S. J., Enck, J. W., Lauber, T. B., Curtis, P. D., & Mattfeld, G. F. (2000). The future of hunting as a mechanism to control white-tailed deer populations. Wildlife Society Bulletin, 28(4), 797-807.

Calvo, E., & Escolar, M. (2003). The local voter: A geographically weighted approach to ecological inference. *American Journal of Political Science,* 47(1), 189-204.

Byrne, J. (2012). When green is White: The cultural politics of race, nature and social exclusion in a Los Angeles urban national park. *Geoforum*, *43*(3), 595-611.

Byrne, J., & Wolch, J. (2009). Nature, race, and parks: past research and future directions for geographic research. *Progress in Human Geography*, *33*(6), 743-765.

Clawson, M. (1959). *Methods of measuring the demand for and value of outdoor recreation* . Washington, D.C: Resources for the Future Inc.

Charlton, M., Fotheringham, S., & Brunsdon, C. (2009). Geographically Weighted Regression: White Paper. National Centre for Geocomputation, National University of Ireland Maynooth.

Chavas, J. P., Stoll, J., & Sellar, C. (1989). On the commodity value of travel time in recreational activities. *Applied Economics*, 21(6), 711-722.

Chubb, M., & Chubb, H. (1981). *One third of our time? An introduction to recreation behaviour and resources*. New York: John Wiley & Sons Inc.

Cordell, H. K. (2004). *Outdoor recreation for 21st Century America*: a report to the Nation, the National Survey on Recreation and the Environment. MA: Venture Pub.

Cordell, H. K. (2012). *Outdoor recreation trends and futures: a technical document supporting the Forest Service 2010 RPA Assessment*. Washington, D.C: US Department of Agriculture, Forest Service, Southern Research Station.

Cook, R. L., & Mccleary, K. W. (1983). Redefining vacation distances in consumer minds. *Journal of Travel Research*, 22(2), 31-34.

Cliff, Andrew D., & J. Keith Ord. (1981). *Spatial Processes: Methods and Applications*. London: Pion.

Craig, W. (1972). Recreational activity patterns in a small negro urban community: The role of the cultural base. *Economic Geography*, *48*(1), 107-115.

Cressie, Noel. (1991). *Statistics for Spatial Data*. New York:Wiley.

Crompton, J. L. (2002). The rest of the story. *Journal of Leisure Research*, 34(1), 93-102.

Cui, Y. (2010). *A comparison of ordinary least square and spatial regression models for estimating recreational boat ownership in florida.* (Order No. 3435234, Michigan State University). *ProQuest Dissertations and Theses,* 261. Retrieved from http://lib-ezproxy.tamu.edu:2048/login?url=http://search.proquest.com/docview/816022221?accountid=7082. (816022221).

Decker, D. J., Brown, T. L. & Knuth, B. A. (1996). Human dimensions research: its importance in natural resource management. In *Natural Resource Management*: *The Human Dimension* (Ewert, A. W., ed.), pp. 29–52. Boulder, CO: Westview Press.

Ding, Y., & Fotheringham, A. S. (1992). The integration of spatial analysis and GIS. *Computers, Environment and Urban Systems*, *16*(1), 3-19.

Doucouliagos, H., & Hall, J. (2010). Park Visitation, Constraints, and Satisfaction: A Meta-Analysis. School working paper Faculty of Business and Law, Deakin University.

Driver, B. L., & Brown, P. J. (1978). The opportunity spectrum concept and behavioral information in outdoor recreation resource supply inventories: a rationale. In:

Integrated inventories of renewable natural resources: proceedings of the workshop, January 1978, Tucson, Arizona (Edited by HG Lund et al.).USDA Forest Service, General Technical Report, (RM-55), 24-31.

Driver, B. L., & Rosenthal, D. H. (1983). Measuring and improving effectiveness of public outdoor recreation programs. Harpers Ferry, W.V. Dept. of Human Kinetics and Leisure Studies. Dec.11–141979. Report on a recreation output resources workshop, Washington, D.C.: George Washington Univ.

Dubin, R. (2008). *Spatial weights.* London: SAGE Publication

Dupuis, S. L., & Smale, B. J. (2000). Bittersweet journeys: Meanings of leisure in the institution-based caregiving context. *Journal of Leisure Research*, 32(3), 303-340.

Eagles, P. F. (2004). Trends affecting tourism in protected areas. In *Policies, methods and tools for visitor management–Proceedings of the second International Conference on Monitoring and Management of Visitor Flows in Recreational and Protected areas*.

Eitzen, D. S., Zinn, M. B., & Smith, K. E. (2011). *Social problems* (10th edition). Boston: Allyn and Bacon.

ESRI. (2009). *Methodology statement: ESRI Data–Market Potential*. CA: Redlands.

ESRI. (2010). Tapestry Segmentation Reference Guide. Retrieved from http://www.esri.com/library/fliers/pdfs/tapestry_segmentation.pdf.

Faraway, J., (2002). Practical Regression and ANOVA Using R. URL: http://cran.r-project.org/doc/contrib/Faraway-PRA.pdf, last accessed December, 2012.

Fan, Y., French, S. A., & Das, K. V. (2012). Family structure and park use among parents. *American Journal of Preventive Medicine*, *43*(5), 520-526.

Feather, P., Hellerstein, D., & Tomasi, T. (1995). A discrete-count model of recreational demand. *Journal of Environmental Economics and Management*, *29*(2), 214-227.

Fesenmaier, D. R. (1988). Integrating activity patterns into destination choice models. *Journal of Leisure Research*, 20(3): 175–191.

Florida, State of (2008). Outdoor Recreation in Florida: Florida's Statewide Comprehensive Outdoor Recreation Plan. Department of Environmental Protection, Division of Recreation and Parks.

Floyd, M. F., Gramann, J. H., & Saenz, R. (1993). Ethnic factors and the use of public outdoor recreation areas: The case of Mexican Americans. *Leisure Sciences*, 15(2), 83-98.

Floyd, D. (1999). Race, ethnicity and use of the National Park System. *Social Science Research Review,* 1(2): 1–24.

Floyd, M. F., Nicholas, L., Lee, I., Lee, J. H., & Scott, D. (2006). Social stratification in recreational fishing participation: Research and policy implications. *Leisure Sciences*, 28(4), 351-368.

Foody, G. M. (2003). Geographical weighting as a further refinement to regression modelling: an example focused on the NDVI–rainfall relationship. *Remote sensing of Environment*, *88*(3), 283-293.

Fotheringham, A. S., Charlton, M., & Brunsdon, C. (1997). Measuring spatial variations in relationships with geographically weighted regression. In*Recent developments in spatial analysis* (pp. 60-82). Springer Berlin Heidelberg.

Fotheringham, A. S. (1997). Trends in quantitative methods 1: stressing the local. *Human Geography*, 21, 88-96.

Fotheringham, A. S., Brunsdon, C. F., & M. E. Charlton, (1998), Geographically Weighted Regression: A Natural Evolution of the Expansion Method for Spatial Data Analysis. *Environment and Planning* A, 30:1095–1927.

Fotheringham, A. S., & Brunsdon, C. (1999). Local forms of spatial analysis.*Geographical Analysis*, 31(4), 340-358.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically weighted regression.* New York: Wiley.

Fotheringham, A. S., & Rogerson, P. A. (Eds.). (2009). *The SAGE handbook of spatial analysis*. CA: Sage Publications

Fotheringham, A. S., & Wong, D. W. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and planning A*, *23*(7), 1025-1044.

Fox, J., (1997), *Applied Regression Analysis, Linear Models, and Related Methods, Thousand Oaks*, CA: Sage Publications.

Freysinger, V. J., & Chen, T. (1993). Leisure and family in China: The impact of culture. *World Leisure & Recreation*, 35(3), 22-24.

Garber-Yonts, B. E. (2005). *Conceptualizing and measuring demand for recreation on national forests: A review and synthesis*. US Department of Agriculture, Forest Service, Pacific Northwest Research Station. Gartner, W. C., & Lime, D. W. (Eds.). (2000). Trends in outdoor recreation, leisure and tourism. CABI.

Getis, A. & D. A. Griffith, (2002), Comparative Spatial Filtering in Regression Analysis," *Geographical Analysis*, 34, 130–140.

Getis, A. & J.K. Ord (1992). The analysis of spatial association by use of distance statistics, *Geographical Analysis* 24, 189-206.

Gómez, E. (2008). Race, ethnicity, recreation, and leisure: An assessment of research gaps. *Recreation Visitor Research: Studies of Diversity*, 75-84.

Goldbloom, A. (1991). *Citizen survey on Texas state parks. Austin, TX: Statewide Planning and Research,* Public Lands Division, Texas Parks and Wildlife Department.

Goldsmith, J. (1994). Designing for diversity. *National Parks*, 68(5/6), 20-21.

Goodchild, M., Haining, R., & STEPHEN, W. (1992). Integrating GIS and spatial data analysis: problems and possibilities. *International Journal of Geographical Information Systems*, 6(5), 407-423.

Gordon, C., Gaitz, C. M., & Scott, J. (1976). Leisure and lives: Personal expressivity across the life span. *Handbook of Aging and the Social Sciences*, 310-341.

Griffith, D. A., & Layne, L. J. (2000). *A casebook for spatial statistical data analysis: a compilation of analyses of different thematic data sets*. Oxford University Press, USA.

Grimshaw, D. J. (2000). *Bringing geographical information systems into business*. New York:Wiley.

Haas, P. L. (1981). *Leisure activity as a stress-coping behavior*. Pennsylvania State University. Department of Recreation and Parks.

Haas, G. E. (2001). *Visitor capacity in the national park system*. US Government Documents (Utah Regional Depository), 426.

Haas, G. E., Wells, M. D., Lovejoy, V., & Welch, D. (2007). *Estimating Future Recreation Demand: A Decision Guide for the Practitioner*. US Department of the Interior, Bureau of Reclamation.

Hall, C. M., Jenkins, C. L., & Singh, S. (2009). Archetypal approaches to implementation and their implications for tourism policy. *Tourism Recreation Research*, 34(3), 235-245.

Hall, C. M., & Page, S. (2006). *The geography of tourism and recreation: environment, place and space*. Kentucky: Psychology Press.

Hanemann, W. M. (1978). *A methodological and empirical study of the recreation benefits from water quality improvement*. Department of Agricultural and Resource Economics, University of California.

Hanemann, W. M. (1984). Discrete/continuous models of consumer demand. *Journal of the Econometric Society*, 52, 541-561.

Hanemann, W. M. (1999). *Welfare analysis with discrete choice models.Valuing Recreation and the Environment*, Edward Elgar.

Hardman, A. (2003). *Physical Activity and Health: An Evidence-Based Assessment*. London: Routledge.

Harris, P., Fotheringham, A. S., & Juggins, S. (2010). Robust geographically weighted regression: a technique for quantifying spatial relationships between freshwater acidification critical loads and catchment attributes. *Annals of the Association of American Geographers*, *100*(2), 286-306.

Hatry, H.P. and Dunn, D.R. (1971). Measuring the effectiveness of local government services. Washington, D.C.: The Urban Institute.

Hayslip, B. J. R., & Panek, P. E. (1989). *Adult Development and Aging,* New York: Harper and Row.

Henderson, K. A., Bialeschki, M. D., & Liemohn, W. (1991). Women and the meaning of physical recreation. In *Abstracts of research papers presnted at the San Francisco, California Convention of American Alliance for Health, Physical*

Education, Recreation and Dance in the Research Consortium Meetings, 1991.. American Alliance for Health, Physical Education, Recreation and Dance.

Ho, C. H., Sasidharan, V., Elmendorf, W., Willits, F. K., Graefe, A., & Godbey, G. (2005). Gender and ethnic variations in urban park preferences, visitation, and perceived benefits. *Journal of Leisure Research*, 37(3), 281-306.

Howard, D. R., & Crompton, J. L. (1980). Financing, managing, and marketing recreation and park resources. *Journal of Parks and Recreation Administration,* 3(1), 33-40.

Hung, K. (2003). *Achieving cultural diversity in wilderness recreation: A study of the Chinese in Vancouver*. Unpublished master's thesis, the University of Waterloo, Waterloo, Ontario, Canada.

Hussain, A., Armstrong, J. B., Brown, D. B., & Hogland, J. (2007). Land-use pattern, urbanization, and deer-vehicle collisions in Alabama. *Human-Wildlife Interactions*, 1, 89-96.

Jackson, E. L., & Henderson, K. A. (1995). Gender-based analysis of leisure constraints. *Leisure Sciences*, 17(1), 31-51

Jackson, E. L., & Scott, D. (1999). Jackson, E. L., & Scott, D. (1999). Constraints to leisure. *Leisure studies: Prospects for the twenty-first century*, 299-321. PA: Venture Publishing Inc.

Jelinski, D. E., & Wu, J. (1996). The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology*, *11*(3), 129-140.

Jensen, C. R., & Guthrie, S. (2006). *Outdoor recreation in America*. Human Kinetics.

Johnson, C. Y., Horan, P. M., & Pepper, W. (1997). Race, Rural Residence, and Wildland Visitation: Examining the Influence of Sociocultural Meaning1.*Rural Sociology*, *62*(1), 89-110.

Johnson, D. H., & O'Neil, T. A. (2001). *Wildlife-habitat relationships in Oregon and Washington*. Oregon State Univ Press.

Karlis, G., & Dawson, D. (1995). Ethnicity and recreation: Concepts, approaches and programming. *Canadian Ethnic Studies*, 27, 166-179.

Katz, A. (2002). Where I come from we don't talk about that. Exploring sexuality & culture among blacks, Asians and Hispanics. *Obstetric and Neonatal Nurses*, *6*(6), 533.

Kelly, J. R. (1980). Outdoor recreation participation: A comparative analysis. *Leisure Sciences*, 3(2), 129-154.

Kelly, J. R. (1983). *Leisure identities and interactions*. New South Wales: George Allen & Unwin.

Kelly, J. R. (1996). *Leisure (3rd).* Boston: Allynand Bacon.

Kleiber, D. A. (1999). Leisure experience and human development: A dialectical interpretation. New York: Basic Books.

Kim, S. I., & Fesenmaier, D. R. (1990). Evaluating spatial structure effects in recreation travel. *Leisure Sciences*, 12(4), 367-381.

King, M., Ruggles, S., Alexander, T., Leicach, D., & Sobek, M. (2010). Integrated Public Use Microdata Series, Current Population Survey: Version 2.0.[Machine-readable database]. *Minneapolis: University of Minnesota*.

Kochhar, R., Suro, R., & Tafoya, S. M. (2005). *The new Latino south: The context and consequences of rapid population growth*. Washington, DC: Pew Hispanic Center.

Kwate, N. O. A., Loh, J. M., White, K., & Saldana, N. (2012). Retail redlining in New York City: Racialized access to day-to-day retail resources. *Journal of Urban Health*, 90(4), 1-21.

Kyle, G. T., Graefe, A. R., & Absher, J. D. (2002). Determining appropriate prices for recreation on public lands. *Journal of Park and Recreation Administration*, 20(2), 69-89.

Le, L. (2012). Hispanic and White Visitors in US National Parks: Meta-Analysis of Visitor Use Survey. *Journal of Park and Recreation Administration*, 30(4), 1-20.

Lee, B., Graefe, A., & Burns, R. (2003). Analysis of demographic segmentation of local residents in the Columbia River Gorge National Scenic Area: A GIS approach for the potential park visitors. 2003 Annual Conference of Population Association of America, Minneapolis, MN.

Lee, J. H., Scott, D., & Floyd, M. F. (2001). Structural inequalities in outdoor recreation participation: a multiple hierarchy stratification perspective. *Journal of Leisure Research*, 33(4), 427-449.

Leeworthy, V.R., & Bowker, J.D. Hospital, and E.A. Stone. (2005). Projected Participation in Marine Recreation: 2005 & 2010. NOAA, SEA Division, National Ocean Service.

Leitner, M. J., & Leitner, S. F. (2004). *Leisure in later life*. NY: Haworth Press.

Leung, Y., Chang-Lin, M., & Wen-Xiu, Z. (2000). Testing for spatial autocorrelation among the residuals of the geographically weighted regression. *Environment and Planning*, 32(5), 871-890.

Leung, Y. F., Shaw, N., Johnson, K., & Duhaime, R. (2002). More than a database: Integrating GIS data with the Boston Harbor Islands carrying capacity study. In *The George Wright Forum* 19 (1), 69-78.

Légaré, A. M., & Haider, W. (2008). Trend analysis of motivation-based clusters at the Chilkoot Trail National Historic Site of Canada. *Leisure Sciences*, 30(2), 158-176.

Li, L., Chang, H. J., Yeh, H. I., Hou, C. J. Y., Tsai, C. H., & Tsai, J. P. (2010). Factors associated with leisure participation among the elderly living in long-term care facilities. *International Journal of Gerontology*, *4*(2), 69-74.

Lin, C. H., & Morais, D. B. (2008). The Spatial Clustering Effect of Destination Distribution on Cognitive Distance Estimates and Its Impact on Tourists' Destination Choices. *Journal of Travel & Tourism Marketing*, 25(3-4), 382-397.

Longley, P. A., Goodchild, M. F., & Rhind, D. W. (2008). *Geographical Information Systems and Science.* NJ: John Wiley & Sons.

Lu, J., & Tang, G. A. (2011, July). The spatial distribution cause analysis of theft crime rate based on GWR model. In *Multimedia Technology (ICMT), 2011 International Conference on* (pp. 3761-3764). IEEE.

Lucas Jr, R. E. (1990). *Supply-side economics: An analytical review*. Oxford Economic Papers, 293-316.

Ludlow, K. (2007). *Recreation site-facility master planning process overview and summary*. United States department of agriculture forest service general technical report, 698, 81.

Manfredo, M. J., Driver, B. L., & Brown, P. J. (1983). A test of concepts inherent in experience based setting management for outdoor recreation areas. *Journal of Leisure Research*, 15(3), 263-283.

Manning, R. E. (2011). *Studies in outdoor recreation: Search and research for satisfaction*. Oregon State University Press

Manning, R. E., & Anderson, L. E. (2012). *Managing outdoor recreation: Case studies in the national parks*. Oxford: CABI.

Mansfeld, Y. (1992). From motivation to actual travel. *Annals of Tourism Research*, 19(3), 399-419.

Marcouiller, D. W., Prey, J., & Scott, I. (2009). The regional supply of outdoor recreation resources: Demonstrating the use of location quotients as a management tool. *Journal of Parks and Recreation Administration,* 27(4), 92-107.

Martin, D. (2004). Apartheid in the Great Outdoors: American advertising and the reproduction of a racialized outdoor leisure identity. *Journal of Leisure Research*, 36(4), 513-535.

Marvasti, A. (2012). Estimating Outdoor Recreation Demand with Aggregate Data: A Revealed Preference Approach. *Ocean & Coastal Management.* 71, 170-175

McLoughlin, J. B. (1969). *Urban and regional planning: A systems approach*. Faber & Faber, Limited

McDonald, J. M., & Hutchison, I. J. (1987). Minority and ethnic variations in outdoor recreation participation: trends and issues. *Therapeutic Recreation Journal*, 21(1), 26-35.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, 3(4), 303-328.

McKercher, B., & Lew, A. A. (2003). Distance decay and the impact of effective tourism exclusion zones on international travel flows. *Journal of Travel Research*, 42(2), 159-165.

McKinney, S. M. (1999). *Outdoor recreation in American life: a national assessment of demand and supply trends*. London: Sage Publications.

McNiel, J. N. (2011). *The Community of Member-Based Nature Parks: An Integration of Spatial, Socio-Demographic, and Interview Data* (Doctoral dissertation, Texas State University). https://digital.library.txstate.edu/handle/10877/2512

Moon, Z. K., & Farmer, F. L. (2013). Deforestation Near Public Lands: An Empirical Examination of Associated Processes. *Society & Natural Resources*, *26*(5), 605-621.

Morse, W. C., Hall, T. E., & Kruger, L. E. (2009). Improving the integration of recreation management with management of other natural resources by applying concepts of scale from ecology. *Environmental Management*, 43(3), 369-380.

Murphy, P. E., & Keller, C. P. (1990). Destination travel patterns: An examination and modeling of tourist patterns on Vancouver Island, British Columbia. *Leisure Sciences*, 12(1), 49-65.

National Park Service. (2012). Visitor Use Statistics Retrieved from http://www.nature.nps.gov/assets/redirects/statsRedirect.cfm

National Park Service. (2011). The National Park System Comprehensive Survey of the American Public retrieved from http://www.nature.nps.gov/socialscience/docs/archive/IMR_Tech_Rep.pdf

Nicholls, S., & Shafer, C. S. (2001). Measuring accessibility and equity in a local park system: the utility of geospatial technologies to park and recreation professionals. *Journal of Park and Recreation Administration,* 19(4), 102-124.

Noresah, M. S., & Ruslan, R. (2009). Modeling urban spatial structure using Geographically Weighted Regression. In 18th World IMACS congress and MODSIM09 international congress on modeling and simulation (pp. 13-17). Canberra, ACT: The Australian National University.

Odland, John. (1988). *Spatial Autocorrelation*. London: Sage Publications.

Office of National Statistics (2003). *General Household Survey (2003) population estimate.* Great Britain: ONS

Openshaw, S., & Taylor, P. J. (1979). A million or so correlation coefficients: three experiments on the modifiable areal unit problem. *Statistical Applications in the Spatial Sciences*, *21*, 127-144.

Openshaw, S. (1983). *The modifiable areal unit problem* (Vol. 38). Norwich: Geo books.

Ostergren, D., Solop, F. I., & Hagen, K. K. (2005). National Park Service fees: Value for the money or a barrier to visitation. *Journal of Park and Recreation Administration*, 23(1), 18-36.

Perez-Verdin, G., Lee, M. E., & Chavez, D. J. (2008). Use of the recreation opportunity spectrum in natural protected area planning and management. *Recreation Visitor Research: Studies of Diversity*, 23.

Phaneuf, D. J., & Smith, V. K. (2005). Recreation demand models. *Handbook of Environmental Economics*, 2, 671-761.

Pigram, J. (1983). *Outdoor recreation and resource management*. Beckenham: Croom Helm Ltd.

Pigram, J. J., & Jenkins, J. M. (2006). *Outdoor recreation management*. London: Routledge.

Philipp, S. F. (1995). Race and leisure constraints. *Leisure Sciences*, 17(2), 109-120.

Poudyal, N. C., Paudel, B., & Tarrant, M. A. (2012). A time series analysis of the impact of recession on national park visitation in the United States. *Tourism Management*. 35, 181-189.

Prato, T., & Fagre, D. (Eds.). (2012). *Sustaining Rocky Mountain Landscapes:" Science, Policy, and Management for the Crown of the Continent Ecosystem"*. London: Routledge.

Prey, J., & Kiefaber, K. (2006). *Outdoor recreation in Wisconsin*. Wisconsin Department of Natural Resources.

Prideaux, B. (2000). The role of the transport system in destination development. *Tourism management*, 21(1), 53-63.

Radeloff, V. C., Hammer, R. B., & Stewart, S. I. (2005). Rural and suburban sprawl in the US Midwest from 1940 to 2000 and its relation to forest fragmentation. *Conservation Biology*, 19(3), 793-805.

Robertson, B. (1995). Factors impacting leisure in middle aged adults throughout the world. *World Leisure & Recreation*, 37(1), 30-38.

Rodrigue, J. P. (2006). Transportation and Urban Form. In *The Geography of Transports Sistems*. Rodrigue, J.P., C. Comtois, and B. Slack (eds.). 171-203. NY: Routledge, Taylor and Francis Group.

Sallis, J. F., & Owen, N. (1999). Physical activity intervention with individuals. In *Physical Activity and Behavioral Medicine*. Sallis JF, Owen N, Eds. Thousand Oaks, CA: Sage, 135-152.

Sessoms, H. D. (1993). Leisure services. Englewood Cliffs, NJ: Prentice Hall Professional.

Schuett, M. A., Jacob, J. S., Lu, J., & Respess, L. (2008). Keeping our charm: Residents, growth, and quality of life issues in a small but growing Texas coastal community. *Journal of Extension*, 46(6).

Schwartz, Z., & Lin, L. C. (2006). The impact of fees on visitation of national parks. *Tourism management*, 27(6), 1386-1396.

Scott, D., & Munson, W. (1994). Perceived constraints to park usage among individuals with low incomes. *Journal of Park and Recreation Administration*,12(4), 79-96.

Shaw, S. M., & Henderson, K. (2005). Gender analysis and leisure constraints: An uneasy alliance". In *Constraints to leisure*, Jackson, E. L. 23–34. State College, PA: Venture.

Shaw, S. M. (1994). Gender, Leisure, and Constraint: Towards a Framework for the Analysis of Women's Leisure. *Journal of Leisure Research*, *26*(1), 8-22.

Shaw, S. M., Bonen, A., & McCabe, J. F. (1991). Do more constraints mean less leisure? Examining the relationship between constraints and participation. *Journal of Leisure Research*, 23(4), 286-300.

Sheffield, E. (2005). Parks and recreation trends in California (2002); *an element of the California outdoor recreation plan*. Sacramento, CA: California State Parks.

Shores, K. A., Scott, D., & Floyd, M. F. (2007). Constraints to outdoor recreation: A multiple hierarchy stratification perspective. *Leisure Sciences*, *29*(3), 227-246.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society*. Series B (Methodological), 238-241.

Smith, M. D. (2005). State dependence and heterogeneity in fishing location choice. *Journal of Environmental Economics and Management,* 50(2), 319-340.

Stedman, R. C., & Heberlein, T. A. (2001). Hunting and rural socialization: Contingent effects of the rural setting on hunting participation. *Rural sociology*, 66(4), 599-617.

Stewart, S. I., Radeloff, V. C., Hammer, R. B., & Hawbaker, T. J. (2007). Defining the wildlandurban interface. *Journal of Forestry*, 105(4), 201-207.

Struglia, R., & Winter, P. L. (2002). The role of population projections in environmental management. *Environmental management*, 30(1), 13-23.

Terrón, J. M., da Silva, J. M., Moral, F. J., & García-Ferrer, A. (2011). Soil apparent electrical conductivity and geographically weighted regression for mapping soil. *Precision Agriculture*, *12*(5), 750-761.

Tita, G. E., & Radil, S. M. (2010). Making space for theory: the challenges of theorizing space and place for spatial analysis in criminology. *Journal of Quantitative Criminology*, *26*(4), 467-479.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46, 234-240.

Tomczyk, A. M. (2011). A GIS assessment and modelling of environmental sensitivity of recreational trails: The case of Gorce National Park, Poland. *Applied Geography*, 31(1), 339-351.

Torkildsen, G. (1992). Leisure and recreation management. London: Routledge.

Torkildsen, G. (2005). *Leisure and recreation management*. London: Routledge.

Tremblay, P. (2005). 13 GIS Techniques in Tourism and Recreation Planning: Application to Wildlife Tourism. *Tourism Research Methods: Integrating Theory with Practice*, 163.

Trice, A. H., & Wood, S. E. (1958). Measurement of recreation benefits. *Land Economics*, 195-207.

Trost, S. G., Pate, R. R., Sallis, J. F., Freedson, P. S., Taylor, W. C., Dowda, M., & Sirard, J. (2002). Age and gender differences in objectively measured physical activity in youth. Medicine and Science in Sports and Exercise, 34(2), 350.

Troy, A., Morgan Grove, J., & O'Neil-Dunne, J. (2012). The relationship between tree canopy and crime rates across an urban–rural gradient in the greater Baltimore region. *Landscape and Urban Planning*, *106*(3), 262-270.

Van Velsor, S.,W. (2004). A qualitative investigation of the urban minority adolescent experience with wildlife. (Order No. 3204337, University of Missouri - Columbia). ProQuest Dissertations and Theses, 248-248. Retrieved from http://lib-ezproxy.tamu.edu:2048/login?url=http://search.proquest.com/docview/305162297?accountid=7082. (305162297).

Virden, R. J., & Walker, G. J. (1999). Ethnic/racial and gender variations among meanings given to, and preferences for, the natural environment. *Leisure Sciences*, 21(3), 219-239.

Ward, M. D., & Gleditsch, K. S. (2008). *Spatial regression models (Vol. 155)*. CA: Sage Publications.

Warnick, R. B., Stevens, T., Schuett, M. A., Kuentzel, W., & More, T. A. (2010, July). Changes in national park visitation (2000–2008) and interest in outdoor activities (1993–2008). In *Proceedings of the 2009 Northeastern Recreation Research Symposium* (pp. 204-13).Wearing, S., & Bowden, I. (1999). Tourism and a changing Public Sector culture for Parks. *Australian Parks and Leisure*, 1(3), 6-8.

Wen, M., Zhang, X., Harris, C. D., Holt, J. B., & Croft, J. B. (2013). Spatial Disparities in the Distribution of Parks and Green Spaces in the USA. *Annals of Behavioral Medicine*, *45*, 18-27.

White, T. H. (1975). The Relative Importance of Education and Income as Predictors in Outdoor Recreation Participation. *Journal of Leisure Research*, 7(3), 191-199.

Whiting, J. W., Larson, L. R., & Green, G. T. (2012, March). Valuing state parks: Accounting for diverse visitor perspectives. In *In: Fisher, Cherie LeBlanc; Watts, Clifton E., Jr., eds. Proceedings of the 2010 Northeastern Recreation Research Symposium. Gen. Tech. Rep. NRS* (p. 94).

Wisconsin, State of. (2006). *Wisconsin's Statewide Comprehensive Outdoor Recreation Plan 2005 – 2010. Madison*, WI: Wisconsin Department of Natural Resources.

Wong, D. W. S., & Lee, J. (2005). *Statistical analysis of geographic information with ArcView GIS and ArcGIS (Vol. 1)*. New York:Wiley.

Yetman, N. R. (1999). *Majority and minority: The dynamics of race and ethnicity in American life*. Boston: Allyn & Bacon.

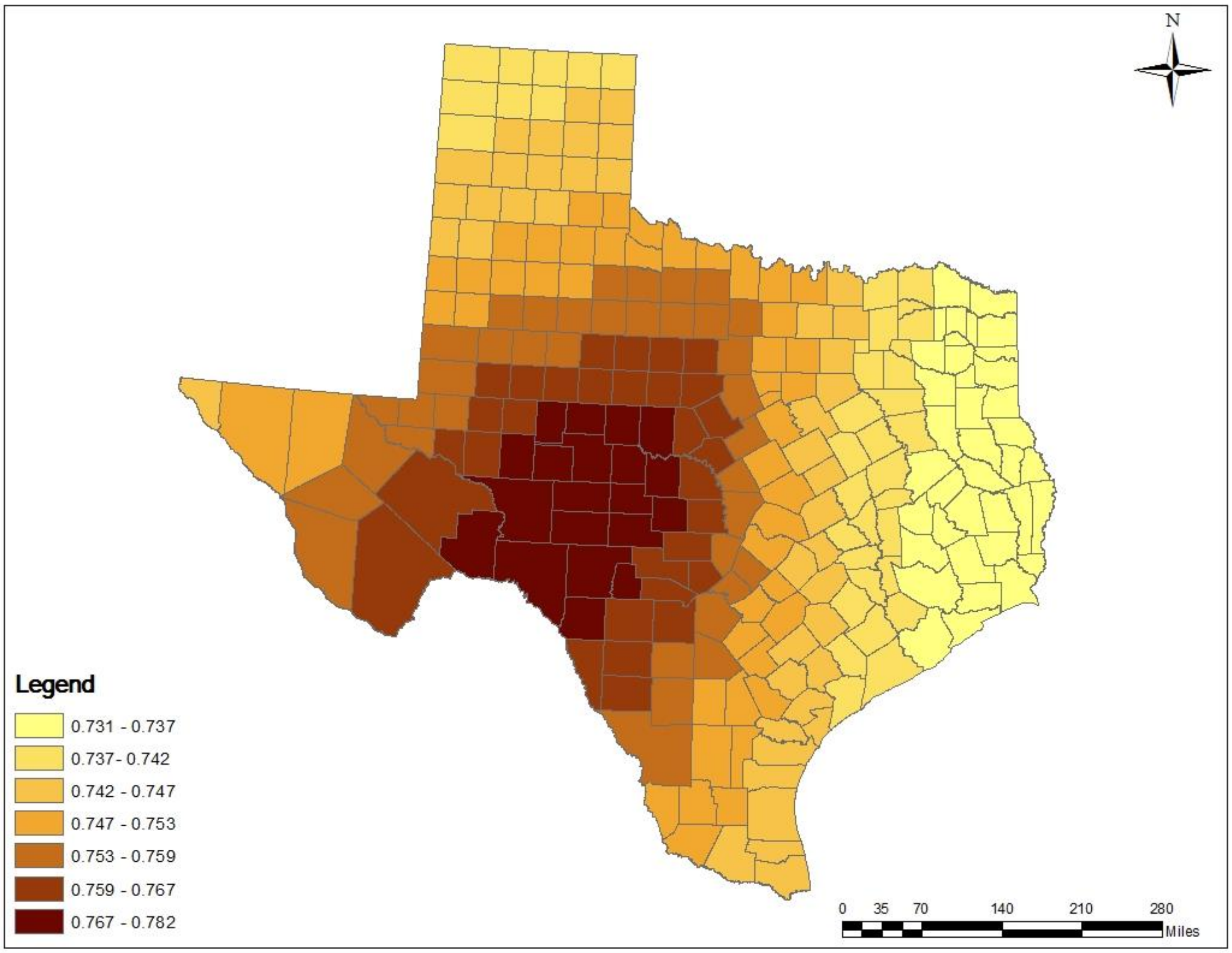
Zanon, D., Doucouliagos, C., Hall, J., & Lockstone-Binney, L. (2013). Constraints to Park Visitation: A Meta-Analysis of North American Studies.*Leisure Sciences*, *35*(5), 475-493.

Zarnoch, Stanley J.; Cordell,H. Ken; Betz, CarterJ.; Langner, Linda. (2010). *Projecting county-level populations under three future scenarios*. Washington, D.C: USDA Forest General Technical Report.

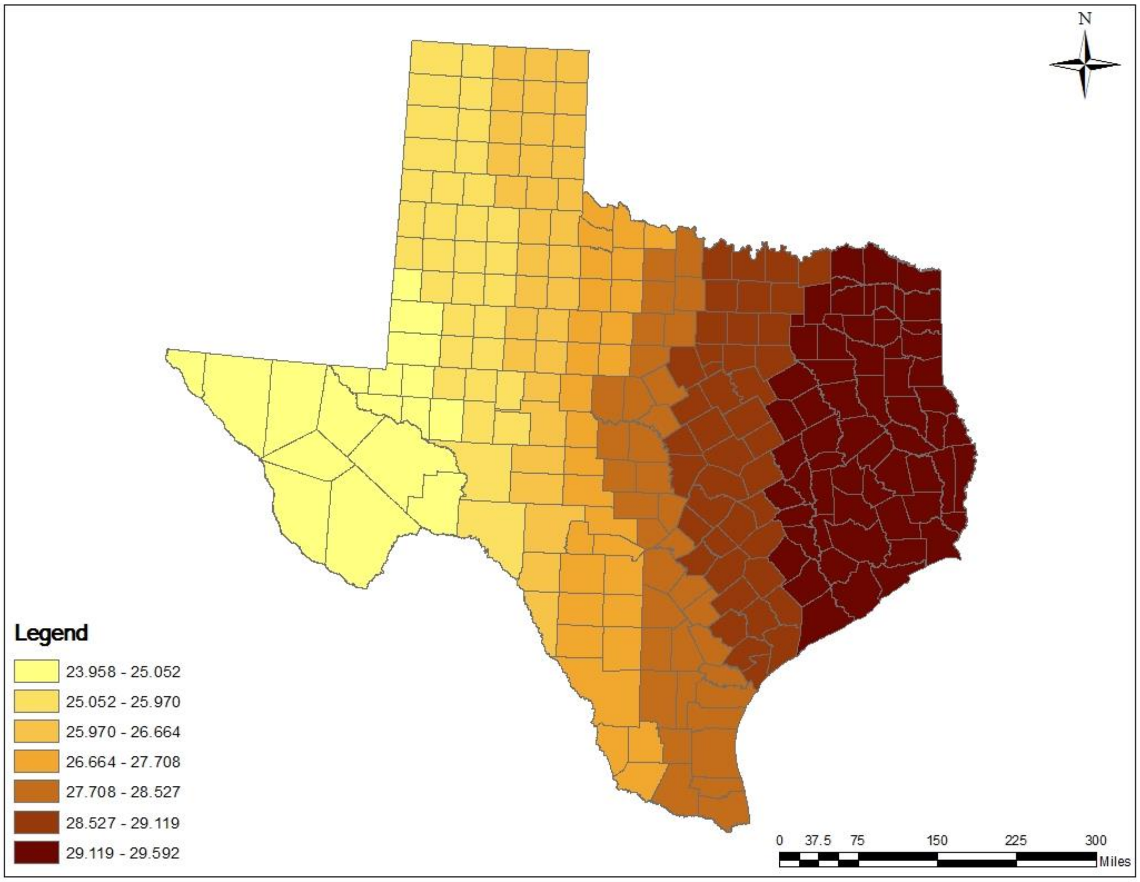
APPENDIX A

GWR RESULTS MAPPING

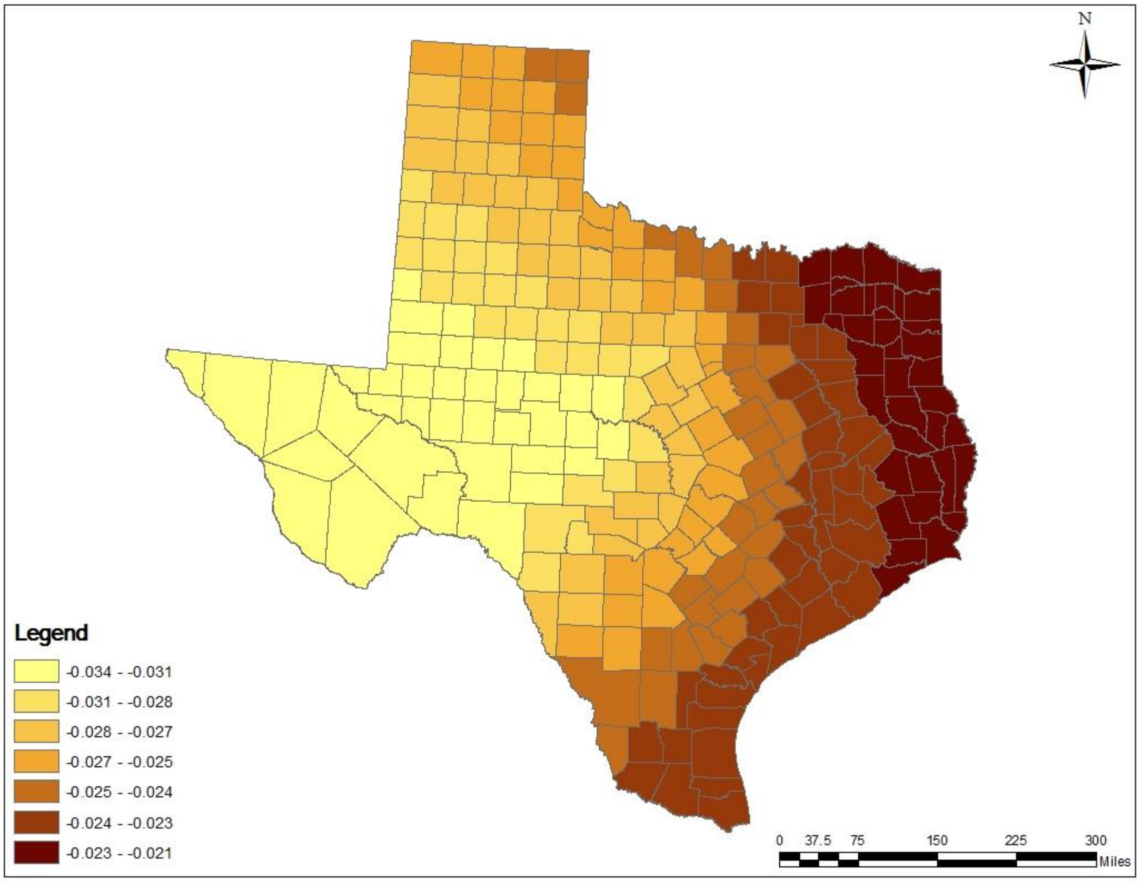
A-1 Spatial variation of local R square



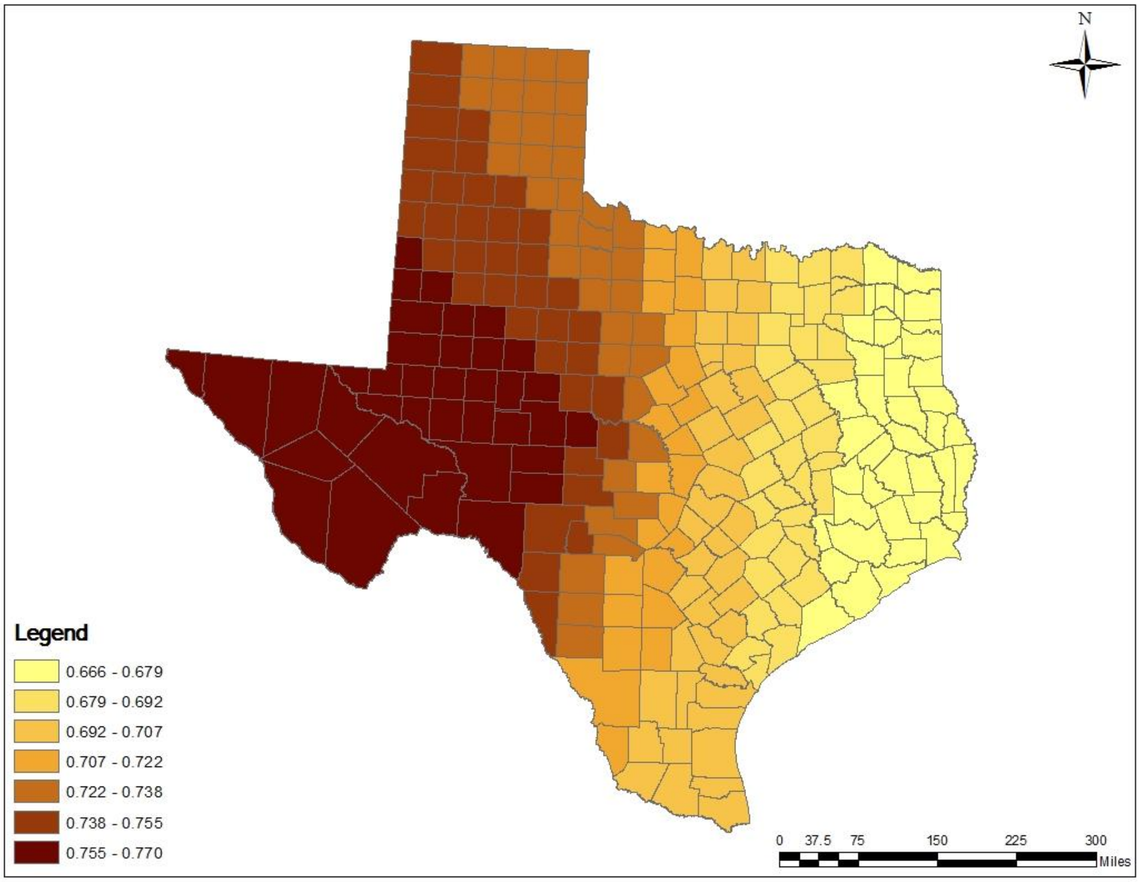
A-2 GWR model stability by county



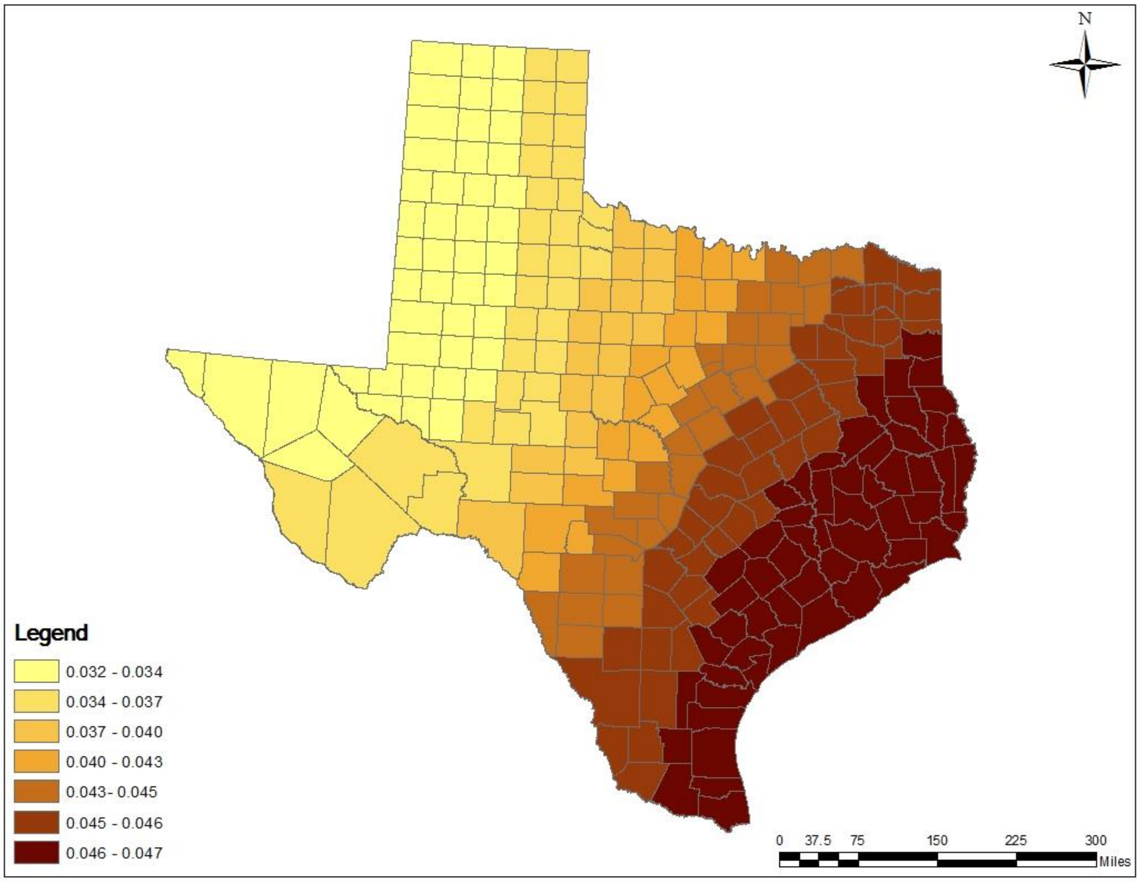
A-3 GWR associations between poverty rate and national park visitation at the county-level



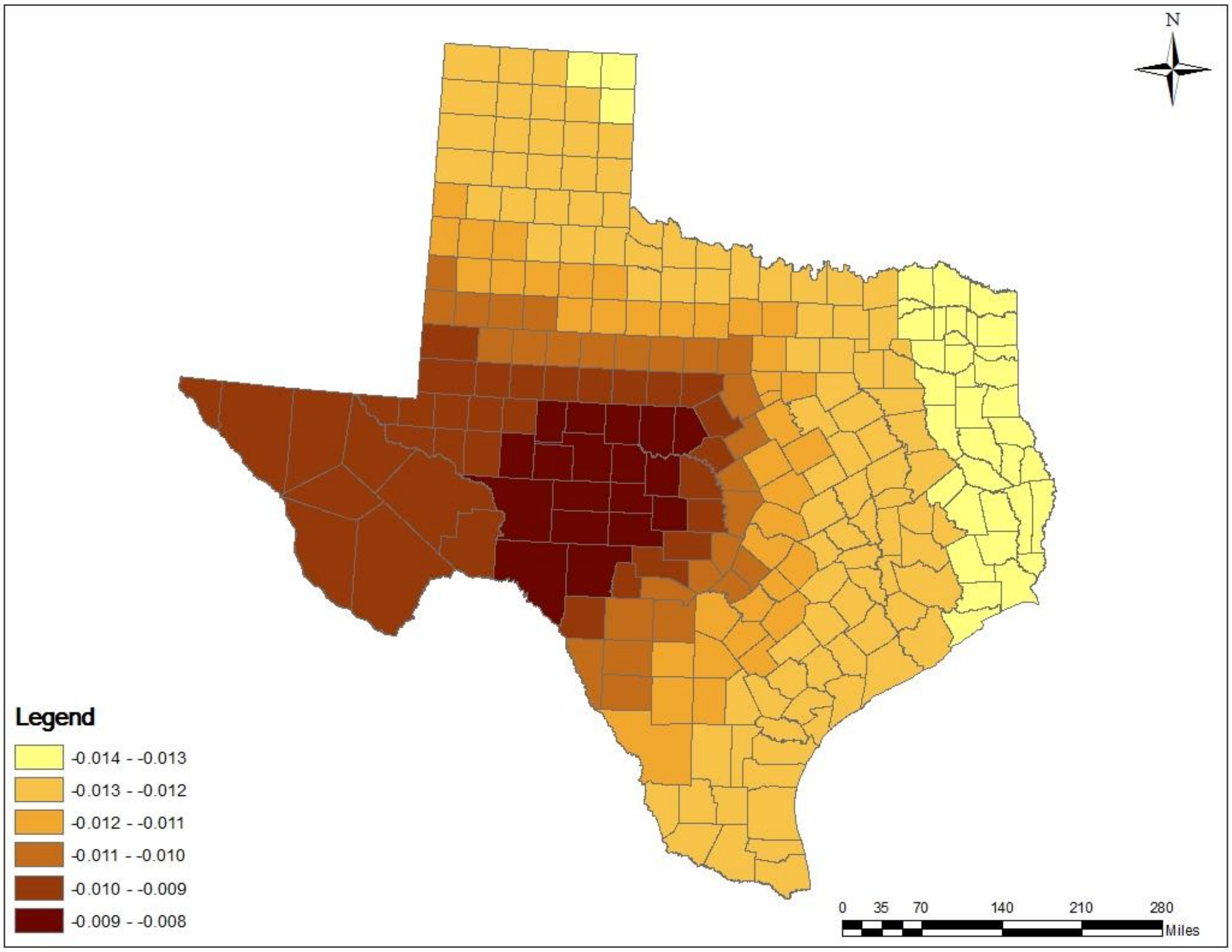
A-4 GWR associations between high recreation spending and national park visitation at the county-level



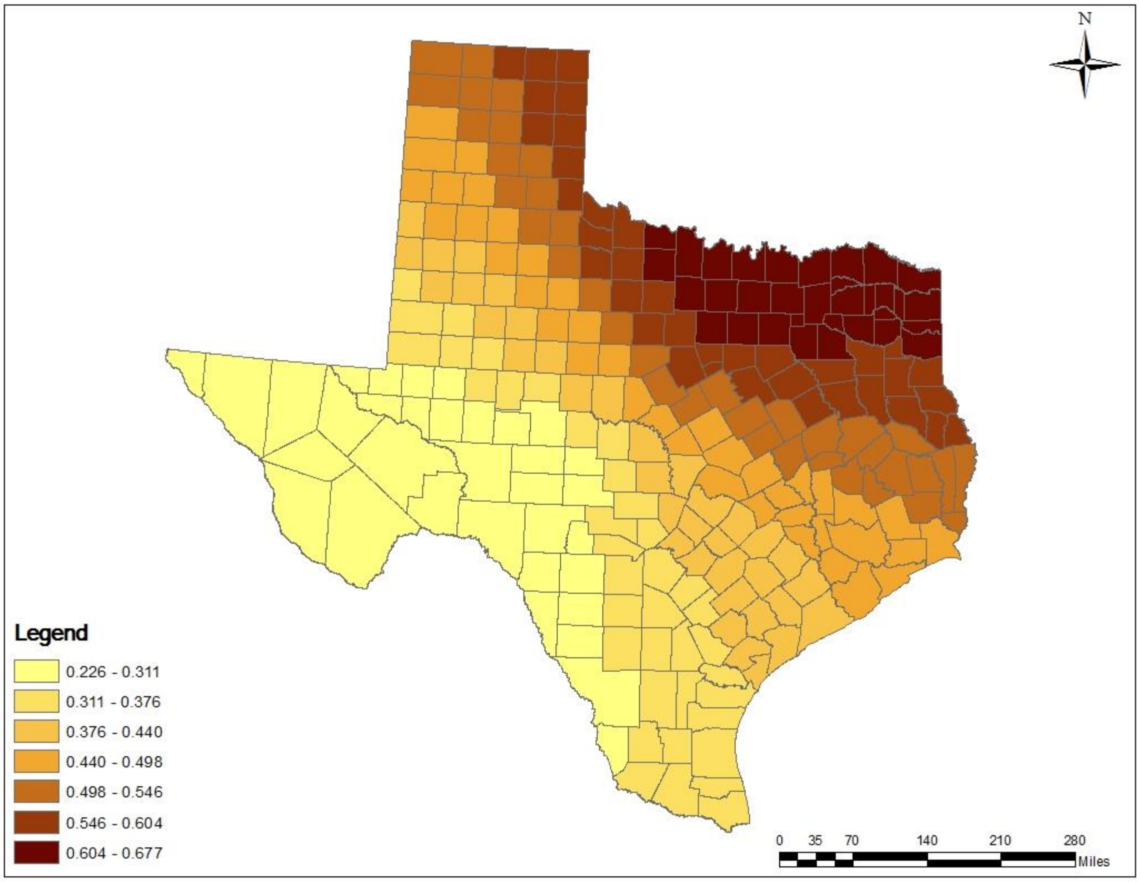
A-5 GWR associations between the proportion of elderly and national park visitation at the county-level



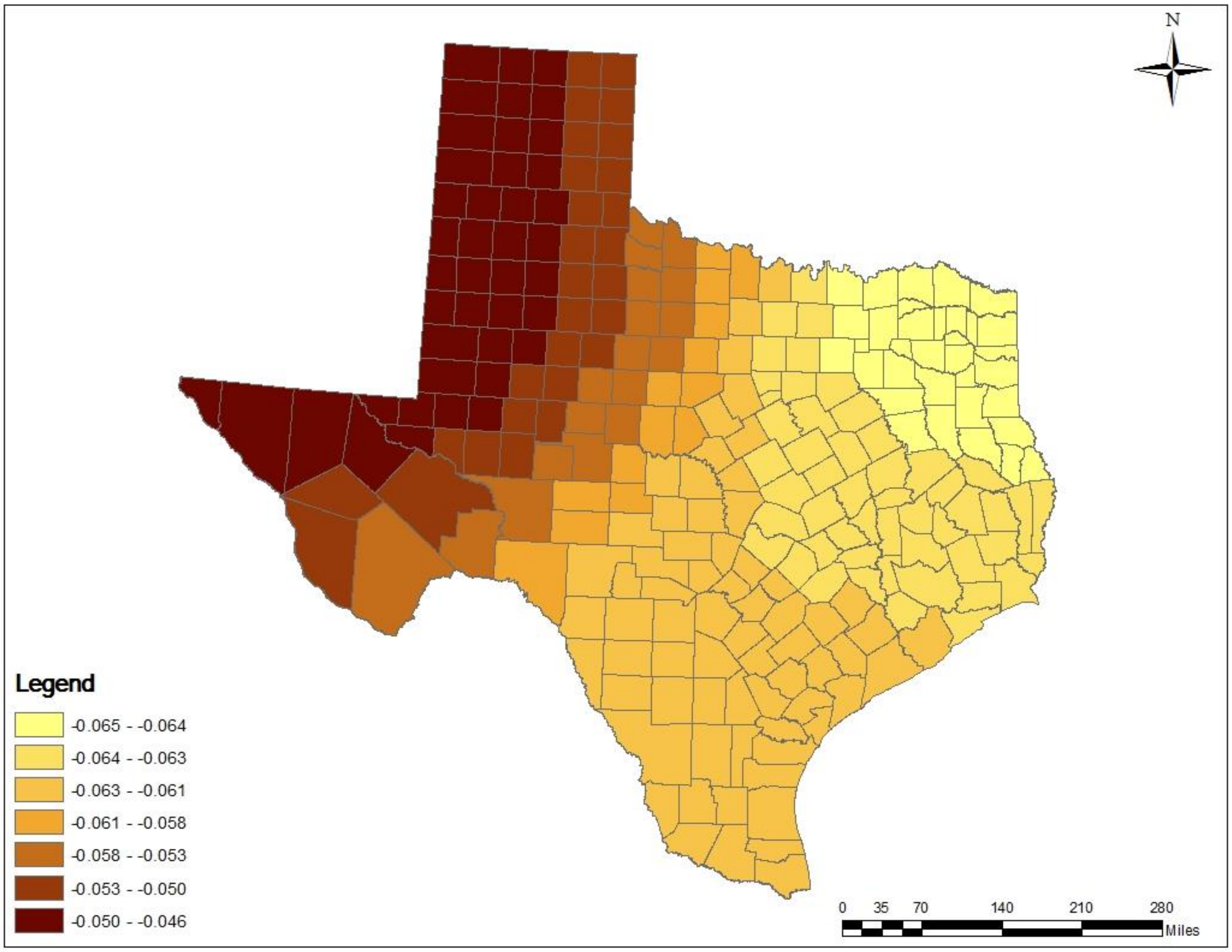
A-6 GWR associations between the proportion of less than high school graduation and national park visitation at the county-level



A-7 GWR associations between the proportion of residents who live alone and national park visitation at the county-level



A-8 GWR associations between the proportion of residents who live alone and national park visitation at the county-level



APPENDIX B

CLUSTER ANALYSIS TABLE FOR NATIONAL PARK VISITATION BY COUNTY LEVEL

Name	LMiZScore	LMiPValue	Type
Denton County, Texas	5.45402000000	0.00000004924	HH
Gillespie County, Texas	10.35260000000	0.00000000000	HH
Burnet County, Texas	6.66122000000	0.00000000003	HH
Travis County, Texas	4.60881000000	0.00000404979	HH
San Saba County, Texas	4.08742000000	0.00004362060	HH
Hays County, Texas	4.16947000000	0.00003053040	HH
Kimble County, Texas	3.72757000000	0.00019333800	HH
Kerr County, Texas	9.11638000000	0.00000000000	HH
Rockwall County, Texas	7.70300000000	0.00000000000	HH
Tarrant County, Texas	2.12367000000	0.03369750000	HH
Bexar County, Texas	2.72326000000	0.00646409000	HH
Bandera County, Texas	6.95224000000	0.00000000000	HH
Blanco County, Texas	11.09650000000	0.00000000000	HH
Williamson County, Texas	3.35261000000	0.00080053400	HH
Comal County, Texas	8.81357000000	0.00000000000	HH
Llano County, Texas	9.96378000000	0.00000000000	HH
Real County, Texas	6.74092000000	0.00000000002	HH
Collin County, Texas	8.30771000000	0.00000000000	HH
Kendall County, Texas	14.11060000000	0.00000000000	HH
Mason County, Texas	4.41778000000	0.00000997200	HH
Edwards County, Texas	3.02341000000	0.00249946000	HH
Waller County, Texas	-2.98220000000	0.00286182000	LH
Wheeler County, Texas	2.30528000000	0.02115080000	LL
Knox County, Texas	4.05871000000	0.00004934360	LL
King County, Texas	4.35340000000	0.00001340410	LL
Hall County, Texas	5.00217000000	0.00000056688	LL
Jones County, Texas	3.30183000000	0.00096057200	LL
Howard County, Texas	1.96945000000	0.04890190000	LL
Lipscomb County, Texas	2.16661000000	0.03026460000	LL
Dickens County, Texas	4.45818000000	0.00000826566	LL
Sterling County, Texas	2.88648000000	0.00389575000	LL
Roberts County, Texas	2.21943000000	0.02645730000	LL
Kent County, Texas	2.93236000000	0.00336396000	LL
Borden County, Texas	2.40002000000	0.01639420000	LL

Motley County, Texas	3.55737000000	0.00037458900	LL
Dallam County, Texas	2.16175000000	0.03063750000	LL
Foard County, Texas	4.05354000000	0.00005044940	LL
Hardeman County, Texas	2.98064000000	0.00287645000	LL
Childress County, Texas	5.15494000000	0.00000025371	LL
Mitchell County, Texas	4.54967000000	0.00000537307	LL
Fisher County, Texas	3.48621000000	0.00048990900	LL
Collingsworth County, Texas	4.39850000000	0.00001090040	LL
Sherman County, Texas	2.28801000000	0.02213720000	LL
Stonewall County, Texas	4.35941000000	0.00001304160	LL
Hartley County, Texas	2.32981000000	0.01981600000	LL
Briscoe County, Texas	3.22369000000	0.00126549000	LL
Hemphill County, Texas	2.83959000000	0.00451715000	LL

Note: High-High (HH): Cluster of high values and Low-Low (LL): cluster of low values.

APPENDIX C

URBANIZATION SUMMARY GROUP BY SEGMENT CODES

U1 Principal Urban Centers I: 08, 11, 20, 21, 23, 27, 35, 44

U2 Principal Urban Centers II: 45, 47, 54, 58, 61, 64, 65

U3 Metro Cities I: 01, 03, 05, 09, 10, 16, 19, 22

U4 Metro Cities II: 28, 30, 34, 36, 39, 52, 60, 63

U5 Urban Outskirts I: 04, 24, 32, 38, 48

U6 Urban Outskirts II: 51, 55, 57, 59, 62

U7 Suburban Periphery I: 02, 06, 07, 12, 13, 14, 15

U8 Suburban Periphery II: 18, 29, 33, 40, 43, 53

U9 Small Towns: 41, 49, 50

U10 Rural I: 17, 25, 26, 31

U11 Rural II: 37, 42, 46, 56

APPENDIX D

LIFEMODE SUMMARY GROUP BY SEGMENT CODES

L1 High Society: 01, 02, 03, 04, 05, 06, 07

L2 Upscale Avenues: 09, 10, 11, 13, 16, 17, 18

L3 Metropolis: 20, 22, 45, 51, 54, 62

L4 Solo Acts: 08, 23, 27, 36, 39

L5 Senior Styles: 14, 15, 29, 30, 43, 49, 50, 57, 65

L6 Scholars and Patriots: 40, 55, 63

L7 High Hopes: 28, 48

L8 Global Roots: 35, 38, 44, 47, 52, 58, 60, 61

L9 Family Portrait: 12, 19, 21, 59, 64

L10 Traditional Living: 24, 32, 33, 34

L11 Factories and Farms: 25, 37, 42, 53, 56

L12 American Quilt: 26, 31, 41, 46

APPENDIX E

TAPESTRY SEGMENTATION OF HH AND LL CLUSTER AREAS

Segment Code	Demographic	Socioeconomic	Residential	Preferences	Cluster type
04	The newest additions to the suburbs, these communities are home to busy, affluent young families.	Residents are well educated: more than 50% of the population aged 25 years and older hold a bachelor's or graduate degree. The median household income is $110,681, more than double that of the US median.	Many work outside their resident county; 35% cross county lines to work (compared to 23% for the United States).	Residents' product preferences reflect their suburban lifestyle. They go online frequently to buy flowers and tickets to sports events, trade and track their investments, do their banking, and make travel plans. Family vacations are a top priority; trips to Disney World, Sea World, and other theme parks are popular destinations. They will readily spend more than $250 a year on high-end sports equipment.	HH
07	*Exurbanite* residents prefer an affluent lifestyle in open spaces beyond the urban fringe. There is little ethnic diversity; most residents are white.	These residents are educated; more than 40% of the population is aged 25 years and older; hold a bachelor's or graduate degree; approximately three in four have attended college. The median household income is $84,522. More than 20% earn retirement income.	Although *Exurbanite* neighborhoods are growing by 1.61% annually, they are not the newest areas.	They are very physically active: when vacationing in the United States, they hike, downhill ski, play golf, attend live theater, and see the sights. This is the top market for watching college basketball and professional football games.	HH

APPENDIX E Continued

Segment Code	Demographic	Socioeconomic	Residential	Preferences	Cluster type
12	A mix of Generation Xers and Baby Boomers with a median age of 31.9 years, this segment is the youngest of the Tapestry Segmentation's affluent family markets. Most of the residents are white; however, diversity is increasing as the segment grows.	*Families* are earning above-average incomes. The median household income is $76,135, higher than the national median.	Most residents live in new single-family housing; more than half the housing units were built in the last 10 years.	Many are beginning or expanding their families, so baby equipment, children's clothing, and toys are essential purchases. They play softball, take the kids to the zoo, and visit theme parks (generally Sea World or Disney World) where they make good use of their digital camera or camcorder.	HH
19	Young, affluent married couples who are starting their families or already have young children. The median age of 33.8 years represents the presence of kids; nearly half of the households include children.	The median household income is $64,880, and the median net worth is $135,190. Fifty-eight percent have attended college; more than 20% hold bachelor's or graduate degrees.	Given the concentration of dual-income families, 71% of households have at least two vehicles. A family with two or more workers, more than one child, and two or more vehicles is the norm for these neighborhoods.	To save time in their busy lives, they frequently buy prepared dinners from the grocery store and fast food. They play video games, go bowling, and visit theme parks such as Six Flags and Sea World.	HH

APPENDIX E Continued

Segment Code	Demographic	Socioeconomic	Residential	Preferences	Cluster type
31	These neighborhoods are found in pastoral settings in rural nonfarm areas throughout the United States. Most residents are white in these low-diversity neighborhoods.	The median household income is $47,120, slightly below the US level. Six percent of those who are employed work at home, twice the US rate. Because so many residents are aged 65 and older, receipt of retirement income and Social Security benefits is common.	The number of households in these small, low-density neighborhoods is growing at 1.5% annually.	Active participants in local civic issues, residents also belong to environmental groups, church and charitable organizations, fraternal orders, unions, and veterans' clubs. They go hiking, boating, canoeing, hunting, fishing, horseback riding, and golfing.	HH
46	50% are older than age 55. Married-couple families dominate these rural neighborhoods. There is little ethnic diversity in the *Rooted Rural* segment; almost 90% of the residents are white.	The median household income for this segment is $37,952; the median net worth is $60,202. The labor force participation of 56% is below the national level.	A higher proportion of seasonal housing contributes to higher vacancy rates in these neighborhoods. Local residents tend to move infrequently.	They hunt, fish, ride horseback, attend country music concerts, and go to car races. They read hunting and fishing magazines and listen to country music and auto racing on the radio.	HH
59	These young families form the foundation of Hispanic life. Children are the center of these households that are composed mainly of married couples with children and single-parent families. Ethnic diversity is high; 83% of the residents are Hispanic.	The median household income for this segment is $28,307. They carefully budget their income month to month to pay for the upkeep of their homes and families.	72% of these households are in Texas. Home ownership is important to these settled, suburban folks; more than two-thirds own their homes. The median home value is $57,028, the second lowest among the Tapestry segments.	They listen to Hispanic and contemporary hits radio. TV and radio are the best media to reach them instead of newspapers or magazines	HH

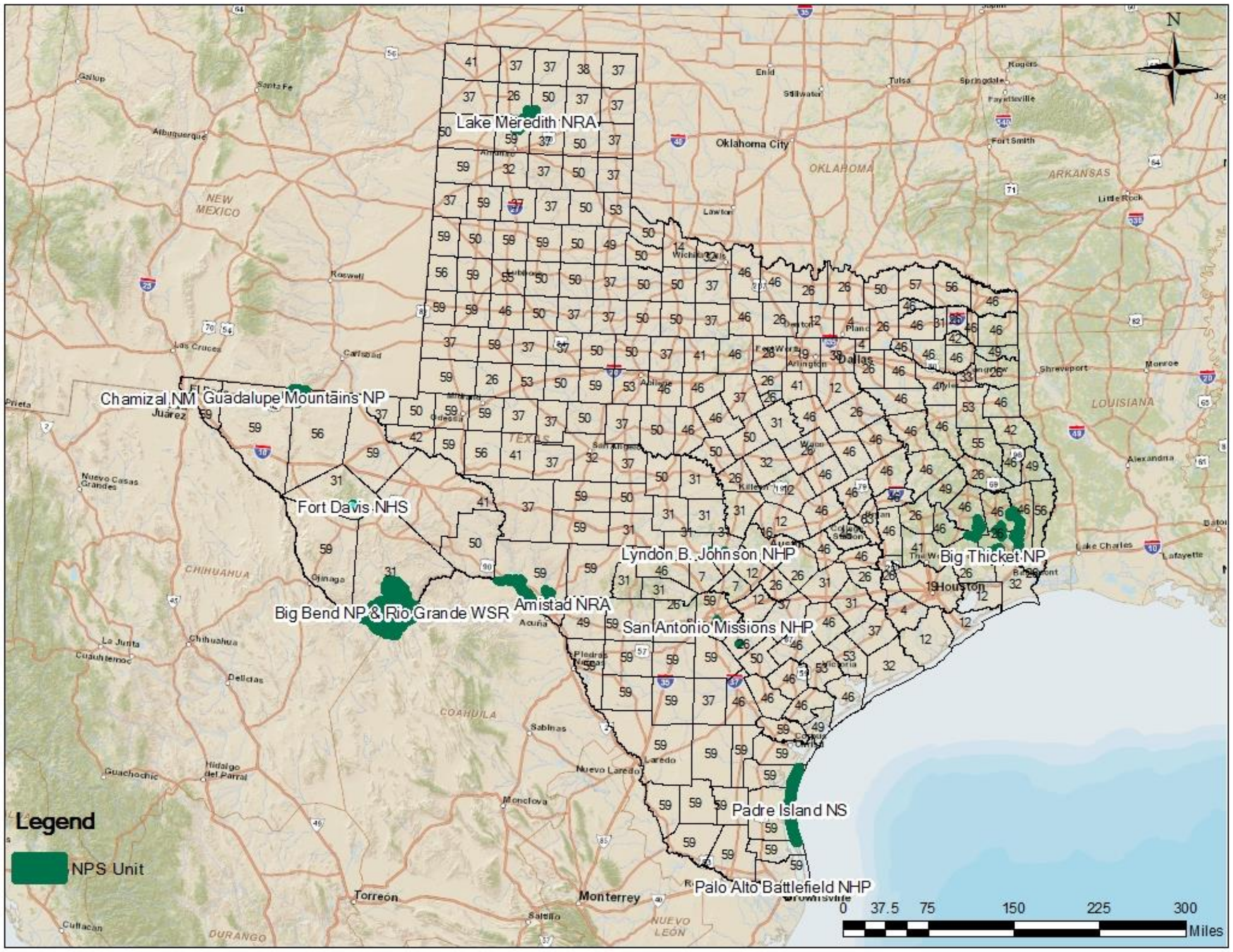
APPENDIX E Continued

Segment Code	Demographic	Socioeconomic	Residential	Preferences	Cluster type
37	Small, family-owned farms in the Midwest dominate this stable market. Two-thirds of these households are composed of married couples with or without children. These residents are slightly older, with a median age of 42.9 years. There is little diversity here; 9 in 10 of these residents are white.	Even though agricultural jobs are important to the local economy, 40 percent of the residents work in white-collar jobs. The median household income is $43,161. Fifty-three percent of the residents aged 25 years and older have graduated from high school; the number of those who hold a bachelor's or graduate degree is below the US level.	Most of these Midwestern neighborhoods are centered in Iowa, Nebraska, Minnesota, and Kansas. Smaller groups are concentrated in the West and South. Eighty-one percent own their homes. Most housing is single-family; however, 11 percent are mobile homes, slightly higher than the US average.	Their purchases reflect their rural lifestyle.Many own satellite dishes because cable TV is not available in many rural neighborhoods. Prairie Living residents are loyal country music fans and tune in to radio and television for their favorite music. They enjoy hunting, fishing, and horseback riding, target shooting, and riding around on their all - terrain vehicles	LL
50	Approximately half of the residents have already retired, many in the same towns where they have lived and worked their whole lives. Nearly half are aged 55 years or older. Almost one-third are singles who live alone, other family types and shared housing are also represented. Diversity is minimal; nearly 9 in 10 residents are white.	The median household income is $34,088. Two-thirds of the households earn wage and salary income, and 39 percent receive Social Security benefits. The percentage of the population aged 25 years or older that has completed high school is higher than the US level; the percentage that has attended col-lege is far lower than the US figure.	Low-density neighborhoods dominate, with older homes in urban clusters and rural, nonfarm areas. More than half of the housing units were built before 1960. Home ownership is at 67 percent. More than three-fourths of the housing is single-family dwellings.	They go hunting and fishing. They also read gardening, fishing, and hunting magazines and listen to country music and auto races on the radio. Many Heartland Communities households subscribe to cable and usually watch news programs and movies on TV.	LL

Source: ESRI Tapestry Segmentation Reference Guide, 2012

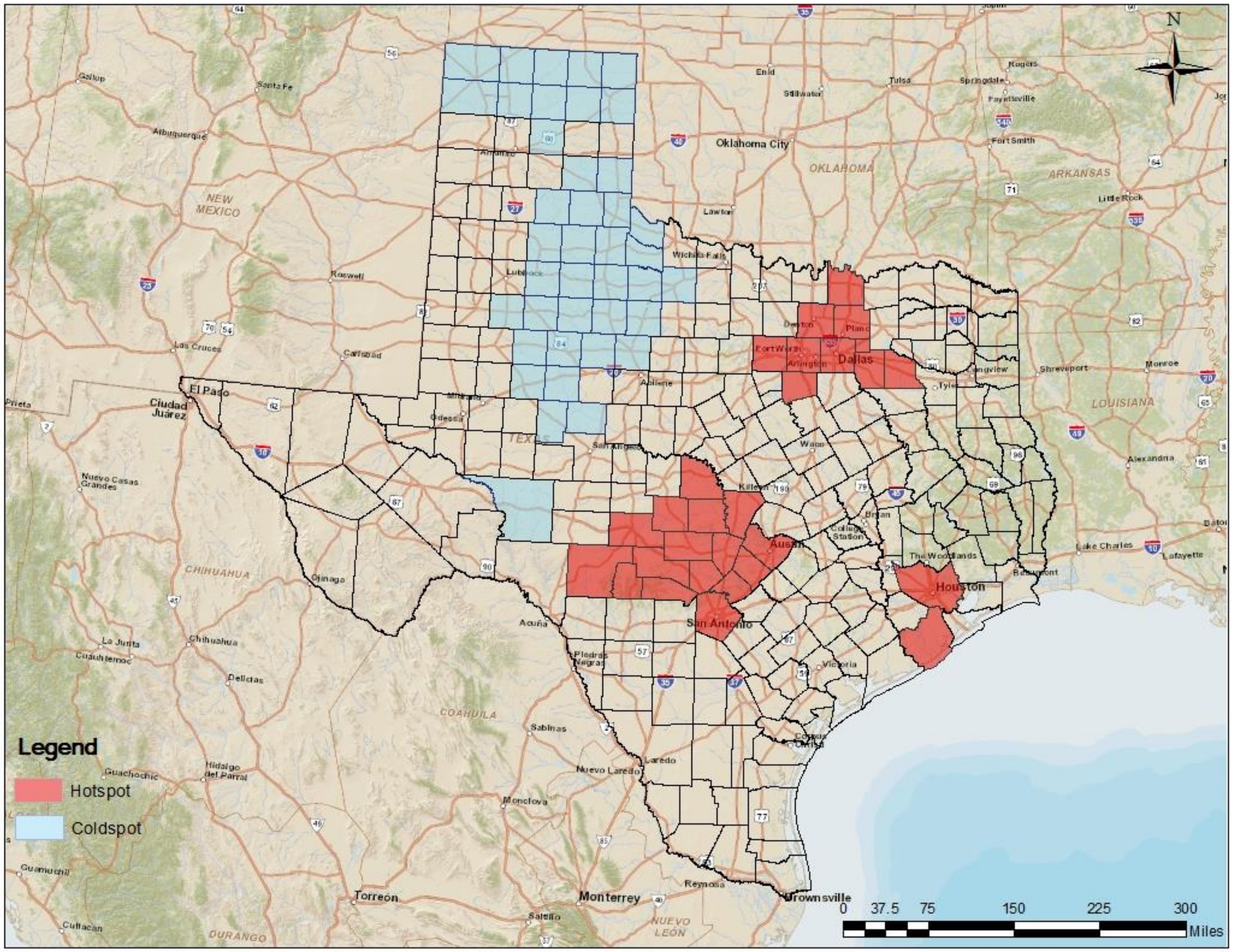
APPENDIX F

DISTRIBUTION OF SEGMENTATION CODE AROUND TEXAS NPS UNITS



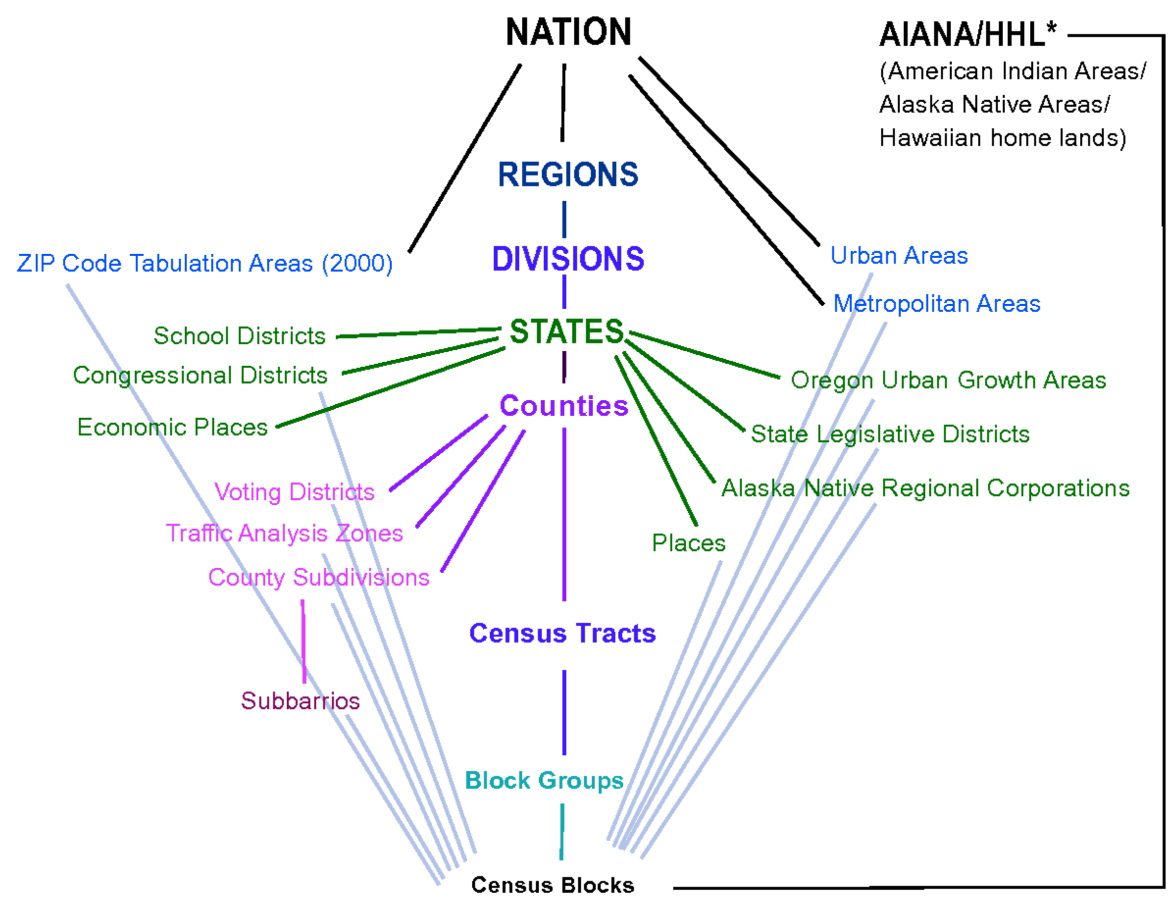
APPENDIX G

HOTSPOT/COLDSPOT AREAS OF NATIONAL PARK VISITATION



APPENDIX H

STANDARD HIERARCHY OF CENSUS GEOGRAPHIC ENTITIES



Source: http://www.census.gov/geo/www/geodiagram.pdf