ABSTRACT

Title of dissertation: Metropolitan Spatial Structure: Measuring the Change

Selma Hepp, 2011

Dissertation directed by: Professor Gerrit-Jan Knaap
Urban and Regional Planning and Design

Since 1990s, the metropolitan spatial structure has been alleged to be growing smarter. Excessive suburbanization trends characterizing urban form since the Second World War are now believed to be reversing in favor of urban environment. The reversal is driven by changing household preferences as well as a series of changes that urban areas have gone through which make them more attractive living environments for some demographic groups.

This is a dissertation consisting of three related essays which examine change in the metropolitan spatial structure over the past two decades to determine if suggested changes are in fact observable in urban form. In measuring the change, I consider a number of measures that characterize urban form, particularly density, concentration, clustering, infill and growth allocation of urban growth. Given the prevalence of foreclosure crisis in the later part of the first millennium decade, I also explore the impact of urban form on accumulation of foreclosures as an indicator of future spatial structure change.

The study finds two different trends at force facing the American metropolitan spatial structure. For the metropolitan areas with weak growth pressures or those loosing population since 1990, suburbanization trends continue to define spatial structure.

However, in the metropolitan areas that are facing moderate and strong population growth pressures and constituting the majority of the largest urban areas in the U.S., the importance of urban center is ever more significant and their spatial structure is greatly dependent on denser urban form. Desirability for urban environment also manifested itself in the spatial distribution of foreclosures in Maryland.

METROPOLITAN SPATIAL STRUCTURE: MEASURING THE CHANGE

By

Selma Hepp

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:

Dr. Gerrit J. Knaap, Chair
Dr. John I. Carruthers
Dr. Alexander Chen
Dr. James Cohen
Dr. Maureen Cropper
Dr. Chengri Ding

© Copyright by

Selma Hepp

2011

ACKNOWLEDGEMENTS

I would like to sincerely thank my advisor Gerrit J. Knaap, for his support and encouragement, and mentoring. Also, I would like to thank John I. Carruthers for his instrumental support and guidance. As well, I would like to thank the rest of my committee, Alexander Chen, James Cohen, Maureen Cropper, and Chengri Ding, for their suggestions and comments.

I would also like to thank two other University of Maryland faculty members for their constructive suggestions, Casey Dawkins and Charles Towe. I am in debt to Robert N. Renner from the Department of Housing and Urban Development for his priceless time and invaluable expertise in Geographical Information Systems and Department's housing data sources. I would also like to thank several colleagues from the National Center for Smart Growth who helped in a number of ways on the dissertation: Rebecca Lewis, Sabyasachee Mishra, and Niu Yi. I am also grateful to the staff of the Maryland's Department of Housing and Community Development for authorizing access to the foreclosure data.

Finally, I would like to thank all my loved ones for their unconditional love and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

The study of urban spatial structure concerns the organization of land uses in urban areas. Urban structure describes the arrangement of housing and businesses, public and private spaces within urban settings and the degree of connectivity and accessibility among them. The spatial structure of metropolitan areas today is the result of interactions among its residents, land markets, job markets, landscape, regulation, infrastructure, and climate through history. The way that spatial structure is arranged greatly determines how the city functions and has consequences for accessibility, environmental sustainability, economics, welfare, social equity, social capital, and cultural innovation. Inefficient spatial structure can lead to increasing distances among people, jobs, and amenities and consequently defragment labor and consumer markets, diminish environmental quality, and generally compromise quality of life (Bertaud and Malpezzi, 2003). In the time since the World War II, the fear has been that excessive suburbanization led to those negative externalities. Yet, recent inquiries into urbanization trends began suggesting that since 1990s metropolitan spatial structure has reversed again towards an increasingly urban future. Now is a particularly interesting period to reflect on the change in urban structure because urbanization patterns have perhaps been experiencing momentous change.

This dissertation consists of three essays which address change in the metropolitan spatial structure over the past two decades using a number of measures of urban form. In the first essay, I investigate how metropolitan areas have changed between 1990 and 2007 using several urban form measures established in the urban form literature. These measures reflect the change in density, concentration, clustering, and growth allocation of urban growth.

In the second essay, I employ a new measure for evaluating the change in the spatial structure. While the method is not new in the study of settlement patterns, it is a new approach to characterizing urban form and measuring the change. In two of the essays, the analysis of urban spatial structure is extended to focus specifically on changes over time in the largest metropolitan areas. The essays will determine whether these trends are consistent with the proposition that U.S. cities are now growing smarter. While growing smart generally suggests use of compact development, mixed uses, and close coordination between transportation and land use policies, the working definition of smart growth employed herein refers to new growth occurring in already existing urban areas.

The third essay considers the spatial problem of foreclosures across the state of Maryland and its two largest metropolitan areas, Baltimore and Washington. By measuring the effect of a number of urban form measures on the accumulation of foreclosures, I aim to delineate the impact of foreclosures on the changing metropolitan spatial structure.

Theoretical Models of Urban Form

Theoretical explanations of urban spatial structure commenced with von Thünen's (1826) theory of agricultural land use. The theory was based on the notion of economic rent to explain how competition for land use among various agricultural activities leads to their spatial organization. The original ideas of this model laid the foundation for many of the following urban spatial models. In von Thünen's city, land use patterns are determined by transportation costs to the central market where the most expensive agricultural product to transport occupies land closest to the consumer. The core hypothesis of the model is that agricultural land uses are patterned in the form of concentric rings around a central city. Although using very restrictive assumptions, this model preceded the age of large-scale industrialization thus assuming that all other land uses placed around the city were agricultural.

Since von Thünen, technological advances, such as telecommunication and transportation, added new dynamics to urban spatial structure and led to reorganization of land uses within cities. Subsequently, manufacturing and commercial land uses also began competing for the use of land closest to the central city. The arrival of electric streetcars, and later of personal automobiles, allowed residents more flexibility in choosing where they lived. Consequently many of upper income residents moved away from the central cities. Businesses on the other hand continued benefiting from economies of agglomeration in the central cities and their process of bidding for rents led to high-rise commercial buildings which now define the skylines of most all metropolitan areas.

Following these changes, in what became known as the Alonso-Mills-Muth (AMM) model, Alonso (1964), Mills (1967), and Muth (1969) modernized von Thünen's theory to include land use, housing services, rent, disposable income, intensity of land use, population, transportation cost, and employment. Similarly to von Thünen, the AMM model postulated a flat, monocentric, continuous and uniform urban area in which central business district (CBD) represents the center for work and shopping. The essential notion of their utility-maximization theory, also borrowed from von Thünen, was the bid-rent function for each household or firm. The bid rent is the maximum amount that a household is willing to pay for rent at different locations in the city such that it maintains the same level of utility. A household's utility depends on the size of housing, distance from the city center reflecting the transportation cost, and all other goods. Given household's preferences, a household allocates its fixed budget among these three components with the aim to maximize its utility. The bidding among households and firms for land closer to the city center results in higher rents in those closer locations. To capitalize on higher land values, there are more residences and/or offices built per unit of land closer to the CBD resulting in higher overall density in central cities. With increasing distance from the CBD, there are fewer bidders for the land causing reduced land prices and falling density. However, the cheaper rent of locations further from the center is offset by corresponding transportation costs.

Another theory of spatial structure developed by Homer Hoyt in 1939 proposed that an urban area grows outward in wedge-shaped sectors rather than concentric rings. Hoyt suggested that particular parts of an urban area are more attractive for some activities than others, either inadvertently or due to geographic and environmental reasons. As the urban area grows, these activities expand outward in a wedge-shaped pattern along railroads, highways, and other transportation routes. Again, better accessibility necessitates higher land rents, thus location of commercial functions remains in the CBD but manufacturing land uses develop in a wedge adjoining transportation routes. Residential land uses also grow in wedge-shaped patterns; however low-income housing adjoins manufacturing and industrial sectors since noise and pollution reduces desirability of the area, while higher income residents locate furthest away from these sectors.

Theories by von Thünen, AMM and Hoyt, also referred to as natural evolution theories (Mieszkowski and Mills, 1993), offer static descriptions and explanations of urban land use, yet they are not explicit theories of urban form change and land use change. These theories explicitly treat the actual amount of land consumed via the bidding process, but the systems of land use change remain implicit, i.e. it the change is understood from the factors, including location preferences and income, which in the model are assumed to determine the shape of the bid rent functions. The factors can change causing the land use system to change into a new equilibrium position. Thus, following this logic, advances in the transportation system allowed not only businesses, as mentioned above, but residents to bid for rent in suburban areas. Higher incomes afforded households to move out to single family homes on larger lots while the lower income residents remained in poorer quality and smaller spaces in the central cities. These theories emphasize the importance of transport costs and incomes in changing urban structure. However, elements which are arguably equally important to one's bidding function, such as social, cultural, and political influences, are not explained by these theories.

A theory, elaborated by Mieszkowski and Mills (1993) following Oates, Howrey and Baumol (1971) and Bradford and Kelejian (1973), examined the suburbanization phenomena by focusing on fiscal and social problems of post- war cities, such as crime, congestion, low environmental quality, high taxes, low quality of government services and public schools. As a result of such an environment, higher income households moved out first as they were able to afford initially expensive means of transportation, a personal automobile. Departure of higher income residents left cities impoverished for tax revenues and led to further deterioration of the quality of life and provision of public services, inducing additional outmigration. The vicious circle was exacerbated as suburbs became hubs of high-achieving school districts which attracted other better-off households. Other land uses, such as commercial, followed afoot. This theory draws greatly on Tiebout model (1956) in which households choose to move to jurisdictions within a region that offer a combination of government services and tax rates that maximize their utility. Through households' choice process and resulting revealed preferences, jurisdictions and residents find an equilibrium provision of local public goods, thereby sorting residents into optimum communities. This theory again is not a land use change theory per se, but deals with the change in its determinants and thus offers explanation for the change in urban form.

<u>A Brief History of Changes in Urban Form</u>

Regardless whether the change in urban form resulted solely from changes in budget opportunities and its allocation or dynamics of socio-political and fiscal influences, studies measuring urban form and change have consistently found that the period following the introduction of the personal automobile is characterized by decentralization, in particular, falling central-city population densities and decreasing rates of decline in density with distance from the central core (Mills, 1972; Edmonston 1975; Maccauley, 1985; and Kim 2007). More specifically, while the first half of the century experienced increasing urban density, the second half experienced both falling urban density and the dispersal of urban populations at the urban fringe. Most frequently, this new urban form has been referred to as “sprawl”. Sprawl, however, has been characterized along many dimensions which are discussed in *Literature Review* section.

In 1990s, however, despite the expected continuation of suburbanization trends and declining importance of central cities, both central cities and suburbs gained in population. As the 2000 census revealed, many Northeastern and Midwestern cities with over 500,000 people gained population for the first time since 1950. Chicago grew by 4 percent; New York City grew by 9 percent. Overall the median growth rate for cities in the 1990s was 8.7 percent, more than double the median growth rate in the 1980s (Glaeser and Shapiro 2001). During the same decade, suburbs elsewhere in the country grew as well, by about 16.5 percent (Lucy and Phillips, 2001). In fast growing cities in the South and West, in particular, most of the growth occurred in the outer ring neighborhoods (Katz, 2002). These “boomburgs”, defined as suburbs with more than 100,000 residents, were growing at double digit rates (Lang 2001). In the 1990s, boomburgs accounted for over half of the growth in cities between 100,000 and 400,000 residents. In 2000s, the same trend continued and central-city populations in the metropolitan areas with more than one million people grew at an annual rate of 0.5 percent between 2002 and 2005. Suburbs of these cities grew at (growth) rates between 1.29 percent and 1.48 percent during the same period (Frey, 2009). By 2006-2007, however, central cities’ growth rate increased to 0.90 percent, and reached 0.97 percent by July 2008. At the same time, the growth rate of suburbs declined by 1.11 percent.

There are two general theories behind urban resurgence beginning in the 1990s. The first theory stems from the increasing importance of knowledge of the economy and the ability of large and dense metropolitan areas to facilitate the flow of knowledge.

Many of the technological advances in the last quarter of the 20th century led to improvements in communications and processing of information, causing a fundamental shift in the perception of space and human interactions. The emergence of this new economy, referred to as “knowledge-based economy”, relied primarily on production, distribution and use of knowledge and information. In the context of the urban spatial structure, knowledge distribution through formal and informal networks became essential to economic performance. But, an even more important component of the new economy was transmittance of tacit knowledge which insured continuous learning and advancement of individuals and firms. Tacit knowledge includes skills to use and adapt codified knowledge. Metropolitan areas, particularly central cities, gained comparative advantage in facilitating transfer of tacit knowledge by reducing the costs of interactions between firms and individuals through proximity (Fujita, Krugman and Venables, 1999).

The second theory of urban resurgence elaborated by Glaeser and Gottlieb (2006) arises out of increased demand for urban amenities beginning in 1980s and resulting from changes in city governance, improvements in law enforcement technology and rising incomes. Since 1980s, crime rates have significantly dropped in many large U.S. cities (Schwartz et al, 2003). Additionally, cities have invested more in quality of life, such as museums, theaters, concerts, restaurants, urban landscape, offering a much richer social life. With general rising of incomes in the United States, individuals’ willingness to pay for these urban amenities increased. And it is not to say that cities were not already supplying such amenities, but scale efficiencies ensure that cities supply a greater bundle of urban amenities. Additionally, individuals were willing to pay more for proximity to the constant bundle of urban amenities and proximity to other people.

Empirical Measures of Urban Form at the Metropolitan Scale

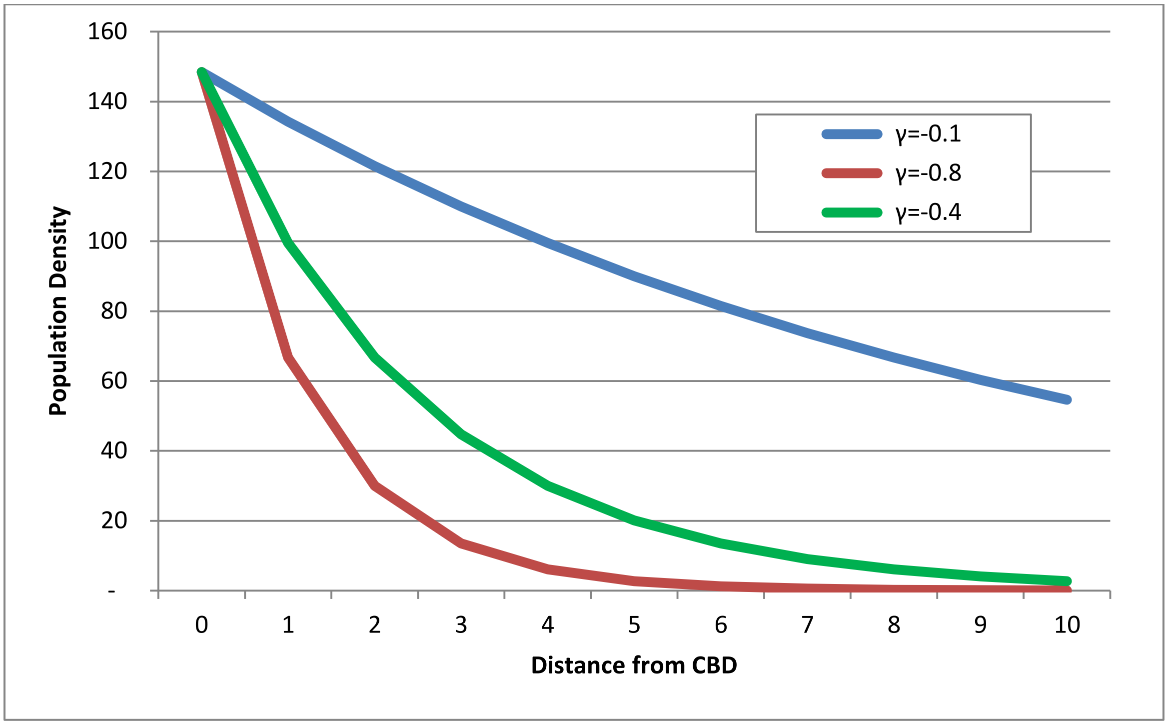
The literature on measures of urban form is widespread across disciplinary boundaries. As Clifton et al. (2008) suggest in their multidisciplinary review of quantitative approaches to urban form, urban form has been examined from various disciplinary approaches, including landscape ecology, economic structure, transportation planning, community design, and urban design, to name a few. These measures differ with each discipline, as do the questions being asked, the targeted audience, and the data sources. Also, urban form has been measured at different geographical scales – from metropolitan area, to city, to neighborhood. The varying measures reflect the distinct public policy issues that occur at each scale. The metropolitan scale, as addressed in this study, urban form questions are concerned with size of cities, location and number of centers of economic activity, as well as type and intensity of development. The measures employed include size of metropolitan populations, size of metropolitan areas, and population density, with density analyses measures prevailing in the literature.

The classic measure of urban form at the metropolitan scale has been the population density gradient. Formally, the population density gradient of a city is expressed as follows:

$$D(x) = D_0 e^{-yx\varepsilon} \quad (1)$$

where $D(x)$ represents population density at distance x from the center; D_0 is the density at the center; and y is the density gradient or the rate at which the population density decreases as one moves away from the center. The final error term, ε, is included when the formulation is stochastic. Figure 1 illustrates an example of three density patterns, for different values of γ.

Figure 1: Alternative Density Gradients



In the seminal work by Clark (1951), urban population densities for 19 cities in United States, Western Europe and Australia were described using the negative exponential function showing that density declines exponentially from the central core towards the outskirts of a metropolitan area. After Clark's initial work, Muth (1969) expanded the inquiry to 46 U.S. cities using the negative exponential function and similarly found "significant tendency of population densities to decline with distance from the CBD at the 0.01 level" (p. 140). In applying the same method, Mills (1972) studied the patterns and causes of suburbanization of population and employment from the 1880's to the 1960's for five metropolitan statistical areas. Later, Mieszkowski and Mills (1993) and Jordan, Ross, and Usowski (1998) also used density gradients to determine the degree to which population is located at a distance from the city centers.

In a comprehensive literature review, McDonald (1989) summarized the density gradient method as straightforward to implement and very flexible, one that can be estimated in virtually any functional form, and expanded to include any number of explanatory variables. McDonald concluded that the complexity of the contemporary metropolitan areas may have moved beyond the negative exponential function, but a superior method had yet to be developed. In addition, similar to the standard urban model theory preceding it, the strength of the density gradient lies in both its simplicity and its ability to describe the general tendencies of land use worldwide (Anas et al, 1998, Bertaud, 2003, DeBorger, 1979, Alperovich, 1980), with the exception of cities in former communist countries (Bertaud and Renaud, 1997).

However, Brueckner (1982), criticizing the method, presented an extended model where housing demand is a function of income adjusted for commuting cost, rather than gross income originally proposed by the model. In another paper, Brueckner (1987) argued that the model insufficiently accounts for vintage effects of the housing production. In other words, Bruckner argued that the model as presented assumes housing capital is perfectly malleable. In reality, producers are not able to adjust their capital and land inputs without costs from one period to another. However, models which explicitly account for housing durability are significantly more complex.

While the metropolitan spatial structure analysis has traditionally assumed monocentricity, research periodically brings into question a linear form of the density gradient by demonstrating the existence of population and employment sub-centers, higher-density neighborhoods on the urban fringe, discontinuities due to open spaces, and the dominance of commercial land use at the center of an urban area. Batty and Kim

(1992), for example, criticized the method suggesting it was only applicable for modeling intra-urban variation, while the inverse power function is more appropriate for analyzing the urban fringe and hinterland. They argued that the decrease from the negative exponential function was too great at the urban fringe. Marshall (2007) further pointed out that the model in theory assumes the population density at city center to be greatest, but it in fact overestimates its true value because the city center is most often occupied by commercial land uses. The author subsequently proposed a linear population density model as an alternative; however the model is even more restrictive and unrealistic and it never caught on in the urban form research.

To better reflect metropolitan spatial structure and its variations, efforts have been made to develop alternative and more complex forms of population density (Casetti, 1969; McDonald and Bowman, 1976; Zielinski, 1979; Eldridge, 1984; Latham and Yeates, 1970; Newling, 1969). However, in testing for the appropriate specification of the density gradient, Kau and Lee (1976), similarly to McDonald and Bowman (1976), concluded that no single functional form is optimal for uniform application across urban areas since urban areas differ in the age of housing stock, transportation modes, and geographical restrictions. In addition, even when the employment monocentricity assumption is relaxed, the negative exponential function can still remain, given higher density of employment at the urban center than elsewhere in the metropolitan area.

Other nonparametric approaches, such as the cubic spline density functions, have been proposed (Anderson 1982, 1985; Munız et al., 2003) for polycentric urban forms. Anderson (1982, 1985) applied the cubic spline model to estimate the changes in Detroit’s urban form between 1960 and 1980 arguing that the model more appropriately identifies urban sub-centers as it offers more flexibility by allowing for nonmonotonic and nonconstant density functions. However, as the spline density function gradient has to be evaluated at various distances, Muniz et al. (2003) defended the negative exponential function saying that the function's single estimate of density gradient allows for easy comparison between different cities and across time. It is also not clear how cubic spline function distinguishes between polycentric and decentralized urban form (Tsai, 2005). Alperovich (1995) also noted the function is unlikely to be useful for testing hypotheses on the processes underlying the determination of population densities, both empirically and theoretically. And, due to multicollinearity among distance variables and its high-order functional form, the cubic spline insufficiently improves the performance of the model.

Another method of analyzing metropolitan spatial structure is via point pattern analysis. Point pattern analysis involves describing patterns of locations of point events and comparing them to theoretical distributions. The location of point events in the case of urban form refers to distribution of settlements. Getis (1983) applied the point pattern analysis to urban form and examined population clusters in the Chicago area. The shortcoming of the method is that it only mathematically describes distribution of points in space. It can then be tested for deviation of the particular spatial distribution from hypothesized patterns; however it is not able to account for any behavioral influences (Carruthers et al, 2010).

Acknowledging the difficulty in modeling urban complexities, Batty and Longley (1987, 1994) and Frankhauser (1994) offered a method using fractal geometry. Fractals have a dimension of between one and two, indicating a one-dimensional line to a two-dimensional polygon, and measure space filling. The greater the fractal dimension, the greater the space filling and the more compact the development pattern. Seemingly simple, the estimation is, however, difficult because there are multiple definitions of fractal dimension, (not all of which agree,) and multiple ways of calculating it. In other words, a fractal has the same shape regardless of the scale employed for viewing it. Because they appear similar at all scales, fractals are often considered to be infinitely complex. The usefulness of fractal geometry in the study of urban form lies in the self-similar attribute of the fractals. Though ostensibly chaotic at the local level, fractals aggregated to larger areas generate the same patterns over time and space rising to an organized and hierarchical structure. Compared, cities are also composed of self-similar phenomena, such as roads networks, neighborhoods and centers which repeat themselves on many levels. Fotheringham et al (1987) and Longley and Mesev (1997, 2000, 2002) explored the connection between a fractal dimension and density of development. Torrens (2006, 2008) used the approach to measure sprawl. However, in generating fractal dimensions of 20 large U.S. cities along with their surrounding urbanized areas, Shen (2002) found that different cities may have virtually the same fractal value but be very dissimilar in population sizes. The author concluded that fractal dimension itself says little about the specific orientation and configuration of an urban form and is not a good measure of urban population density.

Most recently, researchers have explored survival analysis methods to examine urban form and the change. The survival analysis methods were primarily developed in medical and biological sciences, but they are also broadly used in social and economic sciences and engineering. They are used to characterize occurrence and timing of events.

For example, in social sciences, they analyze time to events such as job changes, marriage, and birth of children. In economics, survival analyses are used to measure risks and timing of mortgage default. When applied to a spatial setting, the time variable is replaced with distance variable and measures the conditional probability of a distance between two points ending (Waldorf, 2003). For instance, Odland and Ellis (1992) first applied the method to measure spacing of urban settlements in Nebraska. Also, Irwin and Bockstael (2007) used the spatial hazard models to study the timing of land use change in Howard County in Maryland for the period between 1973 and 2002, while An and Brown (2008) explored how the method, in conjunction with GIS and remote sensing data, can better inform parcel level land change analysis. At the metropolitan scale, Carruthers et al (2010) used the method to characterize urban form of 25 largest metropolitan areas in 2006.

In sum, the exponential density function has the advantage of being derivable from a simple model of a city, a monocentric city; however it is a measure with several restrictive assumptions, more specifically constant returns Cobb-Douglas production functions for housing, consumers with identical tastes and incomes, and unit price elasticity of demand for housing. But although the exponential density function is criticized as being a univariate measure, it fits most all American cities. Further, it is a univariate measure which in itself provides a single index of decentralization or urban sprawl. Depending on the degree to which one wants to scrutinize urban form, the simplicity of its use and understanding of the exponential density function serves well needs of many. The alternative measures also hold promise for measurement of urban areas. They are generally labor and data intensive which may limit their use. In the end, the use of a measure highly depends on the research task at hand, its effectiveness in answering the question, and ultimately audience which needs it.

Empirical Measures of Urban Sprawl

Since the introduction of smart growth notions to urban planning policy in 1990s, researchers have actively pursued measures of urban form which identify sprawling urban form features. While definition of sprawl has been debated, Nelson and Duncan summarized the idea into the following definition (1995, page 1): "Unplanned, uncontrolled, and uncoordinated single-use development that does not provide for an attractive and functional mix of uses and/or is not functionally related to surrounding land uses and which variously appears as low density, ribbon or strip, scattered, leapfrog, or isolated development". Ewing (1997), who later identified four characterizations of sprawl, namely low density, strip, scattered, and leapfrog development, acknowledged that these distinctions exists on a continuum and development patterns may not necessarily easily fit into sprawl and non-sprawl categories. In his literature review, Ewing finds poor accessibility to be the common denominator of sprawl, where poor accessibility is identified as scattered or leapfrog development, commercial strip development, uniform low-density development, or separation of land uses.

The literature measuring sprawl views it most frequently in line with the six features outlined by Downs (1997): (1) no limits placed on the outward suburban expansion; (2) divided legal control over land use, local services, transportation, property taxes, and fiscal policy divided among many small entities or jurisdictions, with no central agency responsible for the planning or control of these issues regionally; (3)

extensive "leapfrog" development; (4) fragmented land ownership; (5) different types of land use, spatially separated or zoned into distinct areas; and (6) extensive strip commercial development along larger suburban roads. Following Down's six features, measures examining sprawl can be loosely grouped in three categories, (i) those that measure residential population density, (ii) those based on location and dispersion of jobs, and (iii) those that consider multidimensional land use phenomenon. The sprawl index composites usually contain various components of each of the three categories.

At the outset of the empirical inquiry, the Sierra Club's 1998 report measured sprawl in the U.S. by subjectively ranking U.S. metropolitan regions by the degree to which they sprawled. This sprawl measure was derived from changes in population, land area, traffic congestion, and loss of open space using census data and data collected from institutions such as The Texas Transportation Institute and the American Farmland Trust. The study found that among the biggest cities, Atlanta, St. Louis, and Washington, D.C. were the most sprawling.

Several studies have relied on residential population density alone to examine the extent of sprawl among metropolitan areas (Lopez and Hynes, 2003; Nasser and Overberg, 2001; Lang, 2003) as well as the change in sprawl over time (Fulton et al, 2001, Nasser and Overberg, 2001). Lopez and Hynes (2003) calculated residential densities of the 2000 census tracts and categorized tracts into low-density – with population density between 200 and 3,500 persons per square mile, or high-density – with more than 3,500 persons per square mile; and operationally defined a sprawl index as the difference between the percentage of a metropolitan area's population living in the two categories. While the authors argued that density is the most identifiable feature of metropolitan spatial structure that allows for relatively easy comparison of all U.S. areas, they acknowledge that the measure does not take into account other features of sprawl, including continuity for example. Further, changing their arbitrary cut-off densities would change their results.

Fulton et al. (2001) referred to sprawl as “land resources consumed to accommodate new urbanization,” and measure it as the ratio of growth in land consumption to growth in population of the metropolitan area, then reported the findings as persons per urbanized area. Urbanized areas are considered those with a minimum population density of 200 persons per square mile. Their study focused solely on density by computing average density across entire metropolitan areas and did not assess how density varies across an urban area. On the other hand, Nasser and Overberg (2001) quantified sprawl as the percentage of a metropolitan area’s population that resides within the Census Bureau-defined urbanized area, i.e. contiguous blocks with density of one thousand or more persons per square mile. Lang (2003) expanded Nasser and Overberg’s study and generated two sets of density measures for the fifty largest metropolitan areas using again the Census Bureau-defined urbanized area and the Department of Agriculture’s Natural Resources Inventory (NRI) urban land uses. The study showed that the West often has more densely populated metropolitan areas than the East due in part to the arid and rugged environment in the West. Also, while the Sunbelt is characterized by newer metropolitan areas with lower-density urban form, intraregional differences within the Sunbelt make such general comparisons difficult and deceptive.

Two relatively recent works serve as a useful framework because they analyze urban spatial structure at the metropolitan scale and rely extensively on census data. The first was developed by Galster et al. (2001) using GIS technology and data primarily from the US Census Bureau. To begin, all land parcels were categorized into one of three types: residential, non-residential, and undevelopable with a ¼ square mile grid laid over it. Then, the authors generated eight distinct dimensions of land use patterns: development density, continuity, concentration, clustering, centrality, nuclearity, mixed uses, and proximity. For each UA and each dimension, the authors added up standard deviation a dimension is from the mean of the dimension's distribution into a Z score. At the end, the Z scores for each UA across all six dimensions were summed into a sprawl index. Due to resource and time constraints, the authors forwent operationalization of mixed-use and continuity measures. Unlike studies that used MSAs or PMSAs, Galster et al. used urbanized areas (UAs) boundaries arguing they are preferable because they do not include rural land and more undeveloped land that MSAs or PMSAs do. That is, however, a limitation of Galster's et al. analysis since most sprawl does in fact occur outside urbanized areas. The analysis also only examines thirteen urbanized areas and it does not analyze segregation of land uses at the expense of accessibility. The study also implies that one has to be careful with composite sprawl index as one high scoring dimension of the eight calculated may be driving the overall value of the composite index. Not all dimensions are equally concerned to different analysts. Consequently, a metropolitan area may appear sprawling along some of the dimensions but completely opposite along others. Thus, the dimensions used should be refined and possibly consolidated to specify different types of sprawl.

Cutsinger et al. (2005) extended the Galster et al. (2001) analysis and included additional measures, inter-use proximity, continuity, and mix of uses. Inter-use proximity measures housing proximity, job proximity, and job to housing proximity, while mixed use measures evaluate ratio of jobs to housing and housing to jobs. Continuity measures the extent to which developable land within the metropolitan area has been skipped over and the extent to which development occurs beyond UA boundaries. The study also expands the UAs to Extended Urban Areas (EUAs) to account for issues raised by Wolman et al (2005) of under- or over- bounding biases with standard Census boundaries. The EUAs include UAs plus any additional square-mile cells within an MSA or PMSA that contain 60 or more dwelling units and send at least 30 percent of its workers into the UA on daily work commutes. Since the authors are primarily interested in interrelationship among different indices, 14 indices are reduced via principal components analysis to seven-factor solution: density/continuity factor, proximity factor, job distribution factor, mixed-use factor, housing centrality factor, nuclearity factor, and housing concentration factor. These different conceptual dimensions can characterize any given land use type. However, the results suggests that metropolitan areas often demonstrate both high and low levels of sprawl-like patterns across the seven components and housing and employment sprawling patterns differ in nature.

Another composite index of urban sprawl was developed by Ewing, Pendall, and Chen (2002). Like Galster et al., Ewing et al. work toward a single overall sprawl index for 83 U.S. metropolitan areas and counties. However, the authors contribute to previous work by adding measures of accessibility. Unlike Galster et al., however, they proceed in two steps: first, using principal components analysis, they develop indices of four factors of urban form, and then use these subcomponents to develop an overall sprawl index. Twenty two measurable subcomponents were based on the following four factors: (i) residential density; (ii) neighborhood mix of homes, jobs, and services; (iii) strength of activity centers and downtowns and; (iv) accessibility of the street network. Neighborhood mix index is an interesting addition which includes a ratio of residents with businesses or institutions within ½ a block of their homes, with "satisfactory" neighborhood shopping within one mile, with a public elementary school within one mile, balance of jobs to residents, balance of population-serving jobs to residents, and mix of population-serving jobs. The first three variables are somewhat restrictive in definition as ½ a block for example is a short distance to accommodate many business or institutions. Also, "satisfactory" shopping is a subjectively derived measure. Finally, since those variables were obtained from the American Housing Survey, they are limited to the responses given by a small sample. Correlation analysis among the four factors showed that centrality is largely independent of residential density, suggesting both variables add unique information to the overall sprawl index. While density captures intensity of land use, centrality measures the focus of development on the central business district and presence of subcenters within a metropolitan area. The land use mix factor however is moderately correlated with the density factor, which is expected given that higher densities are needed to support mixed uses. Finally, the street network factor is highly correlated with density, which is also to be expected since higher density requires more street capacity to meet travel needs. The composite approach by Ewing et al. offers several advantages. First, because they use multiple sources of data, including the Transportation Planning Package, they add new information on transportation infrastructure and provide richer measures of density, centrality, and mix. Further, by developing sub-indices and overall sprawl indices they provide measures that include a wealth of information while removing troublesome problems of multicollinearity.

In evaluating most of the studies reviewed herein, Jaret et al (2009) concluded that while some measures might seemingly address the same phenomenon, i.e. density, it is important to pay attention to data used, geographical boundaries, and other components of the measure to make sure it is well suited for the intended purpose. Ewing's et al. (2002) correlation and factor analysis study showed that their four dimensions can essentially be reduced to two: one that measures how strongly centered the metropolitan area is, and the other, sprawl measure, that combines density, mixed land use, and street pattern characteristics. Separating centeredness from sprawl allowed the authors to distinguish why seemingly sprawling areas rank low (not sprawling) using density measures. For example, Los Angeles is a dense but decentered metropolitan area with relatively few people living clustered near major activity centers, such as the CBD.

Tsai (2005) also attempts to distinguish between compact and sprawling metropolitan areas using four matrices: metropolitan size, density, the degree of equal distribution and degree of clustering. To evaluate the degree of equal population distribution, Tsai uses Gini coefficients. Higher Gini coefficients suggest population or employment density is unevenly dispersed across the metropolitan area, while lower Gini coefficients suggest more even distribution in a metropolitan area. As the author discusses, though, Gini coefficients alone do not reveal any spatial relationships and are limited to differentiating between monocentric, polycentric or decentralized sprawl spatial structure. They do address the extent to which development is concentrated in a relatively small number of sub-areas.

Finally, with the growth of GIS technology and remote sensing data, studies have attempted to directly measure land use and land cover change (Irwin and Bockstael, 2002; and An and Brown, 2008; Ji et al, 2006; Torrens, 2008; Burchfield et al., 2006). Although measures using remote sensing data have a promising future, there have been some issues noted. Due to the technology of collecting the data, the data are sensitive to meteorological conditions and annual changes in vegetation growth. This can also lead to underreporting of low density urban forms and can bias thresholds used to classify different land covers (Irwin and Boaksdale, 2002). Alone, remote sensing data are capable of detecting land cover changes; however this information by itself is not sufficient to address metropolitan spatial structure. While the data may suggest a change in land use/land cover, it does not provide information on the intensity of land use. For example, in Ji et al (2006) study, the ‘‘built-up’’ land use classification referred to residential areas of single houses and apartment buildings, shopping centers, industrial and commercial facilities, highways and major streets, and associated properties and parking lots. Also, since in the study of urban form the data have to be delineated into urban and rural uses, categorization of cells is based on thresholds, which arguably may be a random decision by the researcher. On the other hand, some have chosen to impose census boundaries in delineation of urban and rural areas (Ji et al, 2006). Finally, remote sensed images need to be classified into land covers, which is a procedure usually requiring human supervision and accuracy assessment (Ji et al, 2006).

Nevertheless, in an extensive analysis, Torrens (2008) measured change in sprawl in Austin, Texas, between 1990 and 200 on a series of measures – 18 measures - including urban growth, density, social, activity-space, fragmentation, decentralization, and accessibility using high resolution parcel level data. Torrens uses fractal dimension to measure fragmentation and scattering in urban spatial structure. Decentralization characteristics are measured via spatial autocorrelation metrics, global Moran's I index (Moran 1950) on metropolitan level, and local Moran's I and localized Getis–Ord G statistic for per-parcel level analysis. The analysis was based on a very detailed parcel level data and while it offered a very thorough representation of Austin, an inter-metropolitan analysis using these measures would be difficult given the lack of comparable time-space data across a large number of metropolitan areas. The author did conclude by suggesting that sprawl and "smart growth" are found to co-exist and co-evolve.

In summary, it is evident that this body of research is poised with debate over what to measure, how to measure it and what is important to consider. One common conclusion that emerges is that sprawl is a multidimensional phenomenon which exists on a continuum. Each dimension requires a separate examination. Consequently, depending on the way in which it is measured, the same metropolitan area can be typified on a different end of a spectrum. But again, density characteristics are principal traits of sprawl. Being that they are relatively straightforward to measure and across a large number of metropolitan areas, they are often used as the sole indicator of sprawl.

Measures of Changes in Urban Form

Accurately capturing change in urban form is a challenging task. Most studies have relied solely on densities to measure such change. Studies looking as far back as 1890s consistently confirmed that urban densities modestly rose until 1950s but then decreased greatly between 1950 and 2000, whereas density gradients fell monotonically over time (Mills, 1972; Edmontson, 1975; McMillen and McDonald, 1998; Kim, 2007) As discussed in the theoretical background, decreasing cost of transportation is most frequently attributed cause of this trend.

Fulton et al. (2001) relied solely on population densities and density changes to examine the sprawl phenomenon. The authors evaluated relative land consumption to population change for all U.S. metropolitan areas between 1982 and 1997, in 5-year increments. If land is consumed at a faster rate than the population is growing, it is assumed that sprawl is increasing. The analysis revolved around three measures: rate of conversion of undeveloped into urban land, metropolitan area's population density and the change, and difference between the change in population and change in urbanized land. The study reported an increase of 47 percent in urbanized land and only 17 percent in population over the 15 year study period. Following the author's definition, the West is home to some of the least sprawling metropolitan areas in the county. Interestingly also, Honolulu and Los Angeles were rated most compact in 1997, and Las Vegas and Phoenix were both in the top 20 in compactness. However, because the study relied on the USDA's National Resources Inventory (NRI), which is a very small sample, it is a subject to sampling error and considered reliable only at county levels or above. This study examined the change at the metropolitan level and does not account for intra-metropolitan variation. However, the 2009 NRI Summary Report warns against comparing NRI data published prior to 2009 as they may produce erroneous results because of changes in statistical estimation methodology[1].

[1] http://www.nrcs.usda.gov/technical/nri/2007/2007_NRI_Summary.pdf

More recently, Burchfield et al. (2006) examined temporal changes in the American spatial structure using a combination of remote-sensed data to construct a grid covering the coterminous U.S. The data included the 1992 National Land Cover Data (NLCD) and 1976 USGS data attained mainly from high-altitude aerial photographs collected between 1971 and 1982. The study looked at the change in the amount of undeveloped land surrounding an average urban dwelling in the U.S. prior to 1976 and between 1976 and 1992 and concluded that the extent of residential development sprawl has not changed between 1976 and 1992. This did not hold, however, for commercial development which had become more spread out. Developed areas in this period grew at a rate of 2.5 percent annually, amounting to 49 percent over 16 years.

In a follow up study by Irwin and Bockstael (2007), the authors criticize the Burchfield et al (2006) results concluding that the NLCD data they used are systematically biased against recording low-density residential development. Instead, Irwin and Bockstael (2007) used planimetric data for Howard County, Maryland from 1973 and 2000 to quantify land use patterns based on patches. A patch refers to a discrete and contiguous area of the same land use, and finds contrasting conclusions which suggest increasing land fragmentation, and particularly in areas located far from urban areas. Most of the measures generated by authors are commonly used in landscape ecology to capture various dimensions of fragmentation based on patch characteristics. Consequently, the measures can speak to land fragmentation and changes to land use and land cover, but are less meaningful estimates of the metropolitan spatial structure. It is also difficult to make any generalizations based on potential idiosyncrasies of Howard County in Maryland.

Finally, Thomas (2009) evaluated the level of infill across the country by using the U.S. Census residential building permit data for the 50 largest metropolitan regions over the 18 year period from 1990 to 2007. The study compared the number of permits issued by central cities and core suburban communities to clarify if there had been a shift toward redevelopment and in which regions had the shift been most significant. The permit data confirmed drastic increases in several regions and roughly half of them showed larger increases in the urban core. The increase was particularly large in 2000s. In fifteen regions, the central city more than doubled its share of permits, such as New York City, Chicago, Portland and Atlanta. The limitation of this analysis results from the data which are provided at the jurisdiction level and limit spatial knowledge of the building location. Consequently, in suburban communities, development on both undeveloped and previously developed land is grouped into one suburban category and could underestimate the level of infill construction taking place. Also, due to administration boundaries and their changes over this time period, it appears to be difficult to distinguish between redevelopment and new development.

In summary of the literature measuring changes in urban form over time, it is apparent that the quality of studies heavily relies on availability and comparability of longitudinal metropolitan data that are at sufficiently disaggregated levels to account for inter-metropolitan variation and change over time. Much of the data compilation has evolved over time which often makes it difficult to use in comparative analysis of change in urban form. Also, delineation of urban boundaries appears to have led to some criticism for a number of studies. Consequently, studies examining the change in the metropolitan spatial structure and particularly increasing prevalence of sprawling patterns have been limited in use of different measures and again most frequently rely on population densities. Finally, most of the longitudinal studies explore the changes prior to 2000s, and more so prior to 1990s. It is important and timely to now explore the changes in the two decades since 1990 as they may have been significant in reshaping the U.S. metropolitan spatial structure.

Spatial Determinants of Foreclosures

Literature examining mortgage default is certainly not new. Academics have been long interested in understanding determinants of mortgage default. The early studies focused on determining the effects of loan and borrower characteristics, borrowers' decisions to default, and institutional frameworks (Quarcia & Stagman, 1992). Yet, the recent housing crisis, which began showing the signs of distress in 2005, has brought into perspective relatively new concerns related to mortgage default. According to the review by Mayer et al. (2009), the distinguishing feature of the current foreclosure crisis is that mortgage defaults and delinquencies started among borrowers of non-prime types of mortgage products. Non-prime products were generally extended to borrowers who would otherwise not qualify for prime mortgages, because of their compromised credit histories, very little savings available for down payments, or lack of full documentation of assets or income. Immergluck (2009b) effectively summarized how financial innovations and deregulation facilitated the rise of high-risk lending and why these types of loans and their associated regulatory infrastructure failed in substantial ways, harming different populations and communities along the way.

Studies immediately following the onset of the recent mortgage crisis but even those that focused on areas with high incidence of foreclosures prior to the national crisis, found that characteristics of non-prime mortgages were one of the strongest predictors of mortgage delinquency (Apgar and Duda 2005a; Immergluck and Smith, 2005; Coulton et al., 2008; Ding et al., 2008; Gerardi et al., 2007). In comparing prime and non-prime mortgages, non-prime borrowers had much lower credit scores, higher debt-to-income ratios (DTIs) and higher loan-to-value ratios (LTVs) at the time of origination (Okah and Orr, 2010; Amromin, G., and A. L. Paulson, 2009). Generally, the research on both prime and non-prime mortgages singles out the relevance of credit scores, DTIs and LTVs on default probability (Demyanyk, 2009; Foote et al., 2009; Haughwout et al., 2008; Ding et al., 2008; Foote et al., 2009; Haughwout et al., 2008). The problematic role of non-prime mortgage products also stems from their concentration among minority and low-income households, and given the strong probability of those loans to default, the high rate of foreclosure in black and minority neighborhoods (Immergluck, 2004; National Community Reinvestment Coalition, 2003; Coulton et al., 2008; Jiang et al., 2009). Some concluded that race variables are partly picking up the effects of credit history and socio-economic conditions of those neighborhoods (Van Order & Zorn, 2000; Berkovec et al., 1994; Pedersen & Delgadillo, 2007).

In tandem to the research on the role of non-prime loans in the mortgage default crisis, some research suggested that deteriorated underwriting standards were not the culprit of the crisis, but the final problem lay in changing underlying macroeconomic conditions such as housing price depreciation that began in late 2005 and the significant job losses that followed (Mayer et al, 2009). From 2000 through 2005, housing prices appreciated at an average annual rate of 11 percent and then depreciated at an average annual rate of 10 percent from 2006 to 2008. In times of appreciating housing markets, mortgage default is less pronounced because financially distressed borrowers can more easily sell their properties or refinance and prepay their loans (Danis & Pennington-Cross, 2005; Haughwout et al., 2008; Schloemer et al., 2006). Borrowers with negative equity, however, lack those opportunities and are more likely to default on their loans (Foote, Gerardi, and Willen, 2008; Gerardi et al, 2008). Yet, negative equity alone does not necessarily lead to default, rather default is often associated with “shocks” or “double triggers”, particularly unemployment or illness (Bhutta et al, 2010; Foster and Van Order, 1984; Foote et al., 2008).

Despite the rich body of research on foreclosures, very few studies focused specifically on spatial determinants of foreclosures. Some studies that noted spatial patterns, suggest that foreclosures are highly clustered in older urban neighborhoods where residents were predominantly minorities, of lower income, and with higher instances of subprime lending (Gramlich, 2007; Immergluck and Smith, 2005; and Nassar, 2007). In analyzing concentration of foreclosure filings in Atlanta, Apgar and Duda (2005) also showed foreclosure clusters in some suburban areas, including places with high minority populations and some in newly built subdivisions. However, though three-quarters of all foreclosures in the data were in suburbia, they were generally less concentrated than those in the city. The rate of foreclosure filings, on the other hand, was almost twice as high in the city than in the suburbs. Immergluck (2009a) recently described the accumulation of real estate owned (REO) properties within metropolitan areas by grouping the areas according to their initial foreclosures in August 2006 and changes in home value from August 2006 to August 2008. REO is real estate property owned by a lender, typically a bank, government agency, or government loan insurer, after an unsuccessful sale at a foreclosure auction (Roark, 2006). Using descriptive analysis at zip code level, Immergluck observed considerable variation in distribution of foreclosures in urban versus suburban neighborhoods across the groups of metropolitan areas. In traditionally weak housing markets, there were relatively large concentrations of foreclosures in central cities, but those levels were high even before the current crisis. On the other hand, in regions where a high incidence of foreclosures is a newer event, foreclosures are more concentrated in suburban areas. In addition, in markets with severe home value declines, such as often citied in Florida, Arizona, and California, concentrations of foreclosures are highly suburbanized and with longer commute times. In a follow up study, Immergluck (2010) analyzed intra-metropolitan differences in REO accumulation by applying a multivariate analysis and found that after controlling for higher risk lending during the subprime lending boom, suburbanization and commuting had no apparent bearing on REO growth. Other working papers which tested this theory found mixed results. Ong and Pfeiffer (2008), looking at foreclosures in Los Angeles County in early 2008, found that exurban location explained 20 percent of the spatial variation in foreclosure rates. The authors attributed the result to higher speculation on new homes but also to high commuting costs of the exurban areas which makes them more vulnerable to decreases in demand. In a different approach, Rauterkus et al. (2010) modeled the probability of mortgage default in Chicago, Jacksonville, and San Francisco based on differences in location efficiency. Location-efficient homes are located in areas that facilitate lower car ownership. The study found that mortgage default probability increases with the number of vehicles owned after controlling for income. Also, default probability decreases with higher Walk Scores in high income areas but increases with higher Walk Scores in low income areas. These studies are described in greater detail in the last essay which focuses on the spatial determinants of foreclosures.

Overall, the research on determinants of foreclosures provides evidence that lower credit scores, higher debt-to-income ratios and higher loan-to-value ratios at the time of origination, particularly among non-prime borrowers, lead to higher default rates. Also, the communities with lower incomes, higher capitalization rates - lower or uncertain anticipated housing price appreciation, higher credit risk and an older housing stock are more vulnerable to subprime lending. The correlation between largely black neighborhoods and subprime lending, controlling for other factors, is particularly strong. Conversely, the evidence is mixed on the relationship between neighborhood tenure, income, other minority population and subprime lending. Notably, studies looking into spatial distribution provide some insight into the spatial determinants of foreclosures and suggest a link between suburbanization and commuting distance on probability of default, though the link is still debated. Those studies are generally limited by their focus on a few metropolitan areas. Nevertheless, different results confirm the expectation that the effects of spatial determinants on foreclosures may vary in regions with different demographics, housing markets, and geographic patterns, as well as among different time periods.

Essay One: The News On The American Urban Form - Are American Metropolitan Areas Growing Smarter?

Following the decentralization and sprawling of U.S. metropolitan areas in 1970s and 1980s, concerns arose over manageability and sustainability of urban growth processes. While some attacked sprawl for its lack of aesthetics (Mumford, 1961), most critics argue that excessive urban expansion is unsustainable due to loss of open space, traffic congestion, increased air and water pollution, and fiscal costs of infrastructure associated with new low-density development (Duany et al, 2000; Downs, 1999, Brueckner, 2000a; Wu, 2006). Glaeser and Kahn (2003) also argued that the primary social problem associated with sprawl is social stratification between people who can afford cars and live in the suburbs and the abandoned ones that have no access to the variety of jobs and cannot live the car-dependent lifestyle.

In response to these concerns, support has grown for a set of concepts commonly known as smart growth. Smart growth is a toolbox of related land use policies which focus on the following objectives: (i) the location of development – by promoting compact development, preserving farmland and open spaces, protecting natural resources and environmental quality, investing in established communities; (ii) the development design – by providing a range of housing choices and supplying affordable housing, promoting distinctive communities and mix of land uses; (iii) transportation and land use interaction – by creating walkable communities with transportation options; and (iv) the community and stakeholder partnership – by encouraging stakeholder collaboration and making development decisions process transparent and effective (Smart Growth Network, 2011).

Smart growth programs in the context of urban spatial structure largely focus on increasing the density of use of existing urban areas and limiting the conversion of farm land and open spaces to residential and commercial land use. To achieve these goals, the specific policies generally implemented include urban growth boundaries, tax incentives to revitalize the downtown areas, changes in zoning codes to promote infill development, tax incentives to minimize the distance between home and work, transferable development rights and conservation of undeveloped land. Some development restrictions also implicitly lessen marginal negative externalities of some land uses, which could embody aesthetics.

Evaluation of the effectiveness of smart growth programs usually focuses on their ability to increase population density and limit spatial expansion of the urbanized areas. In the empirical evaluation of the programs, both Carruthers (2002) and Anthony (2004) found little evidence in support of smart growth. Though, Carruthers (2002) does find that Oregon's smart growth program led to increases in population density over time. This finding is consistent with arguments made by Burby and May (1997), and reaffirmed by Carruthers (2002a, 2002b) and Dawkins and Nelson (2004) that the institutional framework for growth management is a significant determinant of the program's effectiveness. Howell-Mulroney (2007), responding to the argument, classified state programs into weak, moderate and strong to examine whether the structure of state approaches makes a difference in development outcomes between states. Weak programs were identified in Georgia, New Jersey and Vermont; moderate programs were in Maine, Maryland and Rhode Island; while strong programs were in Florida, Oregon and Washington. The analysis, which included a longer observation period than previous studies, found lower rates of urban spatial expansion and higher population density increases in states with the most stringent smart growth programs, but no effect in states with weaker programs when compared to states with no smart growth programs. The author pointed out that states that do not have smart growth legislation may have lower growth development pressures and consequently do not feel the need for growth planning. Another more recent analysis evaluating development patterns in four smart growth states (Florida, Maryland, New Jersey, and Oregon) and four states without such programs (Colorado, Indiana, Texas, and Virginia) finds that between 1990 and 2000 smart growth states have fared better than the other states on some of the urban form dimensions (edited by Ingram et al, 2009). The smart growth states saw a higher share of infill development in already existing areas, though by an insignificant amount. And, the average rate of decentralization of population and employment in the metropolitan areas of the smart growth states were lower than in metropolitan areas of the other states. Other variables such as population and employment concentration and land consumption were not significantly different between the two groups of states.

This study builds on the literature that has attempted to examine the extent to which U.S. metropolitan areas have been sprawling. Literature suggests significant decentralization among cities in the South, particularly Atlanta, Raleigh and Greensboro, NC, Washington DC, and St. Louis (Galster et al., 2001; Ewing, Pendall and Chen, 2002; Sierra Club, 1998). Density of Western metropolitan regions increased the most; however, those are also regions constrained by natural and imposed geographical barriers (Fulton et al., 2001). In the second half of 1990s and into 2000s, older urban cores, in the Northeast in particular, were seeing some remarkable regeneration through infill development and population growth (Thomas, 2009). While offering interesting insight into dynamic nature of American urban areas, these studies have not probed deeper into changes that have occurred between 1990 and 2007, and particularly between 2000 and 2007. That was a dynamic era for many American urban areas due to the great real estate boom, but also an era in which smart growth policies burgeoned and would have presented themselves in changes in urban form. This study will try to understand how metropolitan areas have changed over the seventeen year period, between 1990, 2000 and 2007, and if the changes are consistent with previously observed patterns of sprawl or whether the trends are consistent with the proposition that U.S. cities are now growing smart.

In the context of this study, growing smart refers to increasing population density of existing urban areas and limiting outward expansion of new development. It is not in the scope of this study to specifically evaluate smart growth programs and their effectiveness. While their existence is important for urban spatial structure, the focus is oriented towards evaluating the general tendencies of metropolitan spatial structure in the United States to grow inward. The following three questions will be addressed: (1) Has the metropolitan spatial structure in thirty five largest metropolitan areas in the United States changed between 1990, 2000, and 2007? (2) Is the change consistent with decentralization trends observed prior to 1990? (3) Is the change consistent across all the metropolitan areas? The hypothesis which assumes that urban areas are growing denser and sprawling at a slower rate arises out of a number of descriptive studies suggesting densification of urban areas in the past decade, although most often measured along one dimension (i.e. Thomas, 2009). Also, as argued in the theoretical background section, households' revealed preferences suggest increased demand for urban amenities and proximity to urban centers. As a result, increase in population density of existing urban areas should be evident in states that do not employ smart growth programs as well as in those that do.

Empirical Strategy

To address these questions, I will compute and interpret multiple measures of urban form for the thirty five largest metropolitan areas in the United States. Metropolitan areas are defined by the United States Office of Management and Budget (OMB) and according to published standards that are applied to Census Bureau data. The general concept of a metropolitan or micropolitan statistical area consists of a core area containing a substantial population core and adjacent communities having a high degree of economic and social integration with that core. The term "core based statistical area" (CBSA) became effective in 2000 and refers collectively to metropolitan and micropolitan statistical areas[2]. The measures generated in this study include: density gradients, concentration indices, clustering indices, density frequency distributions, and growth allocation, and will be computed using normalized census tract data for 1990, 2000, and 2007.

Density Gradients

Density gradients measure the degree to which population density declines as distance to the city center increases. In describing metropolitan spatial structure, density

---

[2] Some metropolitan areas with population of 2.5 million or more are subdivided into metropolitan divisions. In this analysis, such metropolitan areas will be treated as one CBSA and will not be subdivided into metropolitan divisions. http://www.census.gov/population/www/metroareas/files/00-32997.pdf

gradients are used to measure the change in population density from the center to the urban periphery. Following Clark (1951), density gradients describe urban population densities using the negative exponential function form, showing that density declines exponentially from the central core towards the outskirts of a metropolitan area. Negative exponential function is defined as follows:

$$D(x) = D_0 e^{-yx}, \tag{1}$$

where *D(x)* represents population density at distance *x* from the center; D_0 is the density at the center; and *y* is the density gradient or the rate at which the population density decreases as one moves away from the center. After taking the natural logarithm of population density, the equation yields the linear equation and density gradient can be estimated via ordinary least squares:

$$\log D(x) = \alpha + \beta\,(x) + e. \tag{2}$$

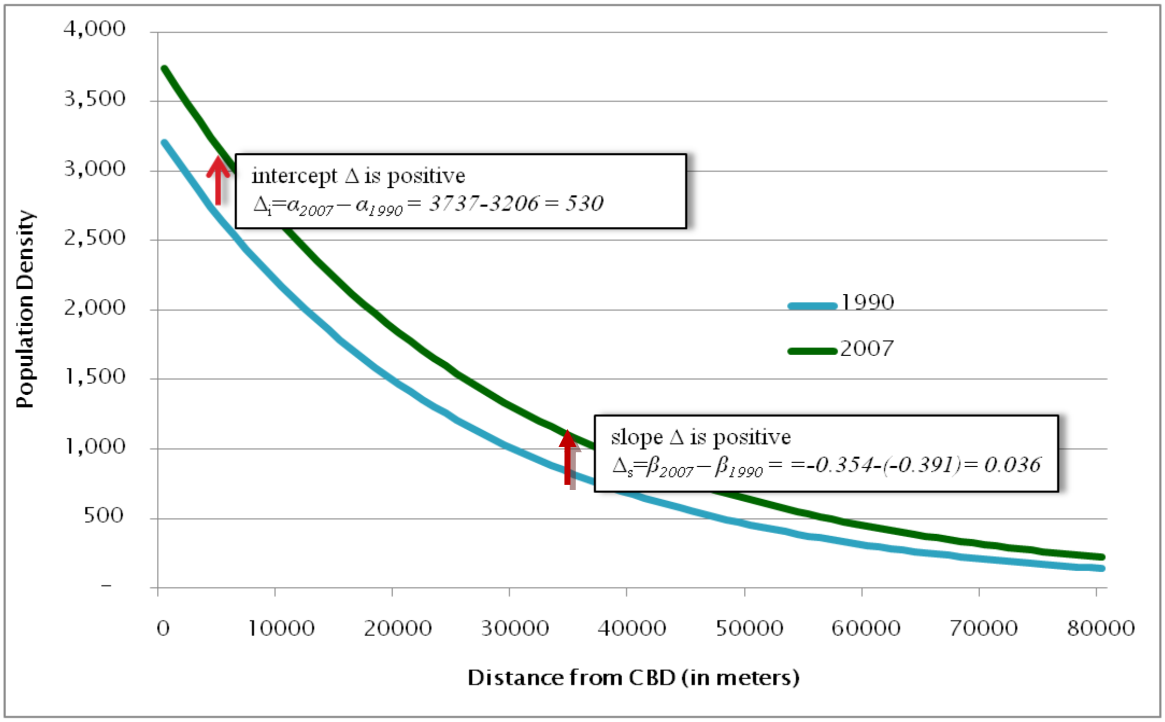
As the study will focus on the changes in central urban densities and density gradients in the period between 1990 and 2007, density gradient models are modified as follows and are estimated for three pairs of years - 1990-2000, 2000-2007, and 1990-2007:

$$\log D(x_{2000}) - \log D(x_{1990}) = (\alpha_{2000} - \alpha_{1990}) + (\beta_{2000} - \beta_{1990})x + e \tag{3}$$

where the dependent variable $\log D(x_{2000}) - \log D(x_{1990})$ is measured as a difference between the logged population density variable in year 2000 and year 1990. Equation (3) is repeated for the 2000-2007 and 1990-2007 pairs. *x* is the explanatory variable measured as distance to the nearest CBSA center in 2007. A constant distance explanatory variable is used for comparability of results. The CBSA centers are defined as the centroids of the census tracts where the Central Business District (CBD) was located according to the 1982 Census of Retail Trade. With specification such as the equation (3), α measures the intercept and change in population density at the core of the CBSA, and β measures the change in slope of the density gradient or the change in the rate at which population density decreases away from the core of the CBSA.

In the analysis, both the difference of the slope and the intercept of the density gradients are interesting. A negative slope coefficient on the gradient and a negative sign for the intercept imply fallen central city density and population densification of existing suburbs. Densification of existing suburbs differs from expansion of urban areas and increasing population density in outer suburbs. Such spatial structure, also described as decentralization and sprawl would be characterized with a positive sign on the slope coefficient coupled with a negative sign on the intercept. A negative sign on the slope coefficient of the gradient coupled with a positive sign on the intercept is indicative of centralization or smart growth. A positive sign on the intercept suggests that population density has increased in the central core, while the negative coefficient on the gradient implies that increases in population density have occurred in areas closer to the core. Finally, both positive coefficients on the slope and gradient suggest increase in central city population density but also greater population density in outer suburbs. In other words, when the coefficient on the change in gradient is positive, it means that gradient in an earlier year, say 1990, was larger than the coefficient in the following year, say 2000, and the gradient flattened between 1990 and 2000. Figure 2 illustrates intercept and slope changes.

Figure 2: Change in Density Gradient

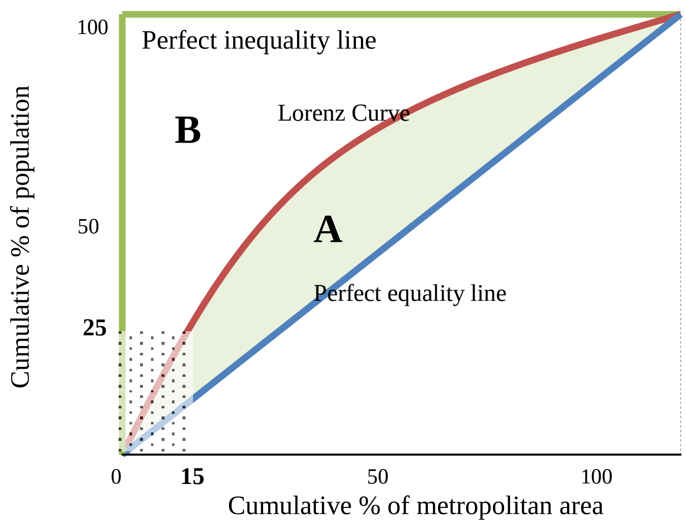


Concentration Indices

Concentration indices include Gini coefficients and Lorenz curves and provide information about the spatial concentration of population within metropolitan areas. They are used to measure inequality of population distribution by spatial units, census tracts, in a metropolitan area. The Gini coefficient has been commonly applied in the study of ecology, sociology, economics, and other sciences to measure statistical dispersion. It is most frequently used as a measure of inequality of income (Gini, 1936).

In the analysis of metropolitan spatial structure, higher concentrations of population are revealed by Gini coefficients that are closer to one and signify that population density is high in fewer sub-areas. A Gini coefficient close to zero means that population is evenly distributed in a metropolitan area. The Gini coefficients are usually mathematically illustrated by Lorenz curves. The Lorenz curve is a graphical representation of the cumulative distribution function of a probability distribution. The graph represented in the Figure 2 shows the cumulative percent of the metropolitan area (noted on the x-axis) that is inhabited by cumulative percentage of population in the same area (noted on the y-axis). For example, the rectangular shape in Figure 3 capturing 25 percent of the population (y-axis) over 15 percent of the area (x-axis) describes a metropolitan area where 25 percent of the population is located on 15 percent of the total metropolitan area.

Figure 3: Gini Coefficient



The Lorenz curve is a function of the cumulative proportion of ordered census tracts' populations mapped onto the corresponding cumulative proportion of their size. Given a metropolitan area with *n* ordered census tracts with x_i' the size of a census tract *i* and $x_1' < x_2' < \ldots < x_n'$, then the Lorenz curve is the polygon joining the cumulative percent population value and the cumulative percent area value for each tract. The Gini coefficient is the area in Figure 2 between the perfect equality line and the observed Lorenz curve as a percentage of the area between the line of perfect equality and the line of perfect inequality. Since the area between the line of perfect equality and the Lorenz curve is A, and the area beyond the Lorenz curve is B, the Gini index is A/ (A+B). Lorenz curves that are more bowed indicate higher concentration of population densities within a metropolitan area.

The implications of the Gini coefficient suggesting growing population concentration on the success of smart growth programs are somewhat ambiguous. Smart growth goals include compact development and rural land preservation. This can be achieved by concentrating development in urban areas and leaving rural areas undeveloped. In a two-region landscape that includes one urban region and one rural region, success in growth management could be measured by how close the Lorenz curve was to L-shaped and how close the Gini coefficient was to one (which indicates that all population was in the urban region and not in the rural region). In the more complicated real world, the optimal concentration of population is less clear. At the metropolitan scale, a more concentrated pattern of development is probably preferred as such a pattern is more likely to allow high-density and mixed use urban cores and a relatively undeveloped urban fringe. Even at the metropolitan scale, however, the optimal pattern of population concentration is difficult to define in cities with multiple centers of activity and relatively tight patterns of development around those centers. Therefore, unequal distribution may be better perceived as a general dimension of metropolitan form, rather than sprawl particularly. A high Gini coefficient indicates an unequal distribution, meaning a large number of people are concentrated in a small area. Subject to the qualifications discussed above, higher coefficients, which imply greater concentration of population, are indicative of smarter growth. It is important to note that these coefficients also reflect the geographic size of census tracts, physiographic constraints, and historical patterns. Also, as the focus of this study is the relative change of metropolitan spatial structure, it is not necessarily as important where a metropolitan area ranks in terms of population concentration, but how the indicator has changed over time.

Density Histograms

Density histograms, as used here, display the frequency of census tracts by density in three periods in time, and convey differences across the three periods. When evaluating urban form through the lens of smart growth, certain critical densities are necessary to accommodate certain types of transit and facilitate the types of development advocated by smart growth. Pushkarev and Zuban (1977) presented data on mode usage and density to make the point that certain levels of density are needed for the viability of types of transportation. In generating density distributions for each metropolitan area, attention is paid to which categories of density within metropolitan areas lost or gained population between 1990 and 2007. Here, histograms of density are generated in intervals of 500 persons per square kilometer for three points in time (1990, 2000, and 2007). Also, histograms of differences between pairs of years (1990-2000, 2000-2007, and 1990-2007) are generated. The shape of these histograms characterizes the distribution of tracts by density and difference histograms show how the distribution of tracts by density changed over time. Difference histograms provide a sense of whether growth occurred primarily in low density tracts, medium density tracts, or occurred in all tracts.

Clustering Index

The clustering index used here includes the global Moran coefficient. To address the question of metropolitan areas being monocentric, polycentric or decentralized sprawling and to complement the other spatial form measures, this study will also compute the global Moran's coefficient for each metropolitan area. Moran's coefficient is a frequently used spatial statistics tool which measures the degree of spatial autocorrelation, or in other words, it measures the extent to which adjacent observations of the same phenomenon are correlated. Values range from −1 indicating perfect dispersion to +1 indicating perfect correlation. A zero value indicates a random spatial pattern. The previous research has shown a high Moran coefficient to be indicative of monocentric spatial form. An intermediate Moran coefficient suggested polycentric form, while a low coefficient suggested decentralized metropolitan form (Tsai, 2005). The Moran's coefficient is defined as:

$$I = \frac{N\sum_i \sum_j W_{i,j}(X_i - \overline{X})(X_j - \overline{X})}{(\sum_i \sum_j W_{i,j})\sum_i (X_i - \overline{X})^2} \tag{4}$$

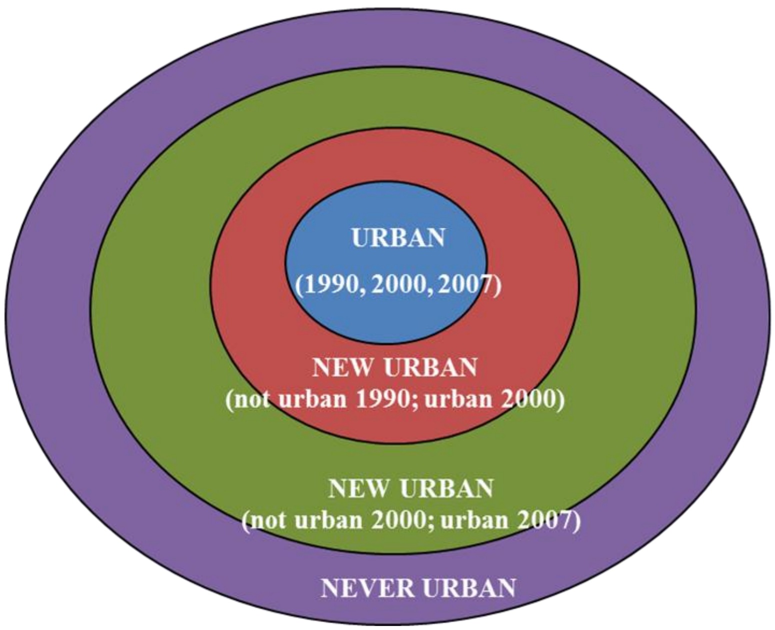
where, N is the number of census tracts; X_i is population density in the census tract i; X_j is population density in the census tract j; $\overline{X}$ is the mean of population density; and W_{ij} denotes the weighting between census tracts i and j. The weighting function is K-nearest neighbor centroids of census tracts, where K=10. K-nearest neighbor weights is used because *X* is a strongly skewed variable. Most of the tracts have relatively lower density while a much smaller share of them have high population density. In the case where the input variable is strongly skewed, some features will have very few neighbors and there may be instances where the Moran value falls outside the bounded range of -1.0 to 1.0. In such cases, each feature should be weighed with at least 8 neighbors (Getis and Ord, 1992). In this study, each tract is weighted with its 10 nearest neighbors and Moran's value of +1 indicates that high density census tracts are closely clustered, while -1 indicates they are scattered or exhibit a 'chessboard' pattern of development (Tsai, 2005).

While, similar to Gini coefficients, Moran coefficients alone do not reveal enough information about the metropolitan spatial structure; however it does characterize the component of the metropolitan spatial structure not addressed with other measures included in this study, namely clustering. As noted in the description of the concentration indices, concentration of population in higher density areas is critical for success of the smart growth policies and while Gini coefficients do not contribute by revealing the degree of clustering, the Moran coefficient will complement the analysis with such information.

Urbanization – Growth Allocation

Growth allocation, as defined here, apportions population growth from 1990 to 2007 in each metropolitan area to: (i) areas urbanized in 1990, (ii) areas urbanized between 1990 and 2000, (iii) areas urbanized between 2000 and 2007, and (iv) areas never urbanized. Growth allocation, also referred to as the urbanization indicator, focuses on the location and density of urban growth relative to the existing urban areas. That is, a census tract is considered urbanized if it had a population density greater than 1000 persons per square mile regardless of where that census tract is located. The 1000 persons per square mile threshold follows Census' delineation of urban and rural areas[3]. A schematic of the approach to this indicator is presented in Figure 4, which presumes, unrealistically, that all urban growth in a metropolitan area is centered around a single urban core. Also note that this is the only indicator which measures population density in people per square mile versus people per square kilometer. This was done to maintain consistency with Census delineation definition.

Figure 4: Urbanization Scheme



In general, smart growth policies prefer development in existing urban areas primarily. The next area of development should be new areas which are hopefully adjacent to the existing urban areas. Finally, the last place of development should be in rural areas. Further, but again with caveats, higher density development is preferred in urban and new urban areas and lower density is preferred in rural areas. Unlike density

[3] http://www.census.gov/geo/www/ua/ua_2k.html

gradients, this measure is not dependent on notions of a monocentric city. Each census tract is defined as either urban 1990, urbanized 1990-2000, urbanized 2000-2007, or never urbanized. The disadvantage of this measure is that it does not account for spatial location of new urbanization, but only identifies in which of the four categories new urbanization was allocated to.

Data

The analysis focuses on the thirty five largest CBSAs of the United States and uses 2000 census tracts as the units of analysis. Figure 1 in Appendix A lists the 35 CBSAs along with population statistics. The data comes from three sources: (i) Census Summary File 3 (SF-3), from the 2000 census of the population; (ii) ESRI 2007 Demographic Update; and (iii) two Geolytics, Inc. products which allocate selected 1990 SF-1 and SF-3 variables from 1990 census block groups and tract boundaries to 2000 block group and tract boundaries. Geolytics products allow for a meaningful comparison as they ensure the underlying geographic boundaries remain constant over time. Technical explanation of the Geolytics tract remapping methodology is available in their Data Users' Guide[4].

The ESRI 2007 Demographic Update provides population estimates for 2007 at census tract geography. To estimate the 2007 population count, ESRI uses three primary sources: (i) residential delivery statistics from the U.S. Postal Service (USPS), (ii) InfoBase database from Acxiom Corporation, and (iii) residential construction data from Hanley Wood Market Intelligence. The USPS publishes monthly counts of residential deliveries for every U.S. postal carrier route. This represents the most comprehensive and

[4] http://www.geolytics.com/Pages/NCDB/NCDB_variables/AppendixJ.pdf

current information available for small, sub-county geographic areas. To allocate a delivery address to block groups, ESRI relies on its proprietary Address-Based Allocation (ABA) method. This allocation method uses the addresses from Acxiom's InfoBase household database which are geocoded with carrier route and block group codes. ESRI tests its results extensively including benchmarking against the 2000 Census. For the small portion of block groups where addresses are not available from the InfoBase database, delivery statistics are allocated from a correspondence file. The correspondence between census block groups and postal carrier routes is developed using quarterly updated data from Tele Atlas. However, given that this analysis focuses on highly urban areas, such block group data would not be utilized. To track new housing developments, especially in previously unpopulated areas, ESRI uses data from Hanley Wood Market Intelligence which tracts new and planned residential construction in the largest 75 U.S. housing markets. This database identifies exact locations of individual construction projects, including, a complex of single-family homes, townhomes, or a condominium building. The database also tracks conversions of apartments into condominiums. The construction information includes: total number of units planned, inventory of units under construction, sold, and/or closed, type of housing—detached homes, townhomes, condominiums, and target markets—families, seniors, empty nesters. Finally, totals for block groups are controlled to the county totals. Again, this analysis focuses on 35 largest metropolitan areas where the input data for estimates are more extensive and complete. Detailed explanation of ESRI's method is available in the ESRI® Demographic Update Methodology[5].

[5]http://www.esri.com/library/whitepapers/pdfs/demographic-update-methodology-2007.pdf.The effectiveness of the ABA method highly depends on the precision of block group assignment to InfoBase

Results

Detailed results of all measures are available in tables and graphs contained in Appendix A.

Density Gradients

Using the data from 1990, 2000, and 2007, the analysis focuses on density gradients among the 35 metropolitan areas during those three periods and the difference in the three periods --- 1990 to 2000, 2000 to 2007, and 1990 to 2007. For density gradients in each of the three periods, the estimated coefficient on intercept, α, identifies population density at the central core of the metropolitan area. The coefficient on the slope, β, measures the density gradient. The difference measuring between three sets of years show how the slope and the intercept of the density gradients changed. In general, a negative slope coefficient on the gradient and a negative sign for the intercept imply fallen central city density and population densification of existing suburbs. A negative sign on the slope coefficient of the gradient coupled with a positive sign on the intercept is indicative of centralization or smart growth. A positive sign on the slope coefficient of the gradient coupled with a negative sign on the intercept implies decentralization and sprawl. Finally, both positive slope coefficient of the gradient and the intercept indicate population growth in both the inner ring and suburbs.

It is important to qualify that the analysis in this essay measures change in a closed city model. A closed city model in the case of this study means that a CBSA

---

addresses. ESRI used improved Dynamap/Address Points database from Tele Atlas, which provides coordinates that are accurate to the building,; however this database currently covers only the most densely populated areas in the United States. Addresses that fall outside the coverage were geocoded with the conventional approach, based on address ranges. Post office delivery counts or address counts provide less coverage in rural areas.

geographic boundary is held constant over time. Also, density gradient measure does not control for population growth in CBSAs. Consequently, in metropolitan areas that gained significant population growth over the observation period, the intercept and the slope of density gradient are naturally expected to show positive change. Thus, the estimated coefficient on the slope and the intercept may be more indicative of the type of change in spatial structure. Again, because this analysis evaluates constant area, all references to suburbs refer to those located inside the constant CBSA boundary. The change beyond this boundary is not accounted for and may not be consistent with the change estimated inside the boundary.

Density gradients in 1990

The analysis of density gradients begins with density gradients of metropolitan areas in 1990. The model specification indicating a natural logarithm of population density in 1990 is regressed on distance from the CBD. In the analysis, the population density variable is expressed as population per square kilometer and distance from the CBD is expressed in kilometers. To get a sense of general trends among the 35 observed metropolitan areas, density gradients were also generated by aggregating all observations. The estimated density gradient for 1990 among the 35 metropolitan areas is summarized as:

$$\textit{log(population density 1990)} = \widehat{8.073} - \widehat{0.391}\ \textit{distance to CBD} \qquad (5)$$

where the first coefficient is the estimated intercept of the gradient for 35 metropolitan areas and the second coefficient is the estimated density gradient, or the rate at which population density falls away from the central core. When the estimated coefficients are exponentiated, they yield the following result:

$$\textit{population density 1990} = \exp(8.073) + (\exp(-0.391)-1) * \textit{distance to CBD} \quad (6)$$

which equals to:

$$\textit{population density in 1990} = 3{,}206 - 32\% \text{ @ } 10km \textit{ from CBD} \quad (7)$$

where the intercept for 35 metropolitan areas yields density of 3,206 people per square kilometer, and falls by almost 32 percent 10 kilometers away from the center. Naturally, the estimates differ significantly among metropolitan areas with New York City having highest intercept, at over 14,000 people per square km and Las Vegas having the smallest intercept, at 400 people per square km. Also, the gradient falls at different rates among the CBSAs with Miami losing about 12 percent of population density and Denver losing over 66 percent of population density 10 km from the CBSA core. Table 1 below contains transformed coefficients for the 35 metropolitan areas for 1990, 2000, and 2007. Appendix A contains estimated output coefficients.

Density gradients in 2000

In 2000, the estimated density gradient among the 35 metropolitan areas suggests that population density has increased in the central core while the gradient has fallen:

$$\textit{population density in 2000} = 3{,}583 - 31\% \text{ @ } 10km \textit{ from CBD} \quad (8)$$

The results again differ among the metropolitan areas. The estimates continue to rank New York City as having highest intercept, but in 2000, Charlotte, NC ranked with the lowest intercept. Miami and Denver maintained their ranking at either end of the gradient continuum, with Miami losing 11 percent of population density 10 km out of the center and Denver losing 66 percent of population density.

Density gradients in 2007

The aggregated result for 2007 for 35 metropolitan areas suggests that central city population density has further increased. Also, density gradient has further flattened:

population density in 2007 = 3,737 – 30% @ 10km from CBD (9)

Table 1 contains estimates for all metro areas and CBSAs are listed in alphabetical order.

Table 1: Density Gradients 1990, 2000, 2007

Metropolitan Areas	1990		2000		2007	
	Gradient (@10km)	Intercept	Gradient (@10km)	Intercept	Gradient (@10km)	Intercept
Atlanta-Sandy Springs-Marietta, GA	-43%	2,089	-41%	2,363	-39%	2,587
Austin-Round Rock, TX	-52%	1,618	-50%	2,244	-48%	2,460
Baltimore-Towson, MD	-53%	4,245	-50%	3,967	-48%	3,882
Boston-Cambridge-Quincy, MA-NH	-37%	4,516	-36%	4,615	-35%	4,568
Charlotte-Gastonia-Concord, NC-SC	-41%	965	-41%	1,251	-41%	1,464
Chicago-Naperville-Joliet, IL-IN-WI	-36%	6,462	-34%	6,590	-33%	6,645
Cincinnati-Middletown, OH-KY-IN	-46%	2,256	-44%	2,194	-42%	2,076
Cleveland-Elyria-Mentor, OH	-43%	3,077	-40%	2,755	-39%	2,656
Dallas-Fort Worth-Arlington, TX	-31%	2,099	-30%	2,676	-29%	2,835
Denver-Aurora, CO	-66%	4,238	-61%	4,679	-58%	4,593
Detroit-Warren-Livonia, MI	-37%	3,880	-34%	3,524	-33%	3,385
Houston-Sugar Land-Baytown, TX	-36%	2,126	-35%	2,497	-34%	2,779
Indianapolis, IN	-55%	2,087	-52%	2,129	-51%	2,113
Jacksonville, FL	-42%	1,281	-39%	1,332	-37%	1,447
Las Vegas-Paradise, NV	-31%	403	-33%	2,296	-33%	2,895
Los Angeles-Long Beach-Santa Ana, CA	-33%	7,875	-30%	7,682	-29%	7,787
Miami-Fort Lauderdale-Miami Beach, FL	-12%	2,242	-11%	2,968	-11%	3,171
Minneapolis-St. Paul-Bloomington, MN-WI	-53%	3,271	-52%	3,698	-49%	3,530
New York-Northern New Jersey-Long Island, NY-NJ-PA	-39%	14,368	-38%	15,172	-38%	16,192
Orlando-Kissimmee, FL	-42%	1,337	-40%	1,766	-37%	1,953
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	-34%	3,431	-32%	3,255	-31%	3,216
Phoenix-Mesa-Scottsdale, AZ	-26%	1,110	-25%	1,979	-24%	2,328
Pittsburgh, PA	-40%	2,393	-39%	2,213	-38%	2,115
Portland-Vancouver-Beaverton, OR-WA	-59%	3,041	-58%	3,716	-58%	3,998
Richmond, VA	-48%	1,343	-46%	1,438	-45%	1,485
Riverside-San Bernardino-Ontario, CA	-18%	1,413	-16%	1,857	-16%	2,133
Sacramento-Arden-Arcade-Roseville, CA	-27%	1,268	-27%	1,929	-27%	2,173
St. Louis, MO-IL	-45%	2,701	-41%	2,355	-40%	2,296
San Antonio, TX	-58%	2,408	-55%	2,670	-54%	2,831
San Diego-Carlsbad-San Marcos, CA	-31%	3,009	-28%	3,537	-27%	3,729
San Francisco-Oakland-Fremont, CA	-37%	6,665	-35%	6,943	-35%	7,088
Seattle-Tacoma-Bellevue, WA	-37%	2,995	-34%	3,282	-33%	3,428
Tampa-St. Petersburg-Clearwater, FL	-22%	1,303	-22%	1,611	-22%	1,819
Virginia Beach-Norfolk-Newport News, VA-NC	-43%	3,271	-40%	3,137	-38%	3,056
Washington-Arlington-Alexandria, DC-VA-MD-WV	-42%	3,768	-38%	3,744	-36%	3,824
All metropolitan Areas	-32%	3,206	-31%	3,583	-30%	3,737

Change between 1990 and 2000

The change in density gradients for three sets of years is measured for all metropolitan areas individually as well as in an aggregated sample. The estimate for change between 1990 and 2000 for the aggregated sample is as follows:

$$log\ D(x_{2000}) - log\ D(x_{1990}) = (\hat{\alpha}_{2000} - \hat{\alpha}_{1990}) + (\hat{\beta}_{2000} - \hat{\beta}_{1990})\ \overline{x} + \hat{e} \tag{10}$$

where

$$(\hat{\alpha}_{2000} - \hat{\alpha}_{1990}) = 0.111,\ and \tag{11}$$

$$(\hat{\beta}_{2000} - \hat{\beta}_{1990}) = 0.021. \tag{12}$$

Positive coefficients on both the slope and the gradient indicate that overall the largest metropolitan areas grew denser in the center while the gradient flattened. Flattening of the gradient indicates that population density decreased at a slower pace, also signifying that suburban areas gained in population density. Over the decade, intercept in the core grew by 12 percent (*(exp (0.111)-1)*100)* while the gradient flattened by 2.15 percent (*(exp (0.021)-1)*100)*. Note that positive value on change of the gradient indicates that the gradient in the later year has lower absolute value than the gradient in the former year.

The change varies some when metropolitan areas are observed individually. Table 2 below clusters the metropolitan areas according to the change in the intercept and gradient between 1990 and 2000. There are three groups of observed changes. The first group, which also covers largest number of metro areas, experienced increases in the intercept and flattening of the gradients. Over seventy percent or 21 of the 35 observed metropolitan areas fall in the first group and have seen significant increasing population density in the urban core. Four more areas saw some marginal increase as well. Among all the 25 metropolitan areas with increasing population in the central core, two thirds also had increasing population density in the suburban communities. The second group, with a positive change in intercept and fallen gradient, include three metropolitan areas, Las Vegas, Sacramento and Tampa. These three areas experienced some centralization, though the estimates are marginal and not significant. Finally, the third group includes metro areas that saw decrease in the intercept and flatter gradient. These areas in fact decentralized or sprawled during 1990s and include mostly older industrialized cities in the rust belt.

The largest positive and significant change on the intercept was noted in Las Vegas, Nevada, where the change was three times as large as the next metropolitan area, Phoenix, Arizona. Sacramento, California, Austin, Texas and Miami, Florida, followed. Among these five, Austin metropolitan area also saw the largest increase in population density in suburban areas, suggested by large and significant β coefficient. Miami and Phoenix grew out but at insignificant levels, while Las Vegas and Sacramento, and Tampa centralized but also at insignificant levels. Remaining metropolitan areas with increasing population in the center and throughout the metropolitan region, the first group, are expectedly newer metros in South and West, but also several in Midwest and Northeast.

The third group, comprising 30 percent of the observed metropolitan areas saw declining population in the central core, with six having significant depopulation. With a positive and significant slope coefficient, it appears the population moved to the suburban areas in all of them. Many are older metropolitan areas that also had overall population loss during this time. The largest and highly significant decrease in central city population density was in St Louis, MO-IL, where also population significantly migrated to suburban areas. Other older, largely rust-belt metropolitan areas followed, such as Cleveland, Detroit, Pittsburg, Baltimore, and Philadelphia. Cincinnati, Los Angeles, Washington DC, and Virginia Beach also had falling central cities but smaller and insignificant, with increasing population density in the suburban communities.

Table 2: Change in Population Density 1990 to 2000

+Intercept & + Gradient	+Intercept & - Gradient	-Intercept & + Gradient
Atlanta-Sandy Springs-Marietta, GA	Las Vegas-Paradise, NV	Baltimore-Towson, MD
Austin-Round Rock, TX	Sacramento-Arden-Arcade-Roseville, CA	Cincinnati-Middletown, OH-KY-IN
Boston-Cambridge-Quincy, MA-NH	Tampa-St. Petersburg-Clearwater, FL	Cleveland-Elyria-Mentor, OH
Charlotte-Gastonia-Concord, NC-SC		Detroit-Warren-Livonia, MI
Chicago-Naperville-Joliet, IL-IN-WI		Los Angeles-Long Beach-Santa Ana, CA
Dallas-Fort Worth-Arlington, TX		Philadelphia-Camden-Wilmington, PA-NJ-DE-MD
Denver-Aurora, CO		Pittsburgh, PA
Houston-Sugar Land-Baytown, TX		St. Louis, MO-IL
Indianapolis, IN		Virginia Beach-Norfolk-Newport News, VA-NC
Jacksonville, FL		Washington-Arlington-Alexandria, DC-VA-MD-WV
Miami-Fort Lauderdale-Miami Beach, FL		
Minneapolis-St. Paul-Bloomington, MN-WI		
New York-Northern New Jersey-Long Island, NY-NJ-PA		
Orlando-Kissimmee, FL		
Phoenix-Mesa-Scottsdale, AZ		
Portland-Vancouver-Beaverton, OR-WA		
Richmond, VA		
Riverside-San Bernardino-Ontario, CA		
San Antonio, TX		
San Diego-Carlsbad-San Marcos, CA		

San Francisco-Oakland-Fremont, CA		
Seattle-Tacoma-Bellevue, WA		

Change between 2000 and 2007

The estimate for change between 2000 and 2007 for the aggregated sample is:

$$log\ D(x_{2007}) – log\ D(x_{2000}) = (\hat{\alpha}_{2007} – \hat{\alpha}_{2000}) + (\hat{\beta}_{2007} – \hat{\hat{\beta}}_{2000})\ \overline{\chi} + \hat{\varepsilon} \quad (13)$$

where

$$(\hat{\alpha}_{2007} – \hat{\alpha}_{2000}) = 0.042, \text{ and} \quad (14)$$

$$(\hat{\beta}_{2007} – \hat{\hat{\beta}}_{2000}) = 0.015. \quad (15)$$

The estimates suggest that although the population density trends of the 1990s decade continued into 2000s, the change from 2000 to 2007 was much less dramatic. The intercept continued to increase, by 4.3 percent, while the gradient flattened by 1.5 percent. Examined individually, most metropolitan areas gained population density in the urban core but also continued densifying in the suburban communities. Las Vegas continued with the highest increase in the intercept, though at lower magnitude than in the decade before. Followed by similar increases were Phoenix and Charlotte. Riverside and Tampa ranked third and fourth. Charlotte was also the only metro area that experienced some insignificant centralization of the suburban population while the other four showed increase in population density on the suburban fringe.

Table 3 groups the metropolitan areas by the type of change seen between 2000 and 2007. The third group, which lost population density in the center between 1990 and 2000, was joined by Boston and Minneapolis after 2000, though Minneapolis' loss was insignificant. The largest significant drop was in Cincinnati and Pittsburg, followed by a marginally significant drop in Cleveland. Expectedly, all these metropolitan areas decentralized with significant increase in population density on the urban fringe. Though intercept in Denver only marginally and insignificantly decreased, Denver had the largest increase in population density at the urban fringe. Following with population growth on the fringe were Minneapolis, Orlando, and Cincinnati. Again, Charlotte was the only area with some marginal centralization. The first group where population growth focused on both inner central areas and suburbs was the largest and joined by several new metropolitan areas. Washington, DC's metropolitan area reversed the 1990s trend from losing population in the central core to small but significant gains after 2000 and continued increasing population density in suburban areas. Las Vegas, Tampa and Sacramento also changed from focusing the population growth inward to expansion on the suburban fringe; however the change in the gradient is only marginally significant or non-significant among all three.

Table 3: Change in Population Density 2000 to 2007

+Intercept & + Gradient	+Intercept & - Gradient	-Intercept & + Gradient
Atlanta-Sandy Springs-Marietta, GA	Charlotte-Gastonia-Concord, NC-SC	Baltimore-Towson, MD
Austin-Round Rock, TX		Boston-Cambridge-Quincy, MA-NH
Chicago-Naperville-Joliet, IL-IN-WI		Cincinnati-Middletown, OH-KY-IN
Dallas-Fort Worth-Arlington, TX		Cleveland-Elyria-Mentor, OH
Houston-Sugar Land-Baytown, TX		Denver-Aurora, CO
Jacksonville, FL		Detroit-Warren-Livonia, MI
Las Vegas-Paradise, NV		Indianapolis, IN
Los Angeles-Long Beach-Santa Ana, CA		Minneapolis-St. Paul-Bloomington, MN-WI
Miami-Fort Lauderdale-Miami Beach, FL		Philadelphia-Camden-Wilmington, PA-NJ-DE-MD
New York-Northern New Jersey-Long Island, NY-NJ-PA		Pittsburgh, PA

Orlando-Kissimmee, FL		St. Louis, MO-IL
Phoenix-Mesa-Scottsdale, AZ		Virginia Beach-Norfolk-Newport News, VA-NC
Portland-Vancouver-Beaverton, OR-WA		
Richmond, VA		
Riverside-San Bernardino-Ontario, CA		
Sacramento-Arden-Arcade-Roseville, CA		
San Antonio, TX		
San Diego-Carlsbad-San Marcos, CA		
San Francisco-Oakland-Fremont, CA		
Seattle-Tacoma-Bellevue, WA		
Tampa-St. Petersburg-Clearwater, FL		
Washington-Arlington-Alexandria, DC-VA-MD-WV		

Change between 1990 and 2007

Finally, the estimate for change between the entire period, 1990 and 2007, for the aggregated sample is as follows:

$$\log D(x_{2007}) - \log D(x_{1990}) = (\hat{\alpha}_{2007} - \hat{\alpha}_{1990}) + (\hat{\beta}_{2007} - \hat{\beta}_{1990})\,\overline{\chi} + \hat{\varepsilon} \qquad (16)$$

where

$$(\hat{\alpha}_{2007} - \hat{\alpha}_{1990}) = 0.153, \text{ and} \qquad (17)$$

$$(\hat{\beta}_{2007} - \hat{\beta}_{1990}) = 0.036. \qquad (18)$$

In general, over the entire observation period, American metropolitan areas grew denser both in the urban core as well as on the urban fringe. Relatively more drastic change can be seen in the central cores than in the suburban areas. Population density in the core grew by almost 17 percent over 17 years as the intercept rose from 3206 people per square km in 1990 to 3737 in 2007. The gradient flattened by 3.7 percent which suggests that density was 3.7 percent higher 10 km outside the center in 2007 than was in 1990.

Twenty seven out of thirty five metropolitan areas experienced increasing population in the center and are included in the first group in Table 4. Twenty two observed significant changes. Thirty two metro areas also expanded their population density on the urban fringe. Among the metropolitan areas that grew smarter, Las Vegas clearly experienced the highest degree of population increase in the urban center, almost three times the magnitude of the second ranking city, Phoenix. Sacramento, Austin and Charlotte followed. Charlotte and Las Vegas did centralize some but not significantly. Austin expanded significantly at the urban fringe, as did Phoenix and Sacramento but only marginally.

Out of the five metropolitan areas with insignificant changes over the observation period, Denver, Boston and Indianapolis increased in central city density between 1990 and 2000 and then decreased between 2000 and 2007. Washington lost density in the urban core prior to 2000 and then gained it significantly after the year 2000. In contrast, Chicago increased significantly denser before 2000, but since then, the change has been insignificant.

Metropolitan areas that did not growth smart and decentralized over the seventeen-year period were again old industrial areas with St. Louis experiencing the greatest degree of decentralization. Following were Cleveland, Detroit, Pittsburg, and Baltimore. Smaller and insignificant changes were observed in Los Angeles and Virginia Beach. Table 4 groups the metropolitan areas according to the type of spatial structure change seen between 1990 and 2007.

Table 4: Change in Population Density 1990 to 2007

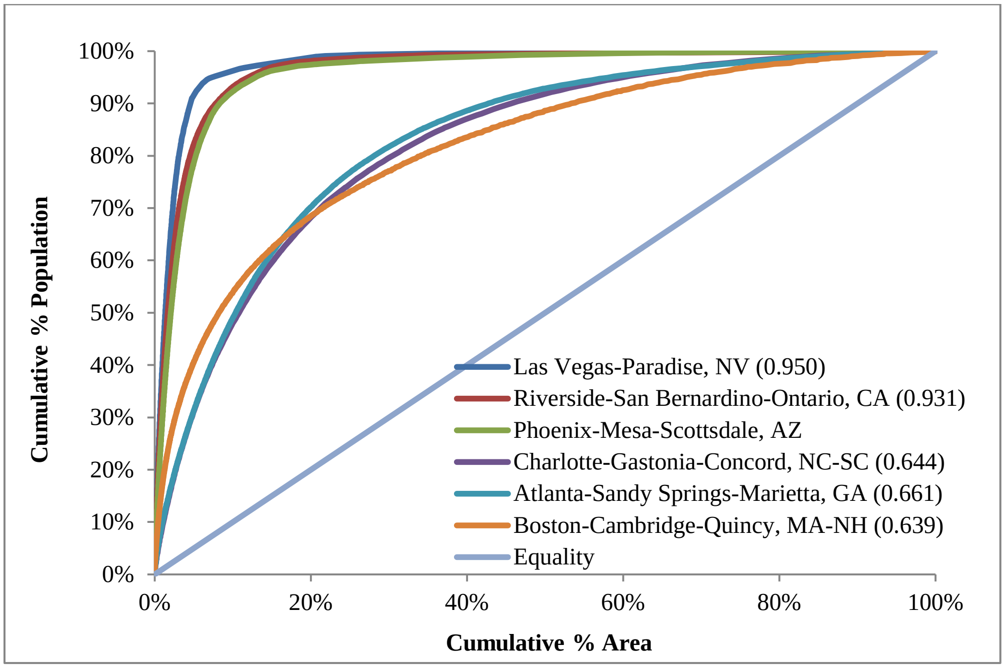
+Intercept & + Gradient	+Intercept & - Gradient	-Intercept &+ Gradient
Atlanta-Sandy Springs-Marietta, GA	Charlotte-Gastonia-Concord, NC-SC	Baltimore-Towson, MD
Austin-Round Rock, TX	Las Vegas-Paradise, NV	Cincinnati-Middletown, OH-KY-IN
Boston-Cambridge-Quincy, MA-NH	Tampa-St. Petersburg-Clearwater, FL	Cleveland-Elyria-Mentor, OH
Chicago-Naperville-Joliet, IL-IN-WI		Detroit-Warren-Livonia, MI
Colorado Springs, CO		El Paso, TX
Dallas-Fort Worth-Arlington, TX		Los Angeles-Long Beach-Santa Ana, CA
Denver-Aurora, CO		Philadelphia-Camden-Wilmington, PA-NJ-DE-MD
Houston-Sugar Land-Baytown, TX		Pittsburgh, PA
Indianapolis, IN		St. Louis, MO-IL
Jacksonville, FL		Virginia Beach-Norfolk-Newport News, VA-NC
Miami-Fort Lauderdale-Miami Beach, FL		
Minneapolis-St. Paul-Bloomington, MN-WI		
New York-Northern New Jersey-Long Island, NY-NJ-PA		
Orlando-Kissimmee, FL		
Phoenix-Mesa-Scottsdale, AZ		
Portland-Vancouver-Beaverton, OR-WA		
Richmond, VA		
Riverside-San Bernardino-Ontario, CA		
Sacramento-Arden-Arcade-Roseville, CA		
San Antonio, TX		
San Diego-Carlsbad-San Marcos, CA		
San Francisco-Oakland-Fremont, CA		
Seattle-Tacoma-Bellevue, WA		
Washington-Arlington-Alexandria, DC-VA-MD-WV		

Concentration

Concentration measured via Gini coefficients varied tremendously among the metropolitan areas over the seventeen-year period. It ranged from high 0.974 in 1990 in Las Vegas to low 0.639 in Boston in 2007. Areas with high Gini coefficients are considered concentrated with highly unequal population distribution across smaller areas. Metropolitan areas with smaller Gini coefficients have more equal distribution of population across the area. Gini coefficients do not address the location of an area's concentration. Thus even if the estimated result suggests an urban area has become increasingly concentrated that could mean around any focal point within a metro area, not only the urban core. For the aggregated sample of 35 metropolitan areas, the 1990 Gini coefficient equaled 0.873, and it fell to 0.858 in 2000 and to 0.844 in 2007. Falling Gini coefficient suggests that most of the areas across the county grew less concentrated.

Figure 5 illustrates differences in Gini coefficients across several metropolitan areas in 2007. The figure shows Lorenz curves which are graphical representations of Gini coefficients. A metropolitan area, such as Las Vegas, which has a very high Gini coefficient, 0.950, is depicted with a very bowed Lorenz curve, while an area that is less concentrated, such as Atlanta with a Gini equaling 0.66, has a less bowed curve. In Las Vegas that means that 83 percent of CBSA's population is located on 5 percent of its area. Conversely, in Atlanta, 30 percent of population occupies 5 percent of the metropolitan area.

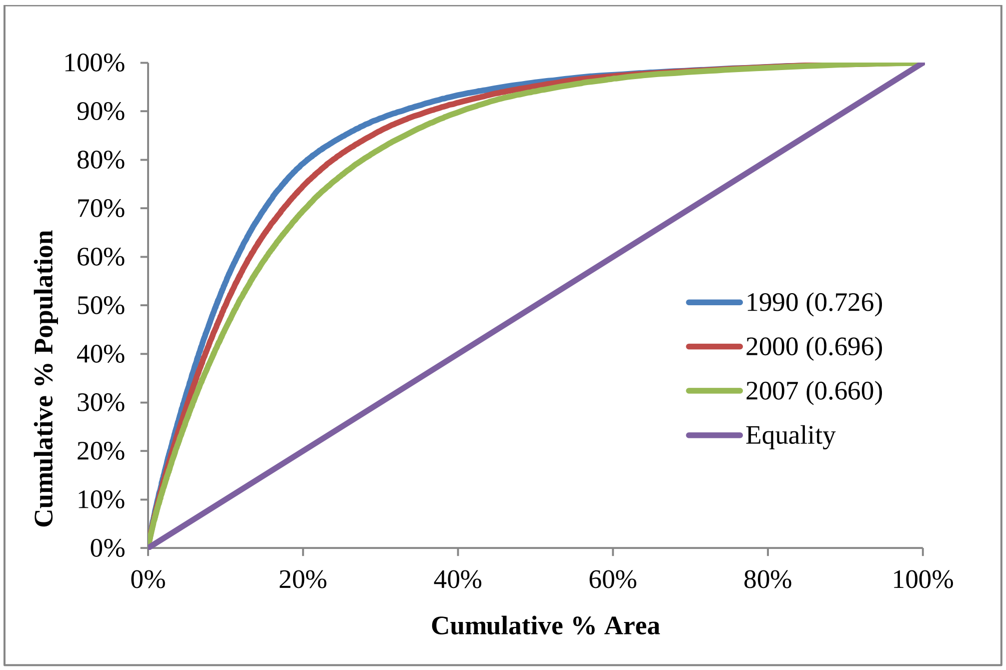
Figure 5: Lorenz Curves – Selected metropolitan areas, 2007



Las Vegas, Nevada, was the most concentrated area in all three periods. Its Gini ranged from 0.974 in 1990 to 0.965 in 2000 and 0.950 in 2007. Riverside, California, and Phoenix, Arizona, closely followed although Phoenix experienced larger drop in overall concentration than Riverside did. Denver and Miami ranked fourth and fifth in 1990, however Miami was replaced with Portland in 2000 and 2007. Boston, Massachusetts, Charlotte, North Carolina, and Atlanta, Georgia were the least concentrated in the three periods with the Gini coefficients ranging between 0.639 and 0.712. These metropolitan areas, thus, have a more even distribution of population across the area. As discussed in *Empirical Strategy* section, Gini coefficient is sensitive to geographic units. Because Las Vegas CBSA consists of large tracts of rural lands surrounding the urban center, Gini coefficient ranks it as the most concentrated. Atlanta is on the other side of the spectrum. Atlanta CBSA is composed of largely equal area census tracts.

Almost all metropolitan areas grew less concentrated over the seventeen years. However, three metropolitan areas continued to grow more concentrated: Portland, Oregon, Houston, Texas, and St. Louis, Missouri. Figure 6 illustrates change in Lorenz curves between 1990 and 2007 in Tampa, Florida. Tampa had the largest absolute change over the period and its Gini consistently declined.

Figure 6: Lorenz Curve – Tampa-St. Petersburg-Clearwater, FL – 1990, 2000, 2007



Between 1990 and 2000, the population of four metropolitan areas grew more concentrated. Concentration of Portland, Oregon, increased the most with an increase in the Gini coefficient of 0.017. Houston, Texas, Minneapolis, Minnesota, and St. Louis, Missouri also grew more concentrated but the coefficient changed only marginally, 0.003 for Houston and 0.002 for Minneapolis and St Louis. The metropolitan areas where population concentration decreased the most over the decade were Detroit, Tampa, Austin, Orlando, and Philadelphia. The greatest decline was in Detroit where the Gini coefficient fell by 0.033. The median change in Gini coefficients between 1990 and 2000 among all metropolitan areas equaled -0.016.

Between 2000 and 2007, the median decrease in Gini coefficient was smaller, -0.013. Four metropolitan areas did grow more concentrated. Houston led with the highest increase in Gini coefficient, 0.008, followed by a 0.007 increase in Sacramento. St. Louis, and Portland also had higher Gini coefficients in 2007 but only marginally, 0.004 and 0.002 respectively. Sacramento however deconcentrated relatively more after 2000 than in the decade before, so the difference over seventeen years still indicated deconcentration of that area. Metropolitan areas that deconcentrated the most in the new millennium were Minneapolis, where the coefficient fell by 0.046, and Orlando and Tampa where the Gini was 0.036 lower.

From 1990 to 2007, metropolitan areas showed varying changes in concentration. Median change reflected a decrease in Gini coefficient, by 0.031. Some metropolitan areas, including Portland, Houston and St. Louis increased in concentration. The increase was largest for Portland where the Gini coefficient increased by 0.019. Houston's Gini increased by 0.011 and St. Louis' by 0.006, which are relatively small and may not suggest any significant change in spatial structure. All other metropolitan areas deconcentrated over the time period. Deconcentration was greatest in Tampa, Orlando, Atlanta, Detroit, and Philadelphia. These decreases in Gini coefficients ranged from 0.066 in Tampa to 0.048 in Philadelphia. Table 5 lists the estimated Gini coefficients for all metropolitan areas as well as changes over the three periods.

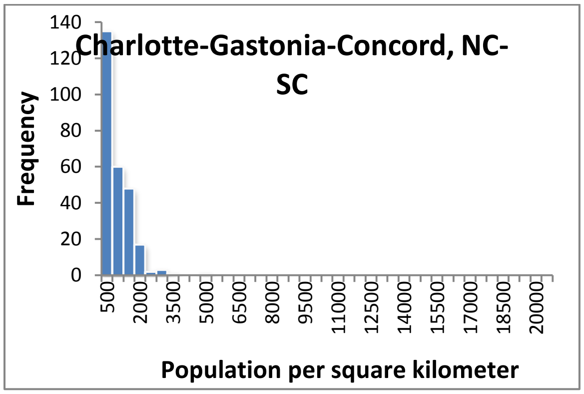
Table 5: Gini Coefficients 1990, 2000, 2007

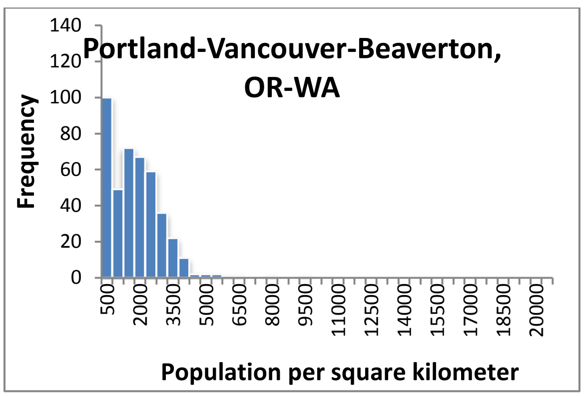
Metropolitan Area	Gini 1990	Gini 2000	Gini 2007	1990-2000	2000-2007	1990-2007
Atlanta-Sandy Springs-Marietta, GA	0.712	0.688	0.661	(0.024)	(0.027)	(0.052)
Austin-Round Rock, TX	0.811	0.782	0.764	(0.029)	(0.019)	(0.047)
Baltimore-Towson, MD	0.751	0.731	0.716	(0.020)	(0.015)	(0.035)
Boston-Cambridge-Quincy, MA-NH	0.671	0.652	0.639	(0.020)	(0.013)	(0.032)
Charlotte-Gastonia-Concord, NC-SC	0.675	0.651	0.644	(0.024)	(0.007)	(0.031)
Chicago-Naperville-Joliet, IL-IN-WI	0.808	0.787	0.765	(0.021)	(0.021)	(0.042)
Cincinnati-Middletown, OH-KY-IN	0.760	0.734	0.721	(0.026)	(0.014)	(0.039)
Cleveland-Elyria-Mentor, OH	0.724	0.698	0.680	(0.026)	(0.018)	(0.044)
Dallas-Fort Worth-Arlington, TX	0.817	0.795	0.771	(0.022)	(0.024)	(0.046)
Denver-Aurora, CO	0.920	0.905	0.892	(0.015)	(0.013)	(0.028)
Detroit-Warren-Livonia, MI	0.748	0.715	0.697	(0.033)	(0.017)	(0.051)
Houston-Sugar Land-Baytown, TX	0.823	0.826	0.834	0.003	0.008	0.011
Indianapolis, IN	0.760	0.745	0.732	(0.015)	(0.013)	(0.028)
Jacksonville, FL	0.794	0.771	0.749	(0.023)	(0.022)	(0.044)
Las Vegas-Paradise, NV	0.974	0.965	0.950	(0.010)	(0.015)	(0.025)
Los Angeles-Long Beach-Santa Ana, CA	0.777	0.777	0.772	(0.001)	(0.005)	(0.005)
Miami-Fort Lauderdale-Miami Beach, FL	0.894	0.871	0.860	(0.023)	(0.010)	(0.034)
Minneapolis-St. Paul-Bloomington, MN-WI	0.795	0.797	0.751	0.002	(0.046)	(0.044)
New York-Northern New Jersey-Long Island, NY-NJ-PA	0.786	0.779	0.773	(0.007)	(0.006)	(0.013)
Orlando-Kissimmee, FL	0.820	0.793	0.757	(0.027)	(0.036)	(0.063)
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.745	0.718	0.698	(0.027)	(0.021)	(0.048)
Phoenix-Mesa-Scottsdale, AZ	0.946	0.934	0.919	(0.012)	(0.015)	(0.027)
Pittsburgh, PA	0.721	0.708	0.702	(0.013)	(0.006)	(0.019)
Portland-Vancouver-Beaverton, OR-WA	0.882	0.899	0.901	0.017	0.002	0.019
Richmond, VA	0.789	0.773	0.767	(0.016)	(0.007)	(0.022)
Riverside-San Bernardino-Ontario, CA	0.942	0.936	0.931	(0.006)	(0.006)	(0.012)
Sacramento-Arden-Arcade-Roseville, CA	0.884	0.871	0.878	(0.013)	0.007	(0.006)
San Antonio, TX	0.879	0.863	0.847	(0.016)	(0.016)	(0.032)
San Diego-Carlsbad-San Marcos, CA	0.888	0.880	0.872	(0.008)	(0.008)	(0.016)
San Francisco-Oakland-Fremont, CA	0.816	0.805	0.795	(0.011)	(0.010)	(0.021)
Seattle-Tacoma-Bellevue, WA	0.866	0.855	0.849	(0.011)	(0.006)	(0.017)
St. Louis, MO-IL	0.830	0.832	0.836	0.002	0.004	0.006
Tampa-St. Petersburg-Clearwater, FL	0.726	0.696	0.660	(0.030)	(0.036)	(0.066)
Virginia Beach-Norfolk-Newport News, VA-NC	0.803	0.783	0.773	(0.020)	(0.010)	(0.030)
Washington-Arlington-Alexandria, DC-VA-MD-WV	0.782	0.760	0.737	(0.022)	(0.023)	(0.045)

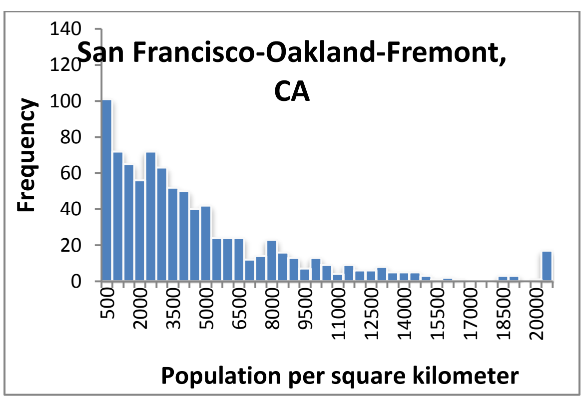
Density Histograms

Frequency histograms measure frequency of densities by category. The first category groups tracts with 0 to 500 people per square kilometer (ppsqkm). Categories increase in 500 ppsqkm increments, such as 500 to 1000, 1000 to 1500, and up to 20,000 ppsqkm. Figure 7 illustrates density frequency histograms for four metropolitan areas which have notably different metropolitan spatial structures. All of the four metropolitan areas are graphed on the same frequency scale to highlight the degree to which they differ. Charlotte, North Carolina, falls in the group of metropolitan areas which is dominated by census tracts where population density does not exceed 500 ppsqkm. Density frequency following the first bin (0-500 ppsqkm) falls precipitously. There are also very few, if any, census tracts exceeding population density of more than 3000 ppsqkm. Some of the other metropolitan areas that follow similar spatial structures are Atlanta, Baltimore, and Houston. Orlando, Florida, falls into the category of metropolitan areas where frequency of tracts with density that is greater than 500 ppsqkm does not decline as precipitously as in the first group. Census tracts with population densities between 500 and 2500 people are still rather frequent. Some metropolitan areas in this category are Austin, Chicago, Riverside, and Washington, DC. The third category describes metropolitan areas where density histograms follow a polynomial-like pattern. First, there are a large number of lowest density tracts, but that frequency falls for density bin between 500-1000 ppsqkm. The frequency then picks up again for tracts with densities between 1500 and 3500 ppsqkm. Some of the metropolitan areas that illustrate this pattern are Denver, Las Vegas, Miami, and Tampa. Finally, large metropolitan areas, such as San Francisco, Los Angeles, and New York, fall into the last group which generally exhibits the right tail of its density distribution extending much further and with tracts containing in excess of 8000 ppsqkm. Appendix A contains histograms for all metropolitan areas for all three years as well as changes between three sets of years.

Figure 7: Density Histograms in 2000





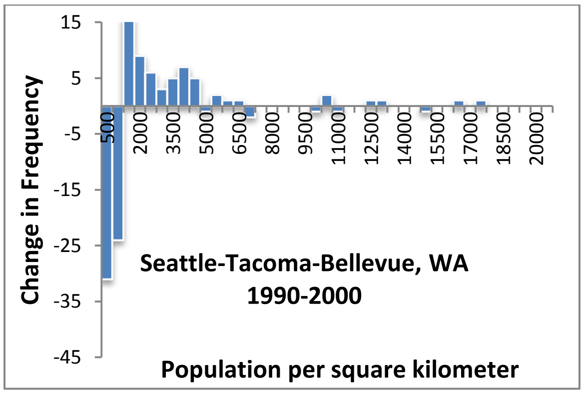


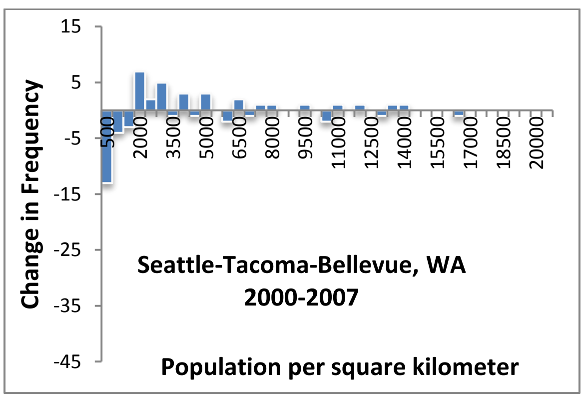
In observing changes in the frequency histograms over seventeen years across the metropolitan areas, there are some general trends that emerge. The number of the lowest density tracts, those with population density of up to 500 ppsqkm, decreased in all metropolitan areas between 1990 and 2007. At the median, there were 39 fewer lowest density tracts. Between 1990 and 2000, Las Vegas was the only area that gained lowest density tracts, 11 of them;, however it lost 15 of them between 2000 and 2007. Between 2000 and 2007, Baltimore and Philadelphia were the only metros areas that gained in lowest density tracts. Median change for medium density tracts was positive suggesting that the number of medium density tracts increased over the period. Medium density tracts are considered those with population density exceeding 500 ppsqkm but below 5000.

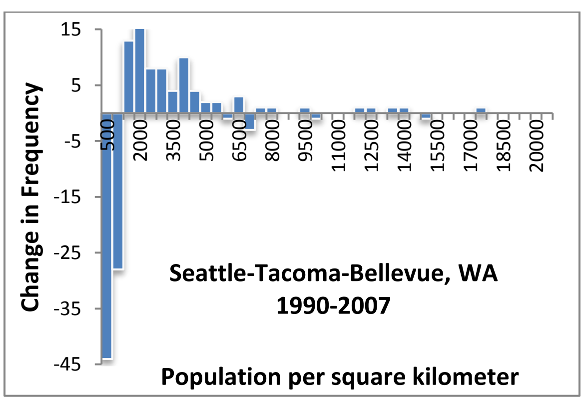
Another trend evident across most of the metropolitan areas is that they exhibit higher activity in density changes in 1990s. After 2000, magnitude of changes slows some except for Charlotte, Los Angeles, Riverside and Las Vegas. Los Angeles shows a large decrease in the number of low and medium density tracts, while Riverside experienced high increase in the number of medium density tracts. Las Vegas continually had increased activity during the whole observation period.

When looking at changes across metropolitan areas in the number of high or low density tracts, four categories emerge. Falling in the first category, “high infill”, are metropolitan areas for which the total number of lower density census tracts declined, but the number of medium and high density tracts increased. These histograms are characterized by tails that extend far to the right. Metropolitan areas in this category include Los Angeles, Miami, San Diego, San Francisco, and Seattle. Figure 8 illustrates density histogram changes for the "high infill" metropolitan areas with Seattle serving as a representative case. In Seattle, there were 44 fewer tracts with population density of 0-500 ppsqkm and 28 fewer tracts with 500-1000 ppsqkm. As shown in the figure, the bulk of the change occurred during 1990s

Figure 8: Density Histograms for High Infill Regions

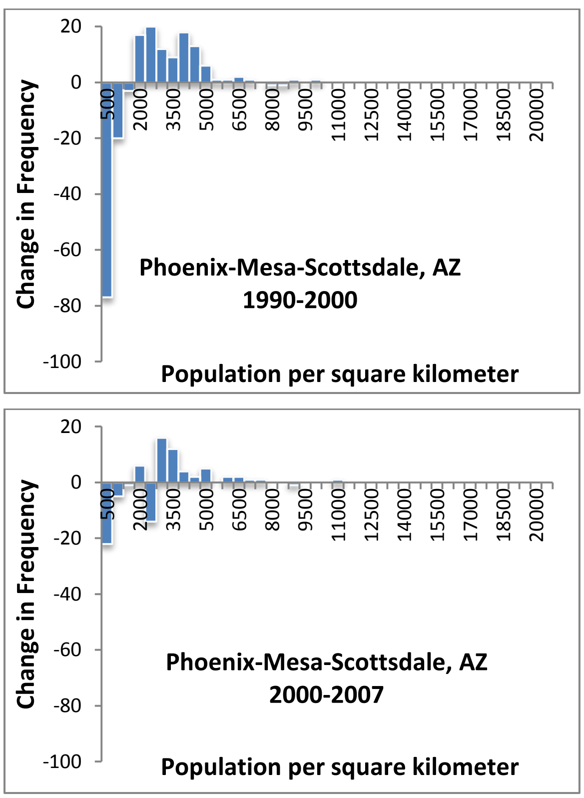


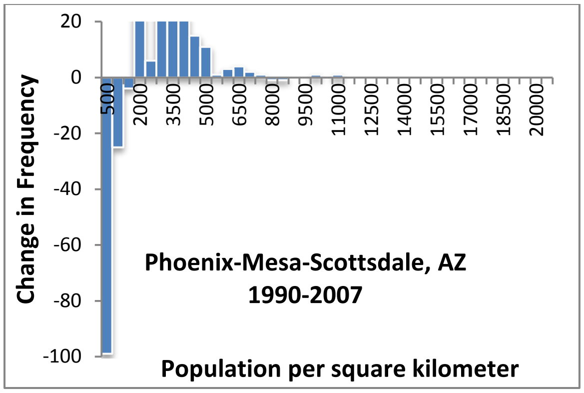




In the second category, “medium infill”, are metropolitan areas for which the total number of low density tracts, below 1500 ppsqkm, decreased, and the number of medium density tracts increased. In these metropolitan areas, little change happened in the high density tails and most of the change occurred below densities of 8000 ppsqkm. Dallas, Denver, Houston, Phoenix, Portland, and Virginia Beach are considered as “medium infill” metropolitan regions. Figure 9 depicts “medium infill” category via changes in Phoenix, Arizona.

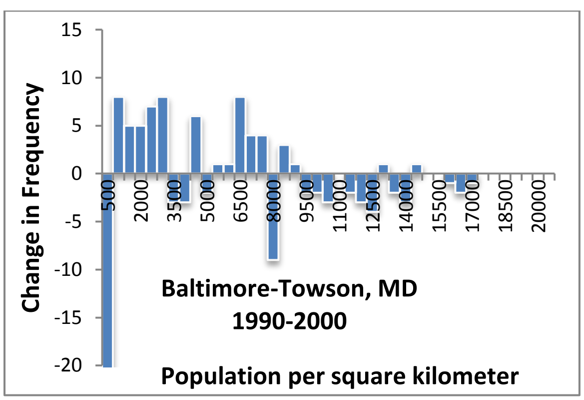
Figure 9: Density Histograms for Medium Infill Regions

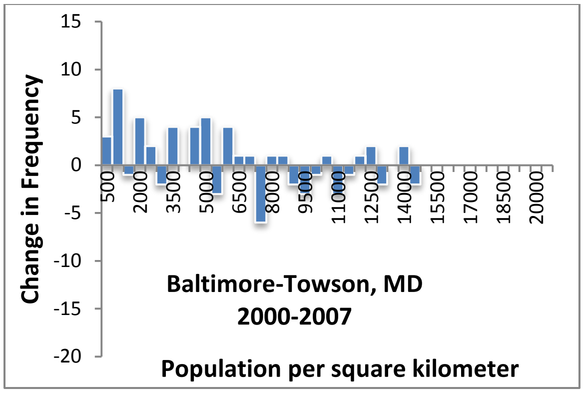


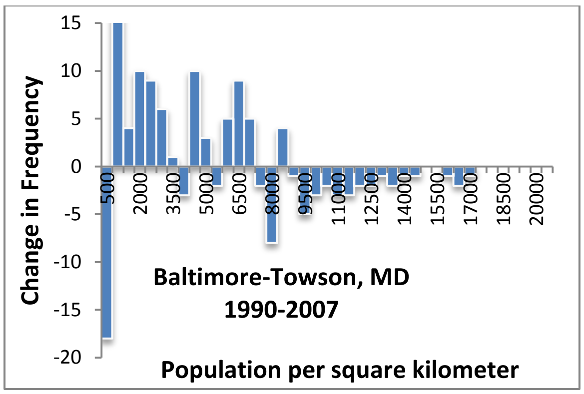


In the third, "mixed", category are metropolitan areas with no distinct trend towards infill or sprawl. While these metropolitan areas generally lost lowest density tracts, gains and losses in higher density tracts do not follow a specific pattern. Falling in this category are older, rust-belt regions in addition to some West Coast cities. Baltimore, Boston, Chicago, Cincinnati, Cleveland, Detroit, Philadelphia, Pittsburgh, and Washington, DC, all fall in the mixed category. Figure 10 illustrates changes in Baltimore as an example of the "mixed category".

Figure 10: Density Histograms for Mixed Regions







In the fourth category, best characterized as "sprawl", are metropolitan areas for which increases in frequency of census tracts occurred primarily in lower density tracts. These metropolitan areas had fewer tracts with population density under 500 ppsqkm, but more tracts with population density between 1000-2000 ppsqkm. These metropolitan areas also had little activity in the right tail of the difference distribution where population densities are high. Unsurprisingly, the majority of metropolitan areas fell into this category including Atlanta, Austin, Charlotte, Indianapolis, Jacksonville, Las Vegas, Orlando, Richmond, Riverside, Sacramento, St. Louis, and Tampa. Figure 11 depicts changes in Las Vegas. Las Vegas is a unique example even within the "sprawl" category because sprawling changes occurred mostly during 1990s whereas the other metropolitan areas saw that period as mostly densification phase. Las Vegas in 2000s experienced rather mixed changes; however, the region at the end gained more of the lower density tracts.

Figure 11: Density Histograms in Sprawling Regions

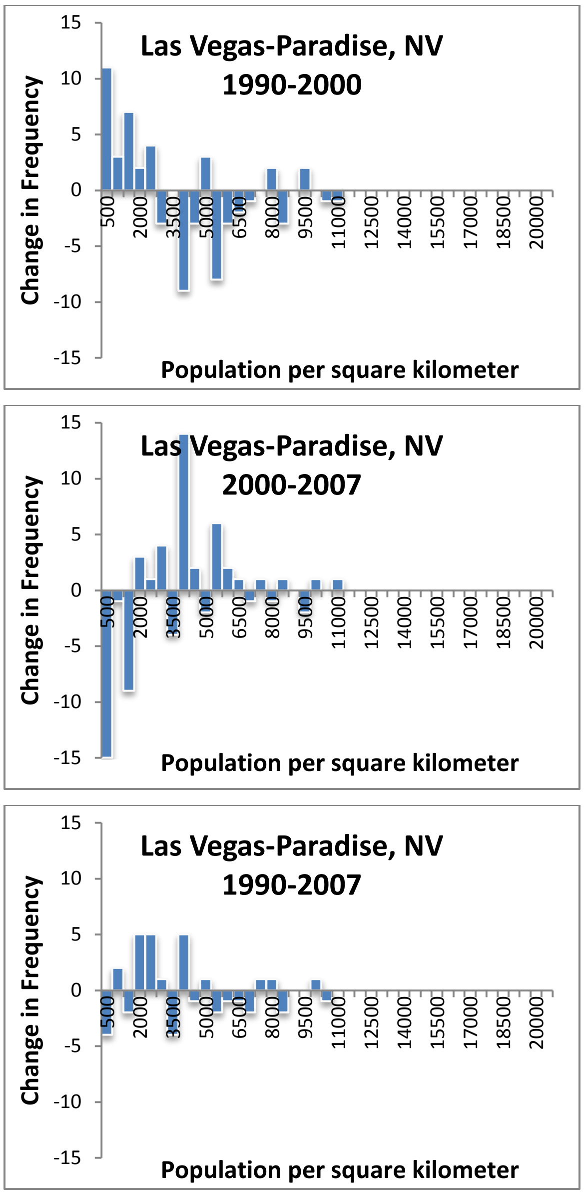


Table 6 distinguishes metropolitan areas in four groups based on the changes in density histograms they have experienced. At the end, the category named sprawl covers the largest share of metropolitan areas. These are also older industrials areas, as well as some high growth metros in the South. Metropolitan areas in California dominate the "high infill" category.

Table 6: Categories of Density Histogram Changes

High Infill	Medium Infill	Mixed	Sprawl
Los Angeles-Long Beach-Santa Ana, CA	Austin-Round Rock, TX	Boston-Cambridge-Quincy, MA-NH	Virginia Beach-Norfolk-Newport News, VA-NC
Miami-Fort Lauderdale-Miami Beach, FL	Arlington, TX	Chicago-Naperville-Joliet, IL-IN-WI	Atlanta-Sandy Springs-Marietta, GA
Riverside-San Bernardino-Ontario, CA	Dallas-Fort Worth-	Jacksonville, FL	Baltimore-Towson, MD
Sacramento-Arden-Arcade-Roseville, CA	Denver-Aurora, CO	Las Vegas-Paradise, NV	Phoenix-Mesa-Scottsdale, AZ
San Francisco-Oakland-Fremont, CA	Houston-Sugar Land-Baytown, TX	New York-Northern New Jersey-Long Island, NY-NJ-PA	Cincinnati-Middletown, OH-KY-IN
Tampa-St. Petersburg-Clearwater, FL	Orlando-Kissimmee, FL	Washington-Arlington-Alexandria, DC-VA-MD-WV	Charlotte-Gastonia-Concord, NC-SC
San Antonio, TX	Portland-Vancouver-Beaverton, OR-WA		Cleveland-Elyria-Mentor, OH
San Diego-Carlsbad-San Marcos, CA	Seattle-Tacoma-Bellevue, WA		Detroit-Warren-Livonia, MI
			Indianapolis, IN
			Minneapolis-St. Paul-Bloomington, MN-WI
			Philadelphia-Camden-Wilmington, PA-NJ-DE-MD
			Pittsburgh, PA
			Richmond, VA
			St. Louis, MO-IL

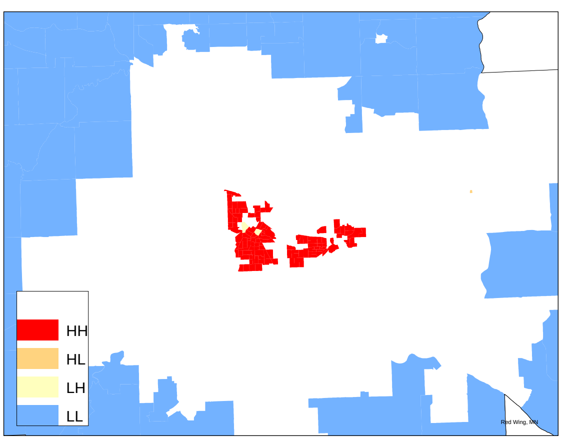
Clustering Index

The Moran's I index generated for each of the 35 metropolitan areas in this study is intended to show degrees of clustering of similar density tracts. Moran's value of +1 indicates that similar density census tracts are closely clustered, while the value of -1 suggests they are scattered or exhibit a 'chessboard' pattern of development. Moran's value closer to zero suggests random scattering. As expected following urban economic theory, all metropolitan areas exhibit at least some level of population density clustering suggested by the positive and highly significant value of the Moran's I. The estimated Moran's I in this sample range from 0.28611 to 0.72525 and encompasses the level of clustering from decentralized sprawling for low Moran's I values to monocentric for high Moran's I values. The values in the middle of the range indicate polycentric spatial structure. The reason for this categorization is that if high density tracts are completely clustered, producing high Moran's I value, they would describe monocentric urban form. If high density tracts are randomly distributed, there would resemble decentralized sprawling form. Polycentric form would have some level of concentration characterized by Moran's values in the middle of the range. It is important to emphasize that this analysis focuses on population distribution in a metropolitan area which may or may not be consistent with employment distribution. Thus, when an area is considered polycentric, it contains multiple population clusters and not necessarily employment clusters.

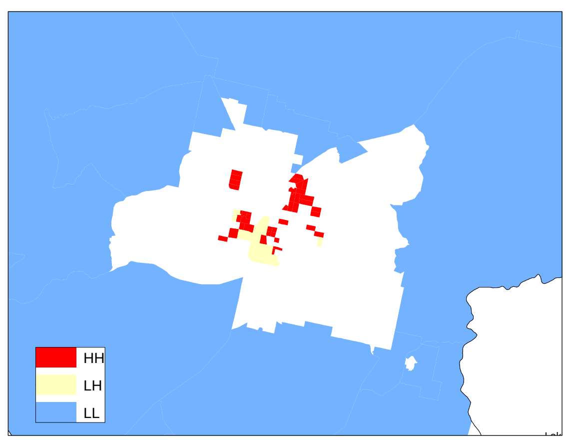
Moran's I index referred to so far indicates Global Moran's I, though the index can be disaggregated to provide a series of local indices, called Local Moran's I or Local Indicators of Spatial Association (LISA). The local statistics provide a spatial autocorrelation measure for each census tract in a metropolitan area and indicate if a tract has population density that is correlated with values in surrounding tracts. While it was more interesting to focus on Global Moran's I for the purposes of this study, the following figure illustrates how LISAs translate into a global one and how monocentric, polycentric and decentralized sprawling spatial structures differ. Figure 12 shows census tracts in three metropolitan areas thematically mapped based on census tracts' cluster value. A census tract surrounded by similar population densities will result in high positive z-score and thus be assigned HH for a statistically significant (0.05 level) cluster of high density values and LL for a statistically significant (0.05 level) cluster of low density values. If the tract has high density but is surrounded by low density tracts, it is labeled HL. And, if a low density tract is surrounded by high density tracts, it is labeled LH. The areas in white are population tracts that do not exhibit any significant degree of density clustering. Since Minneapolis was the most clustered metropolitan area, based on the highest estimated Moran's I, it is a representative sample of monocentric spatial structure. The second featured metropolitan area is Las Vegas, which, given its lowest Moran's I in 2000 and 2007, represents decentralized sprawling spatial structure. Finally, the last metropolitan area, which represents polycentric spatial structure, is Indianapolis. Indianapolis was chosen because its 2000 Moran's I value fell in the middle of the range of 2000 Moran's values.

Figure 12: Moran's I for Monocentric, Polycentric and Decentralized Spatial Structure

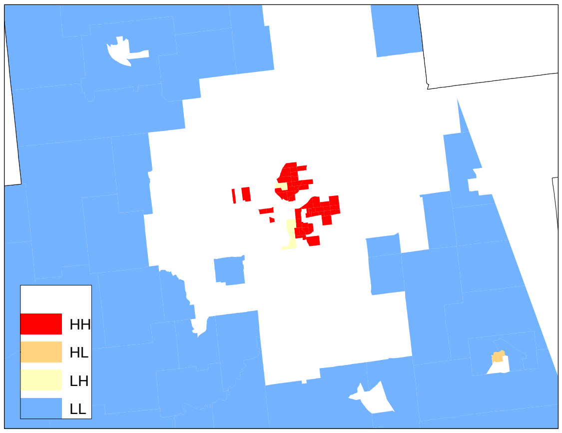
Minneapolis



Las Vegas



Indianapolis



In 1990, the most clustered metropolitan areas were Minneapolis, with Moran's I of 0.716, Philadelphia with 0.701, New York with 0.676, St. Louis with 0.670, and Atlanta with 0.66874. The relative ranking of metropolitan areas by their levels of clustering changed some over the observation period. In 2000, Baltimore with Moran's I of 0.652 replaced Atlanta as the fifth most clustered metro area. The top four remained the same with only a small change in ranking and estimated Moran's Is. In 2007, St. Louis was replaced with Seattle with value of 0.656, while the other four, Minneapolis, New York, Philadelphia and Baltimore remained as the most clustered metropolitan areas. Not all of them, though, changed in the same way over the seventeen years. While Baltimore and Philadelphia grew less clustered, New York, and especially Minneapolis, grew more clustered during the period.

The metropolitan areas that exhibited the least spatial autocorrelation among tract densities in 1990 were Riverside with Moran's I of 0.349, Las Vegas with 0.389, Charlotte with 0.422, Dallas with 0.436, and Orlando with 0.439. In 2000, Boston (Moran's I of 0.333) replaced Charlotte among the five least clustered metro areas, while again the other four lowest ranking metropolitan areas remained the same. In 2007, the ranking did not change. For these least clustered metropolitan areas, Moran's I over the seventeen-year period continuously decreased suggesting these areas grew less clustered. The largest decrease in Moran's I was in Boston metropolitan area, followed by Las Vegas. Using Tsai's (2005) classification, these metropolitan areas would be classified as decentralized sprawl.

The metropolitan areas with Moran's I index falling in the middle of the range for at least two out of three observation periods and classified as polycentric include Indianapolis, San Diego, Portland, San Antonio, Pittsburgh, Cleveland.

Between 1990 and 2007, most metropolitan areas grew less clustered. Six out of 35 metropolitan in contrast grew more clustered. Metro areas that grew more clustered include Chicago, Cincinnati, Minneapolis, New York, Richmond, and Seattle. The median Moran's I value also decreased, from 0.5582 in 1990, to 0.50034 in 2000 and 0.48906 in 2007. During the 1990s, Cincinnati, Chicago, and New York had the largest increase in Moran's' I. The highest decrease in Moran's I was observed in Atlanta, Boston, and Jacksonville. After 2000, seven metropolitan areas grew more clustered while the rest continued to the trend of decreasing clustering. Seattle had the largest increase in the Moran's I coefficient between 2000 and 2007, followed by Charlotte, Baltimore, Los Angeles, New York, and San Francisco, . In contrast, among the metropolitan areas that grew less clustered, Boston and Tampa had the largest change in the Moran's I.

Although clustering trends changed for some areas between 1990 to 2000 and 2000 to 2007, six metropolitan areas grew more clustered over the entire observation period, Chicago, Cincinnati, Minneapolis, New York, Richmond, and Seattle. For four of these though, most of the change occurred before 2000 followed by a smaller reversal after 2000. Only New York grew more clustered over the entire period. For the metropolitan areas experiencing the largest decrease in clustering, the trend was continuous over the whole period. The metropolitan areas with largest decrease in Moran's I included: Atlanta, Boston, Jacksonville, Las Vegas, and Tampa, with decrease ranging from -0.20 to -0.11.

Table 7 provides the estimated Moran's Index for 1990, 2000, and 2007, and change in index for the three set of years.

Table 7: Clustering Index and Change, 1990, 2000, 2007

Metropolitan Area	1990	2000	2007	2000-1990	2007-2000	2007-1990
Atlanta-Sandy Springs-Marietta, GA	0.669	0.485	0.471	(0.184)	(0.014)	(0.198)
Austin-Round Rock, TX	0.445	0.420	0.403	(0.025)	(0.017)	(0.042)
Baltimore-Towson, MD	0.663	0.653	0.657	(0.011)	0.004	(0.006)
Boston-Cambridge-Quincy, MA-NH	0.519	0.382	0.333	(0.137)	(0.049)	(0.186)
Charlotte-Gastonia-Concord, NC-SC	0.423	0.413	0.423	(0.010)	0.010	(0.000)
Chicago-Naperville-Joliet, IL-IN-WI	0.608	0.634	0.633	0.027	(0.001)	0.026
Cincinnati-Middletown, OH-KY-IN	0.560	0.601	0.593	0.042	(0.009)	0.033
Cleveland-Elyria-Mentor, OH	0.561	0.542	0.531	(0.019)	(0.011)	(0.030)
Dallas-Fort Worth-Arlington, TX	0.437	0.361	0.351	(0.075)	(0.010)	(0.085)
Denver-Aurora, CO	0.493	0.471	0.476	(0.022)	0.005	(0.017)
Detroit-Warren-Livonia, MI	0.666	0.629	0.612	(0.037)	(0.017)	(0.053)
Houston-Sugar Land-Baytown, TX	0.483	0.437	0.433	(0.046)	(0.004)	(0.050)
Indianapolis, IN	0.558	0.510	0.489	(0.048)	(0.021)	(0.069)
Jacksonville, FL	0.508	0.417	0.406	(0.091)	(0.011)	(0.102)
Las Vegas-Paradise, NV	0.389	0.317	0.286	(0.072)	(0.031)	(0.103)
Los Angeles-Long Beach-Santa Ana, CA	0.610	0.585	0.592	(0.024)	0.007	(0.018)
Miami-Fort Lauderdale-Miami Beach, FL	0.485	0.458	0.440	(0.027)	(0.018)	(0.045)
Minneapolis-St. Paul-Bloomington, MN-WI	0.716	0.725	0.725	0.009	(0.000)	0.009
New York-Northern New Jersey-Long Island, NY-NJ-PA	0.676	0.702	0.709	0.026	0.007	0.033
Orlando-Kissimmee, FL	0.439	0.388	0.370	(0.052)	(0.018)	(0.069)
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.702	0.691	0.687	(0.010)	(0.004)	(0.015)

Phoenix-Mesa-Scottsdale, AZ	0.495	0.429	0.405	(0.066)	(0.025)	(0.091)
Pittsburgh, PA	0.574	0.549	0.540	(0.025)	(0.009)	(0.034)
Portland-Vancouver-Beaverton, OR-WA	0.567	0.537	0.524	(0.031)	(0.012)	(0.043)
Richmond, VA	0.588	0.591	0.590	0.004	(0.002)	0.002
Riverside-San Bernardino-Ontario, CA	0.350	0.333	0.319	(0.017)	(0.014)	(0.030)
Sacramento-Arden-Arcade-Roseville, CA	0.456	0.434	0.427	(0.022)	(0.007)	(0.029)
San Antonio, TX	0.574	0.541	0.520	(0.033)	(0.021)	(0.055)
San Diego-Carlsbad-San Marcos, CA	0.499	0.500	0.498	0.002	(0.002)	(0.001)
San Francisco-Oakland-Fremont, CA	0.612	0.601	0.607	(0.011)	0.005	(0.005)
Seattle-Tacoma-Bellevue, WA	0.615	0.586	0.656	(0.028)	0.070	0.041
St. Louis, MO-IL	0.671	0.636	0.629	(0.035)	(0.006)	(0.041)
Tampa-St. Petersburg-Clearwater, FL	0.549	0.485	0.441	(0.063)	(0.044)	(0.108)
Virginia Beach-Norfolk-Newport News, VA-NC	0.458	0.399	0.388	(0.058)	(0.012)	(0.070)
Washington-Arlington-Alexandria, DC-VA-MD-WV	0.645	0.610	0.607	(0.036)	(0.003)	(0.038)

Urbanization - Growth Allocation

The results for growth allocation measures are very interesting and vary considerably among metropolitan areas. As defined here, a census tract is considered urban if its population density is greater than 1000 persons per square mile regardless of where that census tract is located. The 1000 persons per square mile threshold follows the U.S. Census Bureau' delineation of urban and rural areas[6]. The growth allocation indicator examines whether new growth occurring between 1990 and 2007 was allocated to already existing urban areas or was it placed in previously non-urbanized areas.

Between 1990 and 2007, 47 percent of population growth went into the areas that were urban in 1990; 19 percent of growth went into the areas that became urbanized between 1990 and 2000; 11 percent of growth went into the areas urbanized after 2000; and 23 percent of growth went into the areas that are not considered urban. Figures 13

[6] http://www.census.gov/geo/www/ua/ua_2k.html

through 15 contain three pie charts which illustrate distribution of population growth between 1990 and 2007 among 4 types of lands: (i) urban in 1990, (ii) urbanized between 1990 and 2000, (iii) urbanized after 2000, and (iv) never urban. The first graph illustrates distribution of growth during 1990s, the second graph illustrates growth distribution between 2000 and 2007 and the bottom graph illustrates distribution during the entire 17-year period.

Figure 13: Growth allocation, 1990-2000

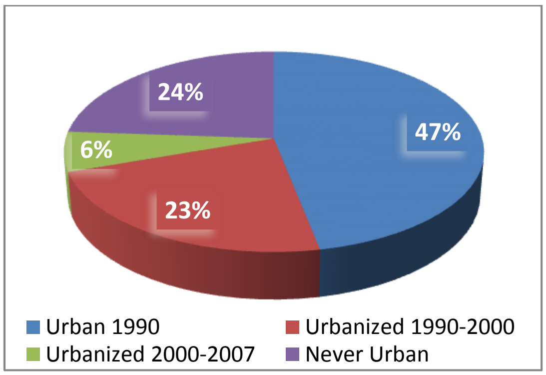


Figure 14: Growth Allocation, 2000-2007

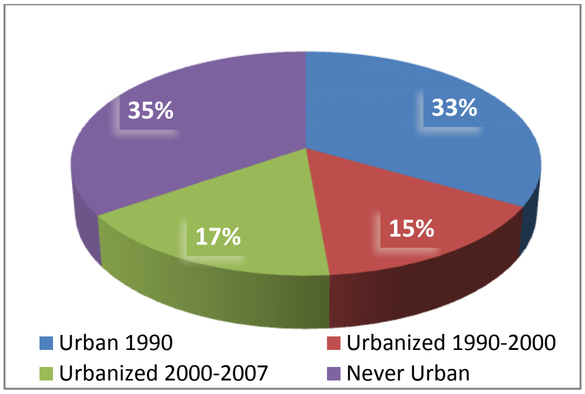
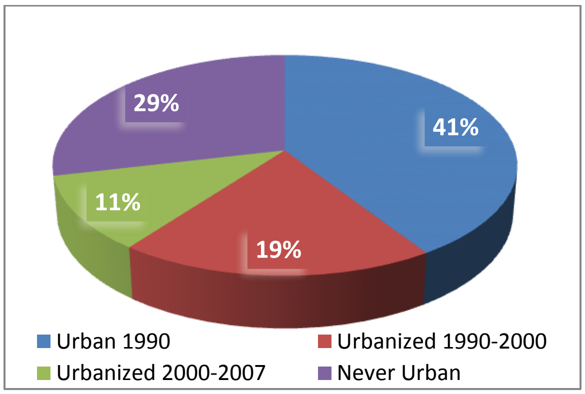


Figure 15: Growth allocation, 1990-2007



Almost a half of population growth during 1990s went into areas that were already urban in 1990, which means that their population density was 1000 people per square mile or greater. During the same period, a little less than a quarter of growth went into the areas that became urbanized between 1990 and 2000 (red pie) and also to the areas that never became urban (purple pie). About 6 percent went into the areas that became urbanized after 2000. In the years after 2000, about one-third of population growth was allocated to already urban areas and 15 percent went into those areas that were urbanized during 1990s. That means that almost half of post-2000 growth went to already urban areas. Still, a large share, over one-third of growth, went into never urbanized areas.

The figures illustrating the aggregated sample of metropolitan areas are not necessarily indicative of change that occurred in the areas that were already urban in 1990. For example, while in Los Angeles and New York, 88 percent and 81 percent respectively, of population growth between 1990 and 2007 went into areas that were urban in 1990, Cleveland and Pittsburg lost population in those areas with -163 percent[7] and -168 percent of total population change. In contrast, they gained a significant share of population in never urban areas, 163 percent and 65 percent correspondingly. In other words, Cleveland gained a little over 48,000 people between 1990 and 2007. However, by 1990 urban areas lost over 78,000 people, while never urbanized areas gained over 78,000 people. The areas urbanized between 1990 and 2000 gained around 22,000 people and areas urbanized after 2000 gained almost 25,000 people. In the aggregate, that constitutes change of 48,000 people observed in Cleveland. The results for all metropolitan areas are summarized in Appendix A.

Although changes vary among metropolitan areas, they can be loosely grouped into four categories. In the first group, the period between 1990 and 2000 can be generalized by often significant population loss in the areas defined as urban in 1990. These include Pittsburgh, Cleveland, St. Louis, Detroit, and Cincinnati. These areas lost between 8 percent of population in already urban areas in Cincinnati to 168 percent in Pittsburg. The loss of population in already urban areas has been matched with almost equal gain in the new urban areas. After 2000, St Louis gained some population in already urban areas, while Indianapolis lost 1 percent in the same areas. The rest of the group continued to lose population in the urban core.

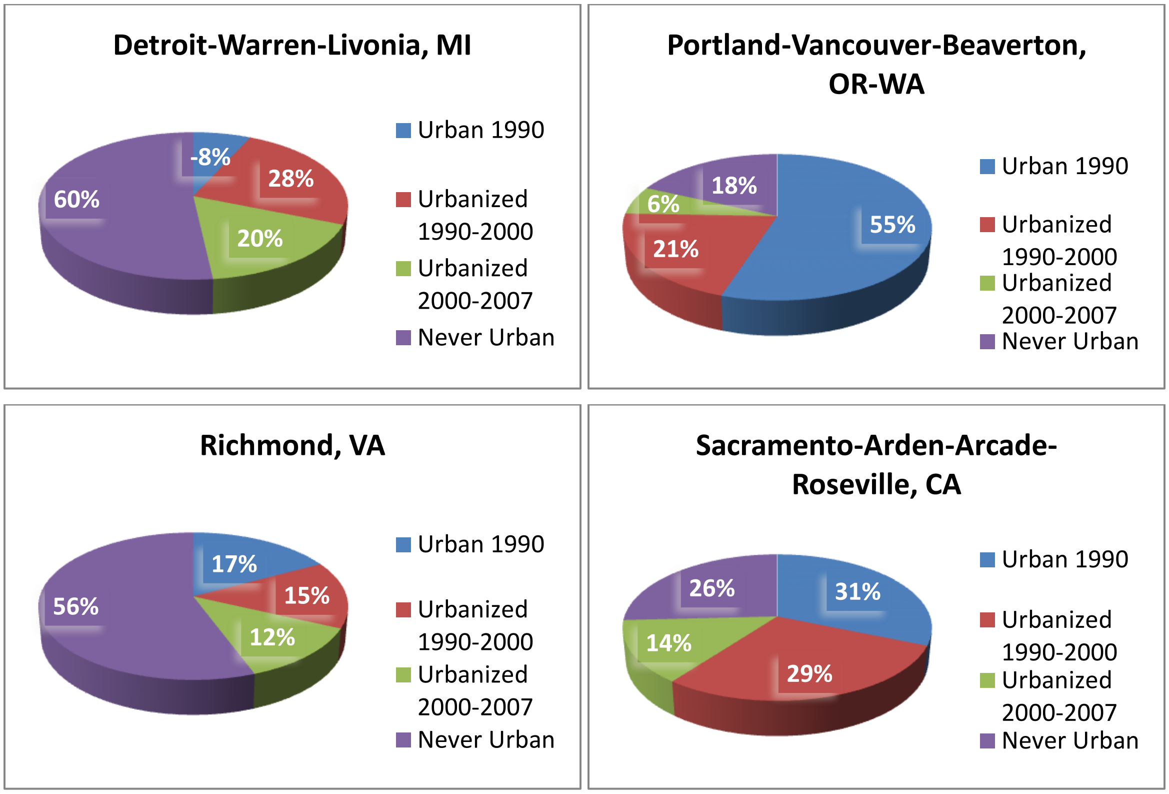
[7] For Cleveland, for example, 163 percent decrease represents 78,132 population loss in 1990 urban areas out of the total gain of 48,038 people between 1990 and 2007in the entire metro area (-78,132/48,038=-163%).

The second category gained almost all of the population growth in urban areas by 1990. Those were Los Angeles, Miami, New York, Portland, Seattle, and San Francisco, with over 55 percent of population growth moving into the existing urban areas. This second group of metros consistently gained over half and up to 88 percent of population growth in the existing urban areas. Denver also gained most of its population growth in 1990s in already urban areas; however that trend shifted to never urbanized lands after 2000.

The third category is characterized by the dominant share of population growth in the never urbanized areas. This category does not include older industrial areas where population growth in the never urbanized areas was the result of population loss in the urban core. Some of these are in the South or South West, though there are several Northeast metropolitan areas that followed the same trend. Falling in this category are: Atlanta, Austin, Baltimore, Boston-, Charlotte, Indianapolis, Jacksonville, Minneapolis, Orlando, Philadelphia, Richmond, San Antonio, and Virginia Beach.

In the fourth and final category are metropolitan areas that gained rather equal population growth across all categories. In this category is Las Vegas which converted the greatest percentage, 40 percent, of its land to urban between 1990 and 2000. Other metropolitan areas in this category are: Chicago, Dallas, Denver, Houston, Las Vegas, Phoenix, Riverside, Sacramento, San Diego, Tampa, and Washington, DC. Some of these are still dominated by largest growth in already urban areas, yet the share does not exceed 50 percent. Riverside and Houston, for example, gained 45 percent of their growth in already urban areas, but also about one quarter in never urbanized areas. Figure 16 illustrates with pie charts the four categories of growth allocation.

Figure 16: Growth Allocation for Select Metropolitan Areas



Discussion

Overview of the Findings

The excessive expansion of American metropolitan areas and the impact of spatial development patterns on open space, traffic, air and water, and social interactions, have led to increased attention to metropolitan spatial structure and its sustainability. Several descriptive studies examining spatial patterns in 1990s and particularly beyond 2000 have suggested that urban areas are growing denser and sprawl at a slower rate. Some theoretical work also argues that households' revealed preferences suggest an increased demand for urban amenities and proximity to urban centers. However, the studies that did examine changes in metropolitan spatial structure focused less on the period after the year 2000 and often measured spatial structure along one dimension and on a larger geography, such as counties.

The purpose of the present study was to address these questions by investigating three hypotheses: (1) metropolitan spatial structure in thirty five largest metropolitan areas in the United States changed between 1990, 2000, and 2007; (2) change is not consistent with decentralization trends observed prior to 1990; (3) change is not consistent across all the metropolitan areas. To investigate the hypotheses, this study has measured metropolitan spatial structure and its change along five dimensions: (i) density gradients, (ii) concentration indices, (iii) clustering indices, (iv) density frequency distributions, and (v) growth allocation, for 1990, 2000, and 2007. Each of these measures capture different dimension of spatial structure. This is further evident as each of the measures group slightly different cluster of the metropolitan areas in the identified categories. For example, while density gradient may suggest centralization of Las Vegas metropolitan area, the clustering index may indicate decrease of clustering. Nevertheless, several clear trends have emerged which confirmed the original hypotheses.

First, the 35 metropolitan areas studied in this research can be generally grouped in two groups: smart growth winners and losers. The winners are the metropolitan areas where population density increased in the central core, most of the population growth went into already urban areas, and the number of medium to high density areas increased while the frequency of low density areas decreased. The losers are those metropolitan areas on the other side of the spectrum. They primarily lost population in the central core and gained most of the population in never urbanized areas and at low density. Naturally, some metropolitan areas do not follow the clear distinction between the two groups with some mixed results along one of the dimensions. The "winners" and "losers" categories also follow some regional groupings. The areas exhibiting sprawling characteristics are largely old industrial cities in the rust-belt and the cities in the South. However both regions are not sprawling due to the same pressures. The old industrial areas, which have also seen only trivial increases in population growth, have been continuously sprawling, i.e. depopulation of central core and urban flight, and loss of high density areas. These are consistent with the trends observed prior to 1990s. Sprawl of the metropolitan areas in the South arises out of different circumstances. Those areas underwent strong population growth pressures during the observation period and the sprawling trends are not evident across all measures. While showing an increasing number of low density neighborhoods, these metropolitan areas have also densified in the central core and gained a notable portion of the population growth in their already urban areas. The metropolitan areas pegged as "winners" and characterized by densification characteristics also loosely follow regional boundaries. These are the metropolitan areas of South West and West, and several in Florida. Notably, these are the areas such as Las Vegas, Sacramento, Seattle, Portland, and Tampa, Miami, and Orlando. They have gained most significant population growth during the observation period. Consequently, not all of their population growth could have gone to already existing areas and in the central core, but a remarkable share did. For example, Las Vegas, which increased in population by 85 percent between 1990 and 2000, allocated 37 percent of that growth to already urban areas and 52 percent to areas that are regarded as urbanized in 1990s. Thus those new urbanized areas were reaching population densities in excess of 1000 people per square mile. By contrast, the “losers” were allocating same shares of new growth to never urban areas. In other words, their densities never reached 1000 people per square mile.

Second, the majority of the metropolitan areas reveal trends towards infill development. This trend is most clearly evident from the aggregated growth allocation measure which shows that almost half of population growth in 1990s went into already urban areas while the same areas gained a third of 2000 to 2007 growth. The density frequency measure similarly shows that the frequency of the lowest density tracts, with 500 or less people per square kilometer, fell most drastically, particularly in 1990s. And likewise, density gradients showed that in the aggregated sample population density in central core increased, however, by more during 1990s than after 2000, 12 percent versus 4 percent, respectively.

This leads to the third trend which brings attention to differences between the 1990s and the 2000s. The decade of 1990s emerges as one with a dominating focus on repopulation of the central core and generally increasing density of existing urban areas. It is in that sense not consistent with trends observed prior to 1990s which primarily embodied decentralization and depopulation of central core. In 1990s, the majority of metropolitan areas actually saw increasing population density in the central core consistent with the urban resurgence theory presented by Glaeser and Gottlieb (2006). The theory of urban resurgence refers to increased demand for urban amenities in explaining the renewed importance of American downtowns. After 2000, the focus of population growth shifts somewhat back to areas that have not been urbanized yet. However it is not clear that there was a reversal in demand for urban amenities, but rather that the unprecedented appreciation in housing prices led to housing stock overbuilding on the urban edge. While it is not in the scope of this study to measure determinants of urban form change, the results presented here suggest such a study may offer some interesting insight. Again, the post-ize2000 results show reorientation of focus from already existing areas in 1990s to never urbanized areas in the 2000s. Many metropolitan areas that have seen gains of 50 percent or more of population growth in already urban areas saw those shares drop by sometimes half of the percentage points.

And finally, while the objective of this study is not to evaluate the effectiveness of smart growth programs, it is important to examine if the increase in population density of the existing urban areas is more evident in states that do employ smart growth programs than those that do not. However, since this study covers only the 35 largest metropolitan areas, there are states with growth management programs whose metropolitan areas are not covered in this study. Nevertheless, there is a representative sample of areas from Florida, Oregon, Washington, Georgia, California, Colorado and Maryland where the programs are either implemented statewide or by localities. The estimates do show that the metropolitan areas in well-known smart growth states, such as Oregon and Washington, show increasing population density in the areas where such growth is desirable, i.e. in already existing urban areas. Portland, in particular, performs well along most of the dimensions and it is also the only area which grew more concentrated during the observation period. Consistent with finding by Howell-Moroney (2007), the states with more stringent smart growth programs, Florida, Oregon, and Washington, also saw larger increases in density than Georgia and Maryland, for example, where it is hard to say that the metropolitan areas grew smarter. But also, it is critical not to underestimate the restriction geographical barriers place on expansion of urban areas. So, naturally, the areas such as Miami which has natural barriers on three sides of the metropolitan region will grow inward. The same is true for the areas in Arizona or California, where either water capacity or federally-owned lands pose as barriers to excessive urban expansion. Atlanta, on the other hand, does not have any geographical barriers to growth and the land is plentiful and relatively less expensive.

Limitations of the Study

Focusing on the five measures of metropolitan spatial structure has the advantage of providing a more detailed analysis of urban form change. However, though these indices measure change in population density which is the most frequently used method for evaluation of smart growth programs, these measures, except density gradients, do not account for spatial location of population change. While it is possible to determine if an urban area grew denser, it is difficult to say where densification occurred. The same applies to the measures of concentration and clustering.

Another limitation of this study concerns the use of census tracts boundaries since Gini coefficient measure, in particular, is sensitive to the size of the geographic unit. And while census tracts usually encompass between 2,500 and 8,000 persons, they can vary significantly in size, particularly between the metropolitan areas in the Northeast and the South West. Preferably, an analysis would employ a population grid with consistent grid size across the country. Even with that method, one would have to decide on a consistent yet meaningful grid cell size and distribute population accordingly. However, given that census population count is recorded in areal units, the smallest one being a census block, which still varies in size across the country, some areas would have more accurate estimates than the others.

This study also only examines changes in population densities alone. When talking about changes in metropolitan spatial structure, research often measures changes in employment patterns as well. The location of employment is critical for spatial structure and particularly for measures such as clustering and centrality. The focus herein was on population patterns, but research needs to be conducted to see how employment density has changed within and among metropolitan areas.

Essay Two: A Spatial Hazard Analysis of Urban Form Changes in America

Because of the vast geographic scope of the United States, metropolitan spatial structure is innately different across the country. American metropolitan areas not only vary in geophysical features, but have been shaped by different histories, cultures, markets, regulations, geographic constraints, natural resources, and unique events. Today, as in the past, they host various economic activities and face different obstacles which further place demands on their urban spatial structure (Perloff et al, 1960). Separately, labor markets are influenced by economic shocks, locally and in world demand. The shocks are passed from the labor market to household income, land rents and ultimately to urban spatial structure. However, given that events, such as economic shocks, are based on some probabilities rather than certainty of occurring, the resultant outcome on urban spatial structure follows those processes in a stochastic fashion (Harris, 1968; Capozza and Helsley, 1990). Additionally, urban form is durable, and when new development takes place, it does not replace the existing built environment but attempts to complement it in also often in a stochastic manner (Brueckner, 2000b). As a result, when attempting to measure urban form, some measures may perform well in an analysis of a single metropolitan area, but be less reliable in a comparative analysis of multiple metropolitan areas (Malpezzi and Guo, 2001).

Researchers have struggled through the years to find ways of characterizing urban form in a way that enables a consistent, objective analysis of similarities and dissimilarities across regions and particularly over time. The classic economic theory of urbanization (Alonso 1964; Muth 1969; Mills 1972), and the ensuing density gradient model, continue to be dominant tools in explaining the general tendencies of urban form, even across the vast area as the United States is (Glaeser and Kahn 2004; Bogart 2006). The draw of this model is that it takes the complex urban spatial structure and reduces it into a few simple relationships which explain general metropolitan spatial structure. However, the theory is limiting in that it is highly deterministic while the urban spatial structure is not. Urbanization does follow a general trend of decreasing density away from the urban center; however that process is not monotonic and often occurs in a seemingly disordered way (Carruthers et al, 2010). Like the theory, the weakness of the density gradient model is that it is deterministic and can mischaracterize the inherent complexity of urbanization (Brueckner 1982, 1987, Kau and Lee 1976a, 1976b, 1977; Johnson and Kau 1980; Kau et al 1983).

The research puzzling with the complexity of urbanization attempted to offer a number of alternative and complementary measures. It argued that development of the urban spatial structure is a chaotic process which can be only defined as a complex structure. To quantify such complexity, one needs spatial patterns which show the irregularity of their configuration. One such approach uses fractal geometry. Fractal analysis looks at the spatial complexity by treating it as dynamic, nonlinear, disperse, open structure that produces unstructured, elaborate geometry in space which resembles urban sprawl (Batty and Longley, 1994; Batty and Xie, 1996). Using fractal analysis, Torrens (2006, 2008) evaluated urban sprawl as a kind of space filling process and assigned a fractal dimension to each urban area. Fractal dimension measured the extent to which a city fills its two-dimensional area. The critique of fractal analysis suggested that it would be difficult to use the measure to compare metropolitan areas given that virtually the same fractal value may result from metropolitan areas with different population sizes and densities (Shen, 2002).

Further exploring the complexity of characterizing urban form, Carruthers et al. (2010) proposed survival analysis methods. Unlike the density gradients which assume urban form to unfold monotonically, the survival analyses are probabilistic by design. The probabilistic nature of survival methods allows for stochastic processes observed in the development of urban spatial structure. The survival analysis models also referred to as longitudinal or duration models, are popular in engineering, economics, and other disciplines, and are generally used to characterize occurrence and timing of events. For example, they were first used by engineers concerned with failure of products, and then applied in biomedical research to time passing away following the beginning of a disease. In social sciences, the method has been used to understand the temporal dimensions of questions such as the length of unemployment spells and residence tenure (e.g., James, 1989; Narendranathan and Stewart, 1993; Clark, 1992; Odland, 1997; Davies Withers, 1997; Glavac and Waldorf, 1998). In urban form research, recent applications by Irwin and Bockstael (2007) and An and Brown (2008) used survival analysis to study timing of land use change. These studies were, however, concerned with parcel level land use changes which are a different scale of the urban form phenomena than the study presented here attempts to examine.

The survival analysis deals with measuring the duration of some state or the length of time prior to a terminating event. The method generates the conditional probability of an event happening at a particular time t, given that the event termination has not happened up to that time. To apply survival analysis to spatial setting, one can think of distance, similarly to duration, as a nonnegative random variable. This property of distance has allowed research to apply a mathematical framework of survival analysis to spatial setting and use distance as the endogenous random variable. While Odland and Ellis (1992) and Esparza and Krmenec (1996) were first to apply survival analysis to a spatial setting, Waldorf (2003) more recently presented a framework for using the method in urban form studies. Waldorf argued that it allows for better understanding of spatial processes in contrast to methods which only examine spatial patterns and linkages, as is the case with spatial point patterns pioneered by Diggle (1983) and Boots and Getis (1988). The limitation of the spatial point pattern analysis is that it is not a true behavioral approach to hypothesis testing (Odland and Ellis 1992). It examines the degree of spatial patterns or randomness among a sample of points, for instance urban settlements or housing, by testing the observed pattern against theoretical or hypothesized patterns. The method is appropriate for evaluating economic theories of location and it has been used in urban form analysis for many years. As early as 1960s, Getis used point pattern analysis to examine commercial and residential land use succession in Lansing, Michigan (Getis, 1964), and, afterward, to identify population clusters in Chicago, Illinois (Getis, 1983). But, spatial point pattern analysis is not able to account for behavioral variables that are critical in developing urban spatial structure. It can describe the degree of compactness versus sprawl, but it cannot explain how a particular pattern evolved, or to identify how to alter its course (Carruthers et al, 2010).

Odland and Ellis (1992) first applied survival analysis to measure spacing of urban settlements in Nebraska. The study found that the pattern of settlements in Nebraska shows heterogeneity and interdependence. While the spacing of settlements increases from east to west within Nebraska, there is less variation in the distances between nearest neighbors within different regions. The study used distance between points as a mathematical correspondent of duration labeling survival models accordingly as *spatial* survival models or *spatial* hazard models. The term hazard arises out of the hazard rate which is an important concept in survival analysis. The hazard rate is the probability of an event occurring at time *t* given that there is a risk of the event occurring. For example, if the hazard rate is constant over time and it is equal to 1.2, this would imply that 1.2 events would be expected to occur in a time interval that is one unit long. More precisely, if a person has a hazard rate of 1.2 at time *t* and a second person has a hazard rate of 2.4 at time *t*, then the second person's risk of an event would be two times greater at time *t*. Thus, the hazard rate is an unobserved variable which controls both the occurrence and the timing of events. It is the primary dependent variable in survival analysis. In the context of spatial analysis, the hazard rate describes spatial hazard instead of temporal hazard. While temporal hazard measures the timeframe coming to an end, spatial hazard measures distance coming to an end. For example, in the Odland and Ellis (1992) study, the authors examined distance intervals separating neighboring settlements and how those varied across the state. While the focus in that study was location of settlements in east-west and south-north directions, the method could be applied, for example, to examine spacing of neighborhoods as a function of their location and distance to the urban core of a metropolitan area.

Hazard models are also designed to estimate the conditional probability of timeframe ending, or distance ending in a spatial analysis case. The traditional regression models, in contrast, focus on estimating the unconditional probability density functions.

And even though the unconditional probability density function and hazard function are mathematically equivalent, there are two advantages with estimating conditional hazard functions. First, the observations for which the exact duration is not known do not have to be discarded. And second, changes in exogenous variables during the observation period can be accounted for (Waldorf, 2003).

This study builds on the literature that uses survival analysis to characterize urban form and extends the analysis to examine longitudinal changes in the U.S. metropolitan spatial structure between 1990 and 2007. In using survival analysis to measure urban form, Carruthers et al (2010) estimated spatial hazard models for the 25 largest core-based-statistical-areas (CBSAs) of the United States and showed that hazard functions are particularly effective in describing the stochastic nature of urban form. The study relied on 2006 housing unit count at census block level to illustrate how urban development patterns unfold across a metropolitan area. In a follow up study, the same group of authors examined the ability of spatial hazard models to detect changes in urban form between 1990 and 2006 (Carruthers et al, forthcoming). In that study, the change in urban form was measured via changes in population. The 2006 population in a census block group was estimated by multiplying the 2000 average household size and the 2006 housing unit count. This method effectively assumed that all 2006 housing units were occupied. This study again established that spatial hazard models behave as expected in the context of urban economics and are an effective tool for analyzing spatial structure change.

While the two studies mentioned focus on establishing spatial hazards and their application in the study of urban form, the analysis presented here extends beyond them by applying the method to measure change between 1990 and 2007 in the same 35 metropolitan areas as in the first essay. Given that the data for 2007 population count comes from the ESRI 2007 Demographic Update, it more accurately measures population change and is therefore consistent with the analysis performed in the first essay. The population estimate used by Carruthers et al. (forthcoming) assumed that all 2006 housing units were occupied. This is a particularly important issue in metropolitan areas that have in fact lost population in some central core areas, such as Detroit. By assuming full occupancy in 2006, the analysis may show that those areas have repopulated since 2000 when that may not be the case. Another divergence from the previous work is the location of the center of a metropolitan area. Consistent with the first essay, the center of a metropolitan area in this essay is where the Central Business District (CBD) was located according to the 1982 Census of Retail Trade.

The two specific questions addressed by this study include: (1) Do spatial hazard models suggest changes in metropolitan spatial structure? (2) Are changes consistent within a metropolitan area? (3) Are changes consistent with traditional measures of metropolitan spatial structure as observed in the first essay? To answer these questions, the study estimates a series of spatial hazard models characterizing urban form in the 35 largest CBSAs of the United States areas in 1990, 2000, and 2007.

Empirical Strategy

Hazard models are a group of longitudinal or survival analysis models used to characterize the occurrence and timing of events, and more specifically, modeling time to event data. However, in contrast to the hazard function framework observing duration as a nonnegative random variable, the spatial hazard function that is used instead observes distance between two spatial observation points as a nonnegative random variable. As such, the spatial hazard function, *h(d),* noted as following:

$$h(d) = \lim_{\Delta d \to 0} \frac{\Pr(D \in [d, d + \Delta d] \mid D \geq d)}{\Delta d} \in (0, \infty) \tag{1}$$

describes the conditional probability of a random distance variable, *D,* terminating at $d+\Delta d$ given that it lasted up to *d*. For $\partial h(d)/\partial d > 0$, the hazard increases, or accelerates, also indicating that the probability of terminating increases with distance. In other words, given that this study examines metropolitan spatial structure, the two hypotheses essential to this model which arise directly from the urban economic theory (Alonso 1964; Muth 1969; Mills 1972) are: (1) the conditional probability of distance between two nearest neighbor points terminating increases with distance between them; and (2) the probability of terminating decelerates with distance from the urban core.

Survival analysis uses nonparametric, parametric, and semiparametric estimations. In social sciences, estimating the effect of independent variables on the hazard is most often done via a semiparametric model such as the proportional hazard model. In the proportional hazard model, the effect of covariates has a multiplicative effect on the hazard rate. The proportional hazard model is the most frequently used survival model because it does not have to be based on any assumptions about the nature or shape of the underlying survival distribution. The proportional hazards function is described as following:

$$h(d \mid X) = h_0(d) * \exp(X * \Phi), \tag{2}$$

where the hazard function consists of two components: (1) a baseline hazard, $h_0(d) = \lambda d^{\lambda-1}$, where λ is a shape parameter giving the instantaneous rate at which the distances between points terminate when $X = 0$ ($X = X_{ik}$ is the vector of covariates); and (2) a function $f(X, \Phi)$ which is independent of the distance and is specified as an exponential function of exogenous variables, X, and an exponential scale parameter, Φ. Exogenous variables, X, have proportional and distance-independent effects on the conditional probability of terminating the distance, while scale parameter, Φ, accelerates or decelerates the baseline hazard. The shape parameter, λ, determines the asymptotic nature of the hazard function. For shape parameters < 1, the hazard is monotonically declining; and for shape parameters > 1, the hazard is monotonically increasing.

The baseline hazard may remain unspecified and then estimated via a partial log-likelihood function. Otherwise, the baseline hazard may assume a particular distribution and be estimated via maximum likelihood procedures. While there are several well-known and well-behaved distributions used in parametric estimations, the Weibull distribution is used most often because it allows for a flexible shape of the hazard function. When $\Phi = 1$, the Weibull distribution becomes the exponential distribution, and its hazard is constant. In other words, the probability of distance terminating is the same irrespective of the distance between the points. For $\Phi > 1$, the probability of distance terminating monotonically increases, and for $\Phi < 1$, the probability of distance terminating monotonically decreases with increasing distance.

In this analysis, the primary dependent variable is the distance between population mean centers of nearest neighbor census tracts. Population mean center of a census tract is defined as the population-weighted average Cartesian $\{x, y\}$ coordinate of all the block group centroids in a given tract. The mean center is calculated using the mean center tool in the ArcGIS Spatial Statistics Toolbox. For each of the relevant years – 1990, 2000, and 2007, block group population count is used as the "weight" field and the tract identification number is used as the "case" field. This allows the block groups to be grouped into their respective tract and evaluated accordingly. The mean center tool thus produces a point within each tract that can be thought of as the population "center of gravity". The center of gravity is generated for 1990, 2000, and 2007 for each tract. And, distance between the nearest neighbor tract mean centers is then calculated for each of the years.

To address the urban form change, the base function (2) is extended by adding two temporal fixed effects for 2000 and 2007 to the vector X_{ik}. The fixed effects variables measure the change in the conditional probability of distance between nearest neighbors terminating in 2000 and 2007 in contrast to the 1990. If the metropolitan spatial structure has grown smarter, i.e. more densely populated, distance between nearest neighbors will decrease between 1990 and 2007 and conditional probability of distance between them terminating will increase. The econometric specification is as follows:

$$h(d_{ij} \mid X_{ik}) = h_0(d_{ij}) \cdot \exp(\phi_{d_{ic}} \cdot x_{d_{ic}} + \phi_{d_{il}} \cdot x_{d_{il}} + \phi_{d_{iD00}} \cdot x_{d_{iD00}} + \phi_{d_{iD07}} \cdot x_{d_{iD07}} + X_{ik} \cdot \Phi_k) \quad (3)$$

where $h(d_{ij} \mid X_{ik})$ indicates that the baseline hazard for distance between nearest neighbors *i* and *j*, $h_0(d_{ij})$, is scaled by X_{ik}, a vector of *k* independent variables, including $x_{d_{ic}}$, the distance from *i* to the regional center; $x_{d_{il}}$, the distance from *i* to the local center; and $x_{d_{iD00}}$ and $x_{d_{iD06}}$, the 2000 and 2007 temporal fixed effects. Regional and local centers are defined in the *Data* section. Scale parameter, Φ_k, including $\phi_{d_{ic}}$, $\phi_{d_{il}}$, $\phi_{d_{iD00}}$, and

$\phi_{d_{iD07}}$, measures influence of independent variables on the conditional probability of distance between nearest neighbors terminating. As described, the proportional hazard model implies that explanatory scale parameters multiply hazard via the hazard ratios estimated by the model. A positive parameter value of temporal fixed effects ($x_{d_{iD00}}$ and $x_{d_{iD07}}$) indicates accelerating baseline hazard and increasing probability of distance between two nearest neighbors terminating in respective years. It, in essence, indicates the nearest neighbor tracts are closer in 2000 or 2007 than they were in 1990. As the objective of this research is to determine urban form change over the 17-year timeframe, the decomposition of the proportional hazard model could be viewed as: (1) a common baseline hazard, or the fundamental part, is specific to the beginning of the observation period, namely 1990; (2) the effect of urban form change is measured with estimated scale parameters for 2000 and 2007; and (3) the effect of economic impacts is captured with additional explanatory variables and isolate (??)impact of those from the temporal fixed effects. The common baseline hazard allows for effective comparison between years as the proportionality among groups is required and is guaranteed with the common shape parameter.

<u>Data</u>

This essay focuses on the same 35 largest core based statistical areas (CBSAs) of the United States as those observed in the first essay and similarly uses 2000 census tracts as the units of analysis. The data comes from four sources: (*i*) 2007 population estimate from the ESRI 2007 Demographic Update Methodology; (*ii*) a nationwide count of housing units at the census block level in 2006[8], (*iii*) Census Summary File 3 (SF-3), from the 2000 census of the population; and (*iv*) two Geolytics, Inc. products which allocate selected 1990 SF-1 and SF-3 variables from 1990 census boundaries to 2000 census boundaries. Measuring the change over time is challenging because it requires using constant geographic units and many Census defined boundaries were modified between 1990 and 2000. As a result, Geolytics files enabled a more accurate measure of change between the two years, because the census block groups and census tract polygons are constant. The estimate of the 2007 population count is available from the ESRI's 2007 Demographic Update and the method for estimating population is described in the *Data* section of the first essay.

The three variables necessary for estimating the model are $d_{ij}, x_{d_{ic}}$, and $x_{d_{il}}$, the distance from *i* to its nearest neighbor and the distance from *i* to the regional and local center, respectively. There were six steps needed to create these distance measures. The first step consisted of generating a mean center for each census tract in the 35 CBSAs. To do that, the population count for each census blocks group was used to produce a population weighted center of all census tracts and for the three observation periods - 1990, 2000, and 2007. As noted, the mean center was calculated using the mean center tool in the ArcGIS Spatial Statistics Toolbox. The population count for each respective year was used as the "weight" field and the tract identification number was used as the "case" field, which organizes block groups into the correct tract and evaluates that tract accordingly.

[8] Provided to the Department of Housing and Urban Development by the Census Bureau. The count represents the universe for the American Community Survey, an annual survey of about three million households that is set to replace the so-called "long form" of the decennial census, which will eventually yield census tract level data on an annual basis.

The second step involved generating 35 CBSA centers. CBSA centers are defined as the centroids of the census tracts where the Central Business District (CBD) was located according to the 1982 Census of Retail Trade. The CBSA centers can be thought of as the core center, or central business district (CBD), of a metropolitan area. Similarly, the third step involved generating the housing weighted centers of the local centers defined by the boundaries of the county subdivisions, or county divisions in cases in which there are no subdivisions. The housing unit count came from 2006 nationwide count of housing units at the census block level. Inclusion of local mean centers in the analysis allows for accounting not only of local spatial homogeneity, but also for policentricity of many urban regions in the United States.

In the fourth step, each tract was assigned to its nearest neighbor and distance measure between them was calculated. To do this, GeoDa was used because the ArcGIS Toolbox does not have a routine that will identify a feature's nearest neighbor that is within the same shapefile and calculate a distance to that feature. Distance between nearest tracts was repeated for each year.

The fifth step included assigning tracts to their nearest CBSA's CBD center point and to the nearest local mean center point and obtaining two additional sets of distance measures: (i) between tract mean centers and their respective CBSA's CBD, and (ii) between tract centers and their respective local mean centers. Distance measure calculations were repeated for each of the three years resulting in three sets of nearest neighbors' distances, for 1990, 2000, and 2007.

The last, sixth step consisted of mapping rays connecting each tract to its CBSA center and nearest neighbor tract using an ESRI user-written extension, Desire Line. This tool creates a line between a point of origin and a point of destination. The results of this step are shown in the Figure 17. The map shows CBSAs CBDs and their spheres of influence for the 35 CBSAs that are the focus of the analysis shown in dark gray. Figure 18 shows a map of spatial point patterns in San Francisco, Atlanta, San Antonio, and Miami metropolitan regions. In the second exhibit, both the rays connecting tracts to their CBD center and the rays connecting nearest neighbor tracts are visible.

Figure 17: Desire Lines

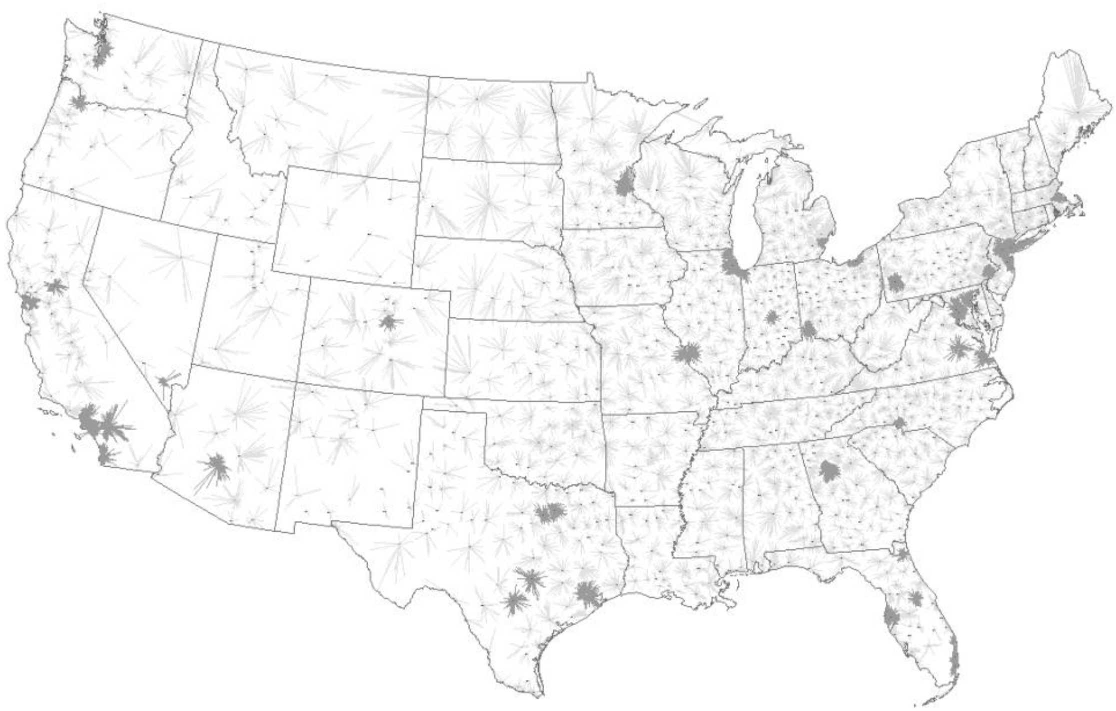
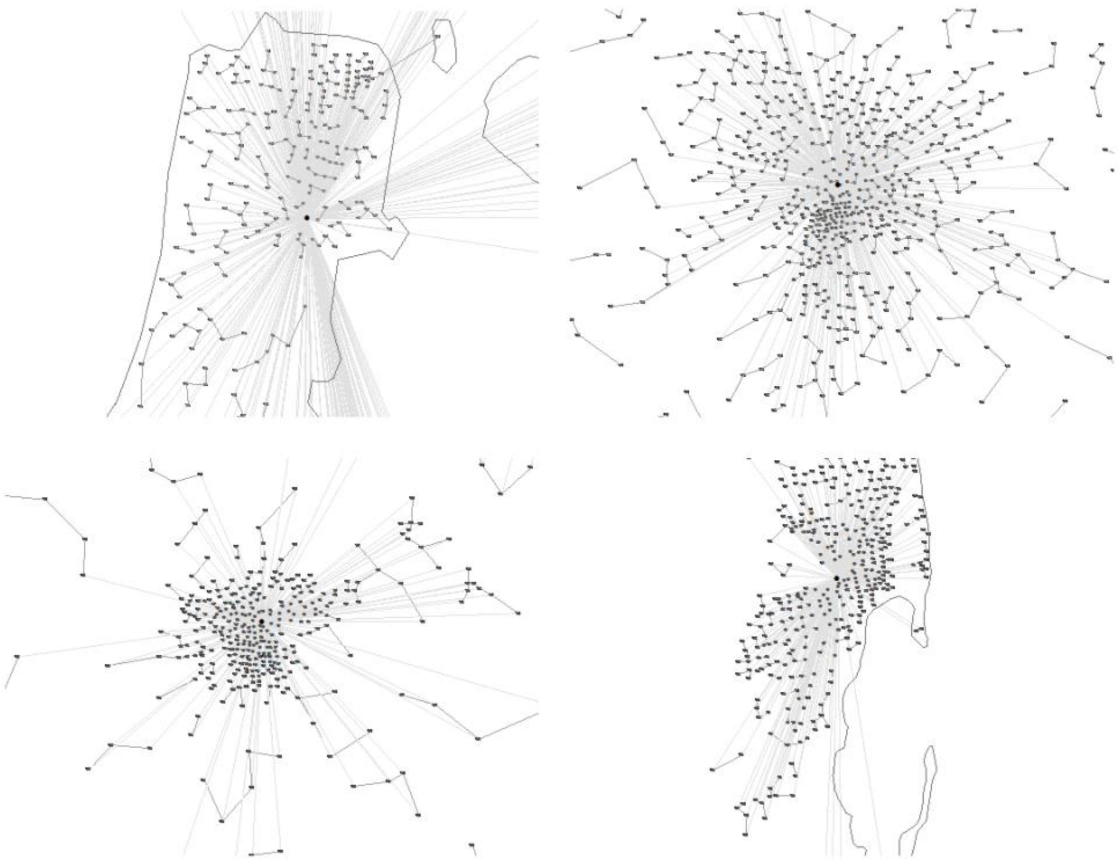


Figure 18: Desire Lines for San Francisco, Atlanta, San Antonio, and Miami



The remaining vector of independent variables, X_{ik}, is composed of explanatory variables directly obtained from the theoretical models used to explain urban form . More specifically, household income and commuting costs are the determining factors of households' location decisions. Given that land is a normal good, households are anticipated to consume larger size lots with greater income. Consequently, income is expected to decelerate the hazard function. Commuting costs, which indicate regions' accessibility, are constrained by the household budget function. In this analysis, commuting costs are measured by the share of workers with travel time to work less than 25 minutes. The parameter estimates are thus expected to decelerate the hazard function. Finally, as discussed above, urban area evolves over time and taking into account the vintage effects, or aged housing, is critical. Though aged development could either accelerate or decelerate the hazard depending on the density at which it was developed in comparison to the current market conditions. Since the control variable used in this model accounts for the percentage of housing units build before 1939, the estimated parameter is expected to be positive given that the density of development prior to 1939 was still uninfluenced by the automobile and was built at higher densities than is the case of the modern market conditions. Beyond these explanatory variables, population count is included in order to control for the size of the census tracts and this variable is expected to decelerate the hazard of the distance between points terminating. Figure 19 provides specific definitions and the source of data of each variable. Figure 1 in Appendix B provides descriptive statistics for distance measures.

Figure 19: Variables Definition and Sources of Data

	Variables	**Source**	**Definition**
Dependent	Distance from Nearest Neighbor	Authors' calculations, U.S. Census and *Geolytics*	Distance from population weighted center to the population weighted center of the nearest tract, 1990, 2000, 2007
Independent	Distance from CBD	Authors' calculations, U.S. Census and *Geolytics*	Distance from population weighted center to the CBD of the nearest CBSA, 1990, 2000, 2007
	Distance from Local Center	Authors' calculations, U.S. Census and *Geolytics*	Distance from population weighted center to the population weighted center of the nearest county subdivision, 1990, 2000, 2007
	Household Income	U.S. Census Bureau, and *Geolytics* — SF-3, Table P68	Median household income, in 1990 and 2000
	Travel Cost	Author's calculations, from U.S. Census Bureau and *Geolytics* — SF-3, Tables P31 and P33	Percent of workers 16+ years old with travel time-work less than 25 minutes, in 1990 and 2000
	Age of Housing Units	U.S. Census Bureau and *Geolytics* — SF-3, Table H35	Percent of homes built before 1939, in 1990 and 2000
	Population	U.S. Census Bureau and *Geolytics*	Estimated population, 1990, 2000, 2007

Estimation Results

The spatial hazard estimation results for 35 CBSAs are listed in alphabetical order in Figure 20. All the parameter estimates are positive because they are hazard ratios that scale the baseline hazard. So, the parameter value of less than 1 decelerates the baseline hazard, while the values greater than 1 accelerate it.

The first three columns following CBSA names refer to the CBSAs' shape parameter, λ, its significance and the z-value. The shape parameter for every region is positive and statistically significant at 99 percent confidence level. This confirms the idea that urbanization patterns described by nearest neighbor tracts exhibit positive spatial dependence, or in other words, the probability of the distance between nearest neighbors ending increases with increasing distance between them.

The following two sets of columns are parameter estimates, significance and z-values for the two temporal fixed effects, 2000 and 2007. For both temporal effects, most of the estimates are statistically significant and greater than 1, suggesting that most of the metropolitan areas grew more compact during this 17-year period. The section below describes temporal effects estimates in greater detail.

The remaining parameter estimates under the Φ heading all show expected results. First, the parameter estimates on distance from the CBD are less than 1 and highly significant in all regions, suggesting that the probability of the distance between nearest neighbors terminating decreases with the distance from their CBSA center.

Second, the parameter on distance from the nearest local center is largely less than 1 and statistically significant, suggesting as well that the probability of the distance between nearest neighbors terminating decreases with distance from their nearest local center. The significance of the local parameters confirms the theoretical assumption of the growing importance of local business centers in structuring metropolitan urban form. There are several exceptions where the estimated effect is the opposite indicating that the probability of the distance terminating increases with distance from the nearest local center. This is the case in the very large metropolitan areas with dense spatial structure, such as Boston, Chicago, Minneapolis, and Philadelphia. Given the polycentric nature of such large metropolitan areas, it may be that numerous population and employment centers are serving as gravity points for population concentration. In cases with the insignificant parameters, the baseline hazard is spatially invariant and the areas are mostly characterized by strong monocentric urban structure where a dominating central business district plays the most important role. The direction of the influence of distance to the nearest local center again largely depends on the complexity of an urban region.

Third, as expected, the parameter on household income is largely less than 1 and statistically significant indicating that, all else being equal, income decelerates the spatial hazard function. That is, probability of distance between nearest neighbors terminating decreases with higher income.

Fourth, the parameter on travel time has a somewhat mixed effect. Again, this variable is defined as the share of workers with commute time of 25 minutes or less. While the parameter is highly significant in all of the metropolitan areas, the areas with a smaller parameter are more spatially spread out or at least pull workers from a larger geographic radius than areas with a larger parameter. New York is the only area with a parameter estimate less than 1 which is indicative of the size of its commuter shed.

Fifth, the parameter on the age of housing units, or the share of housing built before 1939, also varies across regions. In areas where the parameter is statistically significant, the impact is positive two-thirds of the time, suggesting that older development is generally denser than newer development. The parameter values of less than 1 are recorded in areas with relatively newer housing stock, such as Las Vegas and Phoenix, suggesting that the greater share of newer housing stock decelerates the hazard function – or, that newer homes are spaced further apart. And finally, the parameter on population, a control for the size of the census tracts, is also mostly statistically significant and negative. The last two columns provide information on the sample size for each metropolitan area (n) and log-likelihood of the estimation function (LL). The sample size is $n \times t$ observations in the panel, where n is the number of census tracts and t indicates the year, 1990, 2000, or 2007.

Figure 20: Estimated Spatial Hazard Functions — Distance from Nearest Neighbor

				τ						ϕ																			
	l									Distance from			Distance from			Household			Travel			Age of							
				2000			2007			CBSA Center			Local Center			Income			Cost			Housing Units			Population				
CBSA	Est.		z	Est.		z	Est.		z	Est.		z	Est.		z	Est.		z	Est.		z	Est.		z	Est.		z	n	LL
Atlanta-Sandy Springs-Marietta, GA	2.94	***	(68.79)	1.27	***	(4.01)	1.26	***	(3.71)	0.99992	***	(-46.61)	0.99994	***	(-4.42)	1.00000	***	(-3.69)	6.90	***	(9.07)	1.98	***	(2.76)	0.99995	***	(-6.79)	2,070	-883.38
Austin-Round Rock, TX	2.25	***	(30.31)	2.14	***	(7.64)	2.20	***	(7.68)	0.99993	***	(-18.75)	0.99988	***	(-6.13)	0.99998	***	(-8.74)	46.18	***	(13.75)	0.05	***	(-5.06)	0.99988	***	(-7.20)	771	-553.28
Baltimore-Towson, MD	2.65	***	(59.46)	1.50	***	(6.92)	1.32	***	(4.74)	0.99990	***	(-39.73)	0.99989	***	(-6.17)	0.99998	***	(-17.28)	1.88	***	(3.73)	2.95	***	(7.99)	1.00000		(0.28)	1,926	-1,033.76
Boston-Cambridge-Quincy, MA-NH	2.38	***	(68.21)	1.86	***	(12.40)	1.72	***	(10.90)	0.99996	***	(-34.52)	1.00008	***	(4.78)	0.99998	***	(-22.30)	13.26	***	(12.91)	22.76	***	(27.30)	0.99994	***	(-5.69)	2,760	-1,631.33
Charlotte-Gastonia-Concord, NC-SC	2.51	***	(34.82)	1.60	***	(4.86)	1.66	***	(5.14)	0.99993	***	(-20.57)	1.00001		(0.78)	1.00000		(-0.81)	29.77	***	(8.64)	2.62	**	(2.11)	0.99988	***	(-7.22)	801	-490.36
Chicago-Naperville-Joliet, IL-IN-WI	2.67	***	(108.38)	1.29	***	(7.51)	1.16	***	(4.52)	0.99992	***	(-70.52)	1.00008	***	(11.64)	0.99999	***	(-13.37)	2.21	***	(7.67)	2.20	***	(11.72)	0.99992	***	(-14.86)	6,156	-3,207.25
Cincinnati-Middletown, OH-KY-IN	2.36	***	(43.57)	1.43	***	(5.09)	1.32	***	(3.95)	0.99993	***	(-27.24)	0.99996	*	(-1.64)	0.99998	***	(-8.07)	173.98	***	(20.34)	1.61	***	(3.05)	0.99994	***	(-3.94)	1,458	-994.06
Cleveland-Elyria-Mentor, OH	2.86	***	(65.41)	1.43	***	(6.29)	1.22	***	(3.43)	0.99990	***	(-37.05)	1.00012	***	(6.19)	0.99998	***	(-14.44)	3.19	***	(6.20)	2.13	***	(6.76)	0.99995	***	(-3.37)	2,079	-972.21
Dallas-Fort Worth-Arlington, TX	2.09	***	(59.18)	1.39	***	(6.94)	1.30	***	(5.47)	0.99995	***	(-42.33)	0.99999		(-0.83)	0.99999	***	(-9.76)	89.62	***	(27.92)	0.23	***	(-6.40)	0.99995	***	(-7.10)	3,138	-2,340.88
Denver-Aurora, CO	2.46	***	(52.35)	1.03		(0.44)	0.92		(-1.19)	0.99987	***	(-29.48)	0.99996	***	(-2.73)	1.00000	**	(2.15)	4.00	***	(6.30)	0.18	***	(-9.02)	0.99997	***	(-2.91)	1,560	-893.49
Detroit-Warren-Livonia, MI	3.06	***	(101.05)	1.16	***	(3.49)	1.06		(1.33)	0.99992	***	(-51.50)	1.00002	**	(2.17)	1.00000	***	(-4.93)	8.05	***	(13.49)	0.84		(-1.61)	0.99994	***	(-5.64)	3,867	-1,414.03
Houston-Sugar Land-Baytown, TX	2.34	***	(60.52)	1.14	***	(2.62)	1.02		(0.35)	0.99993	***	(-44.48)	0.99997	***	(-3.36)	1.00000		(-1.02)	12.70	***	(15.91)	0.44	***	(-3.26)	0.99997	***	(-4.06)	2,685	-1,809.97
Indianapolis, IN	2.60	***	(38.57)	1.38	***	(3.79)	1.30	***	(3.06)	0.99990	***	(-22.89)	0.99990	***	(-4.17)	0.99998	***	(-6.28)	21.18	***	(7.98)	1.17		(0.80)	0.99993	***	(-4.50)	945	-569.55
Jacksonville, FL	2.35	***	(29.30)	1.19		(1.61)	1.02		(0.16)	0.99992	***	(-16.62)	0.99994	***	(-5.69)	1.00000		(0.28)	18.31	***	(8.89)	1.53		(0.94)	0.99995	***	(-3.29)	603	-387.54
Las Vegas-Paradise, NV	1.77	***	(24.85)	1.28	***	(2.88)	1.14		(1.41)	0.99994	***	(-20.40)	0.99989	***	(-9.33)	0.99999	***	(-3.24)	2.37	***	(5.37)	0.49		(-0.30)	0.99995	***	(-5.42)	1,041	-968.58
Los Angeles-Long Beach-Santa Ana, CA	2.50	***	(111.00)	1.25	***	(7.38)	1.06	**	(2.02)	0.99994	***	(-56.77)	0.99993	***	(-16.61)	0.99998	***	(-28.89)	2.96	***	(11.52)	0.27	***	(-12.87)	0.99999		(-1.12)	7,113	-3,961.24
Miami-Fort Lauderdale-Miami Beach, FL	2.34	***	(71.77)	1.36	***	(6.05)	0.94		(-1.16)	0.99999	***	(-19.77)	0.99982	***	(-12.32)	0.99999	***	(-7.83)	11.12	***	(14.47)	3.37	***	(3.18)	0.99992	***	(-10.87)	2,670	-1,482.88
Minneapolis-St. Paul-Bloomington, MN-WI	2.91	***	(69.18)	1.49	***	(6.93)	1.37	***	(5.38)	0.99989	***	(-42.73)	1.00003	*	(1.72)	0.99998	***	(-12.17)	3.77	***	(6.18)	1.74	***	(4.98)	0.99993	***	(-5.11)	2,238	-1,014.31
New York-Northern New Jersey-Long Island, NY-NJ-PA	2.14	***	(133.91)	1.29	***	(11.46)	1.19	***	(7.62)	0.99995	***	(-76.58)	1.00000		(-0.92)	0.99999	***	(-35.48)	0.29	***	(-23.52)	4.46	***	(33.84)	0.99998	***	(-5.43)	13,188	-9,013.50
Orlando-Kissimmee, FL	2.45	***	(39.23)	1.40	***	(3.99)	1.29	***	(2.91)	0.99991	***	(-23.76)	0.99986	***	(-6.48)	0.99999	***	(-5.67)	5.66	***	(6.30)	6.21	***	(3.13)	0.99995	***	(-4.41)	984	-590.23
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	2.20	***	(76.86)	1.50	***	(10.49)	1.29	***	(6.69)	0.99995	***	(-43.88)	1.00006	***	(6.15)	0.99998	***	(-22.31)	3.03	***	(10.07)	3.03	***	(14.54)	1.00000		(0.34)	4,668	-3,276.48
Phoenix-Mesa-Scottsdale, AZ	1.82	***	(39.72)	1.09		(1.57)	1.04		(0.75)	0.99995	***	(-30.83)	0.99996	***	(-6.38)	1.00000		(-1.26)	1.63	***	(3.28)	0.57		(-1.26)	0.99996	***	(-6.42)	2,115	-1,872.52
Pittsburgh, PA	2.15	***	(48.74)	1.36	***	(5.38)	1.24	***	(3.76)	0.99993	***	(-37.02)	0.99997		(-1.41)	0.99998	***	(-8.07)	19.63	***	(13.05)	3.65	***	(9.90)	0.99989	***	(-7.40)	2,163	-1,631.24
Portland-Vancouver-Beaverton, OR-WA	2.64	***	(49.78)	2.06	***	(8.84)	1.76	***	(6.95)	0.99989	***	(-29.12)	0.99999		(-0.82)	0.99997	***	(-12.88)	62.68	***	(12.73)	0.87		(-0.81)	1.00003	*	(1.82)	1,278	-670.87
Richmond, VA	2.39	***	(33.76)	1.34	***	(3.24)	1.22	**	(2.18)	0.99993	***	(-20.82)	0.99992	***	(-4.09)	0.99999	***	(-4.44)	106.24	***	(15.11)	4.77	***	(6.47)	0.99996	**	(-2.46)	849	-575.86
Riverside-San Bernardino-Ontario, CA	1.70	***	(39.33)	1.02		(0.46)	0.90	*	(-1.91)	0.99998	***	(-25.72)	0.99980	***	(-18.03)	1.00001	***	(5.67)	3.53	***	(8.41)	0.04	***	(-8.22)	0.99994	***	(-10.47)	2,520	-2,339.49
Sacramento-Arden-Arcade-Roseville, CA	1.72	***	(27.17)	1.30	***	(3.41)	1.10		(1.25)	0.99997	***	(-18.63)	0.99989	***	(-7.35)	0.99999	***	(-4.34)	24.24	***	(12.12)	0.01	***	(-11.42)	0.99997	***	(-2.99)	1,209	-1,114.21
St. Louis, MO-IL	2.06	***	(43.11)	1.16	**	(2.25)	1.08		(1.19)	0.99994	***	(-33.84)	0.99990	***	(-4.62)	1.00000	***	(-2.73)	5.26	***	(6.85)	1.37	**	(1.96)	0.99990	***	(-7.91)	1,650	-1,230.09
San Antonio, TX	2.49	***	(40.55)	1.25	***	(2.72)	1.12		(1.31)	0.99990	***	(-26.96)	0.99991	***	(-7.06)	0.99999	***	(-3.82)	32.29	***	(13.59)	0.27	***	(-4.39)	0.99999		(-0.96)	1,014	-604.10
San Diego-Carlsbad-San Marcos, CA	1.88	***	(38.12)	1.14	**	(2.24)	1.07		(1.13)	0.99995	***	(-30.20)	0.99996	***	(-3.66)	0.99999	***	(-7.32)	1.52	**	(2.54)	0.32	***	(-4.67)	0.99999	*	(-1.71)	1,797	-1,512.58
San Francisco-Oakland-Fremont, CA	1.85	***	(49.76)	1.41	***	(6.57)	1.26	***	(4.42)	0.99995	***	(-27.17)	0.99996	***	(-3.11)	0.99999	***	(-12.61)	3.12	***	(6.23)	1.87	***	(6.23)	1.00002	**	(2.05)	2,613	-2,093.67
Seattle-Tacoma-Bellevue, WA	2.18	***	(51.68)	1.58	***	(7.33)	1.42	***	(5.54)	0.99995	***	(-32.49)	0.99986	***	(-9.89)	0.99998	***	(-10.41)	16.63	***	(13.61)	1.00		(0.02)	0.99996	***	(-3.53)	1,992	-1,356.92
Tampa-St. Petersburg-Clearwater, FL	2.55	***	(53.23)	1.24	***	(3.21)	1.03		(0.47)	0.99996	***	(-22.98)	0.99979	***	(-16.50)	1.00000	*	(-1.81)	48.33	***	(15.68)	12.93	***	(8.47)	1.00000		(0.11)	1,641	-923.83
Virginia Beach-Norfolk-Newport News, VA-NC	1.73	***	(25.57)	1.34	***	(3.65)	1.27	***	(2.96)	0.99995	***	(-18.82)	1.00002	**	(2.16)	0.99999	***	(-3.58)	44.53	***	(15.09)	0.35	***	(-3.42)	0.99992	***	(-5.66)	1,104	-1,040.96
Washington-Arlington-Alexandria, DC-VA-MD-WV	2.34	***	(65.78)	1.44	***	(7.62)	1.28	***	(5.23)	0.99992	***	(-45.61)	0.99994	***	(-4.54)	0.99999	***	(-17.22)	4.29	***	(9.62)	1.97	***	(5.30)	0.99997	***	(-3.55)	3,012	-1,935.34

Notes: LL is the log-likelihood; $n \times t$ is the number of observations in the panel; in the event that an observation/s was dropped in the estimation process, $n \times t$ is not symmetric; values in () are z-statistics; all hypothesis tests are two-tailed;
*** denotes significant at 99%; ** denotes significant at 95%; * denotes significant at 90%; and n/s denotes not significant.

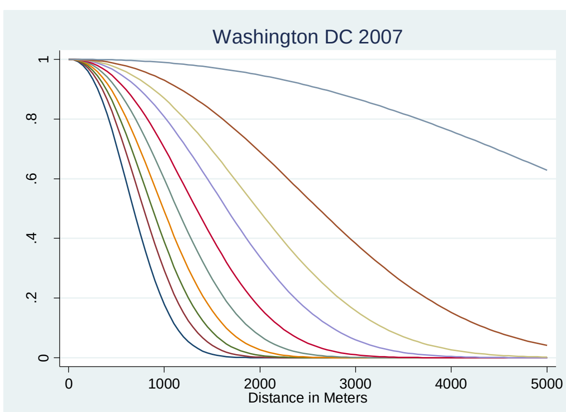
Another way of intuitively illustrating the results is by graphing the survival functions. The survival functions are the opposite of hazard functions and express the conditional probability of distance extending. Instead of expressing the hazard function as $H(d_{ij}) = \Pr(D < d_{ij})$, the survival function is expressed as $S(d_{ij}) = 1 - H(d_{ij}) = \Pr(D \geq d_{ij})$. Survival functions can be graphed by varying distance from the CBSA center and two temporal fixed effects while holding the remainder of X_{ik} constant at the mean $\overline{X}_{ik}$. Also, the survival functions can be generated at radial distances from the CBSA centers that capture ~5%, ~15%, ~25%, ~35%, ~45%, ~55%, ~65%, ~75%, ~85%, and ~95% of each CBSA's total population. In this analysis, distances capturing each successive share of population were generated for each of the three years and applied to the models by substituting relevant values into equation (4):

$$h(d_{ij} / X_{ik}) = \hat{h}_0(d_{ij}) \cdot exp(\hat{\tau}_{2000,\,2007} + \hat{\phi}_{i \Rightarrow center}\, \bar{x}_{i \Rightarrow center} + \hat{\phi}_{i \Rightarrow local}\, \bar{x}_{i \Rightarrow local} + \overline{X}_{ik}\ \hat{\Phi}_k) \qquad (4)$$

where, $\hat{\tau}_{2000,\,2007}$, $\hat{\phi}_{i \Rightarrow center}$, $\hat{\phi}_{i \Rightarrow local}$, and $\hat{\Phi}_k$ are estimated parameters; and $\bar{x}_{i \Rightarrow center}$, $\bar{x}_{i \Rightarrow local}$ and $\overline{X}_{ik}$ are mean values of the vector X. Temporal effects, $\hat{\tau}$, were set to each of the three years: (*i*) 2000 = 0 and 2007 = 0, controlling for 1990; (*ii*) 2000 = 1 and 2007 = 0; and (*iii*) 2000 = 0 and 2007 = 1.

The survival functions for all metropolitan areas are shown in Figure 2 in the Appendix B. The figure contains the survival functions and changes for each metropolitan area in alphabetical order. These survival curves describe the conditional probability of the distance between nearest neighbor tracts extending past a particular distance at specific locations within the metropolitan area. For explanation purposes, Figure 21 illustrates survival functions for Washington, DC for 2007. The *x*-axis records distance between nearest neighbors and ranges from 0 to 5,000 meters. The *y*-axis records the probability that distance between nearest neighbors, d_{ij}, extends and ranges from 0 to 1. Each graph contains 10 different colored curves which represent the distance from the CBSA center that capture ~5% through ~95% of the metropolitan area's population. For example, the curve closest to the y-axis captures urbanization patterns in the ring closest to the metropolitan core that contains ~5% of the area's population. The distance between nearest neighbors in that ring has 18 percent probability of extending beyond 1000 meters. Curves successively follow the share of area's population. The curve at the far right of the graph captures 95 percent of the population and the probability of distance between nearest neighbors extending beyond 5000 meters is 63 percent.

Figure 21: Survival Curves



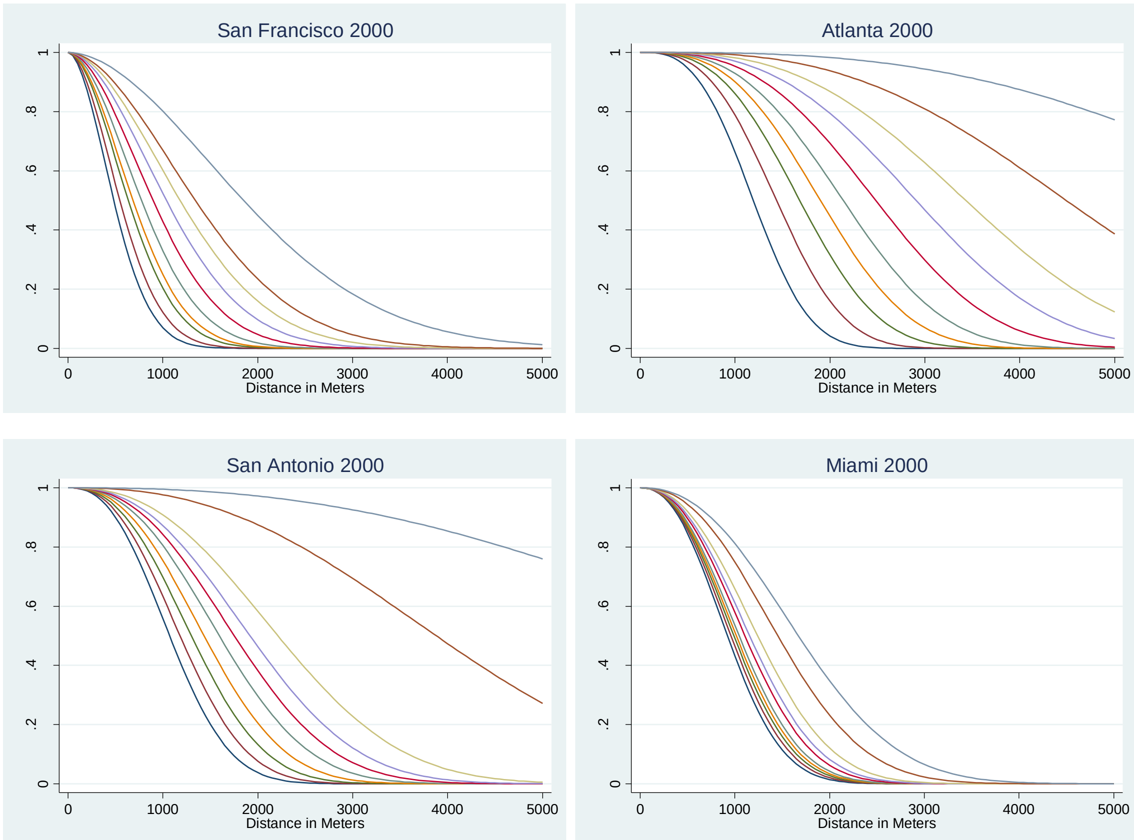
Based on the graphs illustrating urbanization patterns in 1990, 2000 and 2007, metropolitan areas can be subjectively grouped in four categories. Figure 22 illustrates this.. In the first category are regions with high-density, compact patterns of urbanization, like New York, Chicago, and San Francisco. Their estimated survival functions are steeply sloped and tightly bunched together suggesting that the probability of distance between nearest neighbors extending very far is small. This basic pattern is for the most part consistent across the entire region. The last two population rings, capturing 85 percent and 95 percent of the population are often less steep and extend further out on the graph.

In the second category are regions with low-density, sprawling patterns of urbanization, like Atlanta, Phoenix, Charlotte, Riverside, and San Antonio. Their survival functions are more flatly sloped, especially at their tops, and spread out. The last two population rings usually indicate very high probability of distance between nearest neighbors extending beyond 5000 meters. In Atlanta, that probability is around 80 percent.

The areas falling in the third category have high-density core suggested by tightly bunched first (several) population rings. The rings farther from the center, however, suggest sprawling suburban areas and are illustrated by less steep functions. Those areas include Baltimore, Denver, Austin, Cleveland, and Philadelphia. In Baltimore, for example, the outermost population rings have about 60 percent probability of extending beyond 5000 meters.

The last, fourth, category encompasses areas that are nearly spatially invariant at most distances. This is suggested by their survival functions that are clustered together without much variation between inner and outer population rings. Metropolitan areas exhibiting such settlement patters include Miami, Los Angeles, Dallas, Seattle, and Las Vegas.

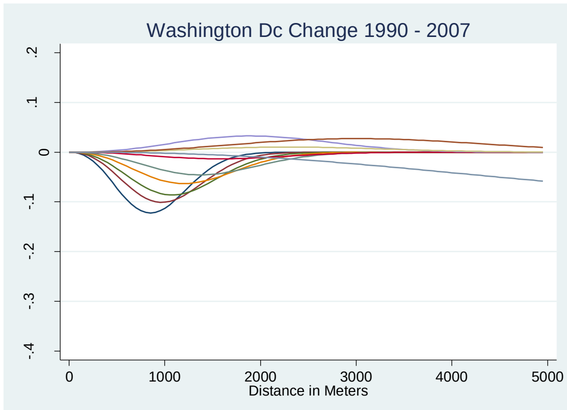
Figure 22: Survival Curves in High Density, Low Density, Mixed, and Spatially Invariant Regions



Metropolitan Structure Changes via Spatial Hazards

To answer the specific questions posed by this study, the survival functions can be graphed to illustrate changes in metropolitan spatial structure. The change graphs are generated by using the numeric output of the survival functions and differencing the various survival functions for each region. Figure 23 serves for explanation. The graph illustrates the change in survival probabilities between 1990 and 2007 for each of the population rings in the Washington, DC metropolitan area. In the change graphs, the x-axis also ranges from 0 to 5,000 meters and records distance between nearest neighbor tracts. The y-axis records the change in the probability that distance extends and ranges from -.4 to .2. The change graphs are much more heterogeneous than survival curves for each of the years as they pick up differing changes within and across regions. When the change curves are positive, survival rate has increased over time, suggesting a sprawling effect on urbanization patterns. In contrast, when the change curve is negative, survival rate has decreased over time implying a compacting effect on urbanization.

Figure 23: Change in Survival Functions



The color codes are consistent across all graphs, so for example, dark blue color always refers to area encompassing initial ~5% of area's population around the urban core. Figure 23 shows that the probability of distance extending beyond 1000 meters in the 5% population ring decreased by 11 percent. A decrease in survival probability means that nearest neighbors are spaced closer together in 2007 than they were in 1990. In contrast, the probability of distance extending beyond 3000 meters in 85% population ring increased by 3 percent. That means that nearest neighbors are spaced further apart in suburban locations of the Washington, DC metropolitan area. That seems to be true for 65% and 75% population rings as well.

Figure 2 in Appendix B contains change graphs for all regions. Changes among all the 35 metropolitan areas suggest a general tendency towards densification. The median change for all population rings is summarized in the Table 8. The first section summarizes changes between 1990 and 2000, the middle section between 2000 and 2007, and the last section between 1990 and 2007. In looking at the last section, we can see that survival function decreased by 3 percent at a 500 meter distance between nearest neighbors in the first population ring. The largest decrease was during 1990s when the probability of survival beyond 1000 meters in the first ring fell by almost 10 percent. The period between 2000 and 2007 is again symbolic of general tendency towards sprawl among most of the metropolitan areas indicated by positive change in survival curves.

Table 8: Classification of Land Use Change

Median 1990 to 2000

	At a Distance from the Regional Center of Gravity Capturing % of Population									
	~5%	**~15%**	**~25%**	**~35%**	**~45%**	**~55%**	**~65%**	**~75%**	**~85%**	**~95%**
500	-4.39%	-3.21%	-2.84%	-2.29%	-1.49%	-1.15%	-0.78%	-0.49%	-0.31%	-0.12%
1,000	-9.67%	-8.63%	-7.14%	-6.97%	-5.33%	-4.55%	-3.36%	-2.58%	-1.54%	-0.62%
2,000	-2.07%	-3.05%	-3.29%	-3.89%	-4.53%	-7.19%	-7.70%	-7.92%	-5.69%	-2.63%
3,000	-0.01%	-0.05%	-0.20%	-0.31%	-0.85%	-0.95%	-2.54%	-5.80%	-7.39%	-4.75%
4,000	0.00%	0.00%	0.00%	-0.01%	-0.02%	-0.03%	-0.48%	-1.36%	-4.85%	-6.16%
5,000	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.01%	-0.09%	-1.81%	-6.26%

Median 2000 to 2007

	At a Distance from the Regional Center of Gravity Capturing % of Population									
	~5%	**~15%**	**~25%**	**~35%**	**~45%**	**~55%**	**~65%**	**~75%**	**~85%**	**~95%**
500	1.20%	1.14%	1.13%	1.11%	1.00%	0.83%	0.66%	0.48%	0.20%	0.06%
1,000	2.57%	2.98%	3.11%	3.54%	3.49%	3.26%	2.54%	2.08%	1.27%	0.25%
2,000	0.19%	0.69%	1.36%	2.42%	3.48%	4.32%	4.71%	4.77%	3.18%	1.14%
3,000	0.00%	0.02%	0.06%	0.12%	0.52%	1.07%	2.51%	3.66%	3.62%	2.19%
4,000	0.00%	0.00%	0.00%	0.00%	0.02%	0.12%	0.35%	1.41%	2.14%	2.23%
5,000	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.03%	0.25%	0.86%	2.31%

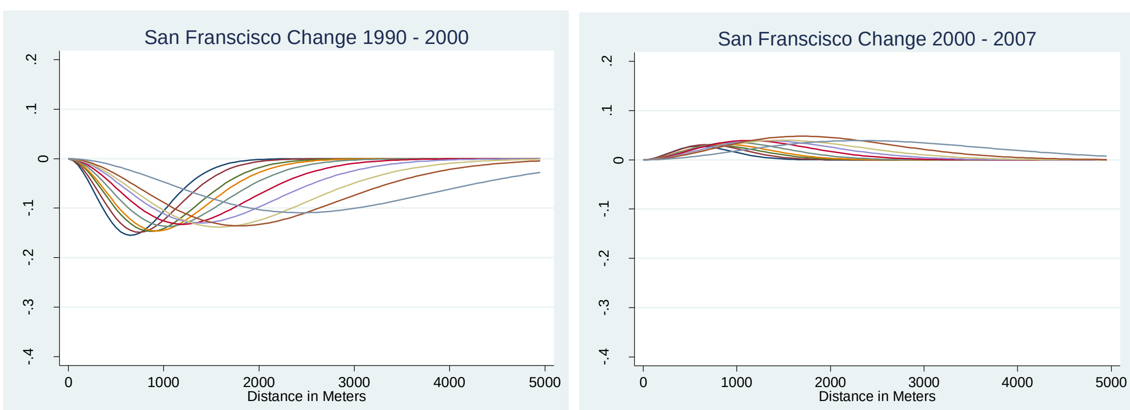
Median 1990 to 2007

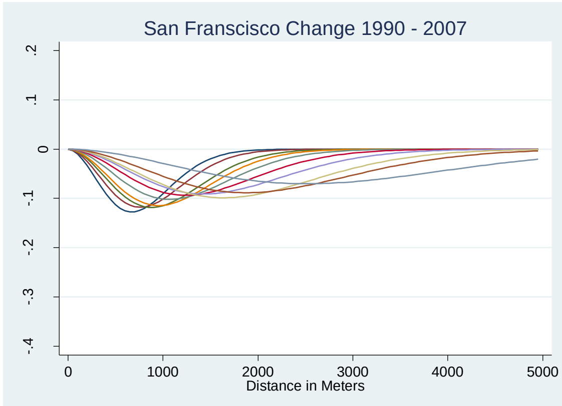
	At a Distance from the Regional Center of Gravity Capturing % of Population									
	~5%	**~15%**	**~25%**	**~35%**	**~45%**	**~55%**	**~65%**	**~75%**	**~85%**	**~95%**
500	-2.59%	-2.25%	-1.56%	-1.11%	-0.36%	-0.23%	-0.13%	-0.07%	-0.09%	-0.04%
1,000	-6.12%	-4.20%	-3.70%	-3.29%	-1.82%	-1.03%	-0.47%	-0.43%	-0.49%	-0.24%
2,000	-0.67%	-1.01%	-1.41%	-2.06%	-1.32%	-1.10%	-0.59%	-1.15%	-1.72%	-1.00%
3,000	-0.01%	-0.01%	-0.02%	-0.03%	-0.14%	-0.02%	-0.30%	-0.01%	-2.22%	-1.83%
4,000	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.02%	0.00%	-1.01%	-2.54%
5,000	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.20%	-2.72%

In looking at changes among the 35 metropolitan areas, several urbanization trends are immediately evident. Between 1990 and 2000, all but seven metropolitan areas saw decreasing survival rates suggesting that the probability of distance extending has fallen and nearest neighbor tracts are spaced closer together. This result implies that most of the metropolitan areas have grown more densified during 1990s. During 2000s, however, every metropolitan area saw an increase in the survival rate suggesting a reversal of the 1990s trends and general suburbanization of metropolitan areas. The final resultant change, between 1990 and 2007, consequently varies a lot among the areas.

Using the four regions illustrated in Figure 22 (above) as examples, it is also possible to categorize CBSAs into four typologies based on their change. The first category, represented by San Francisco, is typified by general compacting over the 17-year period. While the 2000s do show increasing survival rate, the increase is of a smaller degree than the compacting seen during 1990s. On the whole, then, the change reflects decreasing survival of distance between nearest neighbors, namely infill development and compacting of metropolitan areas. Four out of ten metropolitan areas fall into this category, though change is of varying degrees. For example, while the probability of distance between nearest neighbors extending beyond 1,000 meters decreased by 30 percent in the core center of Austin; it only decreased 12 percent in Baltimore. Both CBSAs, however, saw overall compacting of the metropolitan area. Figure 24 illustrates change in the survival function for San Francisco. In the first graph, the survival functions across the whole metropolitan area show a decrease between 1990 and 2000. In the second graph showing change between 2000 and 2007, the survival functions have increased. The cumulative effects in the graph showing change between 1990 and 2007 indicate overall densifying of metropolitan San Francisco, but to a smaller degree than seen in 1990s. The remaining metropolitan areas that fall in the "densification" category include: Portland, Dallas, Austin, Boston, Charlotte, New York, Philadelphia, Seattle, and Virginia Beach. Areas that also densified but to a smaller degree are: Baltimore, Cincinnati, Cleveland, Orlando, Pittsburg, and Sacramento. Change between 1990 and 2007 in the aggregated sample of metropolitan areas that densified indicates that the largest decrease in survival function was for survival of distance beyond 1000 meters between nearest neighbors. In the first and second population ring, that decrease was 12 and 10 percent respectively. The probability of distance extending beyond 500 meters also went down by 7 percent and 5 percent in the first two population rings. For the last population ring, at the suburban fringe, survival function decreased by well over 6 percent for distances beyond 3000, 4000 and 5000 meters.

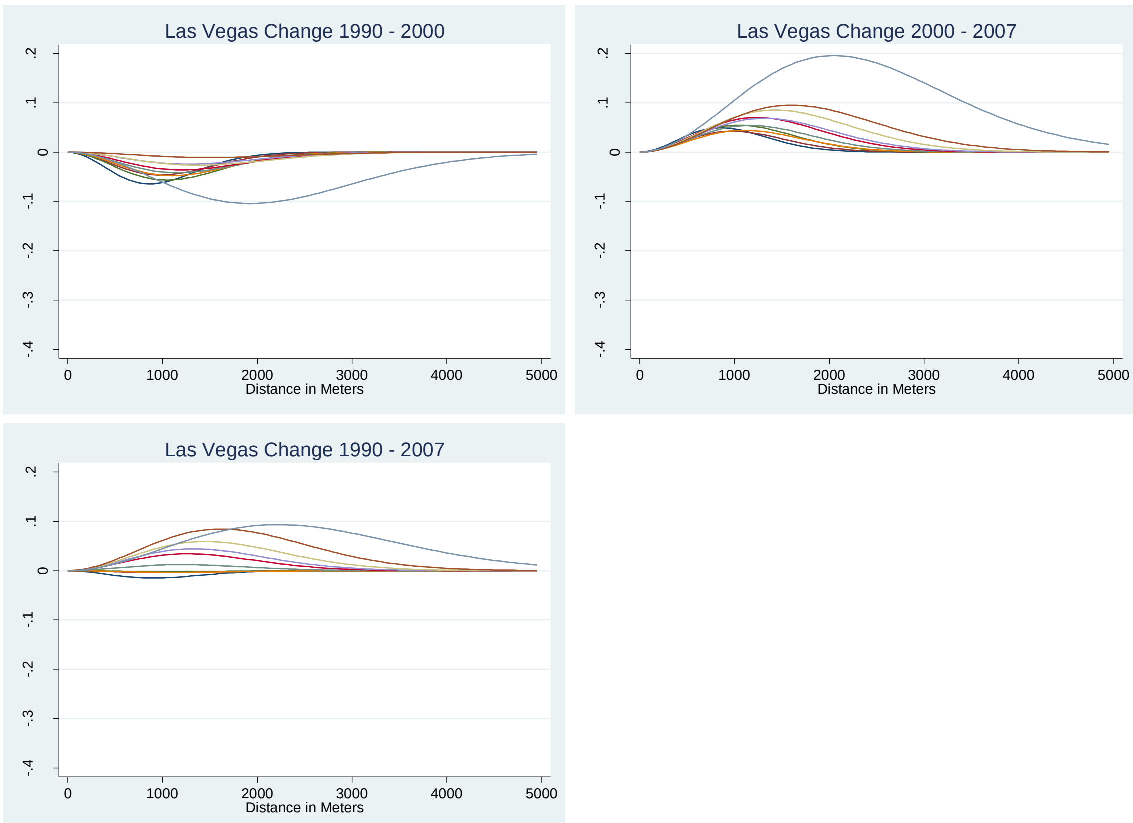
Figure 24: Change Survival Functions in High Density Areas





The second category is typified by areas that overall sprawled out over the seventeen years. While in some of them the 1990s were still characterized by some compacting, the sprawling effect of 2000s dominated the resulting outcome. These metropolitan areas are characterized by positive change in the survival rate, such as Las Vegas, for example in Figure 25. The other five metropolitan areas falling into this category include: Denver, Detroit, Houston, Jacksonville, Phoenix, and Riverside. In Las Vegas, change between 1990 and 2007 shows increase in the survival rates in most of the population rings. The increase was successively larger away from the center. The first ring however saw a minuscule decrease in survival rates. For aggregate sample of sprawling metropolitan areas, that decrease is 0.3 percent and 0.8 percent for extending beyond 500 and 1000 meters respectively. It is also interesting to note that the temporal parameters in this category of metropolitan areas are often insignificant. Again, the change in 1990s differs from that of 2000s. During 1990s, sprawling metropolitan areas in the aggregate still experienced some densification, with the largest relative change being in the first population ring, a 3 percent decrease in survival beyond 1000 meters. The last population ring followed. During 2000s, the increase was of a much larger proportion, thus completely reversing the trend before. Primarily, for the sprawling category of metropolitan areas, median changes were in the magnitude of +3 to +6 percent.

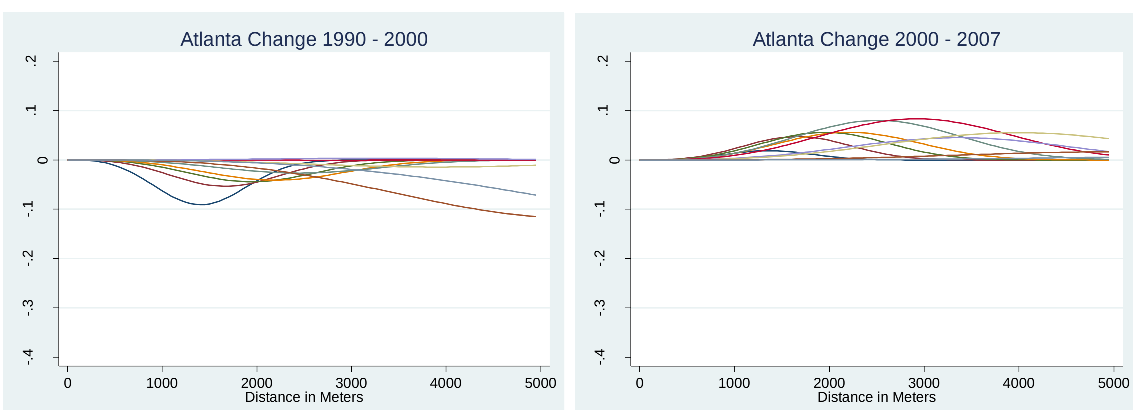
Figure 25: Change Survival Functions in Sprawling Areas

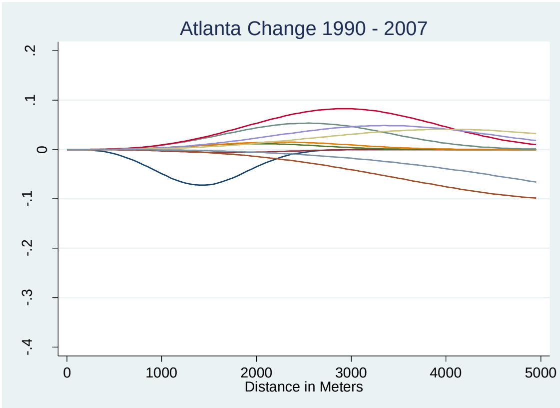


The third category is characterized by mixed outcomes, with densification in some parts of the metropolitan area, usually the central core, and sprawling of the other areas, mostly the suburban areas. Eight metropolitan areas fall into this category: Atlanta, Chicago, Indianapolis, Minneapolis, Richmond, San Antonio, St. Louis, and Washington DC. Again, most of densification occurred during 1990s, while 2000s were significant for decentralization. Figure 26 shows urban form change in Atlanta which typifies changes for the "mixed" category of the metropolitan areas. Atlanta provides an interesting example of conflicting urbanization patterns that occurred in 1990s and 2000s. While the

1990s is generally defined by decreasing survival rates, particularly in the center of the region, the 2000s seem to have reversed that trend. The reversal though is not sudden because the rings that saw the smallest decreases in the survival rates in 1990s also saw the largest increases in 2000s. By the same token, the 5% ring which experienced a large decrease also saw the smallest increase later on. Nevertheless, the overall result is densification of the central core as well as of the two outmost population rings, but the opposite effect occurs in the inner parts of the Atlanta metro region. The median change among the “mixed” category for the entire observation period indicates a 6 percent fall in survival beyond 1000 meters for the first ring, and 4, 3 and 2 percent decrease for the successive rings. The survival rate in the last ring decreased by 3 percent for distances beyond 5,000 meters. For the inner rings, the increase was generally in the 2 percent magnitude.

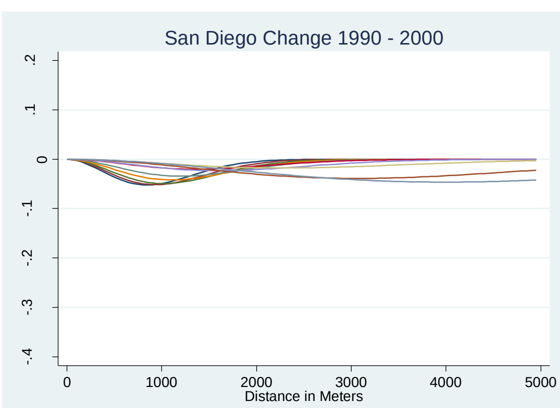
Figure 26: Change Survival Functions in Mixed Areas

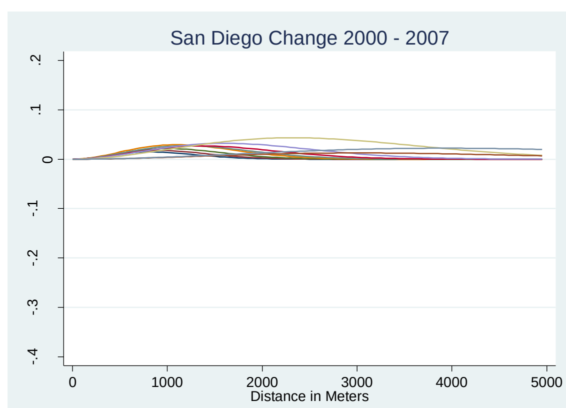


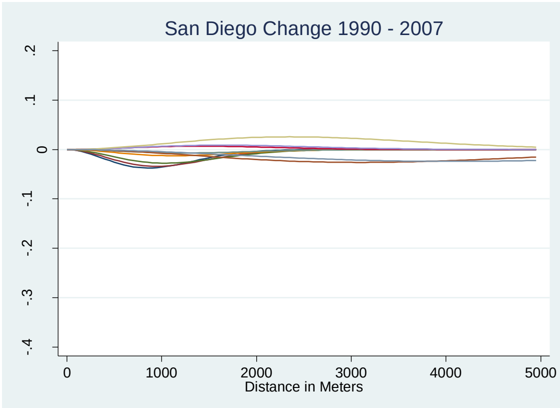


Finally, the last category of metropolitan areas is characterized by minimal overall change in metropolitan spatial structure over the 17-years period. Four metropolitan areas fall into this category: Miami, Los Angeles, San Diego, and Tampa. Figure 27 illustrates the category with San Diego. Some of these metropolitan areas did see a slight difference in 1990s and 2000s; however the two trends offset each other to result in minimal change. Interestingly, still, all of these regions densified during 1990s while they sprawled in 2000s. The median change in this category is a decrease of about 2 percent in the inner rings and 0.5 percent decrease in the outer rings.

Figure 27: Minimal Change Survival Functions







Discussion

Overview of the Findings

Metropolitan spatial structure is ever evolving and intricately complex. The complexity calls for urban measure methods to follow suit. In pursuit of a method which will provide objective yet universal information on the state of urban form and its expansion, researchers have offered a number of alternatives. Carruthers et al (2010, 2011) offered one such alternative via application of survival models to a spatial setting. Also, the dynamics of the last two decades have raised questions about future urbanization trends of the American metropolitan areas. Predominantly sprawling eras may have reversed in favor of urban living. Yet, while the first essay showed that the emerging population density trends validate those assumptions, it also showed that the change is not necessarily universal even among those titled “winners” and “losers”. For a detailed picture of changes within metropolitan structure complex measures are often data intensive. However, detailed comparative data across many metropolitan areas is still hard to come by. Spatial hazard framework used in this study provides a thorough analysis of complex and stochastic spatial structure change across many metropolitan areas and relies on widely available public data. This analysis also goes beyond the first essay and accounts for some of the exogenous factors commonly referred to in urban economics, such as income, travel time, and age of housing.

The purpose of this study was to measure change in urban structure using survival analysis and to answer three questions: (1) Do spatial hazard models suggest changes in the metropolitan spatial structure? (2) Are changes consistent within a metropolitan area? (3) Are changes consistent with more traditional measures of metropolitan spatial structure as observed in the first essay? Application of the spatial hazard functions to 35 metropolitan areas to measure change in metropolitan spatial structure has provided a multihued insight in the way urban form has changed over the 17-year period. The analysis shows that change has not always been consistent over time and even within a metropolitan area. One trend that does emerge again is that majority of the metropolitan areas have densified over the 17 years. About one-half of the areas studied grew denser while one-fifth densified in some parts of the metropolitan area, which was most often in the areas closer to the urban core. One-fifth of the metropolitan areas sprawled over this observation period. Densification trend, though, is of varying degree. For some metro regions, we can say with certainty that urban form is denser. For example, in Austin, change in survival probability decreased by about 30 percent in most parts of the region. In other areas, the change is more subtle. Baltimore, for example, densified in the inner core by about 12 percent and by only 4 percent on the suburban fringe. This type of change may not be easily observable in the built environment.

Separating the effects between the 1990s and 2000s strengthened the conclusion from the first essay that the 1990s era encountered reversal of suburbanization trends towards the urban core. The 2000s in contrast are characterized by general urban expansion. The interesting addition provided through this analysis is how changes varied throughout a metropolitan region. Even in regions that mostly sprawled, there was a dominating effect of densification during 1990s particularly in the inner population rings. Further, when comparing the regions that densified and those that sprawled, the degree of densification is much larger than the degree of sprawl. In some parts, that magnitude is twice as large. What is still difficult to say, and will remain so until the housing market recovers, is whether the reversal of urbanization trends in 1990s was truly a reflection of reversed household preferences for urban living or whether they were the result of some unobserved conditions. The 2000s cannot yet be used as an indication of any long-term trend given that intensified residential construction during the period left vastly spread housing vacancies across many parts of the country.

In comparing changes in metropolitan spatial structure observed via spatial hazard method and those observed in the first essay, the trends prove to be very similar. Both sets of measures indicate that the change has been predominantly towards repopulation and densification of areas closer to the urban core. Evaluating spatial hazards with density gradients provides the most analogous comparison as both measures examine the change in structure relative to the urban core. So, while the density gradient classified most regions in a group for which the population density increased at the core and population density in suburban areas increased as well, the spatial hazard functions grouped them similarly while providing a focused picture of that change. In Dallas, a metro deemed a “winner” by both sets of measures, survival probabilities between 1990 and 2007 show similar densification both in the central core as well as at the outskirts, about 11 percent decrease in survival in the core and 14 percent on the fringe. According to the density gradient measure, the population density at the core increased by about 30 percent while the gradient increased by 4 percent. So, both measures show general tendency of densification. Not all regions though are grouped in the same category. For example, the Baltimore metropolitan area is typified by fallen intercept and flattened gradient between 1990 and 2007, meaning decentralization. Spatial hazards suggest that Baltimore's survival probabilities decreased during the same period, suggestion densification. While the density gradient captures the population change within the same tract and the resultant estimate indicates if the population density in the core increased or decreased, the spatial hazard captures change in distances between population gravity points between two neighboring tracts. The census tract gravity points are based on census block group data. More specifically then, the spatial hazard captures clustering of population centers. It is thus feasible that in Baltimore, the population density in the core decreased while the population gravity points grew more clustered. The clustering referred to here does not measure the same type of clustering as Moran's I estimates in the first essay. The Moran's I from the first essay measures clustering of census tracts with similar population densities.

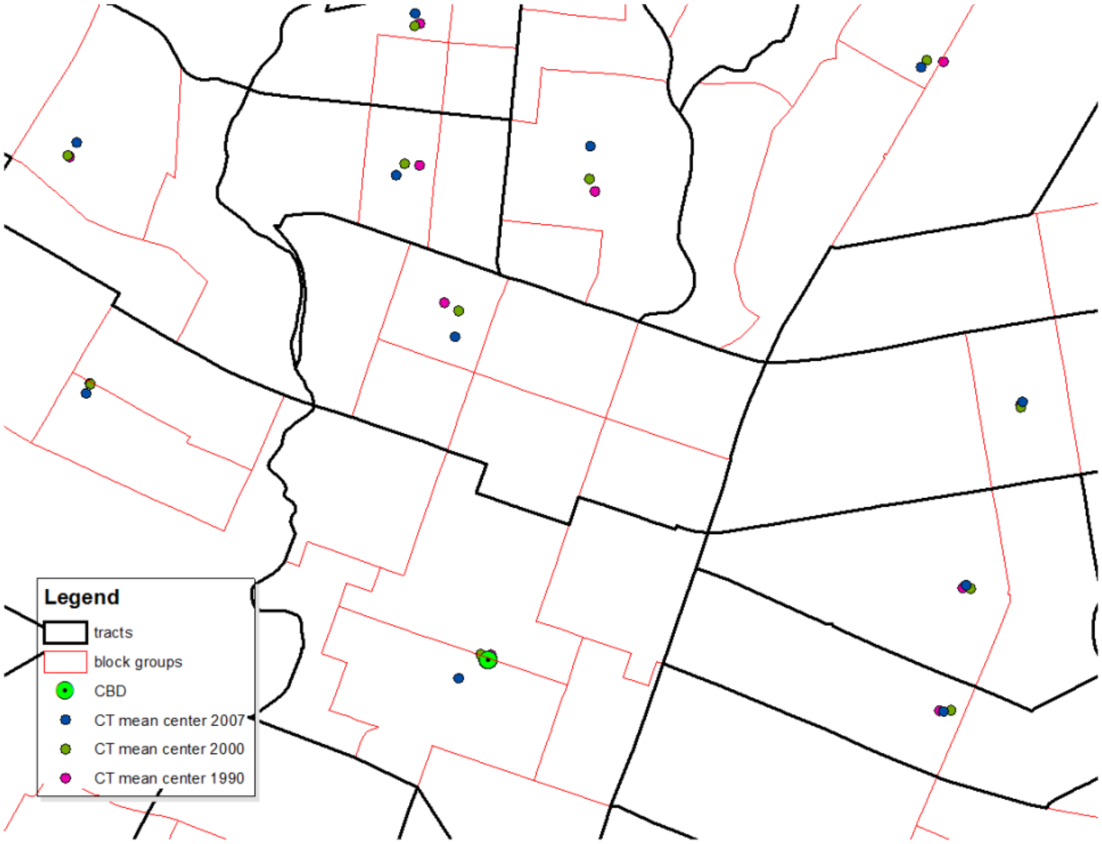
In the end, this analysis shows that demarcation between deemed "winners" and "losers" is a fine line. To obtain the full picture on the spectrum of spatial structure change, multiple measures of urban form are needed. Among the measures presented in the two essays, the estimates do not indicate drastic differences between measures but offer answers to various parts of the change question.

Limitations of the Study

Although the spatial hazard framework is very promising for the study of spatial structure change, it may take some time for the method to gain traction as the method is not as simple to grasp as the density gradient is for example. That continues to be the main advantage of the density gradient model.

Another limitation of this study concerns the use of census tract boundaries to track the movement of the population mean centers. The population mean centers are generated for each tract and for each year based on the population count in block groups encompassed by a tract. Block groups similar to census tracts are built around a constant population count and generally contain between 600 and 3,000 people, with an optimum size of 1,500 people. Census tracts generally have between 1,500 and 8,000 people, with an optimum size of 4,000 people. Thus, there are generally about 3 block groups in each tract. Therefore, constraining the movement of population mean centers by census tract boundaries may not reflect the true change in the population gravity point. And, the movement is constrained by the number of block groups in each tract. Generally, the population gravity point did not move to another block group during the observation period. Figure 28 illustrates the limitation posed by use of census products..

Figure 28: Census Tract Limitation



Essay Three: Spatial Exploration of Foreclosures in Maryland

The decade of the 2000s has also been marked by an unprecedented housing boom followed by a subsequent collapse and a great economic recession. The housing collapse, and more particularly the foreclosure crisis resulting from it, left profound effects not only on neighborhoods, families, and the entire economy, but also on the organization of the metropolitan areas. While the incidence of foreclosures is not a new phenomenon, the recent occurrence is significant due to its magnitude, concentration and suddenness. In metropolitan areas hardest hit by the crisis, foreclosures followed by vacancies have contaminated almost entire neighborhoods.

The beginning of the current housing crisis is inseparably connected to subprime lending. In the Department of Housing and Urban Development's (HUD) "Report to Congress on the Root Causes of the Foreclosure Crisis" in January 2010, the authors discuss the precipitating causes of the crisis and conclude that the significant increase in mortgage delinquencies and foreclosures is primarily the result of hasty growth in high risk loans, both due to loans' terms and the loosening of underwriting oversight and standards. The peculiarity of the current foreclosure crisis is that the economic recession did not initially produce defaults and foreclosures. It was the slowing of house price appreciation that ended lenders' ability to continue extending credit to borrowers. At the outset, high-cost lending allowed borrowers to obtain mortgages that they could not otherwise afford leading to larger house appreciation than would happen under fundamental principles of housing demand. Once the economic slowdown began, both an oversupply of new homes and mounting defaults caused wide destabilization of housing values. Damage rapidly spread throughout the wider financial system nationally and abroad as banks and mortgage companies were highly leveraged with housing securities instruments. What followed was one of the worst economic downturns in the American history.

The HUD report also recognizes spatial heterogeneity of the foreclosure crisis. It differentiates between two types of crisis areas. The areas in the South and West of the country were characterized by frequency of high-cost lending between 2004 and 2006 and larger home price appreciation before the crisis hit, while the areas in the Midwest already suffered from weak economies prior to the current crisis. Prevalently, foreclosures took different spatial forms in the two groups of states. In the fast growing states, foreclosures were primarily concentrated in newly built subdivisions,; while in the established metropolitan areas of the Midwest foreclosures were concentrated in older urban neighborhoods. Within the Midwest, overall worsening of the country's economic conditions further deteriorated already high unemployment rates and exacerbated elevated housing vacancies in older urban cores and among minority groups (Immergluck 2009b, 2010). In the fast-growing metropolitan areas, in the South, Southwest and West, the spatial distribution of foreclosures was different. Due to high home price appreciation, these areas suffered significant decreases in housing affordability. Increased demand for proximity to urban amenities coupled with the rising cost of housing, led to a greater appreciation of housing values in areas closer to the urban centers. Some preliminary spatial analysis suggested that households drove increasingly further away from the employment and service centers into auto-dependent suburban communities that offered more affordable housing (Immergluck, 2009b). However, with loosened credit requirements and consequent availability of more diverse mortgage products, many households overextended themselves. As well, they may have not accounted for the cost of transportation or potential increase in the cost of transportation. Yet, the set of events beginning in 2007, including mortgage interest rates resets on some early adjustable rate mortgage (ARM) products, rising energy costs, and general economic cyclicality, triggered a wave of defaults which escalated to unprecedented proportions. Households who were "on the fringe", both financially overextended and living on fringes of urban areas, found themselves in a very uncertain situation. They were not able to refinance their mortgages into more affordable terms since most did not accrue any equity in their homes. Lack of demand and/or unavailability of mortgage financing on the urban fringe did not allow them to sell or refinance their homes. Consequently, they defaulted on their mortgages (HUD, 2010). To make matters worse, deteriorating of macroeconomic conditions nationally further aggravated an already destabilized housing market. After nearly collapsing in the summer of 2008, the demise of the financial sector threatened to push the entire economy into a deep recession. In October 2008, the federal government responded with a $700 billion bailout of financial institutions. But, the economic problems spread to the other sectors with construction and manufacturing losing significant numbers of jobs. With unemployment reaching 9 percent by early 2009, the foreclosure crisis spread from primarily non-prime borrowers to all borrowers. The delinquency rate among borrowers with prime mortgage products, jumped from historical 2 percent to over 7 percent[9]. The total number of past due mortgages along with those in foreclosure reached an astounding 6.7 million by the end of 2009[10].

[9] Mortgage Bankers Association, National Delinquency Survey
[10] Ibid.

The question on spatial location of foreclosures relevant to urban form is whether their concentration in suburbia stems from reversed preferences for urban centers or was it the spatial distribution of higher risk lending during the subprime boom that led to their concentration in suburbia. According to arguments by William Lucy, there has been a long term demographic and cultural shift away from dispersed and suburban living towards a more urban future (Lucy, 2010). The author saw the foreclosure crisis as a supporting consequence of the urban shift rather than a cause. As the evidence of the shift, the author discuses several indicators, including strong housing prices in many cities versus rapidly declining prices in suburbs, and in those neighborhoods built before 1940s which are in more walkable communities. He also accounts for increased costs of transportation, and demographic shift toward smaller households such as elderly, empty nesters, and singles that tend to favor urban or inner suburban settings. Other experts have made similar suggestions (Dowell and Pitkin, 2009; Birch, 2009; Nelson, 2004).

The empirical examination of spatial distribution of foreclosures, though rather scarce, has shown mixed results. The results from Immergluck's study (2010) suggest that the occurrence of foreclosures in suburban locations may be because of unobserved characteristics of the loans made in such locations. Many of the homes in new suburban communities were likely financed during the peak of the subprime boom. In the study, the author grouped 75 metropolitan areas in two groups. One group included those metropolitan statistical areas (MSAs) that already had relatively high levels of REO in late 2006, before the national foreclosure crisis. The second group of the MSAs included those that had very low levels of REO in late 2006 but saw a steep housing price depreciation and very large increases in REO during the 2006 to 2008 period. This study accounted for spatial variation of REOs via two variables. The first variable which accounts for suburbanization indicates whether at least half of a zip code lies in a primary central city. The second variable accounts for commute efficiency, i.e. proximity to major employment centers, and is the percentage of auto-commuting residents who commute more than 30 minutes to work. After controlling for other determinants, such as lending supply, age of housing, poverty rate, regional housing price, unemployment trends, and race, the two variables measuring intrametropolitan spatial location had no apparent bearing on foreclosure growth. The result was the same for both types of the metro areas. The author concluded that the occurrence of foreclosures in suburban location was due to unobserved characteristics of the loans rather than some spatial disadvantage of these new subdivisions. Ong and Pfeiffer (2008), in contrast, looking at foreclosures in Los Angeles County in early 2008, found that exurban location explained 20 percent of the spatial variation in foreclosure rates. Exurban location was controlled via a binary variable for zip codes located in northern Los Angeles County. The authors believed that those exurban locations suffered from speculation on new home construction and were more vulnerable to decreases in demand due to their high commuting costs and traffic congestion. The only control for the level of subprime lending, though, was the HMDA reported level of first lien, owner-occupied originations in 2006 that were five points or more above treasury rates. There are several other limitations to these studies. Both studies relied on data at zip code level. Zip codes are typically larger than census tracts. Census tracts are mostly used to approximate neighborhoods characteristics since they are aligned with physical boundaries, and have more similar housing and demographic characteristics. Further, when measuring the effect of suburbanization, neither study accounts for actual distances to the central city. Immergluck study also does not account for intrametropolitan variation in housing values and unemployment trends. Finally, it uses median age house built data from 2000 Census and assumes that zip codes with low median housing age in 2000 are more likely to have high levels of post-2000 new construction. That may not be the case universally across MSAs.

Focusing on a different measure of spatial location, Rauterkus et al. (2010) show that location-efficient homes in Chicago, Jacksonville, and San Francisco have a lower probability of mortgage default. To achieve location efficiency, homes should be located in a compact residential development, with transit access, and proximity to schools, shopping, workplaces, and other amenities. The authors hypothesize that residents in location efficient communities save on cost of owning a vehicle because they have alternative modes of transportation: to walk, bike or use public transit. As a result, location-efficient homeowners have a lower probability of mortgage default since they do not have to spend a substantial portion of their household budget on car ownership and are not directly affected by a gasoline price increase. Transportation costs have been the second-largest expenditure for a typical American household, averaging $8,500 per year (Brookings Institution, 2006). In the study, location efficiency was proxied via vehicles per household scaled by income, and the Walk Score. The Walk Score rates the walkability of a specific address on a scale from 0 to 100 by compiling the number of nearby stores, restaurants, schools, parks, etc., within a one-mile radius from the subject location. Higher scores suggest more walkable locations while an address with a score below 50 would be considered car dependent. However, some important walkability factors, such as topography and weather conditions, and proximity to employment centers are not accounted for in the Walk Score, and the distances are measured in straight path distance and not actual distance walked on a street grid. The study finds that mortgage default probability increases with the number of vehicles owned after controlling for income. Also, mortgage default probability decreases with higher Walk Scores in high income areas but increases with higher Walk Scores in low income areas. But, as authors also discuss, the location efficient mortgages may simply perform better not because they are in location efficient areas but the because of some unobserved characteristics of those areas, i.e. amenities, where demand for homes was not as severely impacted by the housing bust. In those instances, homeowners have alternatives to mortgage default, such as selling or refinancing. Based on their findings, the authors promote location efficient mortgages which would presumably reward homebuyers of location-efficient homes with more flexible mortgage underwriting terms, for example higher debt-to-income ratio. They do not discuss the potential capitalization of the more flexible terms into housing prices.

The study presented here examines the spatial distribution of foreclosures in the state of Maryland for the period between the beginning of 2006 and the end of 2009. To date, there have not been any studies specifically looking at spatial distribution of foreclosures in Maryland. In examining the accumulation of foreclosures in Maryland, the study goes beyond previous empirical work by focusing on the relationship between concentrations of foreclosures and their proximity to transit, accessibility to employment centers by automobile and transit, and the proximity to the central business districts. By introducing a richer set of spatial variables, I aim to gather a better sense of the impact of urban form on the foreclosure crisis. Although the analysis is spatially limited to

Maryland, it nevertheless covers a critical period of the foreclosure crisis. The study also focuses on two metropolitan areas in Maryland with seemingly varying housing markets. The Baltimore metropolitan area can be characterized as one of the weaker markets with relatively large concentrations of foreclosures in central city neighborhoods even before the current crisis. In contrast, five Maryland counties included in the Washington, DC, metropolitan area experienced particularly robust housing growth in the boom and were not typified by high levels of foreclosures before the current crisis. Based on the previous literature, I expect to find that subprime lending, the housing bubble and urban form may have had different impacts on the two metropolitan areas. I conduct an empirical analysis to answer the question if the levels of foreclosures across Maryland depend on the proximity to public transit, accessibility to employment centers, or proximity to central business districts.

Foreclosures in Maryland

When a property owner defaults on a mortgage loan, according to the loan terms the lender has the right to foreclose on that property. Across the country, all states have either judicial or non-judicial foreclosure processes. In judicial states, the process is conducted through the court system. In non-judicial states, the foreclose process is defined by state statute, and the lender is only required to publicly file a notice of default. Beyond the type of foreclosure process each state uses, each state also has laws in place that govern the timeline of the process. Ordinarily, these rules are in place to afford troubled homeowners with protections from hasty foreclosures. In Maryland, lenders must file a foreclosure complaint and a lis pendens in the court. Lis pendens is a recorded document that provides public notice that the property is being foreclosed upon. Unlike foreclosure processes in judicial states, a judge is not required to rule on a foreclosure case in Maryland. Before a foreclosure is filed in court, the lender must notify the borrower and property owner that the mortgage is in default under the terms of the loan.

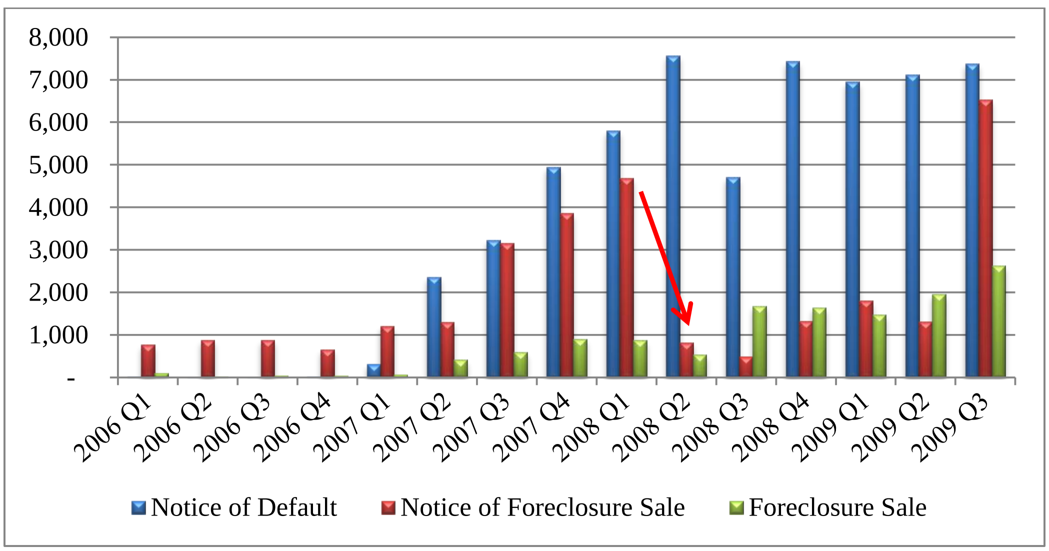
With filings drastically spiking at the beginning of 2008, in April of the same year, Maryland legislators signed an emergency bill which significantly lengthened the foreclosure process from 15 days to 150 days hoping to provide homeowners with more time and notice before a foreclosure sale. The bill required a lender to wait 90 days after default before filing the foreclosure action and to send a uniform Notice of Intent to Foreclose (NOI) to the homeowner 45 days prior to filing an action. It also requires personal service to notify a homeowner of impending foreclosure action and that a sale may not occur for 45 days after service. A lender has to produce a proof of ownership when filing a foreclosure action. The bill codified the right to cure, which allowed a homeowner to stop foreclosure by paying what is owed up until one business day before the sale. Immediately following the bill foreclosures significantly dropped, while filings and the number of properties in distress declined some. The federal government also responded with the Making Home Affordable Program which required that lenders offer all qualified, defaulting homeowners an opportunity to modify their existing mortgages. Though the program itself had a limited impact, it led to a wider effort among lenders to modify defaulted mortgages. An important trend that emerged as a result of the new foreclosure law, which is not specific to Maryland though, is that the average number of days that a mortgage was delinquent before the foreclosure start increased significantly. At the beginning of 2008, it averaged around 233 days in Maryland[11]. That number went

[11] http://www.lpsvcs.com/NewsRoom/IndustryData/Documents/2011%20-01%20January%20Mortgage%20Monitor/LPS_Mortgage_Monitor_January_2011.pdf

to 350 days by the end of 2010. The foreclosure process itself, which is in addition to the number of days a loan is delinquent, is estimated to range from 46 to more than 100 days.

Figure 29 illustrates the rise in foreclosure filings in Maryland. Three different bars indicate three stages of the foreclosure process, (i) Notice of Default, (ii) Notice of Foreclosure Sales, and (iii) Foreclosure Sale. The red arrow illustrates the drop in the Notice of Foreclosure Sale filings following the introduction of the emergency bill in the second quarter of 2008. Nevertheless, mortgage defaults continued mounting.

Figure 29: Foreclosure Filings in Maryland, 2006 - 2009



In 2008, when the foreclosure crisis in Maryland significantly intensified, there were over 37,600 foreclosure filings. The numbers continued increasing ever since ranking Maryland today as the 15th most affected foreclosure state. During the study period, there were 100,666 filings. Notices of Default comprised 58 percent of total filings, while Notices of Foreclosure Sales comprised 30 percent. The remaining 12 percent were Foreclosure Sales. The number of Notices of Foreclosure Sales spiked again in the third quarter of 2009;, however there were no specific changes in Maryland's foreclosure processes to account for that change. According to Maryland Department of Housing and Community Development, the increase in notices of foreclosure sales may be due to improvements in the Maryland's real estate market conditions at the end of 2009 (DHCD, 2009).

The foreclosure filings in Maryland show some spatial clustering across the state. Figure 30 maps the number of distressed properties as a share of total distressed properties in the state. It illustrates the relative concentration of foreclosures across the state. A distressed property is one that has received at least one foreclosure filing between 2006 and 2009. A more detailed definition is provided in the Data section. The counties particularly affected by the foreclosure crisis are those in the central part of the state, with Prince George's county in the lead. The other counties with high shares of foreclosures include Montgomery, Frederick, Washington, Howard, Ann Arundel, and Charles counties. In the areas highly affected by the crisis, distressed properties account for up to 0.5 percent of total distressed properties.

Figure 30: Distribution of Distressed Properties in Maryland

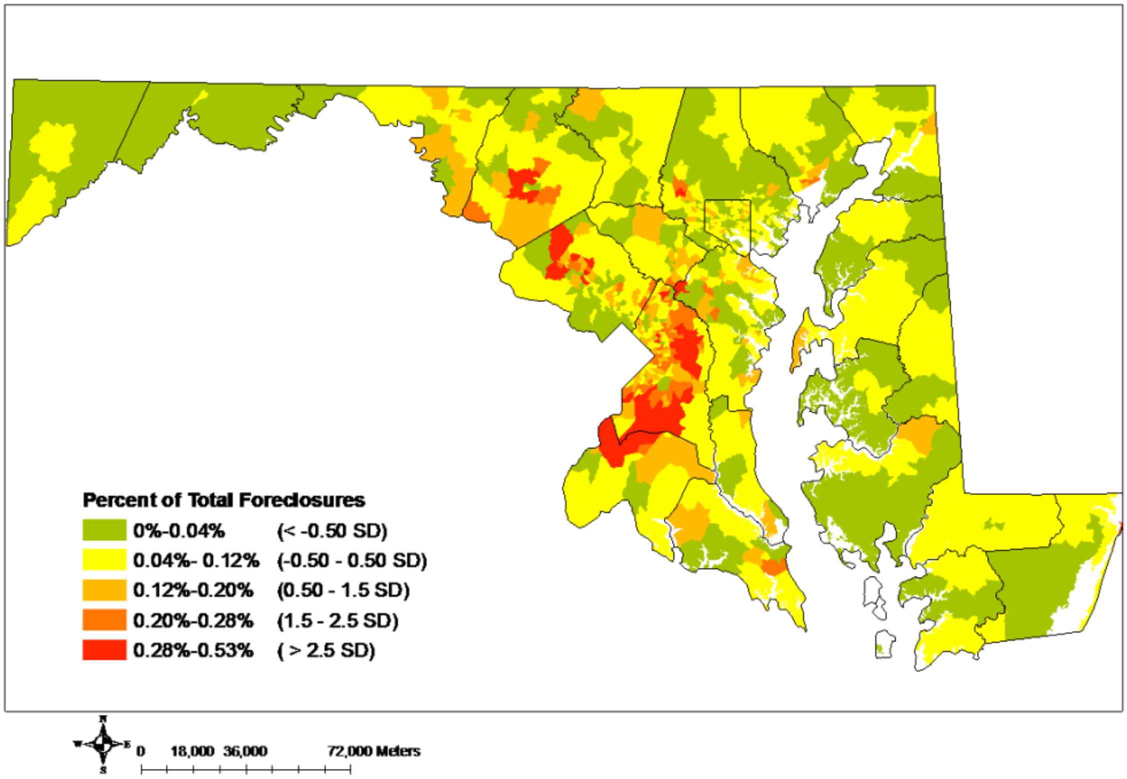
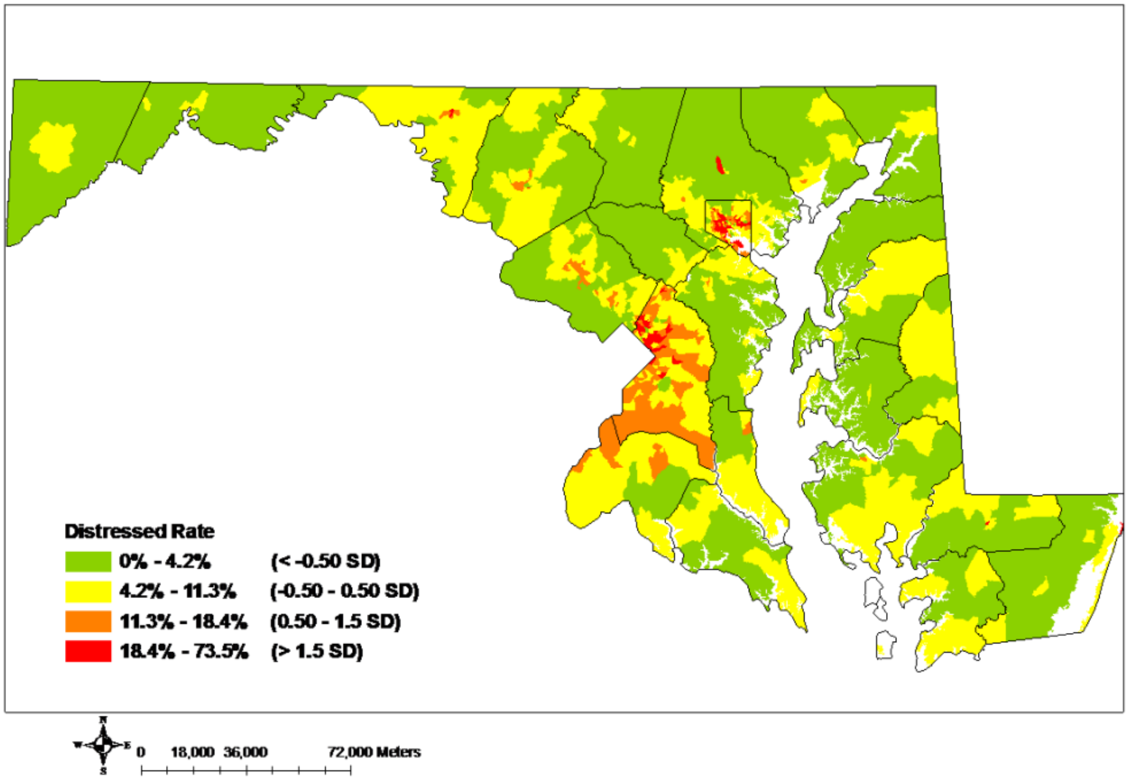


Figure 31 illustrates the rate of distressed properties at the census tract level. Again, the areas with highest rates of distressed properties are in the Prince George's county and Baltimore City, with some pockets of concentration in the other centrally located counties as well. While the median rate of distressed properties per census tract is 5.6 percent, figure 31 illustrates standard deviation distribution by tract. The tracts colored in red have a distressed rate 1.5 standard deviations higher than the mean rate for the whole state and range from 18.4 percent to almost 74 percent.

Figure 31: Distressed Rate across Maryland



Data

The data utilized in the analysis come from several sources which are summarized in Table 9. The data for distressed residential properties is captured in Realtytrac database extending from the first quarter of 2006 until the third quarter 2009. Realtytrac collects foreclosure data from public court records which record one of three activities: 1) a Notice of Default; 2) a Notice of Foreclosure Sale; or 3) a Foreclosure Sale. Since all three stages of the foreclosure process can be recorded for an individual property, the database is cleaned to count one event per property. The properties with at least one foreclosure filing are considered distressed properties. Distressed properties are then aggregated at census tract level. The reason that this analysis accounts for distressed properties instead of only foreclosed or REO properties is because of the prolonged foreclosure process in Maryland discussed in the previous section. Given that the data available tracks information until the third quarter of 2009, it is unclear what happened to defaulted properties after that. Nevertheless, this analysis is not concerned with negative externalities of foreclosed properties; thus it is not critical to know if a property foreclosed or not. According to Maryland emergency foreclosure bill, if a property has at least one filling against it, the borrower has been 90 days or more late on his mortgage. Theory assumes that in a healthy housing market, financially-distressed borrowers can more easily sell their properties or refinance and prepay the remaining balance before seriously defaulting on their mortgage loans (Danis & Pennington-Cross, 2005; Haughwout et al., 2008; Schloemer et al., 2006).

Following the literature on determinants of mortgage default, the first set of variables used in this analysis control for some of the borrowers' characteristics which have historically led to high probability of default: credit scores, loan-to-value ratio (LTV), and debt service-to-income ratio (DTI). The data on credit scores are obtained from Equifax and measure the share of mortgage borrowers who had credit scores less or equal to 639 in 2006. The information on LTV is obtained from the HUD's Neighborhood Stabilization Program (NSP) which was established for the purpose of stabilizing communities that have suffered from high foreclosures and abandonment. To create the scores for NSP neighborhood targeting purposes, HUD used the HMDA database and calculated the share of high cost and highly leveraged mortgages at a census tract originated between 2004 and 2007. In the database, high cost loans are those with a positive annual percentage rate (APR) interest spread of 3 percentage points or more over prevailing Treasury rates at the time of origination. High leverage are "loans with temporary below-market qualifying advantages that pose a risk of payment shock when rates adjust to the fully indexed market rates or payments rise to retire balances from initial negative amortization" (p. 4, HUD, 2008). Based on the loan amount and borrower's gross income reported in HMDA data, HUD has created a method by which to measure the income leverage used by the borrower to obtain the mortgage loan. The method is described in the August 2008 U.S. Housing Market Conditions report (HUD, 2008). The method monitors three ways by which a borrower can increase a loan leverage: (1) by increasing the front-end payment-to-income ratio that determines the size of the payment and associated mortgage allowed, (2) by lowering the interest rate used to calculate the initial qualifying payment and associated mortgage amount, or (3) by reducing the rate at which principal is repaid by extending the term or paying interest only. The two variables developed by HUD, share of high cost and high cost and highly leveraged mortgages, are used as proxies for the extent of subprime lending. Subprime lending peaked between 2004 and 2006. According to Maryland Homeownership Preservation Task Force, between 2000 and 2007, the subprime market share in Maryland increased from 1.5 percent to almost 12 percent of all mortgage loans. "Loans with higher interest rates and "exotic" options that were originated with little to no verification of a borrower's ability to repay made up 60 percent of all foreclosures in Maryland during 2007"[12]. Since mid-2007, only a few subprime loans have been originated. In the two following years, the number of subprime loans outstanding has decreased to 11 percent as loans were cured by default, were prepaid, or were refinanced (Edminston, 2009).

---

[12] Maryland Homeownership Preservation Task Force, November 29, 2007 available at http://www.dhcd.state.md.us/Website/documents/TaskForceReportFinal.pdf

Table 9: Foreclosure Variables and Sources of Data

	Variables:	Source:	Unit of Analysis	Vintage:	Expected Impact:
Dependent	Number of foreclosures in 2006-2009	Realtytrac	Census tract	2006-2009	
	Number of mortgaged properties	ACS	Census tract	2005-2009	+
Subprime Lending	Percent of high cost loans	HMDA	Census tract	2004-2006	+
	Percent of high cost and high leveraged loans				+
	Share of originations that were second liens				+
	Share of loans that were refinances				-
	Share of loans with low credit scores (<639)	Equifax	Census tract	2006	+
Housing Bubble	Change in the number of housing units	Census	Census tract	2000-2007	+
	Increase in home sales from 2002 to 2006	MD Realtors	Zip code	2002-2006	+
	House price appreciation from 2002 to 2006	MD Realtors	Zip code	2002-2006	+
Neighborhood characteristics	Median age of home	Property View	Census tract	2006	+
	Share of households with below median income	Census	Census tract	2000	+
	Percent population age 35 or less				+
	Percent African American population				+
	Percent Hispanic				+
	Percent Asian				+
	Percent renter units				?
	Percent vacancy				?
	Share employed in manufacturing or construction	LEHD Origin-Destination	Census tract	2006	+
	Share employed in finance	Employment Statistics	Census tract	2006	+
Urban Form	Proximity to transit stop (for 0.5, 1, 1.5, and 2 miles)	Tiger Files	Census tract	2000	?
	Job Accessibility by auto	NCSG	SMZ	2007	-
	Job accessibility by transit	NCSG	SMZ	2007	-
	Distance to CBD	Own calculation	Census tract	2000	+
	Median commute time	Census	Census tract	2000	+

Two other subprime lending variables used in the analysis have been shown to have a relationship with rates of foreclosures and both come from the HMDA database. The first variable accounts for the level of leveraging among borrowers and is the number of originations that were second lien mortgages between 2004 and 2006. A body of literature examining the effects of combined loan to value ratio on borrower's likelihood to default showed that piggyback lending was associated with higher default and foreclosure rates during this foreclosure crisis (LaCour-Little et al., 2011; Elul et al, 2010). While second mortgages were often used in combination with other risky features of lending, the use of second mortgage also indicates lower levels of owner equity; thus, this variable is expected to have a positive impact on the number of distressed properties. The second variable is the share of loans that were originated for refinancing. Loans used for home purchase have higher default rates than do refinances, possibly reflecting the fact that those who refinance have longer housing tenure and cannot be first time mortgage borrowers (Chan et al, 2010). The variables used in this analysis observe neighborhood level lending characteristics, but do not account for the individual borrowers. It is expected that all the variables which account for non-prime lending will have a positive impact on the rate on mortgage default. Unfortunately, I am not able to observe from this data the extent to which the borrowers' income and assets have been verified by lenders. I also do not have data on DTI ratios. This type of data is generally available from loan-level mortgage data providers, such as LoanPerformance; however it is not publically available data and can be very costly to access.

The second set of control variables include a series of neighborhood characteristics obtained from the 2000 Decennial Census. These are in line with the studies that showed the impact of neighborhood characteristics on the market dynamics within geographic submarkets. The neighborhood impact is the outcome of households' sorting processes where households with similar socioeconomic and often demographic characteristics sort into neighborhoods that offer similar amenities. Since this study does not question the negative externality of foreclosures on resultant neighborhood change, the use of 2000 Census data allows me to observe neighborhoods before the foreclosure crisis and thus limits the possibility of an endogenous relationship between foreclosures and neighborhood change. The neighborhood characteristics include percent of residents with below median income, percent African American population, percent Hispanic, percent Asian population, percent of units that are vacant, percent of renter-occupied units, age of housing, and age of population. These variables have often been identified in the literature as having some relationship with foreclosure rates. The relationship between income and mortgage default has been well-documented in the literature. High rates of foreclosures are generally associated with low income neighborhoods (Gerardi and Willen 2008) though some research has shown that the link only exists to the degree to which subprime lending has permeated the neighborhood (Mills and Lubuele, 1994). The relationship between race and foreclosure rates has also been well documented with studies generally finding positive relationships between African American or Hispanic populations and foreclosure rates. But similar to the impact of income, after controlling for additional credit information, the relationship between race and foreclosures becomes less clear. I expect in this analysis that lower incomes and higher shares of minority populations are going to have a positive effect on the rate of distressed properties. Because I am not able to observe some of the financial information of the borrowers, such as DTI ratios, and other characteristics of non-prime lending, and given that non-prime lending was primarily concentrated among low income, non-white homeowners, I expect that these variables may pick up some of those unobserved characteristics.

Further, high vacancy rates are generally associated with high levels of foreclosure. And while the relationship can be endogenous, I observe vacancy rates in the year 2000, prior to foreclosure data beginning in 2006. Studies have shown that high vacancy rates lead to depressed property values and neighborhood flight (Immergluck and Smith 2006). Both, in turn, increase foreclosures and further exacerbate low property values and residents' flight. Nevertheless, the areas with high vacancy may be among the drivers of gentrification, particularly if placed in desirable locations (Helms, 2003). Thus, I also account for the change in housing price appreciation from 2002 to 2006. If a neighborhood underwent gentrification, the housing prices will likely appreciate at a higher rate than in neighborhoods that remained plagued with vacancy. This variable is discussed in more detail below.

Another neighborhood variable of interest is owner-occupancy which is generally employed to measure the inverse of investor-owned properties (Edminston, 2009, Chan et al, 2010). The impact of investor-owned properties on foreclosure rates works in several dimensions. First, investors are generally better diversified than owner-occupants and therefore should be better insulated from housing cyclicality. Second though, investors often acquire non-prime loans with typically very low payment options in order to maximize their returns. Thus, as rents fall subsequent to home price declines and investors are not able to sell their properties, investors are more likely to default than homeowners (Brinkmann, 2008). Such a relationship has not always held up with some studies finding higher default rates in primarily owner-occupied neighborhoods (Mayer et al, 2009; Edminston, 2009; Chan et al, 2010). This could happen because owner-occupancy data is self-reported and investors benefit from better mortgage terms if they claim the property as owner-occupied. Based on differing results from previous literature, the expected impact of owner-occupancy on mortgage default rates in this study is ambiguous.

The age of a population is included because some evidence suggests that those younger than 35 have experienced relatively higher rates of default. As affordability deteriorated in the middle of the 2000s, they were more susceptible to targeting for non-prime mortgage products. Additionally, they had less equity built into their homeownership and thus were more vulnerable to economic shocks (Edminston, 2009). Homeownership rates among those aged 35 or less peaked in 2004 at 43.1 percent and fell back to 39.1 percent by 2010. Historically, homeownership in that age category averaged 39.5 percent. For the same reasons shown in Edminston (2009), I expect a higher share of population aged 35 or younger will have a positive impact on the incidence of distressed properties.

The age of housing has also shown some significant relationship with foreclosure rates in previous research although in both directions. The notion is that both older and newer housing stock is more susceptible to riskier lending; the older stock because of higher concentration of minority population; and the newer stock because it is built during the housing boom with proliferation of higher-risk lending practices (Immergluck 2009, 2010). The impact of housing age in this analysis is thus unclear.

With the third set of variables, I account for the housing market bubble. In Maryland, median housing prices increased at an annual rate of almost 15 percent from 2001 to 2006, started declining in 2007, and decreased 25 percent since the peak. The change in housing prices has not played out evenly across the state. I account for housing price appreciation between 2002 and 2006 at a zip code level with data from the Maryland Association of Realtors (MAR). Given that this period observes appreciation before the foreclosure crisis, it aims to distinguish the areas that possibly suffered from a speculative bubble, such as areas on the Eastern Shore or gentrified communities in Baltimore City. Since the MAR captures prices of homes that have been resold during the period, it accounts for change in home prices only for homes that have been resold using a Realtor. Also, the data ranges widely suggested by large difference between the minimum and maximum values in the table of descriptive statistics (see Figure 33). Another source of home price data at lower geography would be the American Community Survey 2005-2009 sample; however that sample overlaps with the period during which home values started to decline which may introduce some endogeneity. Previous research confirmed the strong relationship between default rates of subprime mortgages and house price appreciation (Doms et al, 2007, Gerardi et al, 2008). The increase in housing prices led some borrowers to non-prime and alternative mortgages in search of more affordable payment options. Lenders were comfortable with the increased risk of overextended borrowers since both believed in continued housing appreciation. In fact, Mayer and Pence (2009) showed that areas with high house price appreciation saw growing subprime mortgage originations in the following year. Since housing prices in Maryland did not start falling until the second half of 2007, including a variable which accounts for home price depreciation would overlap with foreclosure data that began in 2006 and again may induce endogeneity between the variables.

Another variable which accounts for the housing bubble is the change in home sales between 2002 and 2006. The areas with more homes sales during those years are expected to a have larger presence of non-prime loans for reasons explained above and may be subject to a speculative bubble. Since some areas are going to have more sales simply because there are more new housing units, I also control for change in the number of housing units between 2000 and 2007. Both of these variables are expected to show a positive relationship with distressed properties. Home sales data also come from the MAR. The 2007 number of housing units at the census tract level comes from the HUD's NSP3 database. The NSP3 program was the 3rd allocation of the NSP grants to communities in need.

Another important macroeconomic factor that played a significant role in the current foreclosure crisis is the widespread increase in unemployment. Generally, when faced with loss of income, homeowners are more likely to default. In the current crisis, falling home prices and increasing unemployment further exacerbated mortgage default. More recent home buyers who were facing loss of income were generally unable to sell their homes because the home prices fell, demand was low, and they were likely facing negative equity. Negative equity occurs when a homeowner's outstanding balance on a home mortgage is more than the value of the home. Unemployment in Maryland spiked in the first quarter of 2009, but since the foreclosure data in this analysis ends in the third quarter of 2009, it is difficult to observe any discernible impact from the changing employment situation. Instead, I control for the share of workers employed in sectors that were significantly impacted in the last economic downturn. Those were jobs in construction, manufacturing, and the financial sector. Shares of unemployment claims by those employed in finance and insurance sectors spiked at 6.4 percent in October of 2007 from the average for the industry of 3.7 percent. The share of construction unemployment claims, though continuously climbing since 2006, peaked at 19.5 percent in September 2009[13]. According to the Bureau of Labor Statistics, Quarterly Census of Employment and Wages, between 2008 and 2009, construction jobs were down 15.2 percent,

[13] U.S. Department of Labor

manufacturing jobs were down 8.5 percent and financial activities were down 5.4 percent[14]. The other industries had less sizable losses. I expect that a higher share of employed in those sectors per census tract will have a positive impact on the number of distressed properties.

Finally, the primary construct of interest in this analysis is the group of variables which control for urban form. I examine several important urban form measures, such as proximity to transit, job accessibility by car and by transit, distance to the central business district, and commute time. These variables are intended to test the arguments that urban areas have resurged in their importance and that the demographic and cultural shift away from dispersed and suburban living is in fact a long-term underlying trend. If so, areas further away from the urban core and poor accessibility to transit and employment will show higher rates of distressed properties because lack of demand and depreciated home values will leave homeowners with fewer options in times of economic shocks.

Proximity to transit are binary variables which indicate if a census tract centroid is located within a half a mile, one mile, one and a half mile, or two miles from a fixed-rail transit station. The proximity to bus stops is not accounted for as that variable alone is not expected to have a significant impact on household location decisions, however it does affect households' transportation costs. The variables for job accessibility by auto and transit indicate the number of jobs by car or transit, respectively, which are within 30 minutes travel time from the centroid of a Statewide Model Zone (SMZ). The description of SMZs is included in Appendix C. Generally and to the extent possible, SMZs conform to census tract geography to best utilize census data products. Job accessibility data at SMZ level were linked to census tracts via spatial join in ArcGIS based on SMZ's

[14] http://beta.bls.gov/maps/cew/MD

centroid falling inside a census tract boundary. Jobs are classified into four categories: industrial, retail, office, other, and all jobs. Figure 32 illustrates, for example, the total number of jobs accessible in 30 minutes by automobile. Each census tract has at least 146 jobs accessible within a 30 minute drive. As anticipated, areas closer to the urban centers will have greater accessibility to jobs than rural areas in the western and eastern part of the state. Transit proximity and job accessibility measures are all expected to have a negative relationship with levels of distress. Figure 33 illustrates job accessibility to all types of employment by transit.

Figure 32: Job Accessibility by Auto in Maryland

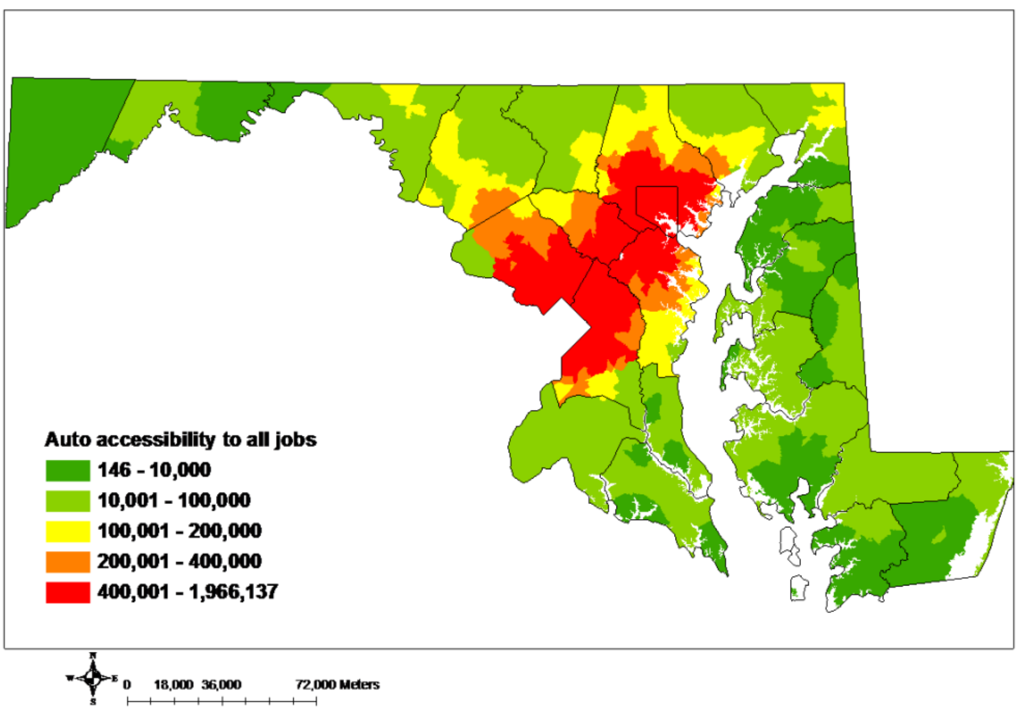
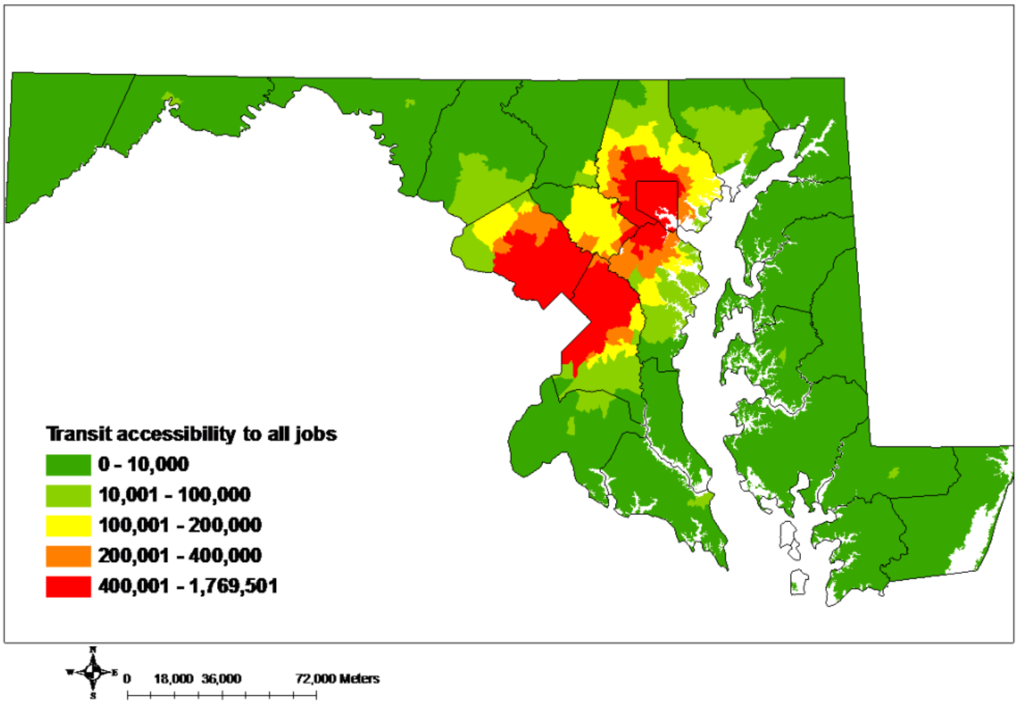


Figure 33: Job Accessibility by Transit in Maryland



The variable controlling for proximity to the CBSA center measures the linear distance from a census tract centroid to the nearest mean housing-weighted centroids of seven CBSAs: Cambridge, Cumberland, Easton, Hagerstown, Lexington Park, Ocean Pines, and Salisbury. For census tracts nearest to Washington DC or Baltimore City CBSA distance is measured to the centroid of the census tracts where the Central Business District (CBD) was located according to the 1982 Census of Retail Trade. With Washington and Baltimore CBSAs, the traditional CBDs are hypothesized to have a greater impact on location decisions than housing-weighted centers of less urban CBSAs because the traditional center provides the amenities at question, such as access to jobs, restaurants, and other cultural activities. If urban centers do have an impact on foreclosures, distance from the center variable is expected to show a positive relationship with levels of distress.

Another variable tested in this analysis is the commute time of auto-commuting residents. This variable is obtained from the 2000 Census. It is intended to measure the proximity of neighborhoods to major employment centers. Other analyses have shown that commuting was not a significant factor is accumulation of foreclosures (Immergluck, 2010). However, if the hypothesis of greater importance of urban areas holds, commute time will have a positive impact on levels of distressed properties.

Summary statistics of all variables are presented in Figure 34. The summary table contains mean values, standard deviation and minimum and maximum values.

Figure 34: Summary Statistics

	Mean	Std. Dev.	Min	Max
n	1212			
Distressed properties (#)	59	57	0	385
Mortgaged properties (#)	922	660	0	5173
Share of high cost originations	0.226	0.134	0	1
Share high cost and high leverage	0.124	0.074	0	1
Share low credit score	0.118	0.104	0	0.667
Share second liens	0.141	0.040	0	0.5
Share refinances	0.580	0.110	0	1
Housing units increase (2000-2007)	0.082	0.558	-0.182	19.2
Increase in home sales (2002-2006)	1.147	6.011	-1	102
Change in median home value (2002-2006)	0.775	0.384	-0.568	2.620
Age of home (=2006-median year built) (#)	43	24	0	131
Share of households with below median income	0.438	0.205	0.000	1.000
Share of population age 35 or less	0.310	0.092	0.069	0.989
Share African-American	0.298	0.322	0.001	0.991
Share Hispanic	0.041	0.069	0.000	0.799
Share Asian	0.035	0.047	0.000	0.482
Vacancy rate	0.070	0.076	0.000	0.857
Share renter-occupied	0.294	0.211	0.000	0.974
Share employed in manufacturing or construction	0.127	0.047	0.043	0.381
Share employed in finance	0.045	0.014	0.000	0.097
Distance to CBD (in meters)	18,017	11,105	365	49,998
Median commute time (minutes)	25.5	6.4	0	45.4
CT centroid within 1/2 mile of transit (0/1)	0.070	0.255	0	1
CT centroid within 1 mile of transit (0/1)	0.204	0.403	0	1
CT centroid within 1.5 miles of transit (0/1)	0.304	0.460	0	1
CT centroid within 2 miles of transit (0/1)	0.399	0.490	0	1

Retail employment accessibility by auto (#)	102,836	73,710	14	285,933
Office employment accessibility by auto (#)	320,200	266,215	0	1,166,101
Industry employment accessibility by auto (#)	64,636	46,366	15	170,923
Other employment accessibility by auto (#)	181,576	127,089	28	454,905
All employment accessibility by auto (#)	669,248	505,265	146	1,966,137
Retail employment accessibility by transit (#)	69,488	70,005	0	282,815
Office employment accessibility by transit (#)	235,232	247,980	0	976,486
Industry employment accessibility by transit (#)	40,804	40,491	0	152,162
All employment accessibility by transit (#)	466,606	468,404	0	1,795,652

Empirical Strategy

To determine the impact of urban form on spatial distribution of distressed properties, I estimate a negative binomial model. A Negative binomial model is generally used for count data when the dependent variable is highly positively skewed. Count variables have certain properties: i) they are never negative; ii) they are integers; and iii) they have tendency to be positively skewed. They are accordingly described via Poisson distribution. Following the method employed by Immergluck (2010), the dependent variable in this negative binomial model is the count of distressed properties by census tract. Immergluck's study focuses on the accumulation of REOs so the author accounts for REO properties exclusively. Using a count model rather than a rate model in this analysis makes more sense intuitively and theoretically. To get at a foreclosure rate, one needs a denominator for the dependent variable. However, there has been some debate in the literature over which denominator is best suited for the foreclosure analyses. Some have used the number of mortgaged properties; some used the number of all residential properties, while others used the number of mortgageable properties. In the analyses where the focus is on the impact of foreclosures, using the foreclosure rate may be more appropriate. However, this analysis is concerned with which factors led to the rise in the number of distressed properties.

The negative binomial regression model is a direct extension of the Poisson model. In Poisson distribution, the variance is expected to equal the mean. Often times, the variance exceeds the mean which is called over-dispersion. The negative binomial models allows for over-dispersion by including a dispersion parameter and allowing for independent specification of the mean and variance. Since the only difference between the Poisson and the negative binomial is in their variances, regression coefficients are often similar, but with different standard errors. In case of an over-dispersed variable, standard errors from the negative binomial model will be larger than those in Poisson regression. Thus using a Poisson regression in case of an over-dispersed variable can lead to low p-values and narrow confidence intervals and can possibly lead to erroneous conclusion of the significance of a parameter. One can test for over-dispersion via the alpha (dispersion) parameter, where a one-tailed z-test of H_0: alpha = 0 indicates when the negative binomial or the Poisson regression model is more appropriate. When alpha is zero, the negative binomial model reduces to the Poisson model (Long and Freese 2006). Additionally, the negative binomial model is estimated via maximum likelihood estimation strategy. Maximum likelihood functions attempt to find the value of regression coefficients that have most likely given rise to the number of distressed properties. The negative binomial model is described via the following form:

$$DP_{(2006\text{-}2009)} = \exp\,[\alpha + \beta \ln(\mathbf{M}) + \gamma \mathbf{S} + \delta \mathbf{B} + \eta \mathbf{N} + \varphi \mathbf{U} + \boldsymbol{\varepsilon}] \tag{1}$$

where $DP_{(2006\text{-}2009)}$ is the count of distressed properties between the first quarter of 2006 and third quarter of 2009 in a census tract. The summary of the dependent variable indicates that the variance is 56 times the size of the mean which suggests that the dependent variable is in fact over-dispersed. The test of the significance of the alpha parameter is provided in the table of results.

M is the estimated average number of mortgaged properties in a census tract between 2005 and 2009 and controls for the total size of the mortgaged inventory in a tract. It is alternatively called an exposure parameter because it constrains the dependent variables to the maximum number of properties that can be distressed during that time period.

S is a vector of control variables which describe the level of subprime lending by census tract. There are five subprime variables: the share of high cost originations, the share of high cost and high leveraged originations, the share of loans with low credit scores, the percent of originations that were second lien mortgages, and the share of loans that were originated for refinancing.

B is a vector of control variables that describe the extent of the housing bubble. There are three bubble variables which include: the increase in the number of housing units between 2000 and 2007 at census tract, the increase in home sales between 2002 and 2006, and the increase in the median home value between 2002 and 2006. The last two variables are generated at the zip code level.

N is a vector of neighborhood control variables at the census tract level. Those variables include median age of homes, percent of households with below median income, percent of population 35 or younger, the share of population that is African-American, the share of population that is Hispanic, the share of population that is Asian, the share of homes that are vacant, and share of homes that are renter occupied. In this category, I also include the variables which control for neighborhood employment characteristics. This includes the share of workers that are employed in construction and manufacturing, or finance.

Finally, **U** is a vector of variables that control for urban structure. These variables include proximity to transit stop, employment accessibility by auto and transit, distance to the central business district, and median commute time.

Results

The results of the negative binomial model are presented in Tables 10, 11, and 12. I estimated three sets of models: 1) for all census tracts in Maryland, 2) for census tracts located in Washington, DC metropolitan area, and 3) those located in the Baltimore metropolitan area. For each of the three sets of models, I first examine the variation in distressed properties using only subprime lending factors. The second model includes the housing bubble effects. The third model includes neighborhood characteristics, while the last set of models examines the impact of urban form characteristics on the level of distressed properties. The tables with results contain untransformed coefficients, p-values, and goodness-of-fit values. All models are estimated with a heteroskedasticity robust estimator in accordance with the Huber-White-Sandwich procedure. With robust estimators, the likelihood ratio chi-square test becomes a Wald chi-square and is based on log pseudo-likelihoods. The Wald test statistic provides a test for the hypothesis that all coefficients in the model except the intercept term are simultaneously equal to zero. In all models, the highly significant Wald test strongly rejects the hypothesis the coefficients are zero. The likelihood-ratio (LP) test for nesting of the models indicates that those models are nested. In other words, the LR tests if coefficients of the variables excluded from the full model are equal to zero. The significant value of LR test rejects the hypothesis that those excluded variables have coefficients equal to zero.

The Bayesian information criterion (BIC) proposed by Raftery (1996) is a measure of overall goodness of fit and a means to compare nested and non-nested models. The more negative the BIC, the better the fit. Raftery (1996) suggested guidelines for the strength of evidence based on a difference between BIC values between two models. When the absolute difference is from 0 to 2, the evidence is weak;, if the difference is between 2 and 6, the evidence is positive;, if between 6 and 10, the evidence is strong;, and absolute difference greater than 10 means the evidence of a better model is very strong. The BIC statistics reported for all three sets of models indicate that the model is improving with the addition of new control variables. For all three models, the absolute values of the difference between the BIC for constrained and fully specified models are well above 100. Another measure of fit is Pseudo R^2 or in this case the McFadden's R^2, also known as the likelihood-ratio index, which compares a model with just the intercept to a model with all parameters. I report adjusted McFadden R^2 for all three sets of models. The increasing value of the adjusted R^2 again suggests that the model is performing increasingly better with the additional variables. To check for multicollinearity among the explanatory variables, I constructed the OLS model by generating a log of the dependent variable and running a simple regression using all explanatory variables, followed by the variance inflation factor (VIF). Generally, a variable whose VIF values is greater than 10 shows collinearity with other variables and can be represented as a linear combination of other independent variables. The VIF values show that there is generally no serious issue with multicollinearity among the variables. The two variables that show increased collinearity are the share of high cost loans and the share of high cost and high leverage loans in Washington metropolitan area sample. The two variables however do not show high collinearity for the rest of the sample. The VIF for the income variable is elevated; however that is expected as income is generally correlated with lending characteristics, type of employment, tenure status, and race.

Table 10: Negative Binomial Results for Maryland

		Dependent: number of distressed properties from 2006 to 2009 in Maryland																		
		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14	Model 15	Model 16	Model 17	Model 18	Model 19
Subprime lending	ln(mortgaged properties)	0.908***	0.912***	1.005***	1.034***	1.032***	1.031***	1.037***	1.031***	1.031***	1.031***	1.024***	1.025***	1.026***	1.024***	1.023***	1.020***	1.028***	1.022***	1.021***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share high cost	3.499***	3.432***	2.832***	2.819***	2.845***	2.791***	2.870***	2.842***	2.836***	2.842***	2.747***	2.746***	2.778***	2.748***	2.800***	2.757***	2.831***	2.838***	2.792***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share high cost and high leverage	3.345***	3.482***	2.655***	2.315***	2.291***	2.431***	2.297***	2.289***	2.290***	2.294***	2.493***	2.505***	2.462***	2.508***	2.458***	2.518***	2.368***	2.423***	2.481***
		(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	Share low FICO	0.864**	0.838*	0.718**	0.416	0.467	0.414	0.493	0.467	0.473	0.470	0.451	0.449	0.450	0.449	0.468	0.462	0.462	0.464	0.463
		(0.048)	(0.055)	(0.028)	(0.193)	(0.143)	(0.192)	(0.122)	(0.143)	(0.140)	(0.140)	(0.160)	(0.163)	(0.162)	(0.163)	(0.146)	(0.152)	(0.150)	(0.151)	(0.151)
	Share second liens	5.633***	5.481***	6.830***	6.459***	6.335***	6.165***	6.287***	6.339***	6.345***	6.338***	6.161***	6.144***	6.211***	6.153***	6.213***	6.035***	6.302***	6.265***	6.150***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share refinances	-2.538***	-2.542***	-1.362***	-1.720***	-1.756***	-1.853***	-1.754***	-1.758***	-1.767***	-1.760***	-1.807***	-1.811***	-1.802***	-1.809***	-1.809***	-1.832***	-1.781***	-1.803***	-1.820***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Housing bubble	Housing units increase (2000-2007)		0.155***	0.160***	0.181***	0.176***	0.176***	0.175***	0.176***	0.175***	0.176***	0.173***	0.175***	0.174***	0.174***	0.174***	0.173***	0.175***	0.172***	0.173***
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Increase in home sales (2002-2006)		0.006*	0.004*	0.003**	0.003*	0.003*	0.003*	0.003*	0.003*	0.003*	0.003*	0.003*	0.003*	0.003*	0.003*	0.003*	0.003*	0.003*	0.003*
			(0.088)	(0.056)	(0.048)	(0.073)	(0.052)	(0.081)	(0.070)	(0.071)	(0.073)	(0.082)	(0.088)	(0.084)	(0.086)	(0.083)	(0.098)	(0.081)	(0.083)	(0.091)
	Change in median home value (2002-2006)		0.013	0.023	0.024	0.023	0.026	0.025	0.023	0.024	0.023	0.027	0.027	0.027	0.028	0.028	0.029	0.026	0.031	0.030
			(0.795)	(0.604)	(0.589)	(0.599)	(0.550)	(0.572)	(0.599)	(0.592)	(0.596)	(0.547)	(0.546)	(0.548)	(0.535)	(0.529)	(0.509)	(0.553)	(0.488)	(0.501)
Neighborhood characteristics	Age of home			0.003***	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***	0.006***	0.005***	0.005***	0.006***
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Percent of households with below median income			0.364*	-0.186	-0.108	0.020	-0.115	-0.110	-0.110	-0.109	-0.178	-0.174	-0.168	-0.172	-0.173	-0.205	-0.140	-0.149	-0.183
				(0.056)	(0.321)	(0.569)	(0.917)	(0.542)	(0.563)	(0.562)	(0.566)	(0.359)	(0.369)	(0.388)	(0.373)	(0.376)	(0.290)	(0.476)	(0.440)	(0.346)
	Percent populations age <=35			0.568*	0.576**	0.562**	0.572**	0.563**	0.561**	0.563**	0.564**	0.539*	0.530*	0.555*	0.537*	0.547*	0.536*	0.561**	0.579**	0.553*
				(0.066)	(0.037)	(0.040)	(0.037)	(0.041)	(0.041)	(0.041)	(0.039)	(0.064)	(0.069)	(0.052)	(0.065)	(0.053)	(0.061)	(0.043)	(0.039)	(0.051)
	Share African-American			0.283***	0.689***	0.674***	0.605***	0.654***	0.677***	0.682***	0.676***	0.733***	0.733***	0.725***	0.735***	0.712***	0.734***	0.695***	0.708***	0.723***
				(0.008)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share Hispanic			1.257***	0.922***	0.949***	0.828***	0.947***	0.952***	0.969***	0.955***	1.152***	1.132***	1.084***	1.139***	1.076***	1.162***	0.991***	1.048***	1.108***
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share Asian			0.974***	1.872***	2.131***	2.058***	2.124***	2.132***	2.138***	2.135***	2.208***	2.216***	2.196***	2.207***	2.248***	2.297***	2.179***	2.229***	2.268***
				(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Vacancy rate			1.865***	1.995***	1.994***	1.950***	1.972***	1.997***	1.998***	1.994***	1.913***	1.921***	1.931***	1.913***	1.919***	1.887***	1.961***	1.910***	1.899***
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share rent occupied			-0.351**	-0.094	-0.141	-0.153	-0.140	-0.140	-0.140	-0.142	-0.112	-0.116	-0.117	-0.116	-0.129	-0.116	-0.133	-0.143	-0.126
				(0.017)	(0.498)	(0.308)	(0.266)	(0.311)	(0.317)	(0.312)	(0.307)	(0.429)	(0.412)	(0.405)	(0.410)	(0.356)	(0.409)	(0.340)	(0.303)	(0.366)
Employment	Share employed in manufacturing or construction				3.638***	3.427***	3.379***	3.405***	3.430***	3.424***	3.423***	3.261***	3.271***	3.297***	3.250***	3.217***	3.163***	3.352***	3.204***	3.185***
					(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share employed in finance				0.002	0.286	0.360	0.141	0.294	0.326	0.297	-0.046	-0.028	0.004	-0.022	0.221	0.174	0.239	0.344	0.231
					(0.998)	(0.789)	(0.735)	(0.896)	(0.783)	(0.760)	(0.782)	(0.965)	(0.979)	(0.997)	(0.983)	(0.836)	(0.869)	(0.822)	(0.746)	(0.828)
Urban form	Distance to CBD					0.033**	0.028**	0.035**	0.033**	0.033**	0.033**	0.030**	0.032**	0.031**	0.030**	0.029**	0.028**	0.032**	0.030**	0.029**
						(0.015)	(0.048)	(0.011)	(0.016)	(0.015)	(0.015)	(0.027)	(0.020)	(0.024)	(0.025)	(0.030)	(0.041)	(0.020)	(0.026)	(0.034)
	Median commute time						0.007**													
							(0.015)													
	Proximity to Transit							0.5 mile buffer	1 mile buffer	1.5 mile buffer	2 mile buffer									
								0.078	-0.006	-0.018	-0.006									
								(0.137)	(0.858)	(0.507)	(0.827)									
	Employment Accessibility by transit											Retail	Office	Industry	Total Emp					
												-0.007***	-0.002***	-0.010**	-0.001***					
												(0.002)	(0.004)	(0.015)	(0.003)					
	Employment Accessibility by auto															Retail	Office	Industry	Other	Tot Emp
																-0.006**	-0.002***	-0.004	-0.003***	-0.001***
																(0.014)	(0.000)	(0.240)	(0.008)	(0.001)
	_cons	-2.820***	-2.847***	-4.771***	-5.131***	-5.171***	-5.281***	-5.200***	-5.168***	-5.162***	-5.166***	-4.997***	-5.003***	-5.037***	-4.998***	-4.991***	-4.910***	-5.098***	-5.020***	-4.959***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	/lnalpha	-1.922***	-1.934***	-2.198***	-2.294***	-2.302***	-2.313***	-2.305***	-2.303***	-2.303***	-2.302***	-2.309***	-2.308***	-2.306***	-2.309***	-2.305***	-2.313***	-2.302***	-2.307***	-2.309***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Log-Likelihood	-5,161.54	-5,154.29	-5,034.96	-4,988.70	-4,985.20	-4,980.86	-4,983.80	-4,985.18	-4,984.99	-4,985.17	-4,979.80	-4,980.45	-4,981.64	-4,979.83	-4,981.62	-4,976.84	-4,984.35	-4,980.97	-4,979.34
	chi2	3,126.333	3,263.374	4,419.135	4,929.813	5,014.386	5,018.720	5,013.584	5,016.213	5,060.617	5,022.227	5,204.143	5,179.099	5,156.080	5,203.591	5,177.160	5,318.748	5,074.585	5,224.490	5,255.019
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	McFadden's Adj R2	0.16	0.161	0.179	0.186	0.186	0.187	0.186	0.186	0.186	0.186	0.187	0.187	0.187	0.187	0.187	0.188	0.186	0.187	0.187
	BIC':	-1942.381	-1935.576	-2117.433	-2195.77	-2195.669	-2197.251	-2191.368	-2188.605	-2188.989	-2188.614	-2199.366	-2198.057	-2195.683	-2199.3	-2195.718	-2205.286	-2190.253	-2197.019	-2200.277
	N	1,212	1,212	1,212	1,212	1,212	1,212	1,212	1,212	1,212	1,212	1,212	1,212	1,212	1,212	1,212	1,212	1,212	1,212	1,212
	note: *** p<0.01, ** p<0.05, * p<0.1																			

Table 11: Negative Binomial Results for Washington Metropolitan Area

		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14	Model 15	Model 16	Model 17	Model 18	Model 19
		Dependent: number of distressed properties from 2006 to 2009 if Washington metro=1																		
Subprime lending	ln(mortgaged properties)	0.955***	0.960***	1.030***	1.052***	1.020***	1.020***	1.022***	1.021***	1.019***	1.020***	1.023***	1.022***	1.022***	1.023***	1.021***	1.022***	1.022***	1.022***	1.022***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share high cost	5.406***	5.408***	4.983***	5.161***	4.998***	5.006***	4.992***	5.006***	4.913***	4.962***	5.048***	5.063***	5.080***	5.067***	4.996***	4.981***	5.059***	5.012***	5.007***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share high cost and high leverage	0.806	0.369	0.167	-0.115	0.035	0.031	0.044	0.033	0.078	0.069	-0.154	-0.124	-0.096	-0.137	0.052	0.019	0.012	0.054	0.035
		(0.640)	(0.828)	(0.879)	(0.915)	(0.975)	(0.978)	(0.968)	(0.977)	(0.945)	(0.952)	(0.887)	(0.910)	(0.931)	(0.900)	(0.961)	(0.985)	(0.992)	(0.961)	(0.974)
	Share low FICO	-0.277	-0.379	-0.846	-0.830	-0.930*	-0.937*	-0.925*	-0.931*	-0.911*	-0.920*	-0.992*	-0.990*	-1.004**	-0.999**	-1.089**	-1.114**	-1.090**	-1.090**	-1.143**
		(0.776)	(0.685)	(0.161)	(0.140)	(0.078)	(0.079)	(0.080)	(0.078)	(0.086)	(0.080)	(0.050)	(0.051)	(0.050)	(0.048)	(0.033)	(0.025)	(0.035)	(0.035)	(0.023)
	Share second liens	7.543***	7.152***	5.640***	5.515***	4.908***	4.911***	4.906***	4.914***	4.869***	4.899***	4.638***	4.602***	4.723***	4.637***	4.397***	3.983***	4.729***	4.758***	4.298***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share refinances	-1.104*	-1.152*	-0.415	-0.860*	-1.285**	-1.290**	-1.274**	-1.277**	-1.361***	-1.303**	-1.339***	-1.339***	-1.323***	-1.333***	-1.384***	-1.432***	-1.320***	-1.325***	-1.393***
		(0.087)	(0.078)	(0.393)	(0.075)	(0.013)	(0.013)	(0.015)	(0.013)	(0.008)	(0.013)	(0.007)	(0.007)	(0.009)	(0.007)	(0.005)	(0.003)	(0.010)	(0.009)	(0.004)
Housing bubble	Housing units increase (2000-2007)		0.218	0.406**	0.364**	0.453***	0.455***	0.449***	0.452***	0.442***	0.445***	0.417***	0.441***	0.439***	0.435***	0.460***	0.454***	0.466***	0.469***	0.465***
			(0.278)	(0.011)	(0.024)	(0.007)	(0.007)	(0.008)	(0.007)	(0.008)	(0.007)	(0.008)	(0.006)	(0.006)	(0.006)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)
	Increase in home sales (2002-2006)		0.015	0.074	0.048	0.056	0.057	0.056	0.055	0.057	0.056	0.074*	0.072*	0.068*	0.073*	0.069*	0.071*	0.062	0.066	0.071*
			(0.782)	(0.108)	(0.274)	(0.160)	(0.147)	(0.160)	(0.163)	(0.143)	(0.157)	(0.067)	(0.076)	(0.093)	(0.073)	(0.084)	(0.074)	(0.121)	(0.105)	(0.078)
	Change in median home value (2002-2006)		0.320***	0.282***	0.215***	0.254***	0.252***	0.254***	0.254***	0.256***	0.255***	0.228***	0.232***	0.242***	0.231***	0.241***	0.237***	0.253***	0.256***	0.245***
			(0.002)	(0.001)	(0.005)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Neighborhood characteristics	Age of home			-0.001	0.000	0.003*	0.003*	0.003*	0.003	0.004*	0.004*	0.004**	0.004**	0.004**	0.004**	0.005**	0.005**	0.004**	0.004**	0.005**
				(0.801)	(0.878)	(0.095)	(0.093)	(0.097)	(0.102)	(0.054)	(0.083)	(0.031)	(0.034)	(0.048)	(0.033)	(0.020)	(0.012)	(0.048)	(0.047)	(0.017)
	Percent of households with below median income			1.223***	0.574*	0.289	0.297	0.288	0.288	0.317	0.307	0.319	0.301	0.291	0.313	0.314	0.298	0.288	0.309	0.306
				(0.000)	(0.052)	(0.361)	(0.354)	(0.362)	(0.362)	(0.314)	(0.335)	(0.299)	(0.328)	(0.346)	(0.309)	(0.293)	(0.296)	(0.348)	(0.317)	(0.297)
	Percent populations age <=35			1.616***	1.391***	1.304***	1.304***	1.306***	1.302***	1.336***	1.322***	1.314***	1.294***	1.339***	1.308***	1.284***	1.310***	1.308***	1.358***	1.332***
				(0.000)	(0.000)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	Share African-American			0.084	0.426**	0.737***	0.734***	0.734***	0.735***	0.759***	0.738***	0.775***	0.773***	0.767***	0.775***	0.787***	0.789***	0.778***	0.769***	0.791***
				(0.570)	(0.014)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share Hispanic			0.793***	0.591**	1.106***	1.101***	1.108***	1.108***	1.114***	1.099***	1.106***	1.101***	1.078***	1.094***	1.128***	1.169***	1.077***	1.082***	1.121***
				(0.002)	(0.020)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share Asian			1.263***	1.823***	2.276***	2.282***	2.274***	2.274***	2.277***	2.280***	2.244***	2.240***	2.256***	2.244***	2.283***	2.244***	2.254***	2.284***	2.262***
				(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Vacancy rate			-0.408	-0.437	-0.964*	-0.971*	-0.959*	-0.966*	-0.942*	-0.962*	-0.894*	-0.913*	-0.938*	-0.911*	-0.910*	-0.809	-0.984*	-0.992*	-0.901*
				(0.518)	(0.496)	(0.056)	(0.053)	(0.056)	(0.055)	(0.067)	(0.057)	(0.085)	(0.074)	(0.066)	(0.076)	(0.085)	(0.121)	(0.054)	(0.052)	(0.084)
	Share rent occupied			-0.790***	-0.457**	-0.428**	-0.430**	-0.429**	-0.429**	-0.437**	-0.432**	-0.387*	-0.390**	-0.394**	-0.389**	-0.403**	-0.398**	-0.395**	-0.417**	-0.399**
				(0.000)	(0.018)	(0.033)	(0.032)	(0.032)	(0.032)	(0.030)	(0.032)	(0.051)	(0.049)	(0.047)	(0.049)	(0.039)	(0.038)	(0.045)	(0.035)	(0.040)
Employment	Share employed in manufacturing or construction				3.305***	1.585**	1.581**	1.571**	1.579**	1.616**	1.588**	2.115***	2.029***	1.957***	2.068***	1.824**	1.943***	1.796**	1.822**	1.939***
					(0.000)	(0.032)	(0.033)	(0.034)	(0.033)	(0.027)	(0.032)	(0.005)	(0.007)	(0.009)	(0.006)	(0.014)	(0.008)	(0.017)	(0.016)	(0.009)
	Share employed in finance				0.867	-2.784	-2.665	-2.829	-2.793	-2.563	-2.651	-1.753	-1.830	-2.106	-1.852	-0.930	-0.584	-1.820	-1.942	-0.914
					(0.726)	(0.164)	(0.207)	(0.162)	(0.163)	(0.199)	(0.193)	(0.374)	(0.357)	(0.288)	(0.349)	(0.649)	(0.769)	(0.369)	(0.334)	(0.650)
Urban form	Distance to CBD					0.102***	0.103***	0.103***	0.103***	0.100***	0.101***	0.061**	0.065***	0.075***	0.064***	0.052**	0.045*	0.076***	0.078***	0.052**
						(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.011)	(0.008)	(0.001)	(0.009)	(0.040)	(0.058)	(0.002)	(0.001)	(0.035)
	Median commute time						0.001													
							(0.857)													
	Proximity to Transit							0.5 mile buffer	1 mile buffer	1.5 mile buffer	2 mile buffer									
								0.024	0.007	-0.058*	-0.020									
								(0.728)	(0.831)	(0.077)	(0.562)									
	Employment Accessibility by transit											Retail	Office	Industry	Total Emp					
												-0.010***	-0.003***	-0.012**	-0.001***					
												(0.002)	(0.009)	(0.031)	(0.007)					
	Employment Accessibility by auto															Retail	Office	Industry	Other	Tot Emp
																-0.013***	-0.004***	-0.011*	-0.004*	-0.002***
																(0.004)	(0.000)	(0.071)	(0.085)	(0.002)
	_cons	-4.101***	-4.256***	-5.589***	-5.768***	-5.279***	-5.309***	-5.296***	-5.286***	-5.239***	-5.270***	-5.123***	-5.120***	-5.194***	-5.138***	-4.989***	-4.912***	-5.181***	-5.211***	-5.016***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	/lnalpha	-2.351***	-2.418***	-2.743***	-2.820***	-2.914***	-2.914***	-2.914***	-2.914***	-2.924***	-2.915***	-2.936***	-2.927***	-2.921***	-2.929***	-2.926***	-2.947***	-2.911***	-2.913***	-2.930***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Log-Likelihood	-1,946.65	-1,936.23	-1,878.26	-1,865.19	-1,845.52	-1,845.50	-1,845.43	-1,845.50	-1,844.08	-1,845.36	-1,839.94	-1,841.03	-1,842.92	-1,840.90	-1,839.47	-1,834.61	-1,843.39	-1,843.53	-1,838.24
	chi2	1,503.103	1,924.807	2,904.619	3,158.673	4,153.274	4,145.886	4,161.723	4,155.114	4,181.146	4,166.908	4,220.530	4,176.997	4,189.111	4,186.239	4,320.661	4,388.651	4,251.784	4,239.580	4,328.013
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	McFadden's Adj R2	0.175	0.178	0.199	0.204	0.212	0.211	0.211	0.211	0.212	0.211	0.214	0.211	0.213	0.213	0.214	0.216	0.212	0.212	0.214
	BIC':	-808.323	-810.956	-878.347	-892.352	-925.621	-919.594	-919.728	-919.592	-922.44	-919.871	-930.719	-928.539	-924.753	-928.792	-931.658	-941.375	-923.813	-923.523	-934.121
	N	432	432	432	432	432	432	432	432	432	432	432	432	432	432	432	432	432	432	432
	note: *** p<0.01, ** p<0.05, * p<0.1																			

Table 12: Negative Binomial Results for Baltimore Metropolitan Area

		Dependent: number of distressed properties from 2006 to 2009 if Baltimore metro=1																		
		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14	Model 15	Model 16	Model 17	Model 18	Model 19
Subprime lending	ln(mortgaged properties)	0.874***	0.878***	0.992***	1.004***	0.999***	0.993***	1.006***	0.998***	0.998***	0.999***	0.996***	0.993***	0.996***	0.994***	1.007***	1.006***	1.007***	1.005***	1.006***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share high cost	3.522***	3.343***	2.768***	2.551***	2.571***	2.559***	2.670***	2.575***	2.583***	2.580***	2.565***	2.553***	2.561***	2.558***	2.541***	2.553***	2.524***	2.540***	2.544***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share high cost and high leverage	1.457*	1.855**	1.803**	1.929***	2.040***	2.120***	2.042***	2.046***	2.060***	2.042***	2.098***	2.176***	2.109***	2.159***	1.955***	1.962***	1.991***	1.983***	1.967***
		(0.051)	(0.017)	(0.019)	(0.009)	(0.007)	(0.005)	(0.006)	(0.006)	(0.006)	(0.007)	(0.005)	(0.004)	(0.005)	(0.004)	(0.009)	(0.009)	(0.008)	(0.008)	(0.009)
	Share low FICO	1.117**	1.140**	0.741**	0.412	0.429	0.421	0.479	0.429	0.412	0.406	0.434	0.438	0.436	0.437	0.422	0.420	0.428	0.425	0.422
		(0.032)	(0.026)	(0.031)	(0.239)	(0.220)	(0.227)	(0.168)	(0.220)	(0.234)	(0.241)	(0.214)	(0.210)	(0.211)	(0.211)	(0.229)	(0.231)	(0.223)	(0.225)	(0.229)
	Share second liens	4.454***	4.478***	5.125***	4.407***	4.327***	4.298***	4.291***	4.316***	4.291***	4.282***	4.348***	4.339***	4.382***	4.347***	4.160***	4.188***	4.033***	4.191***	4.167***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share refinances	-2.353***	-2.351***	-1.288***	-1.692***	-1.734***	-1.770***	-1.770***	-1.731***	-1.705***	-1.697***	-1.742***	-1.747***	-1.750***	-1.748***	-1.679***	-1.700***	-1.682***	-1.695***	-1.692***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Housing bubble	Housing units increase (2000-2007)		0.135***	0.108***	0.121***	0.113***	0.112***	0.111***	0.112***	0.111***	0.111***	0.113***	0.114***	0.115***	0.114***	0.115***	0.114***	0.116***	0.114***	0.115***
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Increase in home sales (2002-2006)		0.004*	0.002*	0.002	0.002	0.002	0.001	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.002	0.001	0.001
			(0.067)	(0.100)	(0.115)	(0.154)	(0.131)	(0.233)	(0.155)	(0.166)	(0.156)	(0.130)	(0.130)	(0.136)	(0.128)	(0.190)	(0.175)	(0.162)	(0.174)	(0.177)
	Change in median home value (2002-2006)		-0.097*	-0.092*	-0.067	-0.065	-0.062	-0.062	-0.065	-0.066	-0.066	-0.064	-0.062	-0.064	-0.062	-0.066	-0.066	-0.067	-0.066	-0.066
			(0.088)	(0.059)	(0.164)	(0.178)	(0.198)	(0.196)	(0.177)	(0.169)	(0.171)	(0.186)	(0.197)	(0.185)	(0.195)	(0.172)	(0.170)	(0.162)	(0.170)	(0.169)
Neighborhood characteristics	Age of home			0.001	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***
				(0.221)	(0.004)	(0.002)	(0.001)	(0.002)	(0.004)	(0.006)	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
	Percent of households with below median income			0.844***	0.438*	0.494*	0.515**	0.465*	0.498**	0.493*	0.495*	0.491*	0.510**	0.478*	0.503**	0.521**	0.521**	0.551**	0.523**	0.525**
				(0.001)	(0.086)	(0.051)	(0.045)	(0.067)	(0.050)	(0.051)	(0.051)	(0.051)	(0.042)	(0.057)	(0.045)	(0.039)	(0.040)	(0.030)	(0.039)	(0.038)
	Percent populations age <=35			0.724**	0.756**	0.792**	0.804**	0.776**	0.798**	0.815**	0.789**	0.782**	0.764**	0.782**	0.770**	0.807**	0.820**	0.797**	0.811**	0.814**
				(0.031)	(0.022)	(0.018)	(0.016)	(0.019)	(0.018)	(0.016)	(0.019)	(0.020)	(0.023)	(0.020)	(0.022)	(0.014)	(0.013)	(0.014)	(0.014)	(0.013)
	Share African-American			0.004	0.364***	0.353***	0.322**	0.307**	0.350***	0.339**	0.342**	0.352***	0.353***	0.360***	0.352***	0.360***	0.350***	0.347**	0.354***	0.352***
				(0.965)	(0.006)	(0.009)	(0.016)	(0.026)	(0.010)	(0.012)	(0.012)	(0.009)	(0.009)	(0.008)	(0.009)	(0.008)	(0.010)	(0.010)	(0.009)	(0.009)
	Share Hispanic			2.300**	2.335**	2.188**	2.274**	2.392**	2.196**	2.152**	2.159**	2.126**	2.066**	2.046**	2.068**	2.078**	2.126**	2.056**	2.126**	2.112**
				(0.019)	(0.021)	(0.025)	(0.023)	(0.016)	(0.026)	(0.030)	(0.027)	(0.030)	(0.033)	(0.037)	(0.034)	(0.029)	(0.027)	(0.028)	(0.026)	(0.027)
	Share Asian			-0.657	0.227	0.456	0.495	0.452	0.458	0.418	0.441	0.450	0.452	0.481	0.443	0.197	0.157	0.033	0.194	0.153
				(0.284)	(0.711)	(0.488)	(0.453)	(0.488)	(0.486)	(0.528)	(0.504)	(0.497)	(0.498)	(0.470)	(0.506)	(0.760)	(0.810)	(0.960)	(0.765)	(0.814)
	Vacancy rate			2.411***	2.193***	2.181***	2.122***	2.066***	2.165***	2.130***	2.162***	2.132***	2.086***	2.114***	2.093***	2.253***	2.253***	2.261***	2.254***	2.257***
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share rent occupied			-0.608***	-0.415**	-0.463***	-0.470***	-0.462***	-0.468***	-0.471***	-0.461***	-0.464***	-0.474***	-0.455***	-0.470***	-0.459***	-0.465***	-0.466***	-0.465***	-0.464***
				(0.001)	(0.016)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.008)	(0.007)	(0.006)	(0.008)	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)
Employment	Share employed in manufacturing or construction				2.609***	2.447***	2.339***	2.350***	2.434***	2.399***	2.416***	2.236***	2.064***	2.200***	2.092***	2.586***	2.555***	2.508***	2.551***	2.559***
					(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.002)	(0.004)	(0.002)	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Share employed in finance				-3.434**	-3.276**	-3.518**	-3.596**	-3.294**	-3.448**	-3.391**	-3.567**	-3.522**	-3.727**	-3.597**	-2.876*	-2.972*	-2.459	-2.915*	-2.887*
					(0.030)	(0.039)	(0.032)	(0.023)	(0.037)	(0.030)	(0.033)	(0.027)	(0.025)	(0.019)	(0.022)	(0.076)	(0.065)	(0.134)	(0.072)	(0.075)
Urban form	Distance to CBD					0.026	0.024	0.030	0.026	0.027	0.028	0.008	-0.009	0.002	-0.006	0.066**	0.065**	0.077***	0.060**	0.066**
						(0.220)	(0.267)	(0.153)	(0.212)	(0.195)	(0.193)	(0.751)	(0.750)	(0.934)	(0.815)	(0.020)	(0.032)	(0.007)	(0.039)	(0.026)
	Median commute time						0.004													
							(0.387)													
	Proximity to Transit							0.5 mile buffer	1 mile buffer	1.5 mile buffer	2 mile buffer									
								0.151**	0.011	0.049	0.038									
								(0.025)	(0.809)	(0.178)	(0.250)									
	Employment Accessibility by transit											Retail	Office	Industry	Total Emp					
												-0.009	-0.005**	-0.019*	-0.002*					
												(0.266)	(0.044)	(0.090)	(0.057)					
	Employment Accessibility by auto															Retail	Office	Industry	Other	Tot Emp
																0.013**	0.004*	0.023***	0.005*	0.002**
																(0.033)	(0.064)	(0.007)	(0.079)	(0.048)
	_cons	-2.470***	-2.452***	-4.415***	-4.325***	-4.347***	-4.354***	-4.358***	-4.347***	-4.345***	-4.359***	-4.220***	-4.115***	-4.185***	-4.128***	-4.629***	-4.587***	-4.656***	-4.570***	-4.603***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	/lnalpha	-2.113***	-2.142***	-2.520***	-2.635***	-2.639***	-2.641***	-2.650***	-2.639***	-2.643***	-2.644***	-2.640***	-2.645***	-2.643***	-2.644***	-2.652***	-2.650***	-2.661***	-2.649***	-2.651***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Log-Likelihood	-2,474.49	-2,466.98	-2,385.08	-2,364.80	-2,363.97	-2,363.51	-2,360.36	-2,363.93	-2,363.03	-2,363.35	-2,363.41	-2,361.97	-2,362.59	-2,362.21	-2,361.91	-2,362.39	-2,360.45	-2,362.52	-2,362.15
	chi2	1,508.448	1,589.785	1,918.926	2,079.241	2,083.447	2,100.238	2,101.722	2,086.790	2,081.836	2,103.266	2,097.141	2,122.166	2,107.211	2,116.745	2,121.029	2,112.047	2,135.876	2,116.636	2,117.352
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	McFadden's Adj R2	0.155	0.157	0.182	0.188	0.188	0.188	0.189	0.188	0.188	0.188	0.188	0.188	0.188	0.188	0.188	0.188	0.189	0.188	0.188
	BIC'	-890.124	-885.854	-998.183	-1025.884	-1021.105	-1015.599	-1021.902	-1014.747	-1016.548	-1015.906	-1015.8	-1018.666	-1017.427	-1018.192	-1018.795	-1017.842	-1021.705	-1017.563	-1018.307
	N	622	622	622	622	622	622	622	622	622	622	622	622	622	622	622	622	622	622	622
	note: *** p<0.01, ** p<0.05, * p<0.1																			

Subprime Lending Variables

As evident by the first and most simplistic model, subprime lending is strongly associated with the number of distressed properties in Maryland. Nearly all variables in all models are highly significant and have the expected direction of influence. The first variable, log of mortgaged properties, serves as an exposure factor and controls for a maximum number of distressed properties in a tract. For the entire Maryland state, the increase in the share of high cost loans, high cost and high leverage loans, increase in the share of borrowers with credit scores below 639, and increase in the share of second liens, all lead to a significant increase in the number of distressed properties.

The share of high cost loans is consistently (highly) significant across all three samples though the effect diminishes slightly when the neighborhood characteristics are introduced to the model. For every one percentage point increase in the share of loans that were high cost, the number of distressed properties is expected to increase by 3.5 percent. Since a standard deviation of the share of high cost loans is 13.4 percent, this effect on distressed properties is rather large. Similarly, a percentage increase in the share of high cost and highly leveraged loans leads to about 3.3 percent higher number of distressed properties. This effect also lessens with the introduction of neighborhood variables. In the Baltimore metropolitan area, the impact of the two subprime variables is consistently significant and of similar magnitude though the latter variable has about half as strong an impact than it does for the entire state. In the Washington metro region, on the other hand, due to increased correlation between the two variables, only the share of high cost loans remains as a significant predictor of levels of distress. The effect of the variable is likewise in the 4 percent range.

Another subprime variable, the share of borrowers with low credit scores, also indicates a positive relationship with the number of distressed properties. The impact of a one percentage increase of low credit score borrowers is around one percent. The variable however loses significance when the model controls for the neighborhood variables such as income and race. This is expected as generally the population with lower incomes and minorities have lower credit scores. The same is true when looking at the Baltimore region only. In Washington, though, this variable exhibits the same magnitude of impact, but it is only marginally significant when the distance to the CBD is teased out in the model as well.

The share of second lien originations is positively and significantly related with the rise in distressed properties. This variable carries the largest impact of the number of distressed properties in Maryland. Across all three samples, one percent increase in the share of second liens increases the levels of distressed properties by 4 to 7 percent. The impact is strongest in Washington metropolitan region though it lessens some when neighborhood characteristics are controlled for. This result suggests that the use of piggybacks might have been more prevalent in Washington due to relatively higher housing prices and appears to have carried more risk.

The last indicator of subprime lending is a variable measuring the share of loans that were refinances. As anticipated, a greater share of originations that are refinances lowers the number of distressed properties. The impact ranges between 1.1 and 2.5 percent and is relatively lower in the Washington metropolitan region. This suggests that while households who refinanced during the housing boom may have refinanced to extract some equity out of their home or obtain better financing terms, they did not necessarily become delinquent as the result of refinancing.

Housing Bubble Variables

The second model accounts for the impact of the housing bubble on the level of distressed properties. The three variables, the increase in the number of housing units, the increase in home sales, and the increase in home values, exhibit expected signs and are generally significant. Housing prices did play some role in the accumulation of distressed properties though the effect dissipates when looking at the full sample. In the Washington region particularly, home prices have significantly impacted levels of distress, with one percent increase in home value leading to about a quarter percent increase in the level of distress. In Baltimore, this impact was negative, i.e. the increase in home value led to lower levels of distressed properties, and it also dissipated when other neighborhood characteristics were accounted for. The impact in Baltimore differs probably due to the different real estate market than in Washington. As noted before, the Baltimore metropolitan area has a weaker real estate market and suffered from relatively large concentration of foreclosures in central city neighborhoods even before the current crisis. Gentrifying and other desired neighborhoods saw larger home price appreciation but also remained stable afterwards. On the other hand, some central city neighborhoods continued with high concentrations of foreclosures and did not see significant changes in home prices.

The variable accounting for change in homes sales between 2002 and 2006 is used as a proxy for a speculative bubble where the relatively higher turnover was not the result of fundamental increase in demand but the expectation of future price appreciation. Those areas may have consequently experienced more unstable housing markets which led to a higher concentration of distressed properties. Similarly, the variable measuring change in the number of housing units between 2000 and 2007 is intended to capture the speculative bubble and also to account for higher incidence of riskier lending during those years. The estimation results suggest that while the areas with higher level of new construction activity did subsequently see more distress, that impact is rather small and only marginally significant. Change in home prices did not have any impact on the level of distress across the state. In the Washington region, increased construction had more than twice the impact on levels of distress than the rest of the state, with each percentage point increase in the number of new housing units leading to almost half a percent increase in the number of distressed properties. The results for Baltimore are consistent with the state sample, though home sales seemed not to have any impact after controlling for neighborhood characteristics.

Neighborhood Characteristics Variables

The third model goes further by accounting for a variety of neighborhood characteristics. All the variables again exhibit the expected signs and the estimated coefficients are highly significant. Income and tenure impacts lose significance when employment is accounted for in the fourth model. It is noteworthy that apart from the share of second liens, subprime lending coefficients are smaller in magnitude when the neighborhood characteristics are added. This finding highlights the importance of neighborhood characteristics in research on mortgage defaults. This result also highlights the tight relationship between subprime lending and neighborhood and racial characteristics. However, in the absence of individual borrower information, the neighborhood variables also capture characteristics of the borrowers in a neighborhood.

As expected, lower income neighborhoods have higher levels of distress. The impact is not consistent though. In the full sample, the income effect is unstable and goes away when employment sectors are accounted for. As previously noted, the income variable is expected to be correlated with some of the other neighborhood and employment characteristics. In the Washington sample, the significance of income variable dissipates when distance from the CBD is added to the analysis. In Baltimore, on the other hand, income maintained significant impact on the level of distressed properties even after accounting for the employment type. One percent increase in the share of population with below the median income led to about a half percent increase is levels of distress.

In line with the previous research, racial composition of a neighborhood had a significant impact of the number of distressed properties. For each of the three races, African-American, Hispanic, and Asian, the impact is positive. While the share of Hispanics has a relatively larger effect, after controlling for spatial location, the effect of the Asian population is more conspicuous. One percent increase in the share of Asian population leads to a two percent increase in distress. The impact is similar in Washington region where the share of African-American population becomes significant after holding the distance to the CBD constant. In Baltimore, the share of Asian population does not appear to bear any significance on accumulation of distressed properties. This may be due to a relatively lower concentration of an Asian population in the Baltimore metropolitan area.

I also account for vacancy status and tenure status as neighborhood variables. As anticipated, higher levels of vacancies led to higher levels of distressed properties and the impact is strongest in the Baltimore region. A one percent increase in vacancies led to over a 2 percent increase in distress. In Washington, though, the impact is reversed. After controlling for distance to the CBD, one percent higher vacancies led to about 1 percent lower distress. This result may contribute further evidence to the difference between the Baltimore and Washington housing markets. While previously vacant areas in Baltimore continued suffering from vacancies, Washington's stronger housing market absorbed the vacancies. The impact of tenure status is also consistent with the previous research. Across the state, tenure status generally does not have a significant impact on the levels of distress, yet impact is more pronounced in Washington and Baltimore regions. One percent more renter-occupants lead to about 4 percent decrease in distressed properties in both regions.

The remaining two neighborhood variables, age of homes and population age, also provide some interesting insights. The age of home variable suggests that older neighborhoods across the state and in the two metropolitan areas experienced higher levels of distress. Though the magnitude of the impact is rather small, it is consistent across space and models. The effect is opposite from the one found by Immergluck (2010) where the areas with newer homes had higher concentration of foreclosures. The differing effect may indicate that non-prime lending in Maryland was heavily concentrated in older neighborhoods with prevalence of minority residents. Immergluck's study encompasses much newer metropolitan areas, in Arizona and Nevada for example, where the construction boom was of much larger proportions than that seen in Maryland.

Age of population is a significant predictor of distress and shows the expected effect. Across the state, one percent increase in population aged 35 or less increases distress by more than a half a percent. In Baltimore that increase is three-quarters of a percent, while in Washington, it is as much as 1.3 percent.

Employment Variables

The employment variables used here are indented to capture employment in sectors which were disproportionally impacted as the result of the economic downturn. As anticipated, employment in construction or manufacturing led to a sizable and significant increase in levels of distress. In the full sample, one percent increase in those sectors led to about a 3.5 percent increase in distress. The effect is somewhat smaller in Baltimore and even more so in Washington, but still accounting for a 2 to 3 percent increase in distress. Employment in finance sector did not exhibit any significant impact on levels of distress across the state or in the Washington region; however the impact was negative in the Baltimore region. This outcome may suggest that those employed in the finance sector may have skills transferable to other sectors which may have insulated them from economic shocks associated with employment layoffs.

Urban Form Variables

Now I turn to the variables which are of primary interest in this analysis. Urban form is accounted for via five variables: distance to the CBD, commute time, proximity to transit, and employment accessibility by transit and employment accessibility by auto. The variable accounting for the distance to the CBD provides particularly interesting information showing that increase in the distance from the CBD results in higher concentration of distressed properties, all else equal. This result is consistent with arguments of those who believe that increased levels of foreclosures on the urban fringe is not merely an outcome of the subprime lending disaster but a shift in preferences for urban living. The effect is evident when looking at the full sample, but with even greater magnitude when looking at the Washington metro communities. In Baltimore, distance to the CBD alone does not affect the level of distress, but distance becomes significant when employment accessibility by auto is kept constant. I suspect that the high correlation between distance to the CBD and accessibility measures in Baltimore impacts the results. In Washington, the correlation is much lower. For the full sample, being 10km further from the CBD increases the level of distress by about 3 percent. In Washington, distress increases by as much as 10 percent before accounting for other accessibility measures.

The second measure of interest is median commute time. This variable is also interesting because it suggests that commute time impacts the accumulation of distress in areas of the state apart from Washington and Baltimore regions. In other words, when the sample is not constrained to either Washington or Baltimore, the areas with longer commute time also experienced higher levels of distress. This result is consistent with the notion that households drove out further away from employment and service centers into auto-dependent bedroom communities that offered more affordable housing.

This notion is further corroborated by measures of employment accessibility by auto and transit which indicate significant relationships with accumulation of distressed properties. In particular, accessibility to employment by transit consistently shows significant impact on properties' distress. Access to retail jobs has a sizable effect of 0.72 percent decrease for 10,000 more jobs, while access to office jobs has a smaller effect of 0.19 percent decrease per 10,000 jobs. The largest impact comes from the accessibility to industry jobs by transit with a decline of 1.01 percent. Access to total employment by transit has the same effect as it does by the car, 0.11 percent. While the anticipated relationship between accessibility and foreclosures is confirmed, the results provide additional information about the relative importance of different employment sectors. In Washington, the relevance of retail jobs and industry jobs is even greater. Ten thousand more retail jobs decrease distress by 1 percent while the same number of industry jobs decrease distress by 1.2 percent. The office jobs reduce distress by 0.3 percent. Access to 10,000 of any type of jobs has a 0.1 percent reducing effect. In Baltimore, the effects are similar. Access to retail jobs leads to 0.9 percent decline though the impact is insignificant. The office jobs have a declining effect of 0.5 percent, industry jobs decrease distress by 1.9 percent and the total employment has a negative 0.2 percent effect.

Accessibility to jobs by auto has similar effects on levels of distressed properties. Ten thousand more retail jobs decrease the number of distressed properties by 0.59 percent. The same accessibility to office jobs decreases distress by 0.24 percent. Accessibility to industry jobs by car does not appear to impact foreclosures, while jobs in the category of "other" lead to 0.35 percent decrease in foreclosures. Employment in the “other” category include: farm, mining, forestry, fishing and agricultural support, construction, health and social services, arts, entertainment and recreation, accommodations, food services, and other services including rental.

Finally, access to 10,000 more jobs in general decreases distress by 0.11 percent. Within the two metropolitan areas, the impact of employment accessibility is still most significant though the coefficients are to some degree smaller and opposite between the two regions. While car accessibility to retail jobs leads to 0.13 percent decline per 10,000 more jobs in Washington, it leads to a 0.13 increase in Baltimore. The same opposite relationship holds for office jobs with a 0.4 percent decreasing effect in Washington and a 0.4 increasing effect in Baltimore. Industry jobs lead to a 1.1 percent decrease in Washington and a 2.3 percent increase in Baltimore, while “other” jobs have a positive 0.4 percent and a negative 0.5 percent impacts for Washington and Baltimore respectively. Finally, access to 10,000 more jobs of any type of leads to a 0.2 percent decrease in distress in Washington and a 0.2 percent increase in Baltimore. The opposite effects in Baltimore and Washington are interesting, though the Baltimore effects are counterintuitive. One would not generally expect the areas with greater accessibility to jobs to also have more foreclosures. The results for Baltimore may be because the accessibility measure is picking up some unobserved neighborhood characteristic that is not accounted for in this analysis. Baltimore’s most blighted neighborhoods are well located near to downtown and with good accessibility to jobs. It may be that the neighborhood variables included in the model may be lacking some distinguishing feature of such neighborhoods.

And, the last set of variables controlling for proximity to transit in general does not appear to have significant and consistent impact on levels of distress. For the full sample, insignificance of transit buffers is expected as transit stops are located in the denser central corridor between the Baltimore and Washington regions. In Washington, being in a 1.5 mile buffer of a transit stop marginally decreases levels of distress but the effect is rather large, 5.8 percent fewer distressed properties. That effect is not conclusive given that proximity buffers of 0.5, 1, or 2 miles do not indicate any significant relationship with distress. In Baltimore, again, the buffer immediately encompassing the transit stops, 0.5 mile buffer, suggests increased levels of distressed properties by as much as 15 percent. The larger buffers are not significant. The result for Baltimore again indicates the peculiarity of the area since one would not expect transit oriented communities to be prone to foreclosures. Same as with the accessibility measure, the transit buffer may be suggesting some unobserved characteristics of those neighborhoods.

Discussion

Overview of the Findings

The current foreclosure crisis is undoubtedly one of the most significant events in the recent history of housing markets in the United States. With the number of distressed households in millions, the future of many still remains uncertain. And while the financial regulatory oversight is still discussing precipitating causes and deliberating the appropriate response, the spatial context of the crisis is decisive for future organization of metropolitan spatial structure. Some have even argued that reversed preference for urban living among the new, non-traditional household structure is among the causes of the foreclosure crisis.

The purpose of the third essay was to examine the importance of urban form in accumulation of distressed properties in Maryland and two metropolitan areas, Baltimore and Washington. I investigated the relationship between distressed properties and their proximity to transit, employment, and central business districts. By introducing a richer set of urban form measures, I aimed to tease out differing findings of the previous research. I do, in fact, find that the majority of these measures have significant impact on the accumulation of distressed properties. Still, as previous research established, the dominating trigger of foreclosures in Maryland was the risky component of subprime lending with excessive mortgage leveraging carrying the most weight. All other subprime lending measures were also important predictors of household distress. The impact of second liens highlights an interesting trend observed in this region during the housing boom, specifically the lack of housing affordability. In order to afford homeownership, households overextended themselves by resorting to "piggyback" loan products.

Following in magnitude was the impact of employment with neighborhoods heavily dependent on jobs in manufacturing or construction suffering the most. These results confirm the expectations that the foreclosure crisis was just as much brought on by the risky lending practices as the subsequent economic downturn. The neighborhood characteristics which are also proxies for borrower characteristics confirm all previous expectations. Interestingly, also, the opposing impact of housing vacancies for the two metropolitan areas draws attention to differing housing markets, with the Washington region being relatively stronger than the Baltimore region. However, this study shows that even after accounting for all expected drivers of the foreclosure crisis, urban form matters. One of the findings suggests that more foreclosures are occurring further from the central business districts. In Washington region, there are about 10 percent more distressed properties 10 kilometers away from the downtown, with the effect only falling down to 5 percent after accounting for accessibility to jobs. The effect is somewhat smaller when looking at the entire state though it is consistent again after accounting for accessibility to employment. This result confirms the value placed on amenities provided by urban living. The outcome on the metropolitan spatial structure works through the standard urban economic theory of changing rent gradients. In competing for locations closer to the urban core, increased demand for closer locations raises the value of that land and leads to higher density development. With increased population density closer to the urban core, the intercept of the density gradient rises as well. That is also the result confirmed by the analysis in the first essay for the Washington metropolitan area.

Accessibility to employment, too, is a significant predictor of mortgage default. With the largest coefficient carried by accessibility to industry jobs, this measure may highlight the vulnerability of those employed in industry jobs to mortgage default if they lose their jobs. The employment accessibility measure may capture both the greater desirability of areas closer to employment as well as the availability of quicker transition of those who face loss of employment. Untangling the two effects may be an interesting subject for future research.

Unexpectedly, I found the contrasting effect of auto accessibility and proximity to transit in Baltimore. There are several differing trends between Washington and Baltimore. First, the higher correlation between accessibility variables and the distance to the central business district in Baltimore leads to the model's instability among the urban form measures. In Washington, the distribution of employment is not as heavily related to the distance to the central business district. Second, some unobserved feature of the Baltimore's most accessible and transit-oriented areas accounts for the increased level of distress.

Finally, the results of this foreclosure study are consistent with the analyses of change in the metropolitan spatial structure from the first two essays. While the first two essays measured change in urban form up to the year 2007, the results on foreclosures provided in this essay may hint at future change in metropolitan spatial structure. Given the causative relationship between foreclosures and dire economic conditions and economic conditions and urban form, foreclosures may be a leading indicator that future urban form is changing.

Limitations of the Study

There are several limitations of this study which are important to acknowledge. The foreclosure data used here captures both mortgage default as well as foreclosures up to the third quarter of 2009. However, due to the extended foreclosure process, I am not able to observe what happened to the defaulted properties thereon out. Also, while I am observing Maryland's side of the Washington metropolitan area, I do not account for the counties in Virginia or the District of Columbia. The sample encompassing the entire Washington metropolitan area may provide a better informational picture of the foreclosure crisis and how it may lead to urban form change. Also, this analysis only looks at the areas in Maryland. It would be interesting to similarly analyze other states and metropolitan regions where forces at hand may have had different effects on the accumulation of distressed properties. Washington's dependency on the federal government and its consequently strong job market insulated the area from the worst of the economic downturn. Finally, it would be valuable to examine the other components of urban form which affect household transportation costs, such as proximity to bus stops and schools.

Final Remarks

In this dissertation, I evaluated changes in urban spatial structure across a number of metropolitan areas in the US between 1990 and 2007. The measures chosen in this work intended to primarily address attributes of urban form synonymous with the idea of sprawl and included a number of density measures, concentration and clustering indices, a measure for allocation of new growth within urban areas, and a new measure of infill development. By focusing on the multi-dimensional nature of urban form, I was able to discern some of the interesting trends at work since 1990.

While I cannot say that suburbanization trends have determinately reversed across the country, the areas experiencing the most growth during this period have certainly showed reversal of pre-1990s trends. The new growth trends which occurred in the metropolitan regions and which can contribute most of their current size to the growth in the last two decades appear to be different than growth trends of older, established regions.

Despite their population pressures, new growth regions focused on increasing density of already existing urban areas. Following the definition of smart growth as presented herein, it is reasonable to conclude that most of the studied metropolitan areas have grown smarter. The growth trends are consistent with the theory of urban resurgence. While it is beyond the scope of this work to examine the underlying causes of increased demand for the urban environment and its amenities, there is an opportunity for future research to address that question. It appears that the urban resurgence trend is more prevalent among the new, incoming population. Also, changes in demographic trends, such as increasing number of single person households and the shrinking household size in general, should be addressed as significant drivers of future change in spatial structure. In areas that remained stagnant in terms of their population size, the previously established trend of decentralization continued. It is in any case difficult to talk about urban resurgence for a region that hardly experienced any population growth. A household with more than one member would rarely relocate from their existing suburban home to a downtown location.

Separately, there also appears to be a clear distinction between the decade of 1990s and of 2000s. The trends of the 1990s can be claimed with some degree of certainty. The results show almost universally the tendency for urban areas to fill in the land previously passed over in the process of development. The decade of the 2000s, while subject to a great deal of innovation, was also the era of what can only be described as irrational exuberance on many fronts. For urban spatial structure, it was manifested in excessive overbuilding across the country and worldwide. The notion of shelter, housing, became tantamount with a perpetually multiplying investment portfolio, or a Ponzi scheme. However, the result of its collapse has been devastating for so many. Today, vacant homes epitomize neighborhoods of many metropolitan areas. Due to these aberrant set of events, it is difficult to ascertain which trends are indeed leading to an underlying reshaping of metropolitan areas and which trends simply caught the tailwinds of the exuberance. The analysis of the foreclosure crisis presented in here also suggests that the theory of urban resurgence at least in the regions like Washington, DC, holds ground. The shift is demographic-specific though. As the 2010 census shows for the District, households in their 20s and early 30s drove almost all of the city's growth since 2000 and today make up almost one third of the District's population. The city also saw an increase in households in their late 50s and early 60s, while the number of children younger than 15 decreased by a fifth (Morello et al, 2011).

Spatial hazard analysis presented in the second essay also illustrated that urbanization and suburbanization trends can coexist within a single metropolitan area. The same applies to change in urban form. Over the last decade, metropolitan spatial structure has not changed consistently or universally. It is evident that some areas grew both denser in the urban core, but also more dispersed on the urban fringe. What all these results may be implying is that the traditional urban areas have gone from being organized based on income and race, to a new urban form which uniquely follows population age distribution, with young and old living close to urban amenities and those middle-aged enjoying the space of suburbs.

Following the findings of this dissertation, I see many opportunities for planners to influence the reshaping of future metropolitan structure. While change in urban form may not seem as drastic when viewed in aggregate, there are clearly opportunities to improve efficiency of spatial structure and revisit collective concerns which led to a call for new urbanism and smart growth in the first place. As previously suggested by Peiser (1989), today's sprawl can turn into compact development in later years as the pace of urban growth leads builders to fill-in previously undeveloped sites. With a high probability of the re-shifting of population within and across metropolitan areas, planners can ensure that the opportunities for infill development are clearly defined and without obstacles. Further, the results illuminate the multi-dimensional nature of the metropolitan areas. When considering public policy responses, it is important to determine which specific dimension of urban form is creating detrimental results.

And finally, the foreclosure crisis, though devastating for many families and highly taxing on cities and neighborhoods, provides a unique opportunity for planners, policy makers, and developers. The location of foreclosures clearly matters; thus, the location of foreclosures needs to be taken into consideration when determining remedial measures. Many have suggested turning vacant homes into long-term affordable housing or redeveloping the areas, given the relatively inexpensive land and housing stock. However, when the housing stock is located further away from the service and employment centers, turning it into affordable housing further challenges the households facing financial difficulties. Also, the current foreclosure crisis is an opportunity for planners to learn from the grave mistakes. Understanding the circumstances that cause neighborhoods to be vulnerable to foreclosure and exposing the planning process and urban form that enable these circumstances is critical in creating more sustainable forms of urban spatial structure.

APPENDIX A

Table 13: Metropolitan Area Population Changes, 1990-2007

Metropolitan Area	Population			Difference in Population			Population change		
	1990	2000	2007	2000-1990	2007-1990	2007-2000	2000-1990	2007-1990	2007-2000
Atlanta-Sandy Springs-Marietta, GA	3,069,427	4,247,981	5,322,915	1,178,554	2,253,488	1,074,934	38%	73%	25%
Austin-Round Rock, TX	850,140	1,256,354	1,577,236	406,214	727,096	320,882	48%	86%	26%
Baltimore-Towson, MD	2,450,250	2,635,625	2,798,417	185,375	348,167	162,792	8%	14%	6%
Boston-Cambridge-Quincy, MA-NH	4,133,897	4,391,344	4,515,779	257,447	381,882	124,435	6%	9%	3%
Charlotte-Gastonia-Concord, NC-SC	1,024,291	1,330,448	1,621,635	306,157	597,344	291,187	30%	58%	22%
Chicago-Naperville-Joliet, IL-IN-WI	8,182,079	9,098,316	9,747,870	916,237	1,565,791	649,554	11%	19%	7%
Cincinnati-Middletown, OH-KY-IN	1,844,793	2,009,632	2,118,580	164,839	273,787	108,948	9%	15%	5%
Cleveland-Elyria-Mentor, OH	2,102,091	2,148,143	2,150,129	46,052	48,038	1,986	2%	2%	0%
Dallas-Fort Worth-Arlington, TX	3,989,291	5,161,544	6,118,183	1,172,253	2,128,892	956,639	29%	53%	19%
Denver-Aurora, CO	1,634,528	2,127,336	2,399,559	492,808	765,031	272,223	30%	47%	13%
Detroit-Warren-Livonia, MI	4,248,698	4,452,557	4,561,522	203,859	312,824	108,965	5%	7%	2%
El Paso, TX	591,610	679,622	751,891	88,012	160,281	72,269	15%	27%	11%
Houston-Sugar Land-Baytown, TX	3,767,463	4,715,407	5,620,734	947,944	1,853,271	905,327	25%	49%	19%
Indianapolis, IN	1,294,217	1,525,104	1,701,870	230,887	407,653	176,766	18%	31%	12%
Jacksonville, FL	925,214	1,122,750	1,359,173	197,536	433,959	236,423	21%	47%	21%
Las Vegas-Paradise, NV	748,000	1,381,417	1,900,147	633,417	1,152,147	518,730	85%	154%	38%
Los Angeles-Long Beach-Santa Ana, CA	10,281,036	11,111,054	11,780,120	830,018	1,499,084	669,066	8%	15%	6%
Miami-Fort Lauderdale-Miami Beach, FL	4,056,100	5,007,564	5,607,038	951,464	1,550,938	599,474	23%	38%	12%
Minneapolis-St. Paul-Bloomington, MN-WI	2,538,831	2,968,806	3,313,789	429,975	774,958	344,983	17%	31%	12%
New York-Northern New Jersey-Long Island, NY-NJ-PA	16,489,798	17,898,155	18,629,884	1,408,357	2,140,086	731,729	9%	13%	4%
Orlando-Kissimmee, FL	1,224,851	1,644,561	2,098,102	419,710	873,251	453,541	34%	71%	28%
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	5,737,662	6,048,477	6,339,267	310,815	601,605	290,790	5%	10%	5%
Phoenix-Mesa-Scottsdale, AZ	2,253,304	3,268,226	4,181,616	1,014,922	1,928,312	913,390	45%	86%	28%
Pittsburgh, PA	2,468,289	2,431,087	2,404,190	-37,202	-64,099	-26,897	-2%	-3%	-1%
Portland-Vancouver-Beaverton, OR-WA	1,523,738	1,927,881	2,162,868	404,143	639,130	234,987	27%	42%	12%
Richmond, VA	992,048	1,168,021	1,316,356	175,973	324,308	148,335	18%	33%	13%
Riverside-San Bernardino-Ontario, CA	3,585,376	4,516,881	5,572,433	931,505	1,987,057	1,055,552	26%	55%	23%
Sacramento-Arden-Arcade-Roseville, CA	1,481,807	1,796,857	2,141,388	315,050	659,581	344,531	21%	45%	19%
St. Louis, MO-IL	2,580,901	2,698,687	2,833,675	117,786	252,774	134,988	5%	10%	5%
San Antonio, TX	1,403,765	1,705,112	1,978,640	301,347	574,875	273,528	21%	41%	16%
San Diego-Carlsbad-San Marcos, CA	2,472,728	2,784,344	3,032,312	311,616	559,584	247,968	13%	23%	9%
San Francisco-Oakland-Fremont, CA	3,686,592	4,123,740	4,316,905	437,148	630,313	193,165	12%	17%	5%
Seattle-Tacoma-Bellevue, WA	2,559,163	3,043,878	3,327,901	484,715	768,738	284,023	19%	30%	9%
Tampa-St. Petersburg-Clearwater, FL	2,067,963	2,395,997	2,765,528	328,034	697,565	369,531	16%	34%	15%
Virginia Beach-Norfolk-Newport News, VA-NC	1,454,603	1,581,449	1,695,860	126,846	241,257	114,411	9%	17%	7%
Washington-Arlington-Alexandria, DC-VA-MD-WV	4,060,750	4,700,926	5,321,363	640,176	1,260,613	620,437	16%	31%	13%

Table 14: Density Gradients 1990, 2000, and 2007

Metropolitan Areas	Ln (population density 1990)			Ln (population density 2000)			Ln (population density 2007)			N
	Distance to CBD	Intercept	Adj R2	Distance to CBD	Intercept	Adj R2	Distance to CBD	Intercept	Adj R2	
Atlanta-Sandy Springs-Marietta, GA	-0.571***	7.644***	0.720	-0.522***	7.768***	0.671	-0.495***	7.858***	0.671	690
	-38.877	178.440		-36.403	164.470		-35.533	170.883		
Austin-Round Rock, TX	-0.736***	7.389***	0.411	-0.694***	7.716***	0.468	-0.656***	7.808***	0.468	257
	-12.314	65.208		-12.765	81.389		-12.372	86.385		
Baltimore-Towson, MD	-0.754***	8.353***	0.575	-0.694***	8.286***	0.562	-0.662***	8.264***	0.549	642
	-28.154	141.525		-27.614	148.636		-27.162	153.118		
Boston-Cambridge-Quincy, MA-NH	-0.459***	8.415***	0.487	-0.448***	8.437***	0.487	-0.435***	8.427	0.476	920
	-23.635	158.470		-23.992	161.217		-23.683	162.626		
Charlotte-Gastonia-Concord, NC-SC	-0.528	6.872***	0.362	-0.527	7.132***	0.431	-0.531	7.289***	0.457	267
	-12.356	61.893		-13.098	68.716		-13.267	71.232		
Chicago-Naperville-Joliet, IL-IN-WI	-0.444***	8.774	0.413	-0.419***	8.793	0.417	-0.402***	8.802	0.431	2,052
	-29.981	168.223		-30.473	184.201		-30.768	201.559		
Cincinnati-Middletown, OH-KY-IN	-0.618***	7.721	0.412	-0.577***	7.693***	0.425	-0.538***	7.638**	0.405	486
	-16.284	95.234		-16.458	114.007		-15.587	115.796		
Cleveland-Elyria-Mentor, OH	-0.564***	8.032***	0.311	-0.509***	7.921*	0.267	-0.490***	7.885***	0.278	693
	-14.949	86.815		-13.405	82.021		-13.708	89.038		
Colorado Springs, CO	-1.215***	7.545***	0.689	-1.136***	7.756***	0.708	-1.097***	7.796***	0.702	123
	-12.753	55.422		-14.896	67.164		-15.841	71.137		
Dallas-Fort Worth-Arlington, TX	-0.377***	7.649**	0.241	-0.358***	7.892	0.296	-0.338***	7.950	0.297	1,046
	-17.522	109.715		-17.517	122.261		-16.762	124.940		
Denver-Aurora, CO	-1.078***	8.352***	0.492	-0.934***	8.451***	0.473	-0.864***	8.432***	0.461	520
	-14.016	72.706		-14.002	85.467		-13.792	91.458		
Detroit-Warren-Livonia, MI	-0.465***	8.264***	0.471	-0.416***	8.167***	0.446	-0.394***	8.127***	0.455	1,289
	-24.716	146.857		-22.758	143.001		-22.856	158.170		
El Paso, TX	-1.167***	8.442	0.358	-0.909***	8.245	0.345	-0.841***	8.209	0.322	126
	-7.690	37.315		-7.128	45.397		-6.482	44.853		
Houston-Sugar Land-Baytown, TX	-0.451***	7.662	0.312	-0.425***	7.823***	0.320	-0.412***	7.930***	0.333	895
	-15.687	99.688		-15.940	108.632		-15.948	116.216		
Indianapolis, IN	-0.795	7.644***	0.538	-0.743	7.663***	0.531	-0.709	7.656***	0.532	315
	-15.717	93.620		-15.538	100.153		-15.257	104.316		
Jacksonville, FL	-0.542***	7.155	0.375	-0.492***	7.195	0.351	-0.462***	7.277	0.345	201
	-7.672	63.554		-7.396	66.735		-7.164	70.698		
Las Vegas-Paradise, NV	-0.373	5.998***	0.095	-0.405***	7.739***	0.338	-0.399**	7.971***	0.411	347
	-6.402	33.979		-7.820	88.660		-7.771	106.066		
Los Angeles-Long Beach-Santa Ana, CA	-0.404	8.971**	0.229	-0.361**	8.947	0.196	-0.339***	8.960**	0.191	2,371
	-17.886	189.828		-16.597	187.520		-16.747	196.798		
Miami-Fort Lauderdale-Miami Beach, FL	-0.128	7.715***	0.089	-0.122	7.996***	0.160	-0.112*	8.062***	0.167	890
	-10.139	98.851		-11.944	148.593		-11.794	164.753		
Minneapolis-St. Paul-Bloomington, MN-WI	-0.764	8.093***	0.593	-0.738***	8.216***	0.661	-0.668***	8.169***	0.618	746
	-20.435	107.812		-28.353	154.598		-21.022	133.316		
New York-Northern New Jersey-Long Island, NY-NJ-PA	-0.487***	9.573***	0.378	-0.482***	9.627***	0.349	-0.481***	9.692***	0.387	4,396
	-48.069	262.434		-47.458	255.291		-48.886	274.798		
Orlando-Kissimmee, FL	-0.537	7.198***	0.258	-0.508**	7.477***	0.336	-0.462	7.577***	0.353	328
	-9.097	64.491		-9.552	85.348		-9.282	96.174		
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	-0.420***	8.141***	0.277	-0.392***	8.088***	0.245	-0.371***	8.076***	0.234	1,556
	-21.238	124.124		-19.430	119.860		-18.925	122.256		
Phoenix-Mesa-Scottsdale, AZ	-0.298***	7.012***	0.188	-0.286***	7.590***	0.279	-0.279***	7.753***	0.300	705
	-6.195	59.420		-6.828	78.484		-6.898	84.451		
Pittsburgh, PA	-0.505***	7.780**	0.373	-4.881***	7.702***	0.373	-0.479***	7.657***	0.370	721
	-18.330	119.286		-18.213	124.696		-17.944	124.513		
Portland-Vancouver-Beaverton, OR-WA	-0.901**	8.020***	0.575	-0.873***	8.220***	0.581	-0.858***	8.293***	0.572	426
	-16.956	94.487		-16.690	102.640		-16.493	104.399		
Richmond, VA	-0.653	7.203***	0.487	-0.617*	7.271***	0.487	-0.594	7.303***	0.482	283
	-13.390	71.222		-12.391	73.616		-11.573	74.149		

Riverside-San Bernardino-Ontario, CA	-0.199***	7.253***	0.091	-0.180***	7.527**	0.104	-0.172***	7.665***	0.105	840
	-8.474	75.668		-7.412	79.489		-6.920	80.628		
Sacramento-Arden-Arcade-Roseville, CA	-0.315***	7.145***	0.185	-0.320***	7.565***	0.304	-0.315***	7.684***	0.324	403
	-7.824	58.459		-8.358	80.476		-8.332	87.199		
St. Louis, MO-IL	-0.593**	7.901***	0.560	-0.535***	7.764***	0.492	-0.511***	7.739***	0.480	550
	-22.371	102.562		-20.017	94.769		-19.637	99.295		
San Antonio, TX	-0.871***	7.786*	0.531	-0.808*	7.890*	0.546	-0.774***	7.948**	0.542	338
	-18.254	85.671		-18.325	94.573		-17.840	98.009		
San Diego-Carlsbad-San Marcos, CA	-0.365***	8.010***	0.151	-0.326***	8.171***	0.205	-0.316***	8.224***	0.213	599
	-9.268	84.319		-8.719	104.601		-8.535	109.346		
San Francisco-Oakland-Fremont, CA	-0.469	8.805***	0.240	-0.437	8.845***	0.239	-0.429	8.866***	0.240	871
	-14.864	108.013		-14.742	115.299		-15.083	120.946		
Seattle-Tacoma-Bellevue, WA	-0.456***	8.005	0.349	-0.420***	8.096*	0.332	-0.406***	8.140*	0.328	664
	-16.935	120.881		-15.615	122.390		-15.089	123.386		
Tampa-St. Petersburg-Clearwater, FL	-0.246***	7.172	0.065	-0.251***	7.385***	0.101	-0.249***	7.506	0.118	547
	-6.804	69.988		-7.514	87.813		-7.701	93.071		
Virginia Beach-Norfolk-Newport News, VA-NC	-0.557***	8.093	0.232	-0.516***	8.051***	0.221	-0.486***	8.025	0.212	368
	-10.911	62.996		-10.305	66.383		-9.988	69.213		
Washington-Arlington-Alexandria, DC-VA-MD-WV	-0.545***	8.234	0.454	-0.484***	8.228***	0.410	-0.451***	8.249	0.393	1,004
	-21.662	134.883		-20.479	133.858		-20.107	140.879		
All metropolitan Areas	-0.391***	8.073***	0.217	-0.369***	8.184***	0.230	-0.354***	8.226***	0.232	29,475
	-68.903	499.767		-69.101	546.214		-68.267	570.324		

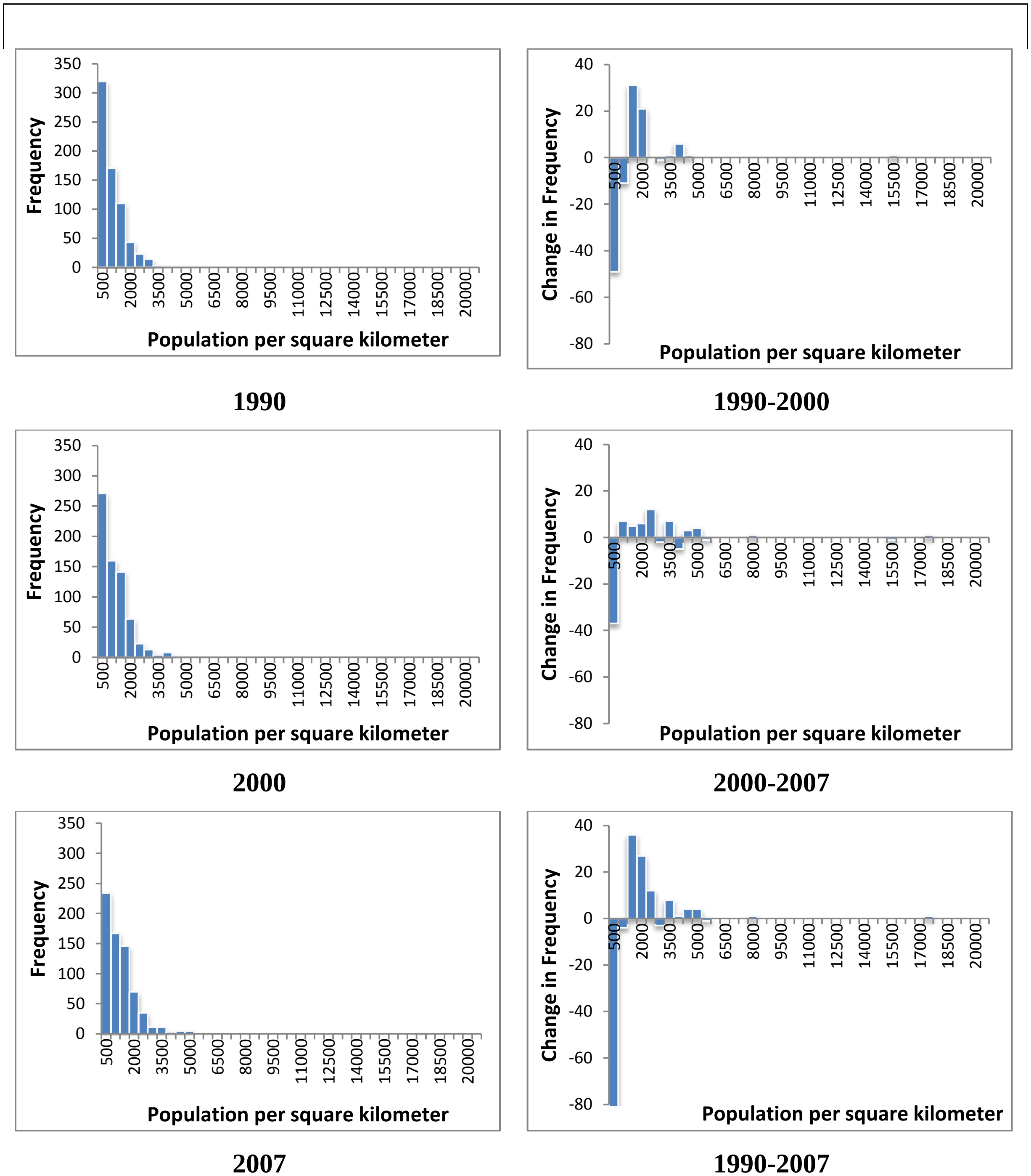
Table 15: Density Gradients Change, 1990-2000, 2000-2007, 1990-2007

Metropolitan Areas	Change 1990 to 2000			Change 2000 to 2007			Change 1990 to 2007			N
	Distance to CBD	Intercept	Adj R2	Distance to CBD	Intercept	Adj R2	Distance to CBD	Intercept	Adj R2	
Atlanta-Sandy Springs-Marietta, GA	0.048***	0.123***	0.062	0.028***	0.091***	0.064	0.076***	0.214***	0.094	690
	7.049	4.895		4.962	7.276		7.463	7.023		
Austin-Round Rock, TX	0.042***	0.327***	0.009	0.038***	0.092***	0.076	0.081***	0.419***	0.030	257
	2.913	6.735		5.212	6.435		4.056	7.602		
Baltimore-Towson, MD	0.060***	-0.068***	0.054	0.032***	-0.022***	0.272	0.093***	-0.089***	0.106	642
	9.265	-3.854		14.043	-4.183		11.574	-4.498		
Boston-Cambridge-Quincy, MA-NH	0.011***	0.022***	0.023	0.013***	-0.010***	0.265	0.024***	0.012	0.088	920
	3.708	2.829		14.947	-3.880		7.252	1.357		
Charlotte-Gastonia-Concord, NC-SC	0.001	0.260***	-0.004	-0.004	0.157***	-0.003	-0.003	0.417***	-0.004	267
	0.062	6.485		-0.734	9.081		-0.213	8.095		
Chicago-Naperville-Joliet, IL-IN-WI	0.025***	0.020	0.016	0.017***	0.008	0.018	0.042***	0.028	0.036	2,052
	4.782	0.908		5.137	0.590		7.080	1.202		
Cincinnati-Middletown, OH-KY-IN	0.041***	-0.028	0.029	0.039***	-0.055***	0.096	0.080***	-0.083**	0.094	486
	3.721	-0.877		6.240	-6.454		7.106	-2.466		
Cleveland-Elyria-Mentor, OH	0.055***	-0.110***	0.034	0.019***	-0.037*	0.016	0.074***	-0.147***	0.066	693
	4.299	-2.854		3.046	-1.948		6.020	-4.099		
Dallas-Fort Worth-Arlington, TX	0.019***	0.243***	0.002	0.020***	0.058***	0.034	0.039***	0.300***	0.008	1,046
	3.575	9.315		7.133	6.993		5.632	10.347		
Denver-Aurora, CO	0.145***	0.100**	0.066	0.070***	-0.018	0.081	0.214***	0.080	0.098	520
	5.248	2.553		4.782	-0.954		5.421	1.483		
Detroit-Warren-Livonia, MI	0.049***	-0.096***	0.039	0.022***	-0.040***	0.056	0.071***	-0.137***	0.072	1,289
	5.027	-3.517		7.326	-3.303		7.062	-5.022		
Houston-Sugar Land-Baytown, TX	0.026***	0.161***	0.009	0.013***	0.107***	0.013	0.039***	0.268***	0.016	895
	2.896	5.003		4.705	10.716		3.886	7.678		
Indianapolis, IN	0.052***	0.020	0.059	0.034***	-0.007	0.077	0.086***	0.013	0.095	315
	5.886	0.912		6.326	-0.639		6.486	0.452		
Jacksonville, FL	0.050***	0.039	0.054	0.030***	0.082***	0.070	0.080***	0.122***	0.084	201
	4.930	1.230		4.632	5.237		5.866	3.188		
Las Vegas-Paradise, NV	-0.033	1.740***	-0.002	0.006	0.232***	-0.002	-0.027	1.972***	-0.002	347
	-1.161	11.218		1.061	7.862		-0.833	12.146		
Los Angeles-Long Beach-Santa Ana, CA	0.043***	-0.025	0.016	0.022***	0.014	0.017	0.065***	-0.011	0.033	2,371
	4.364	-1.438		3.413	1.143		6.066	-0.609		
Miami-Fort Lauderdale-Miami Beach, FL	0.007	0.281***	-0.000	0.010***	0.066***	0.015	0.016**	0.347***	0.002	890
	1.124	5.505		3.663	4.412		2.313	6.201		
Minneapolis-St. Paul-Bloomington, MN-WI	0.026	0.123**	0.005	0.070**	-0.047	0.110	0.096***	0.076**	0.075	746
	0.837	2.160		2.537	-0.959		7.558	2.556		
New York-Northern New Jersey-Long Island, NY-NJ-PA	0.005	0.054***	0.000	0.001	0.065***	-0.000	0.006*	0.119***	0.000	4,396
	1.452	3.653		0.464	9.158		1.734	8.340		
Orlando-Kissimmee, FL	0.029	0.279***	-0.001	0.046***	0.101***	0.039	0.075***	0.379***	0.011	328
	1.300	3.561		3.907	3.307		2.901	4.786		
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.028***	-0.053***	0.012	0.021***	-0.012***	0.087	0.049***	-0.065***	0.036	1,556
	4.521	-2.862		11.990	-2.660		7.786	-3.421		
Phoenix-Mesa-Scottsdale, AZ	0.011	0.579***	-0.001	0.007**	0.162***	0.002	0.019	0.741***	0.000	705
	0.710	10.550		2.112	11.883		1.014	12.190		
Pittsburgh, PA	0.017***	-0.078***	0.033	0.009***	-0.046***	0.071	0.026***	-0.124***	0.060	721
	5.512	-6.534		7.903	-11.178		7.794	-9.652		
Portland-Vancouver-Beaverton, OR-WA	0.028***	0.200***	0.008	0.016***	0.073***	0.029	0.043***	0.273***	0.016	426
	3.063	8.473		3.608	8.239		3.469	9.260		
Richmond, VA	0.036***	0.068**	0.025	0.024***	0.032***	0.109	0.060***	0.100***	0.056	283
	4.853	2.369		5.594	3.019		6.037	3.065		
Riverside-San Bernardino-Ontario, CA	0.019**	0.273***	0.001	0.008***	0.139***	0.007	0.027***	0.412***	0.004	840
	2.532	6.888		2.685	10.994		3.175	9.268		
Sacramento-Arden-Arcade-Roseville, CA	-0.005	0.420***	-0.002	0.005*	0.119***	-0.000	0.001	0.539***	-0.002	403
	-0.640	5.614		1.908	6.035		0.078	6.645		

St. Louis, MO-IL	0.058***	-0.137***	0.125	0.024***	-0.025**	0.112	0.082***	-0.162***	0.196	550
	7.933	-5.354		6.063	-2.461		9.575	-7.177		
San Antonio, TX	0.063***	0.103***	0.038	0.033***	0.059***	0.054	0.097***	0.162***	0.056	338
	6.009	3.360		5.747	4.805		6.311	4.248		
San Diego-Carlsbad-San Marcos, CA	0.039**	0.161***	0.003	0.010***	0.053***	0.005	0.049***	0.214***	0.005	599
	2.346	3.043		2.764	3.772		2.837	3.730		
San Francisco-Oakland-Fremont, CA	0.032***	0.041*	0.015	0.009*	0.021*	0.007	0.040***	0.061**	0.018	871
	3.990	1.824		1.755	1.690		3.799	2.217		
Seattle-Tacoma-Bellevue, WA	0.036***	0.092***	0.043	0.014***	0.043***	0.025	0.050***	0.135***	0.055	664
	7.013	5.226		4.333	4.071		6.994	5.974		
Tampa-St. Petersburg-Clearwater, FL	-0.005	0.213***	-0.002	0.002	0.121***	-0.002	-0.003	0.334***	-0.002	547
	-0.352	3.763		0.323	7.445		-0.199	5.248		
Virginia Beach-Norfolk-Newport News, VA-NC	0.041***	-0.042	0.024	0.030***	-0.026*	0.108	0.071***	-0.068*	0.056	368
	3.818	-1.303		5.471	-1.803		5.228	-1.739		
Washington-Arlington-Alexandria, DC-VA-MD-WV	0.060***	-0.006	0.082	0.033***	0.021***	0.120	0.094***	0.015	0.127	1,004
	8.724	-0.428		9.365	2.899		10.357	0.861		
All metropolitan Areas	0.021***	0.111***	0.004	0.015***	0.042***	0.015	0.036***	0.153***	0.011	29,475
	12.526	18.130		19.290	18.258		19.556	23.341		

Figure 35: Density Histograms 1990, 2000, 2007

Atlanta-Sandy Springs-Marietta, GA

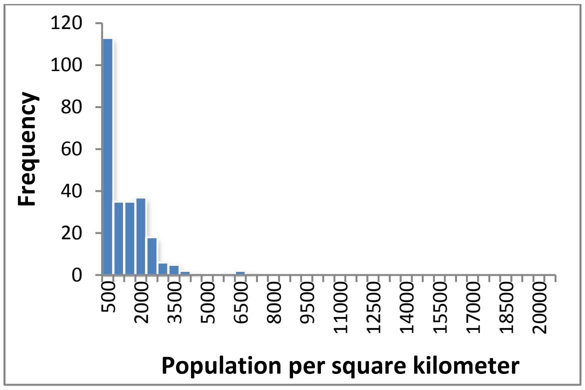


1990 | 1990-2000

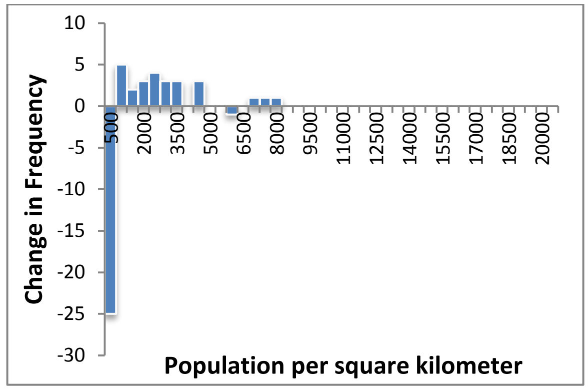
2000 | 2000-2007

2007 | 1990-2007

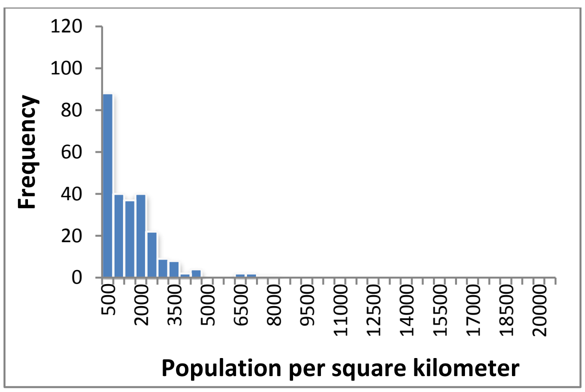
Austin-Round Rock, TX



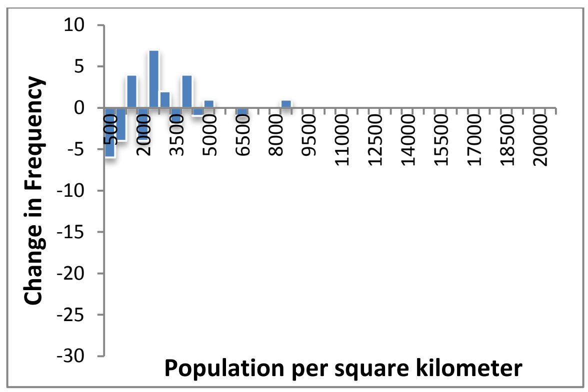
1990



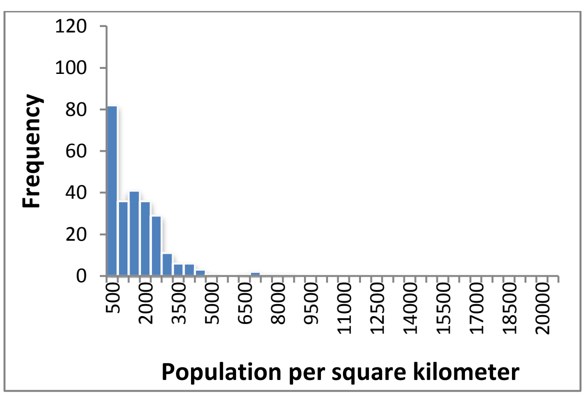
1990-2000



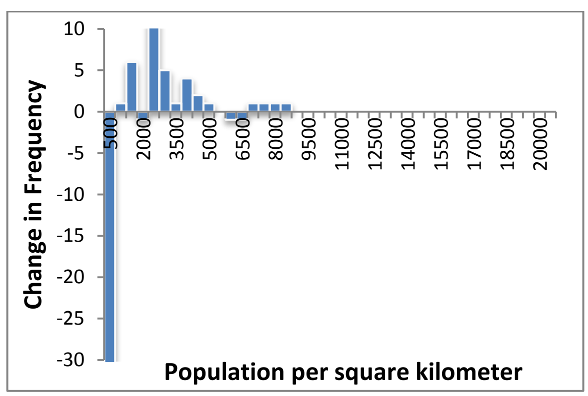
2000



2000-2007

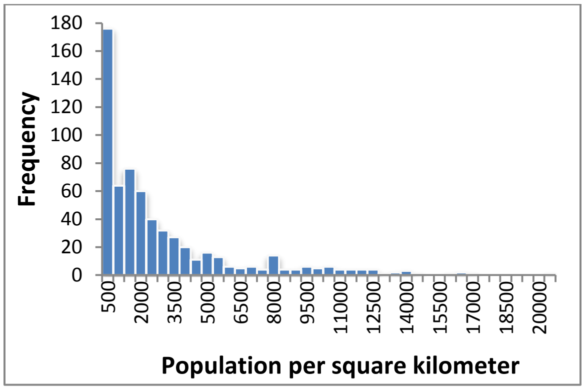


2007

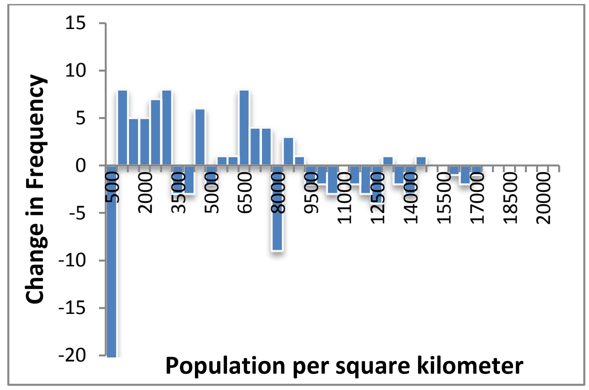


1990-2007

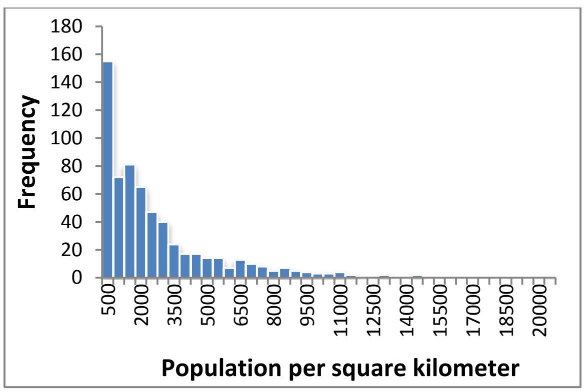
Baltimore-Towson, MD



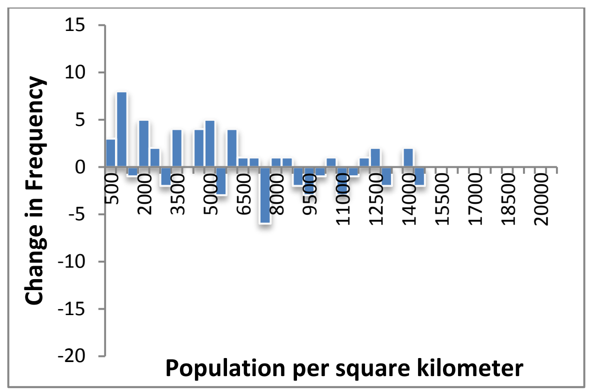
1990



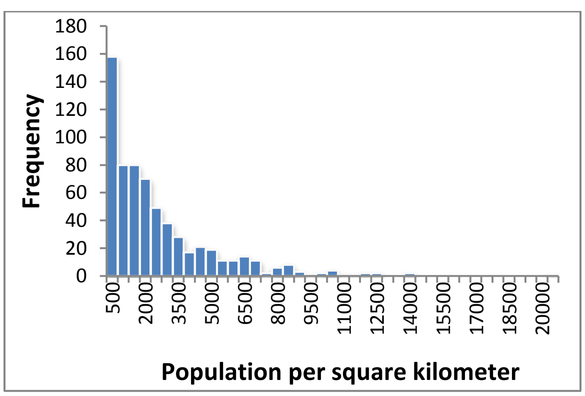
1990-2000



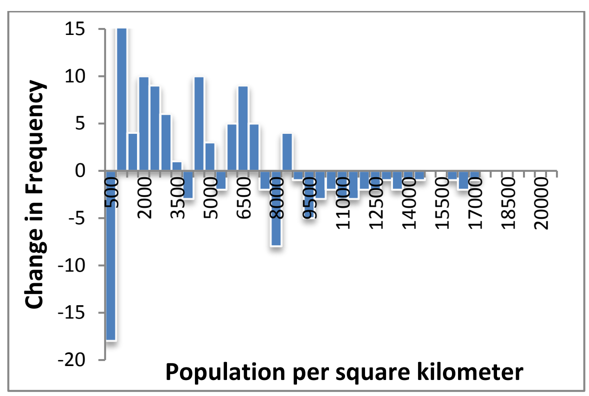
2000



2000-2007

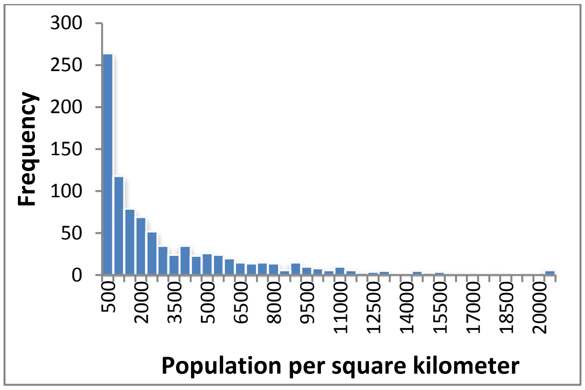


2007

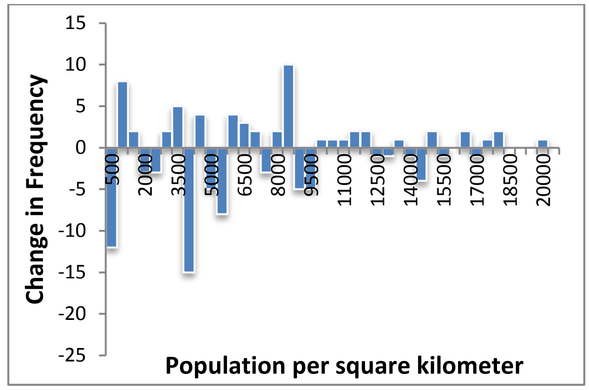


1990-2007

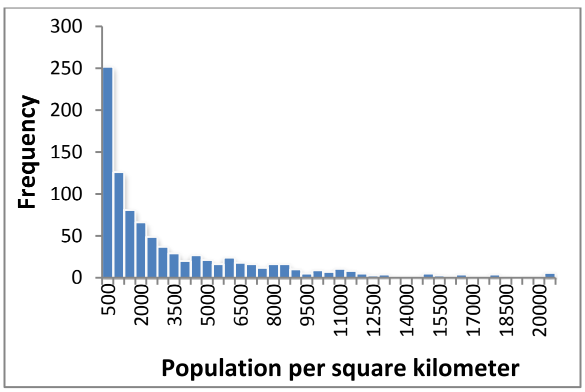
Boston-Cambridge-Quincy, MA-NH



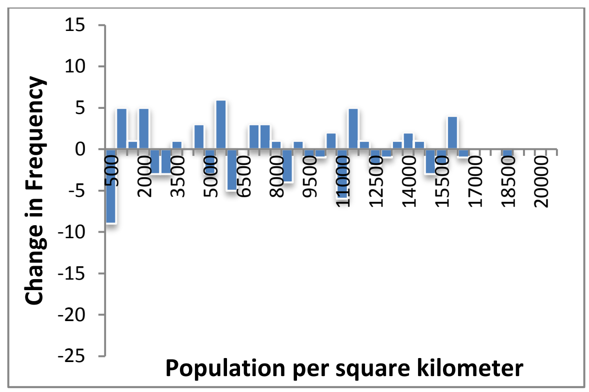
1990



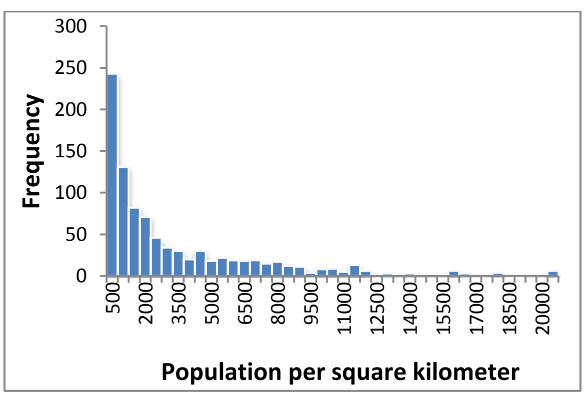
1990-2000



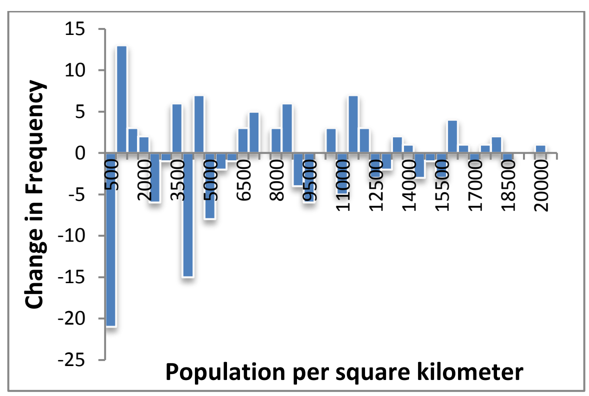
2000



2000-2007

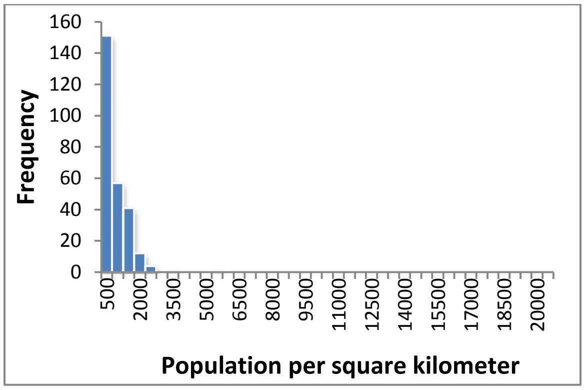


2007

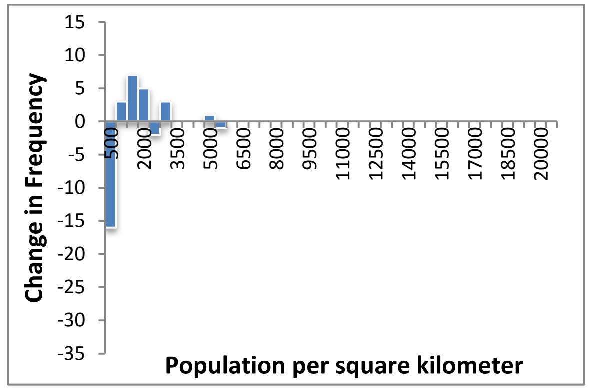


1990-2007

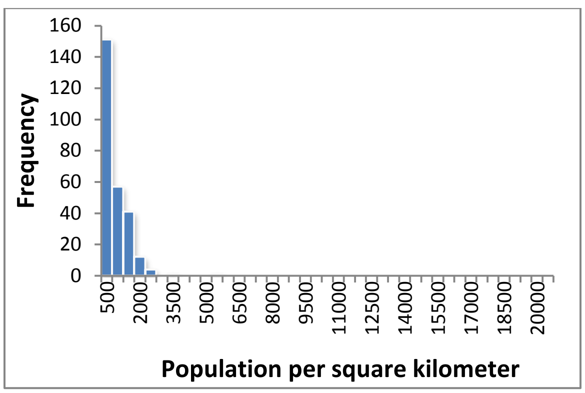
Charlotte-Gastonia-Concord, NC-SC



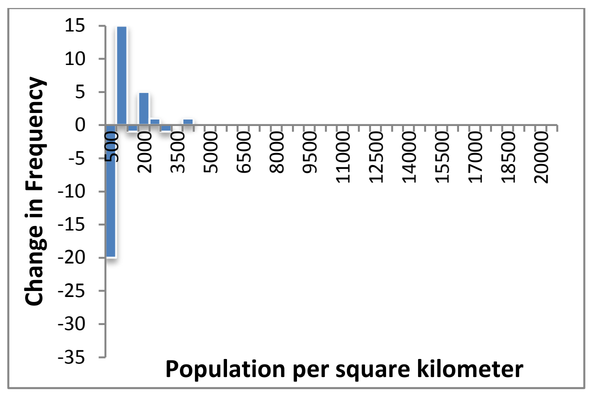
1990



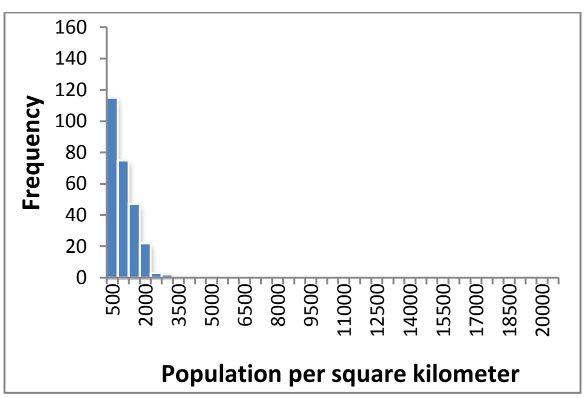
1990-2000



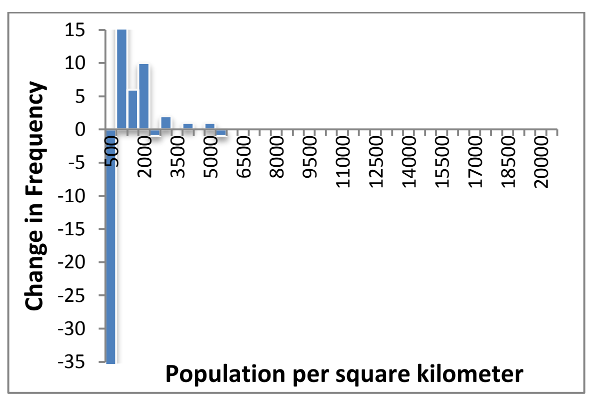
2000



2000-2007

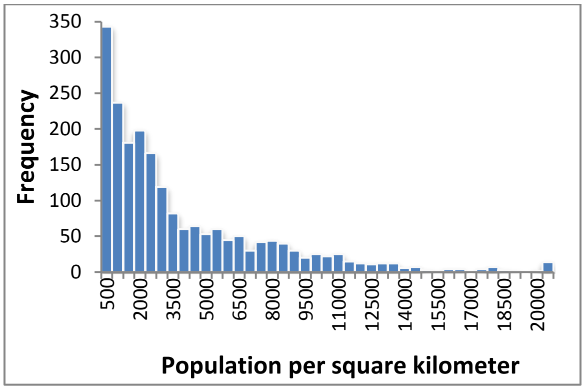


2007

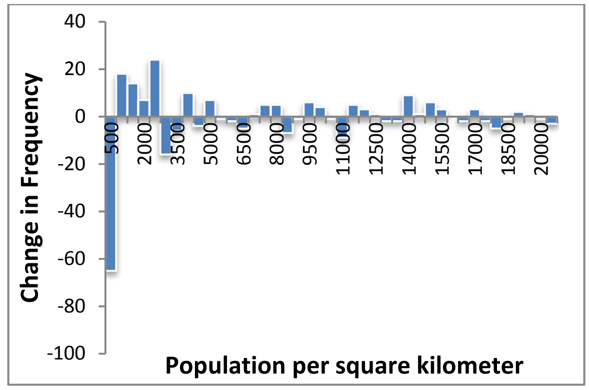


1990-2007

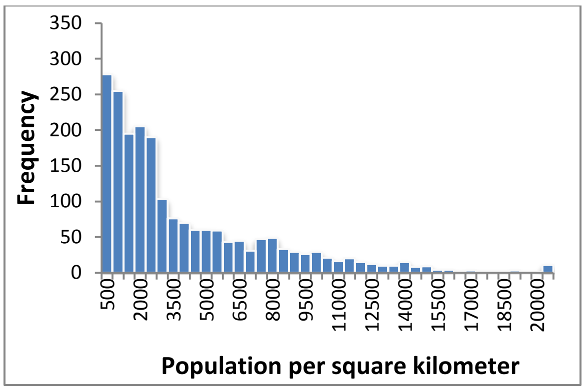
Chicago-Naperville-Joliet, IL-IN-WI



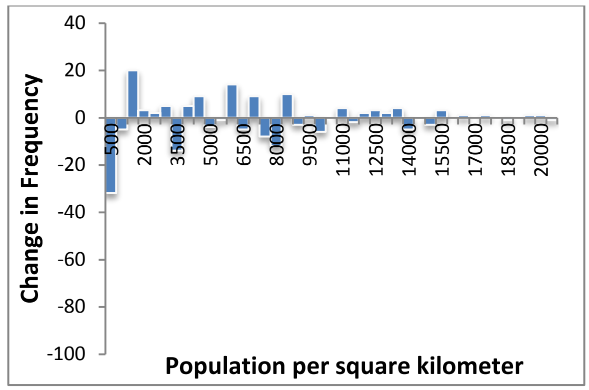
1990



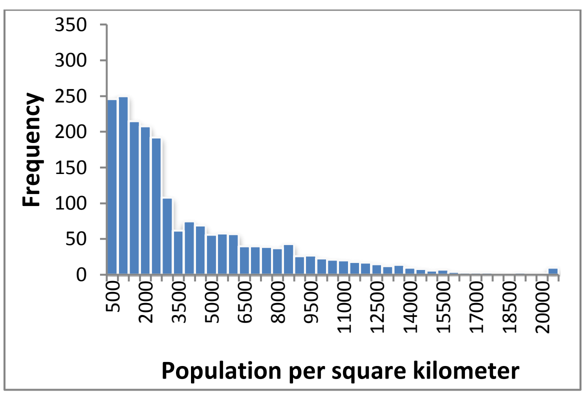
1990-2000



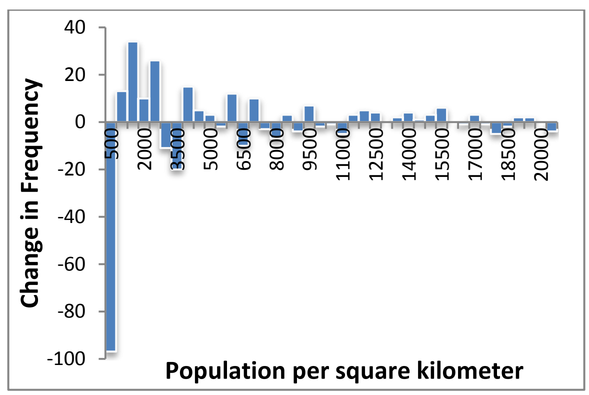
2000



2000-2007

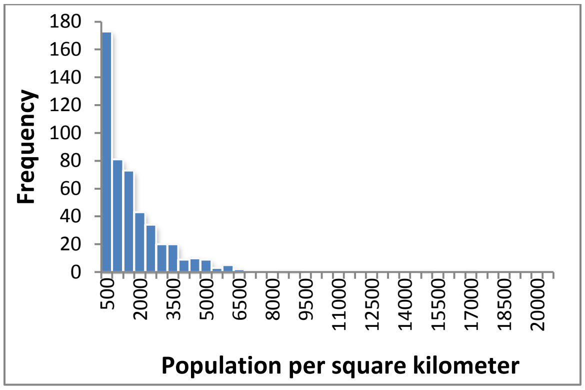


2007

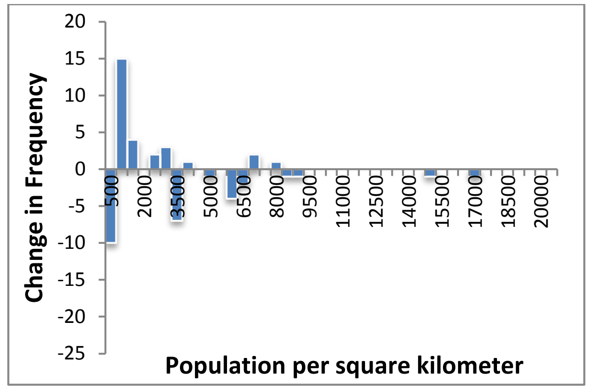


1990-2007

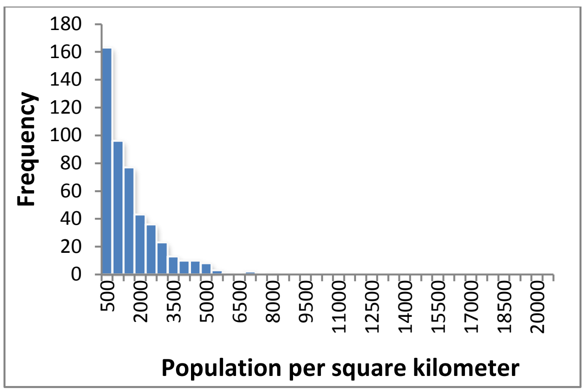
Cincinnati-Middletown, OH-KY-IN



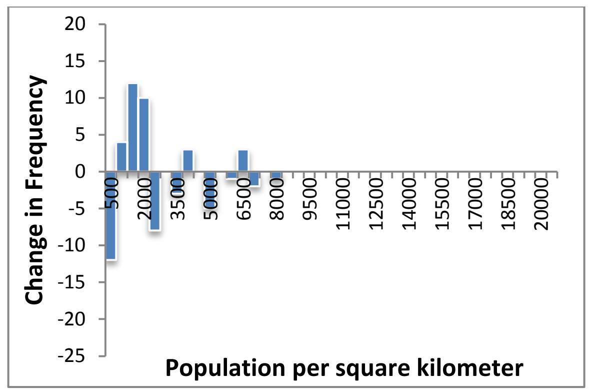
1990



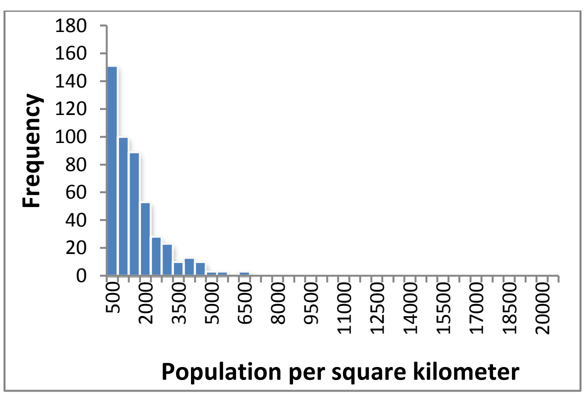
1990-2000



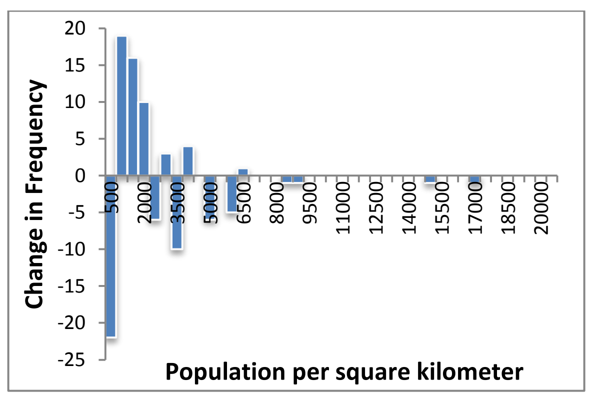
2000



2000-2007

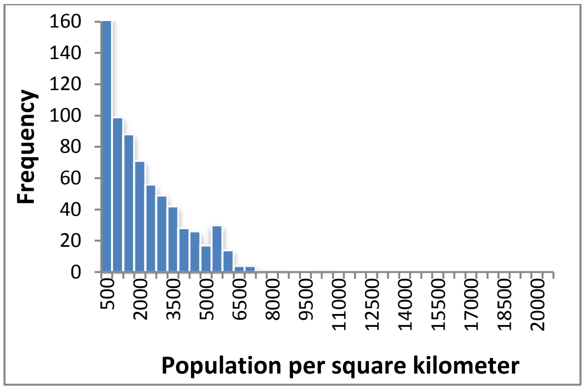


2007

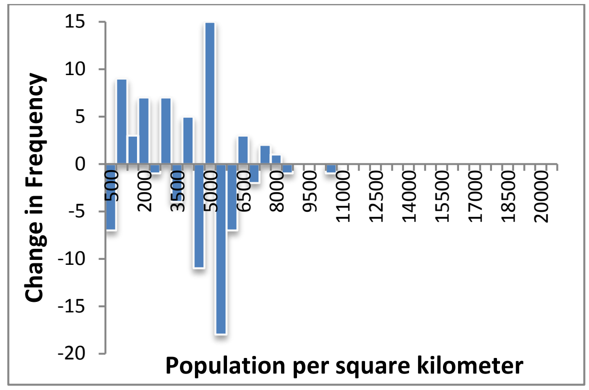


1990-2007

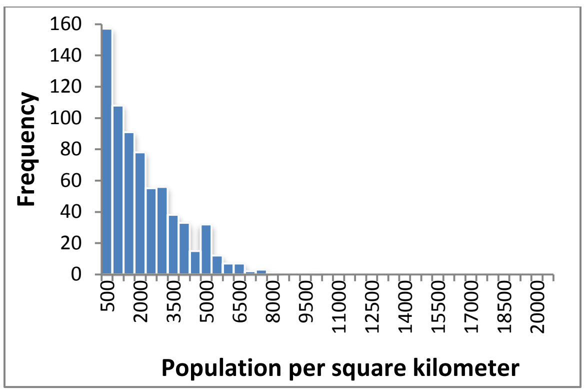
Cleveland-Elyria-Mentor, OH



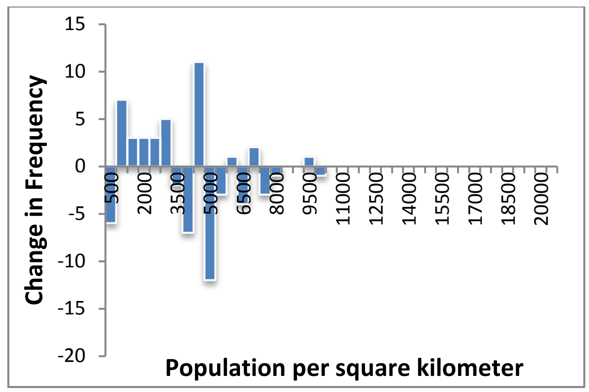
1990



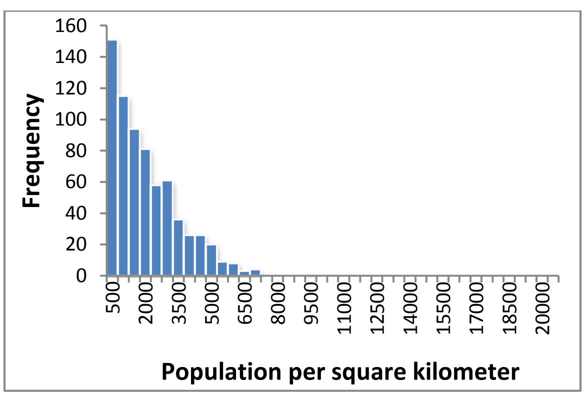
1990-2000



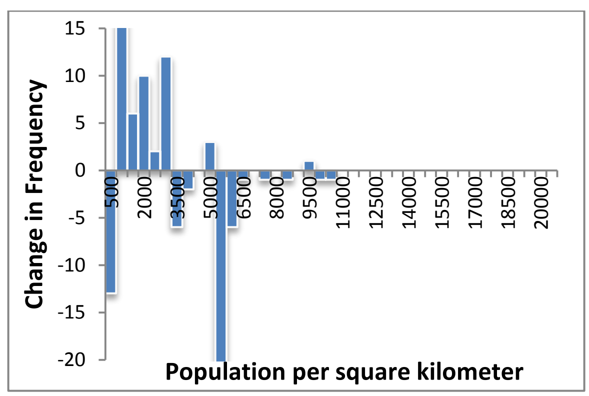
2000



2000-2007

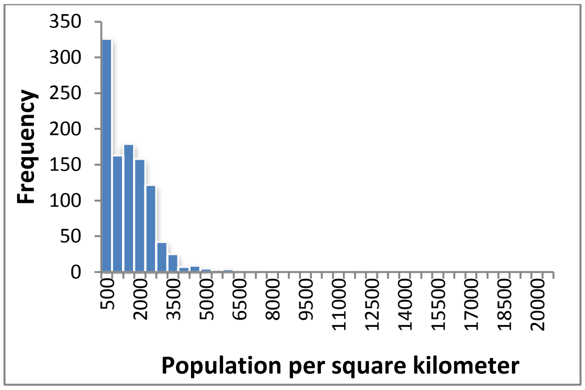


2007

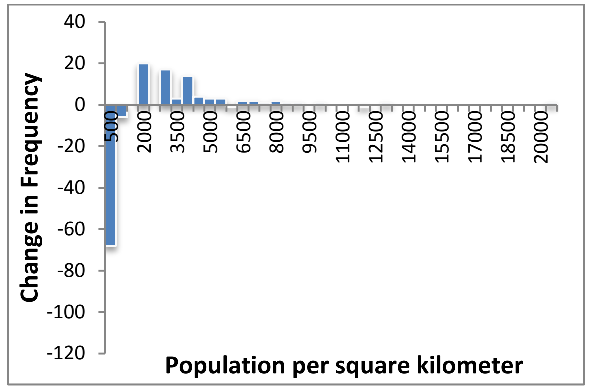


1990-2007

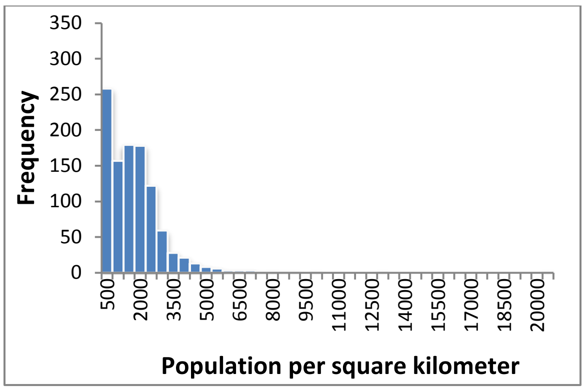
Dallas-Fort Worth-Arlington, TX



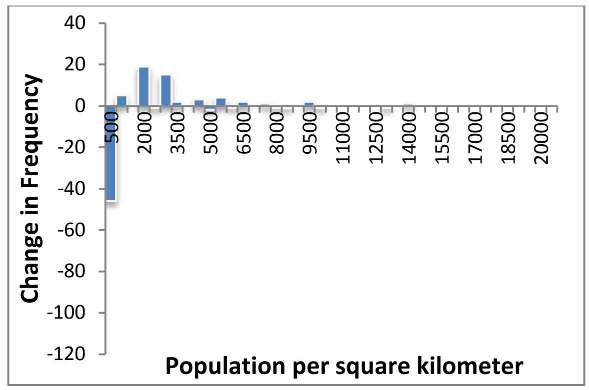
1990



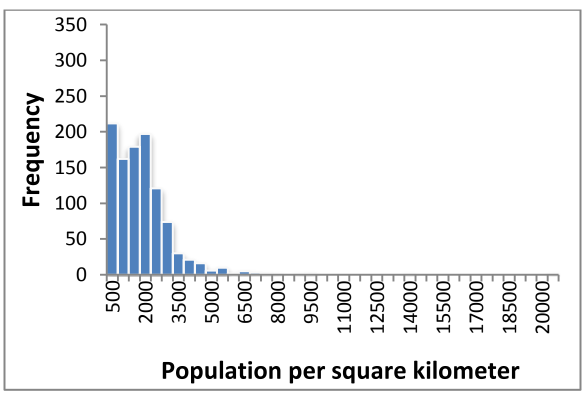
1990-2000



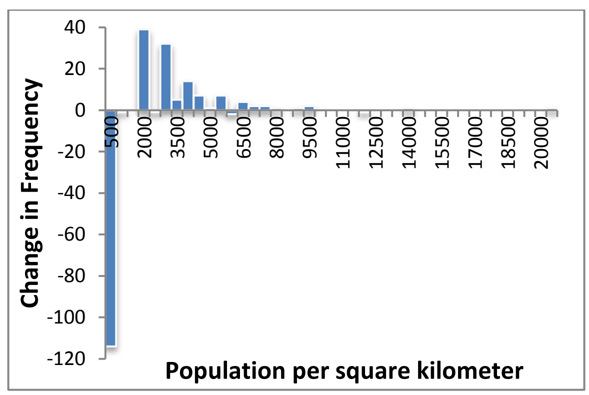
2000



2000-2007

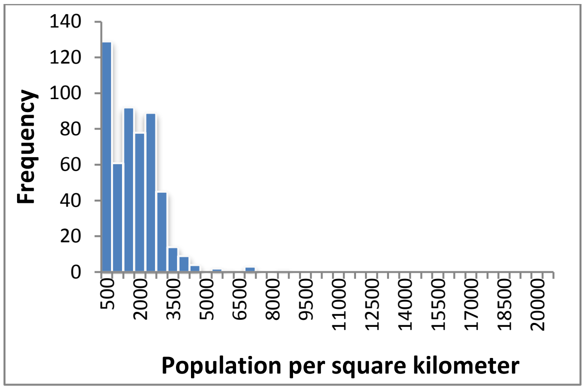


2007

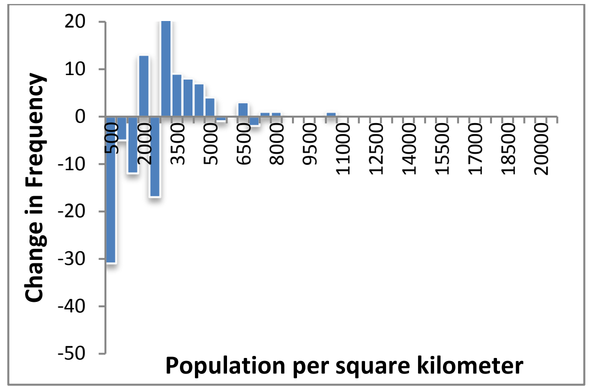


1990-2007

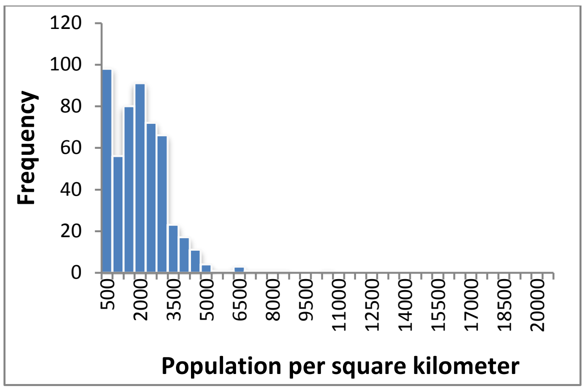
Denver-Aurora, CO



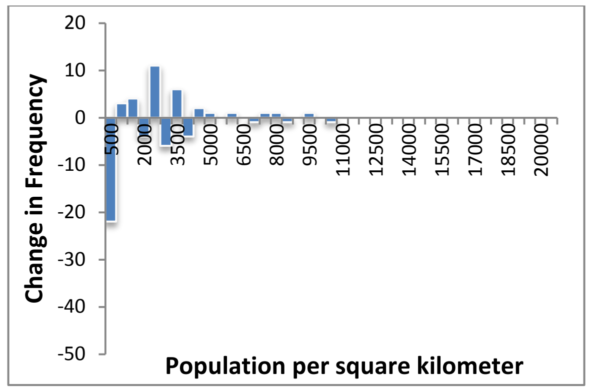
1990



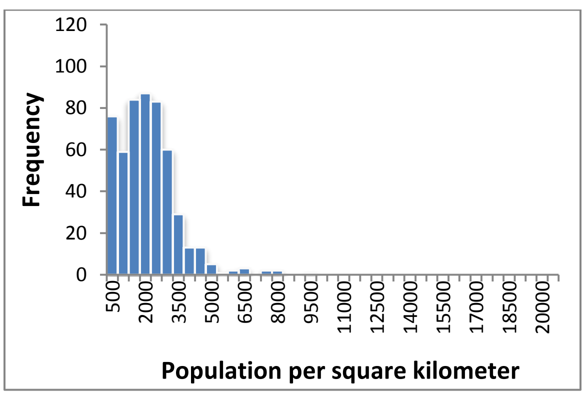
1990-2000



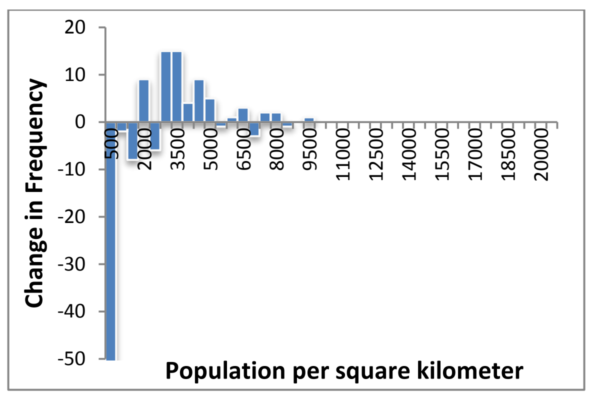
2000



2000-2007

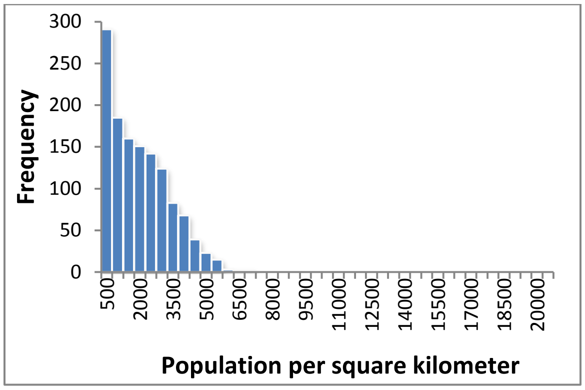


2007

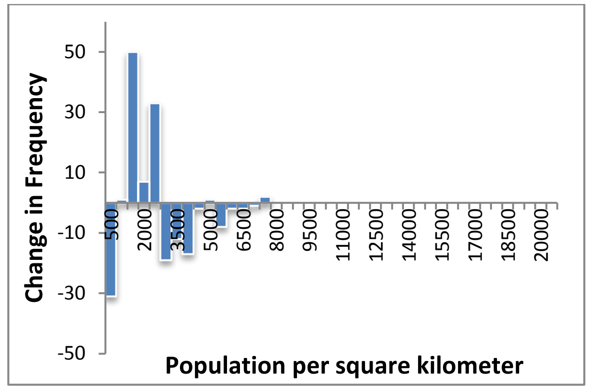


1990-2007

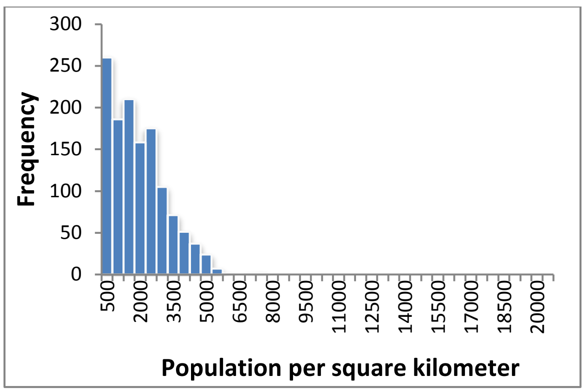
Detroit-Warren-Livonia, MI



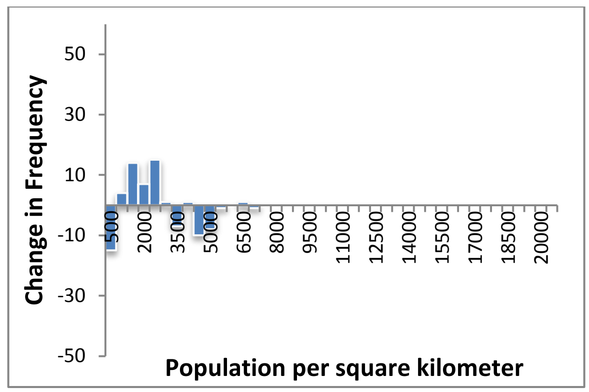
1990



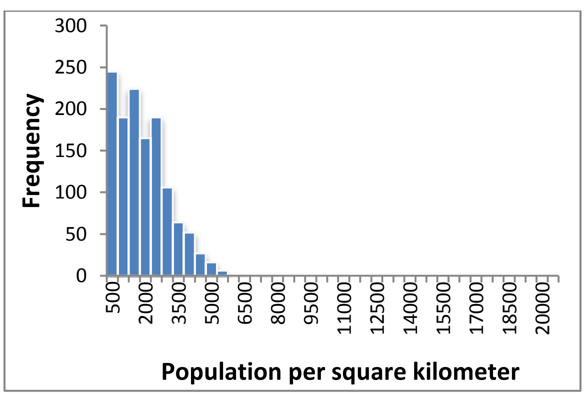
1990-2000



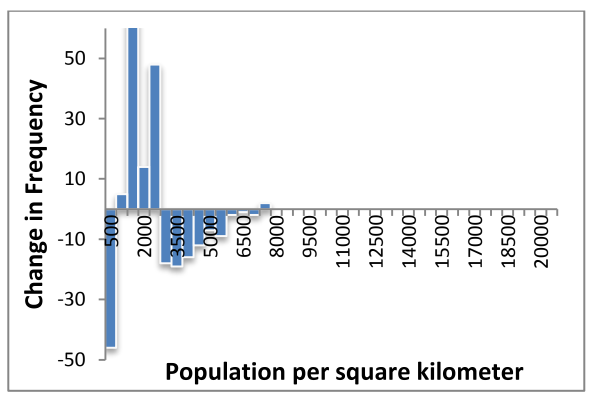
2000



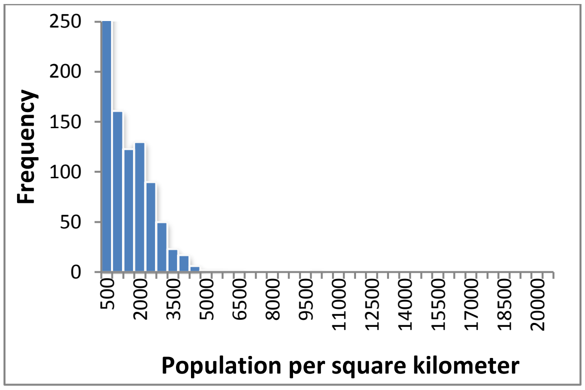
2000-2007



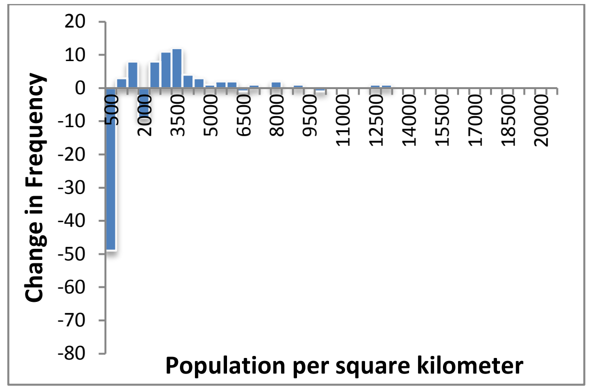
2007



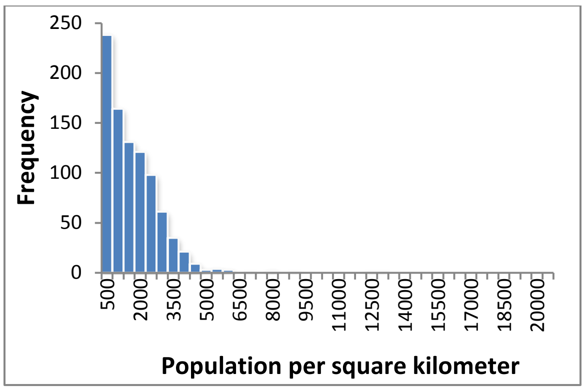
1990-2007



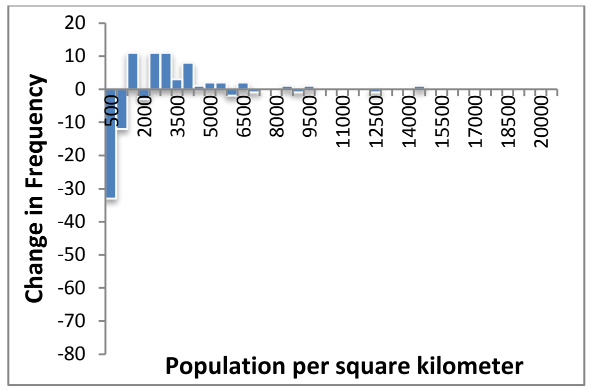
1990



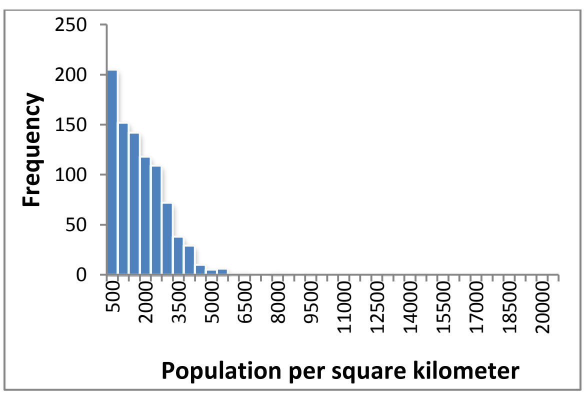
1990-2000



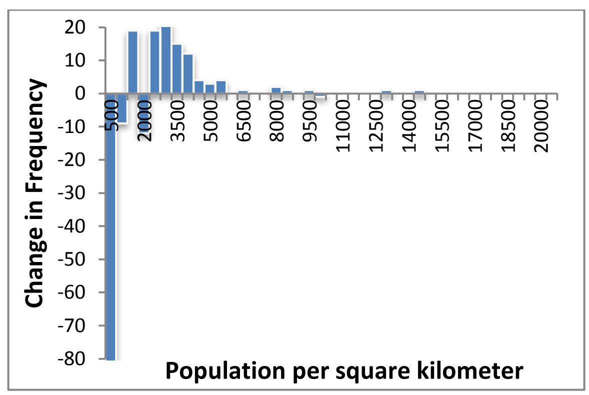
2000



2000-2007

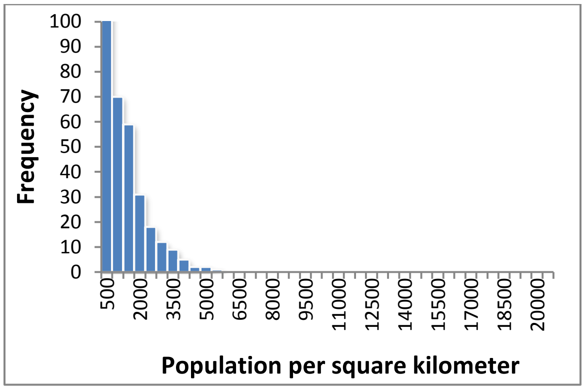


2007

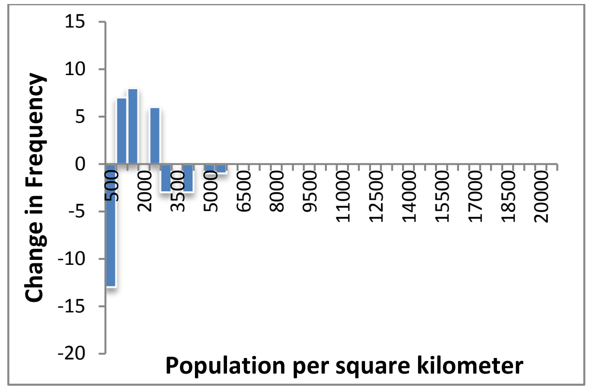


1990-2007

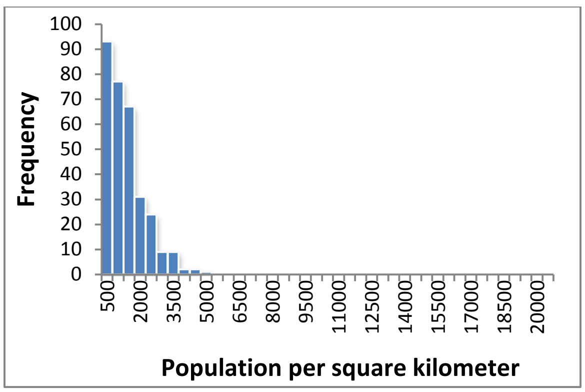
Indianapolis, IN



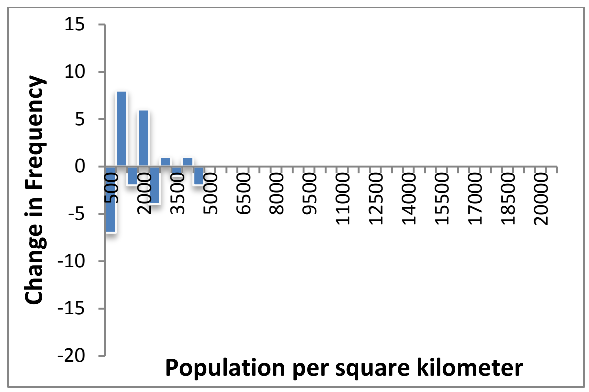
1990



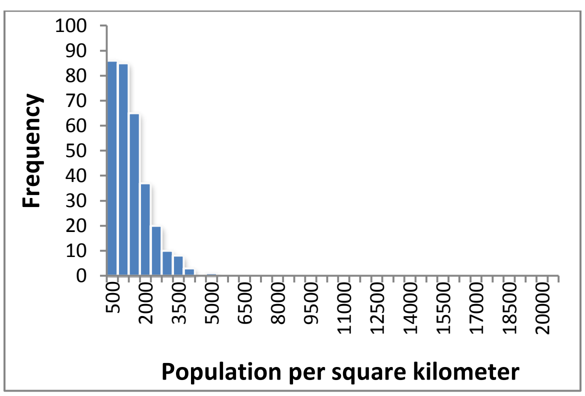
1990-2000



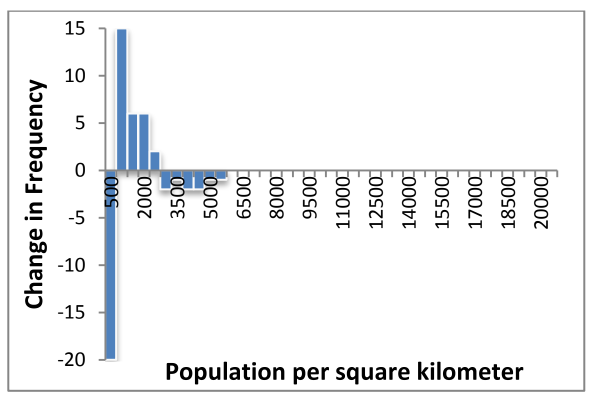
2000



2000-2007

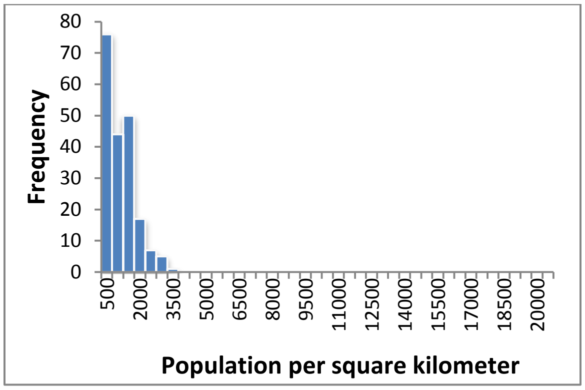


2007

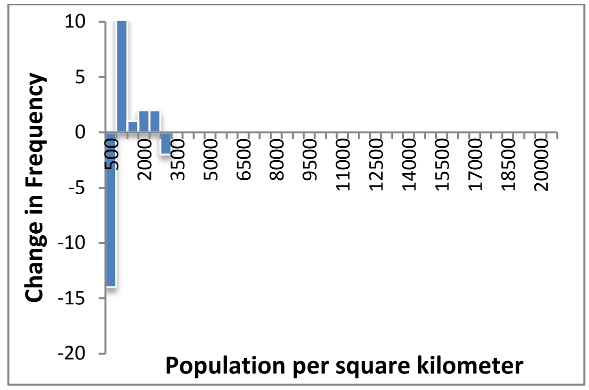


1990-2007

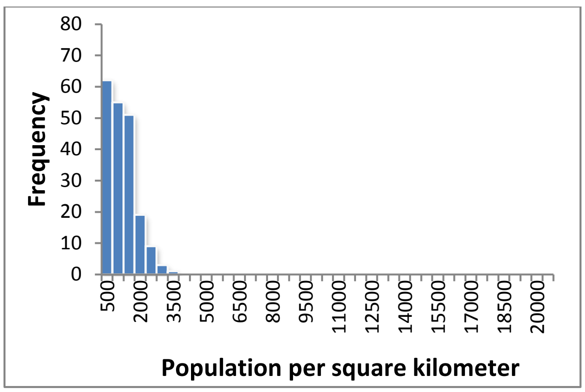
Jacksonville, FL



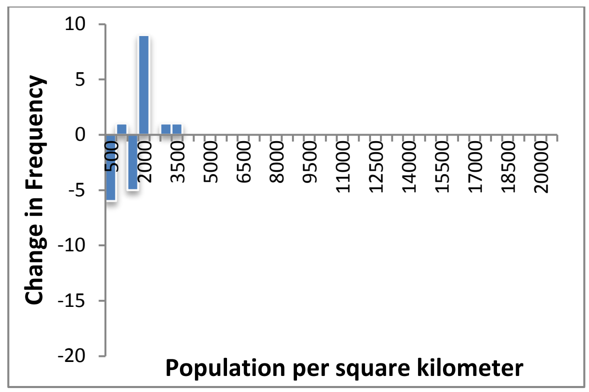
1990



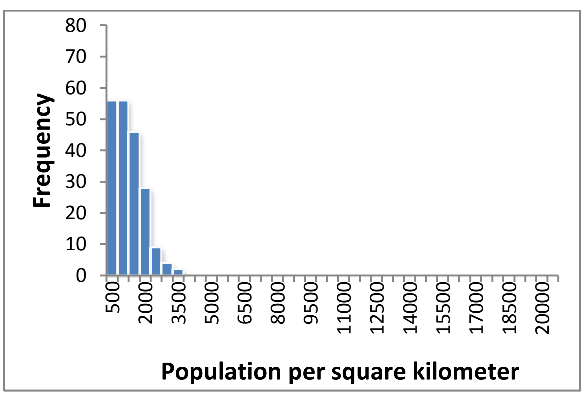
1990-2000



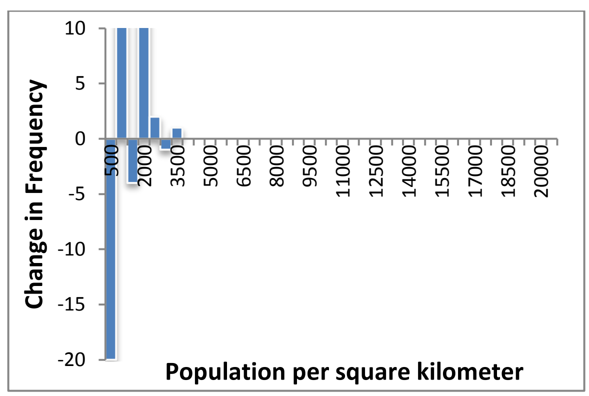
2000



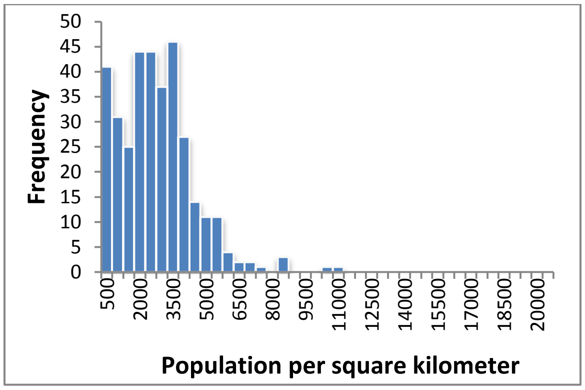
2000-2007



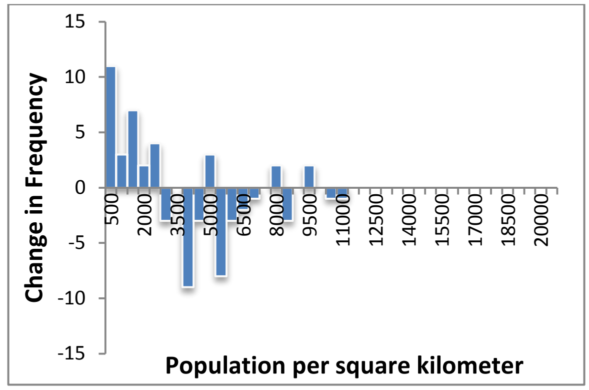
2007



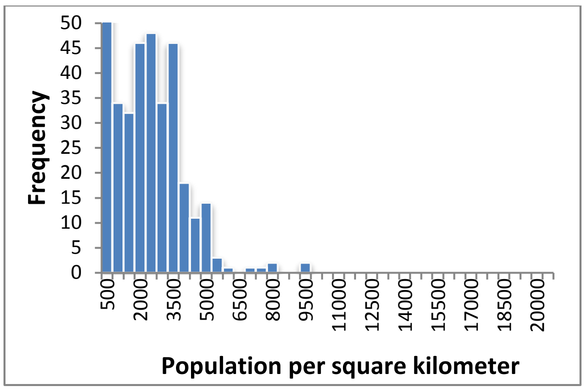
1990-2007



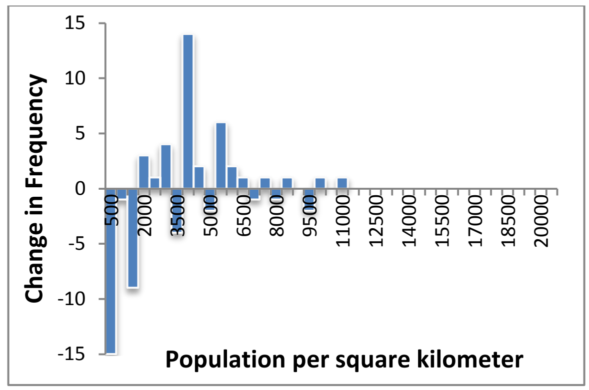
1990



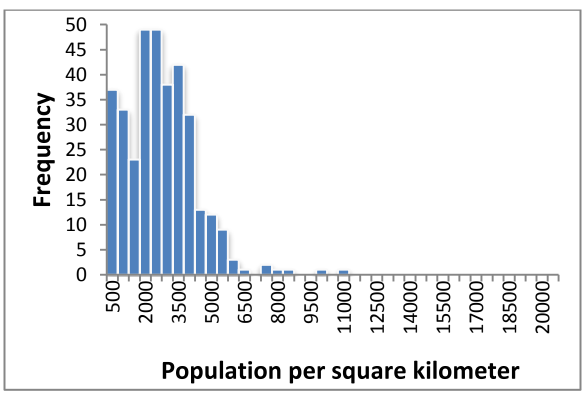
1990-2000



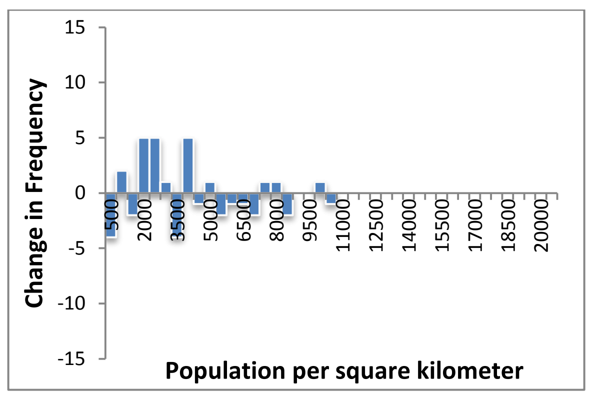
2000



2000-2007

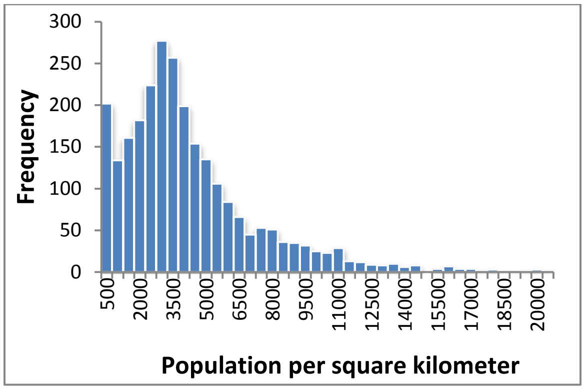


2007

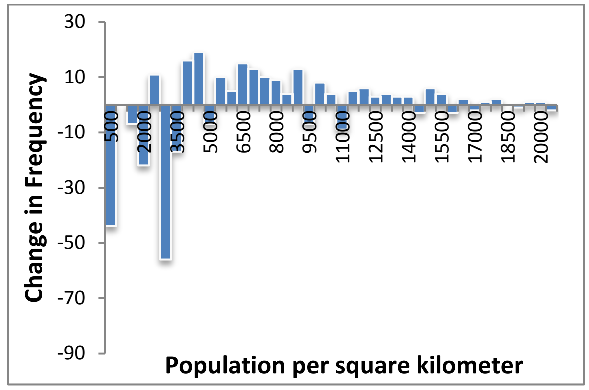


1990-2007

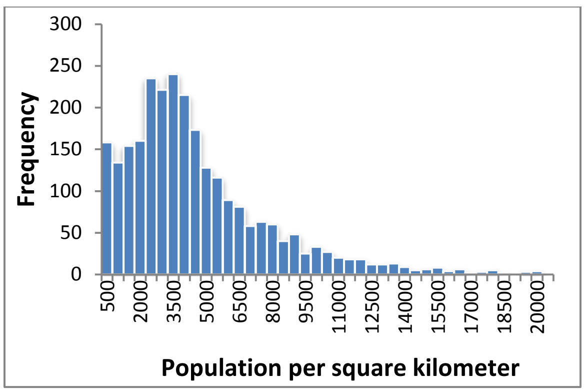
Los Angeles-Long Beach-Santa Ana, CA



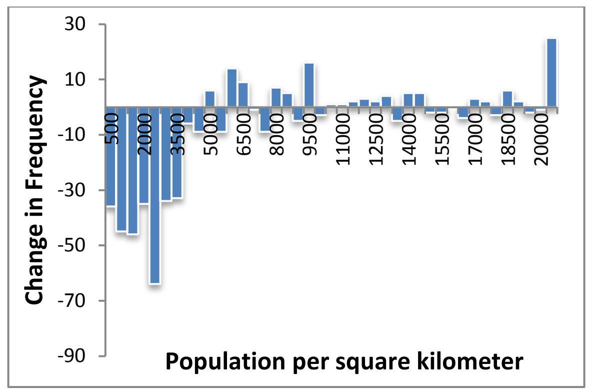
1990



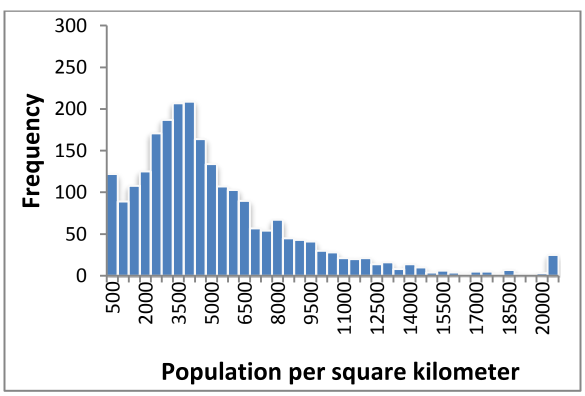
1990-2000



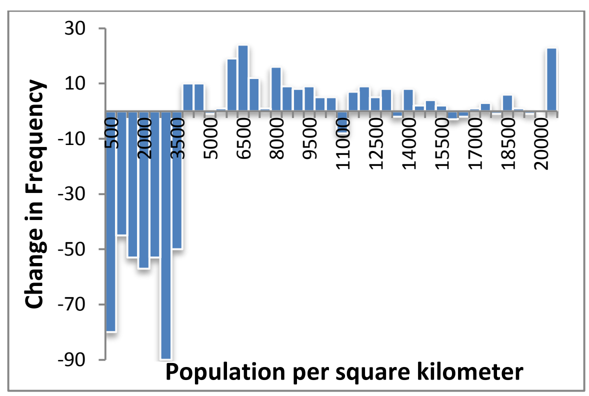
2000



2000-2007

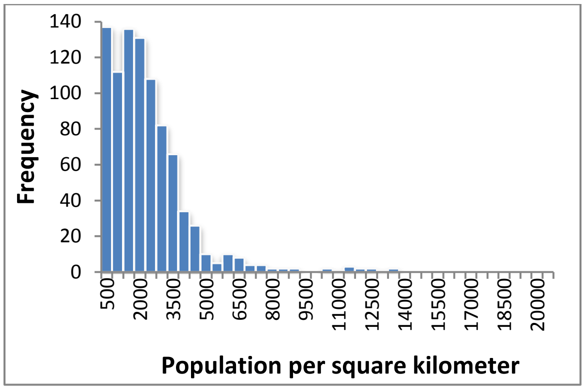


2007

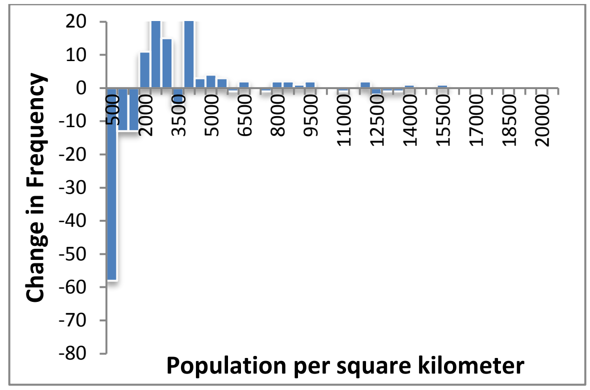


1990-2007

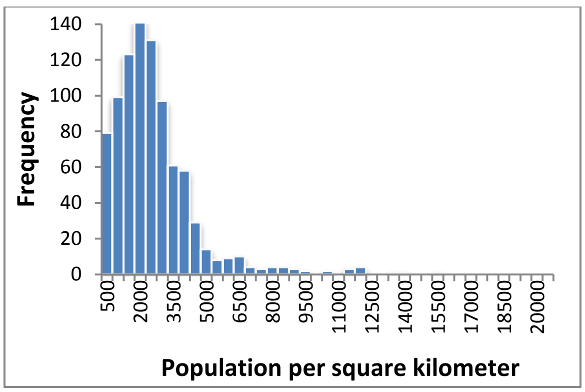
Miami-Fort Lauderdale-Miami Beach, FL



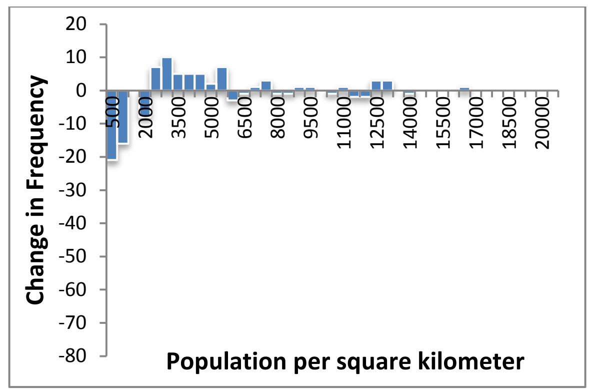
1990



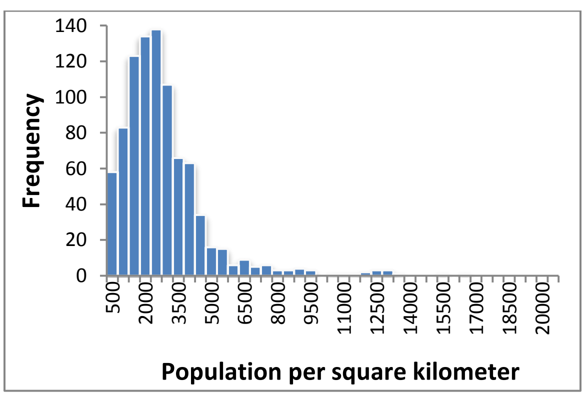
1990-2000



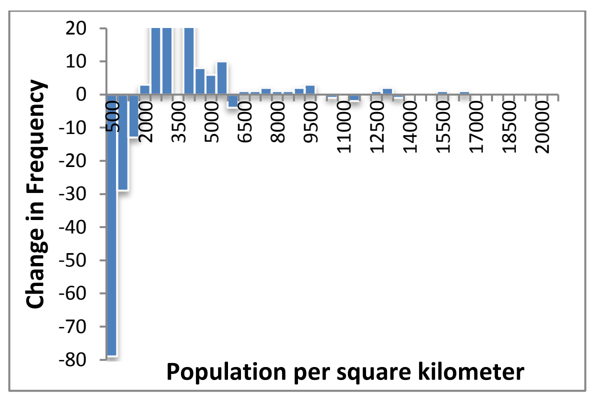
2000



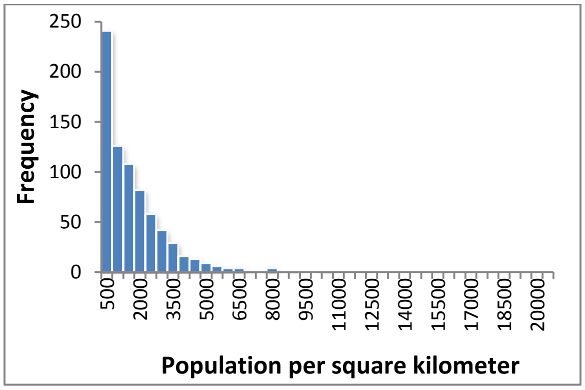
2000-2007



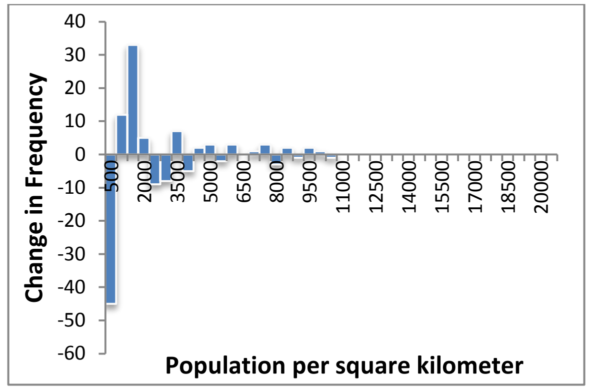
2007



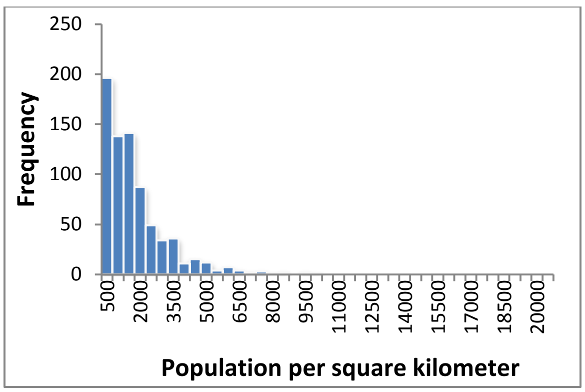
1990-2007



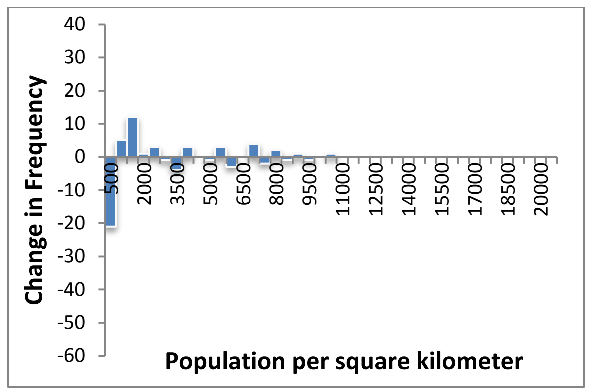
1990



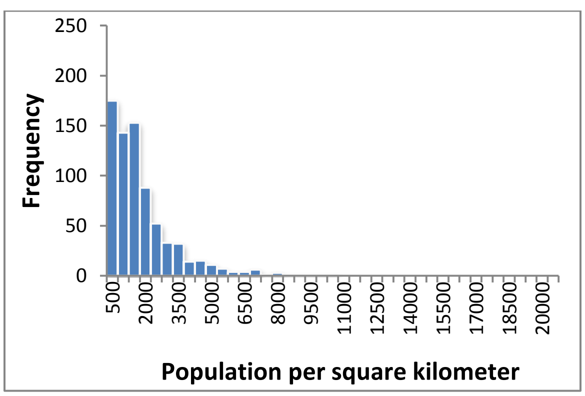
1990-2000



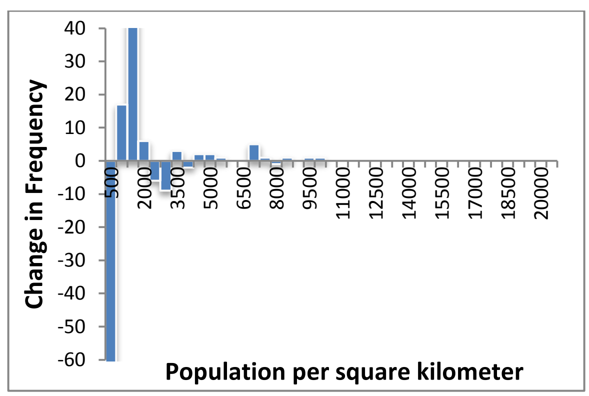
2000



2000-2007

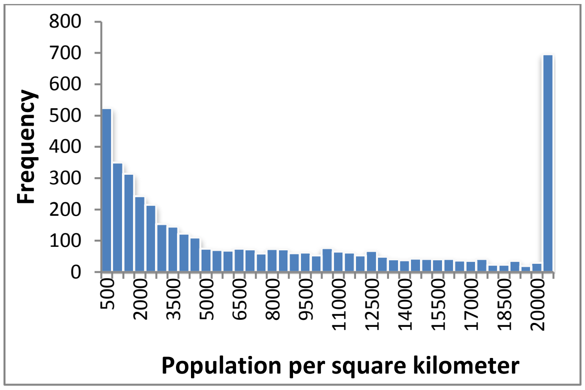


2007

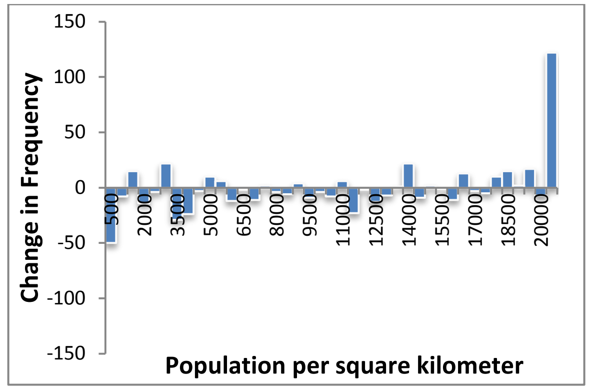


1990-2007

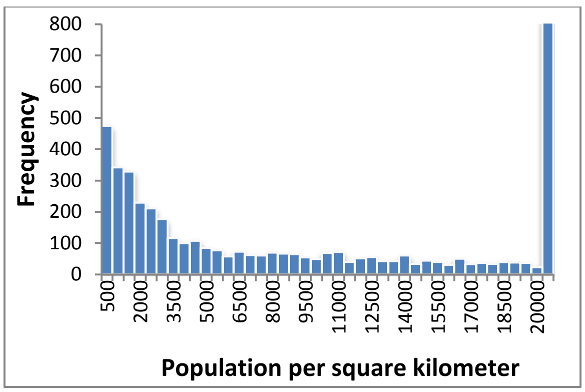
New York-Northern New Jersey-Long Island, NY-NJ-PA



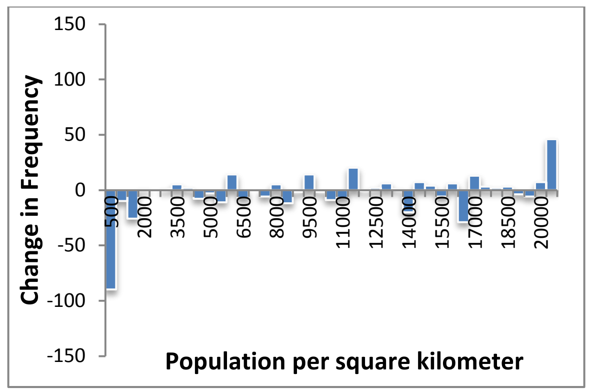
1990



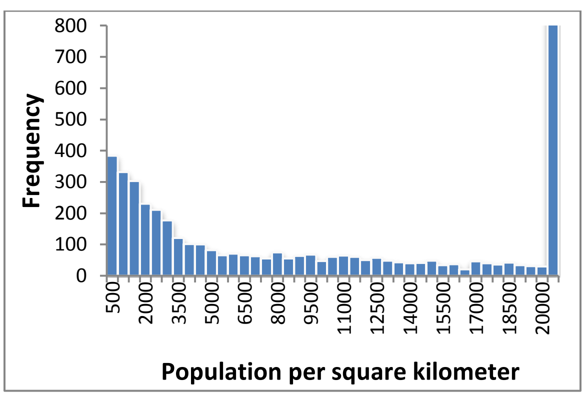
1990-2000



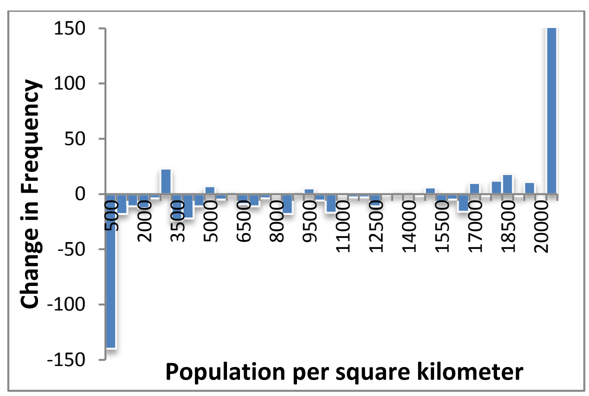
2000



2000-2007

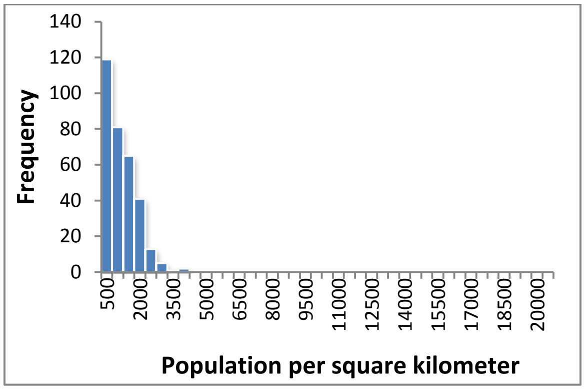


2007

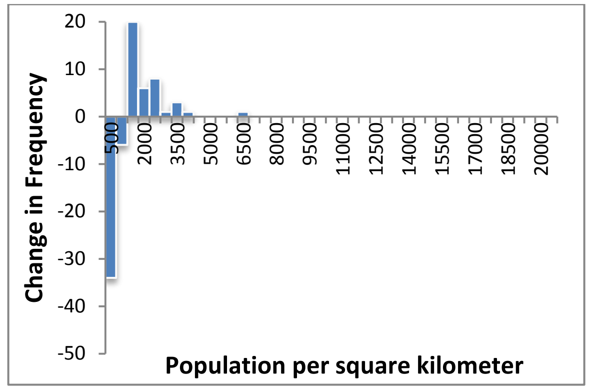


1990-2007

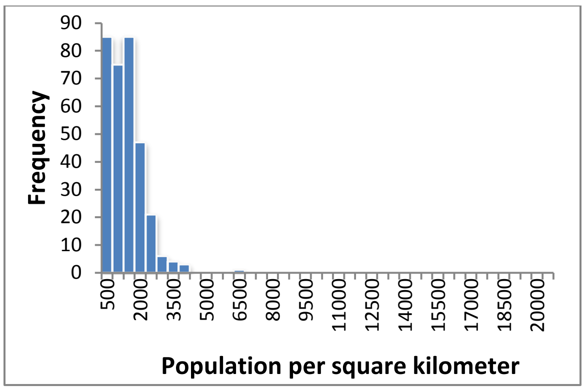
Orlando-Kissimmee, FL



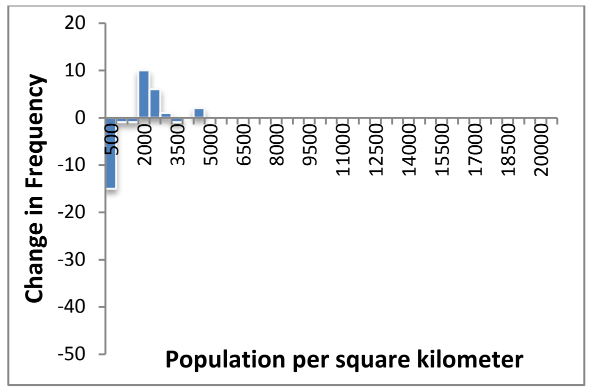
1990



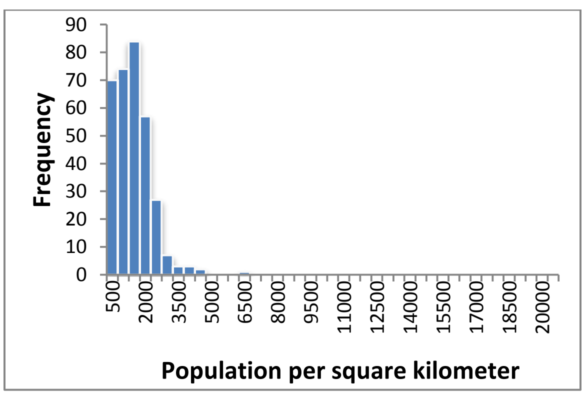
1990-2000



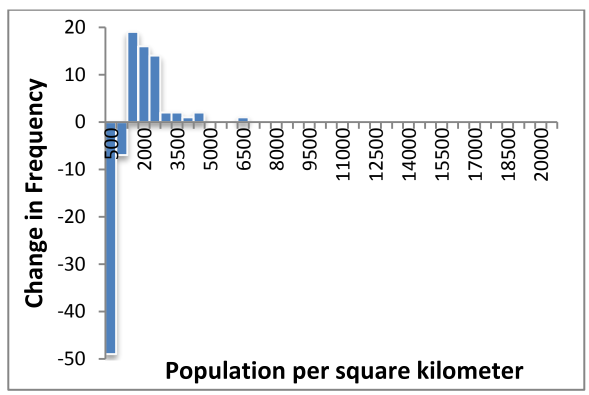
2000



2000-2007

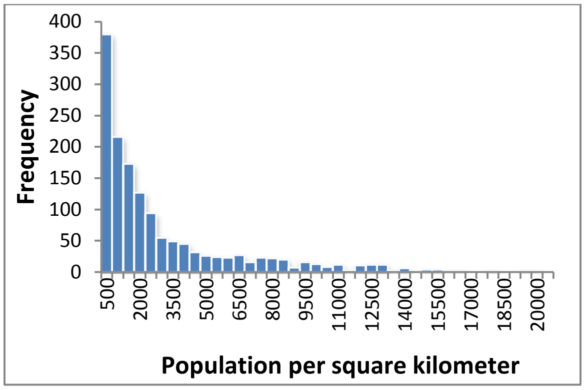


2007

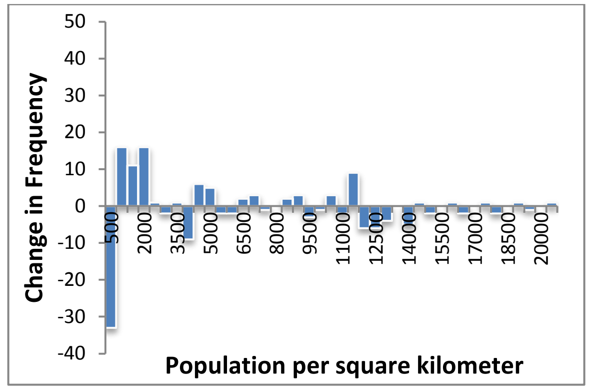


1990-2007

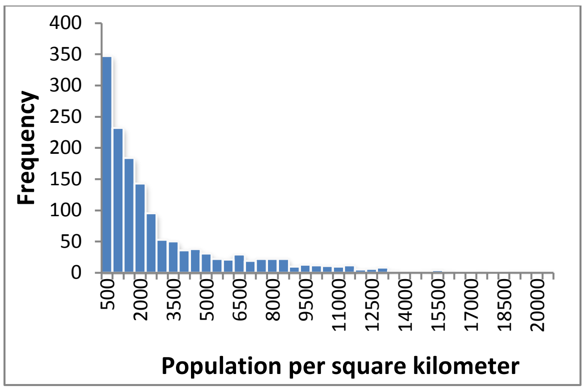
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD



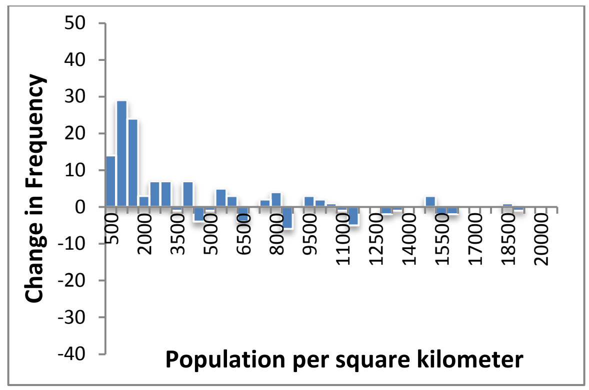
1990



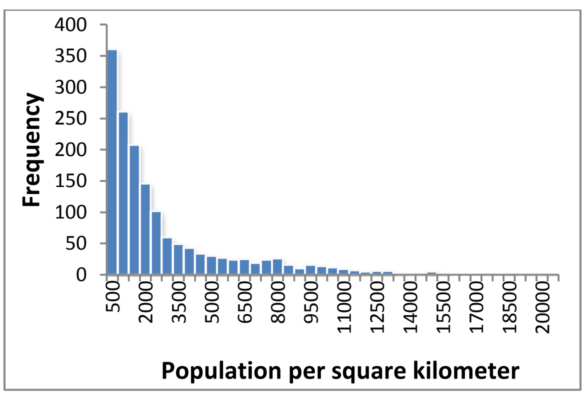
1990-2000



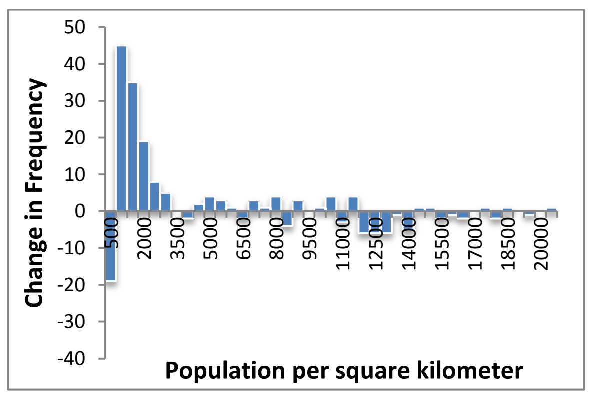
2000



2000-2007

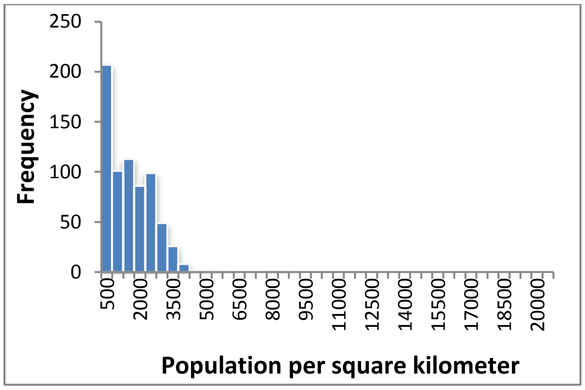


2007

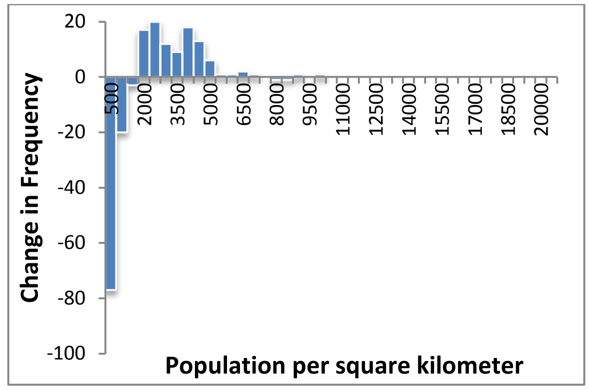


1990-2007

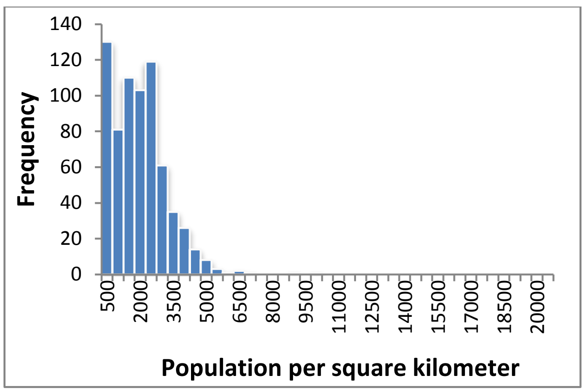
Phoenix-Mesa-Scottsdale, AZ



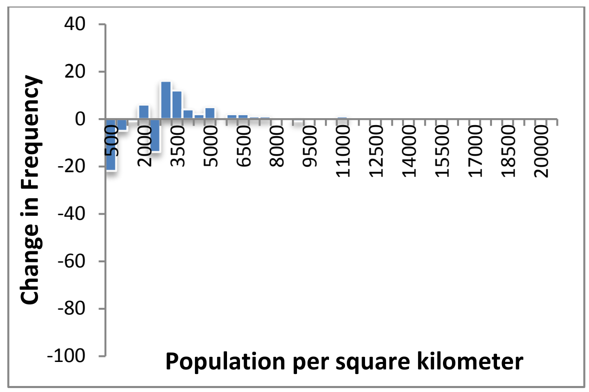
1990



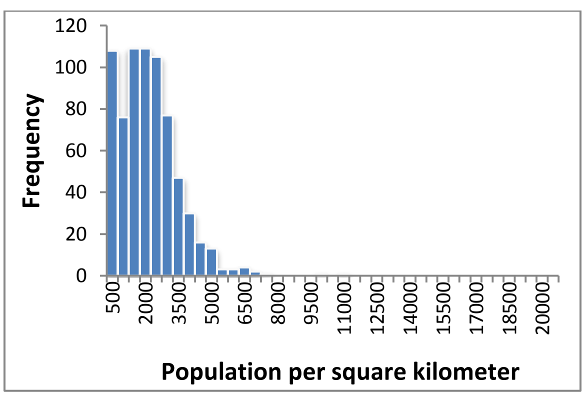
1990-2000



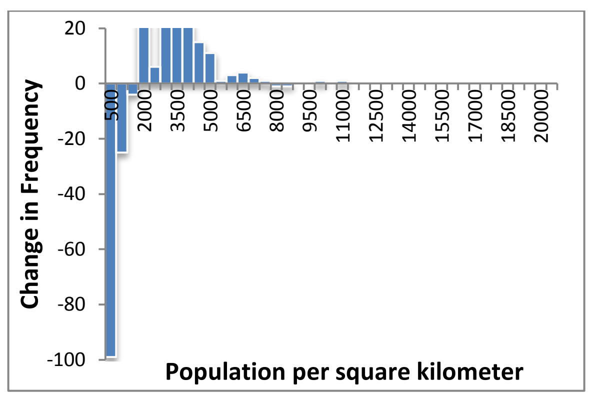
2000



2000-2007

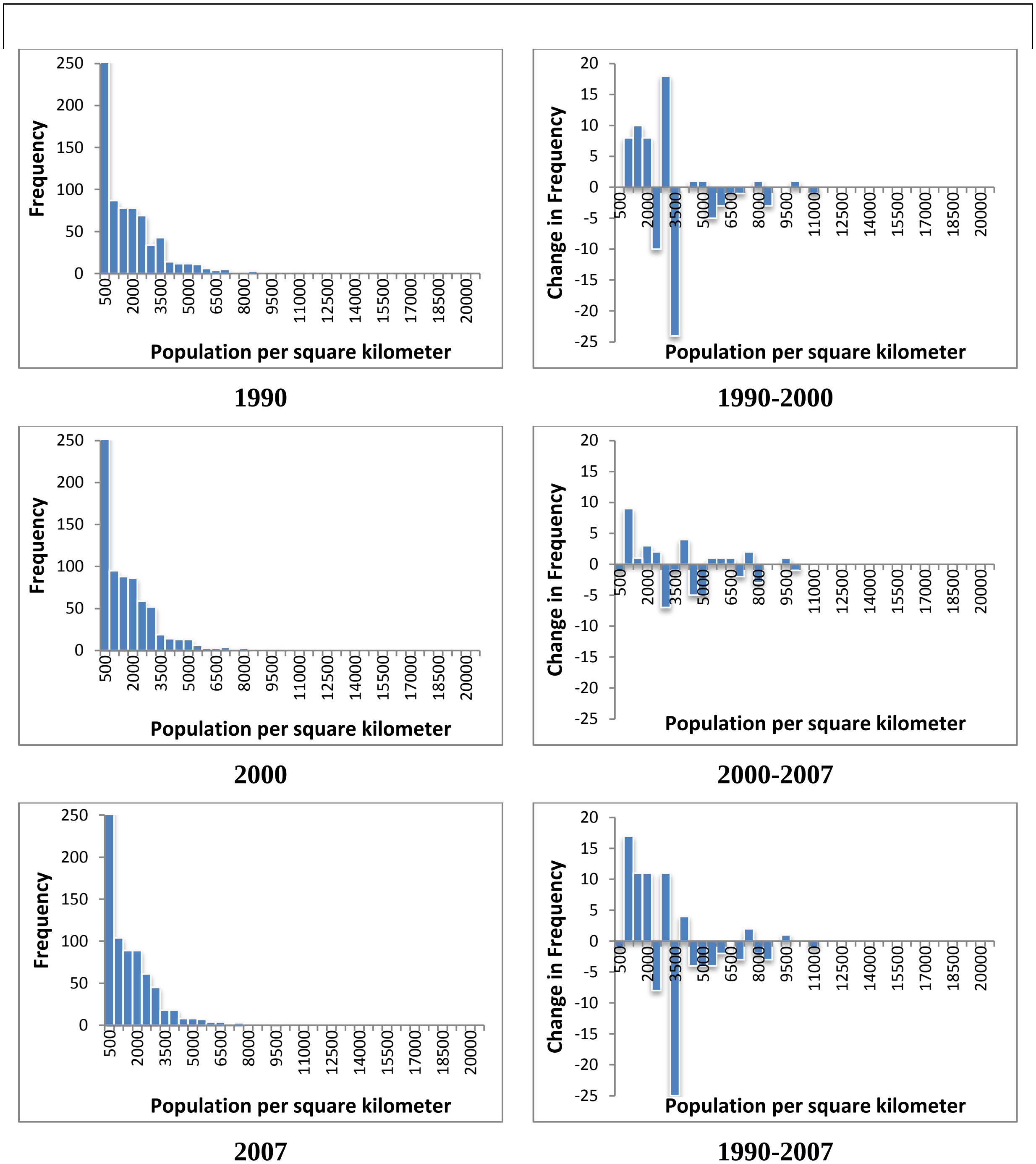


2007



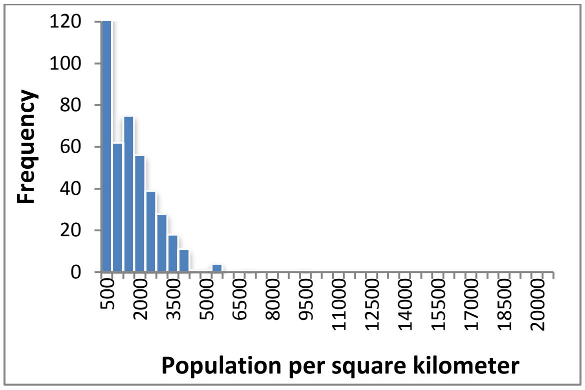
1990-2007

Pittsburgh, PA

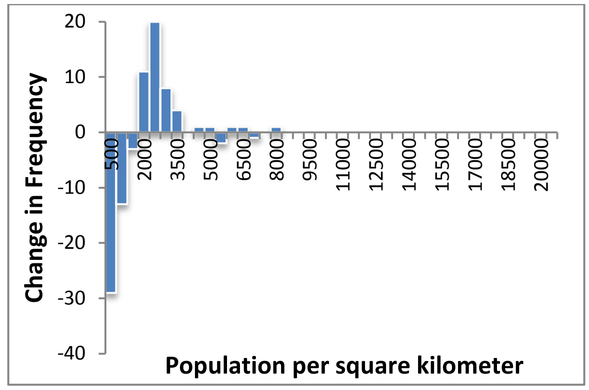


1990 | 1990-2000 | 2000 | 2000-2007 | 2007 | 1990-2007

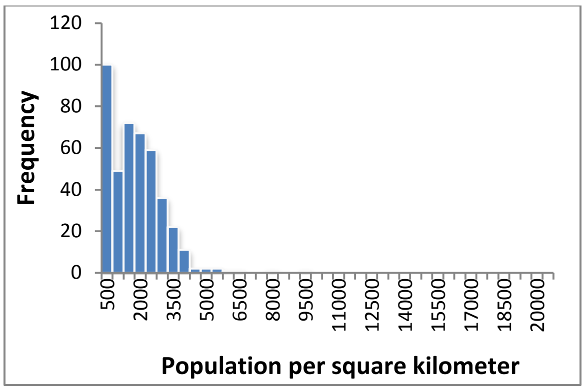
Portland-Vancouver-Beaverton, OR-WA



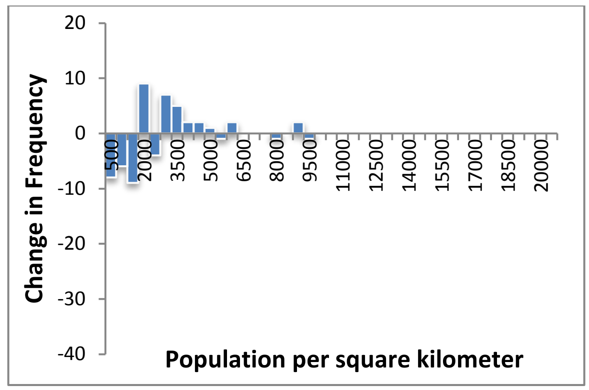
1990



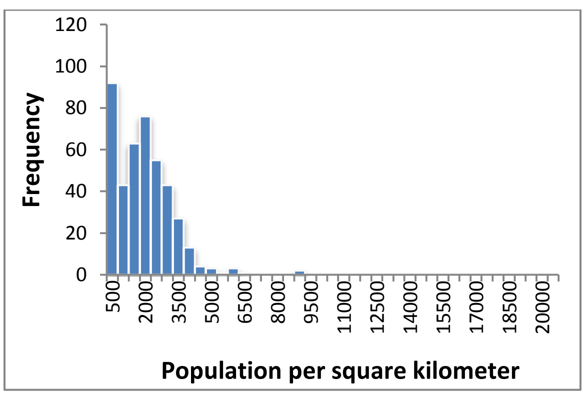
1990-2000



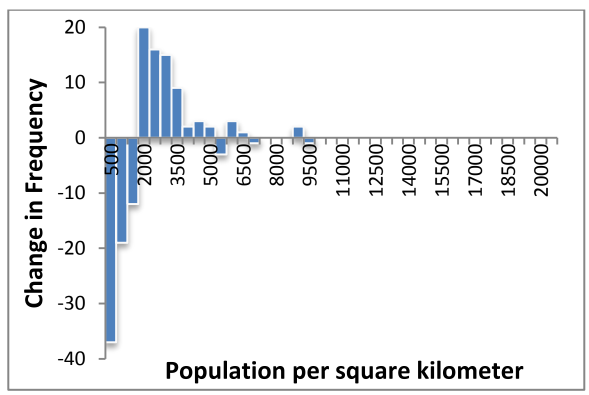
2000



2000-2007

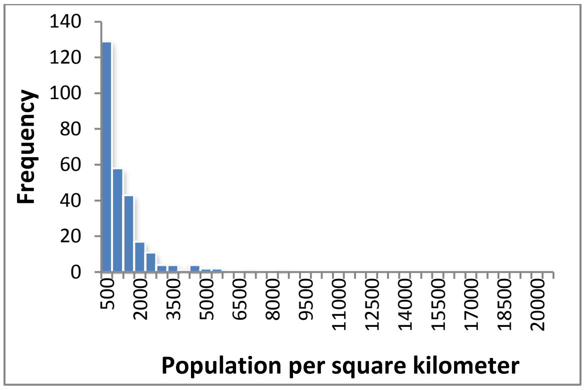


2007

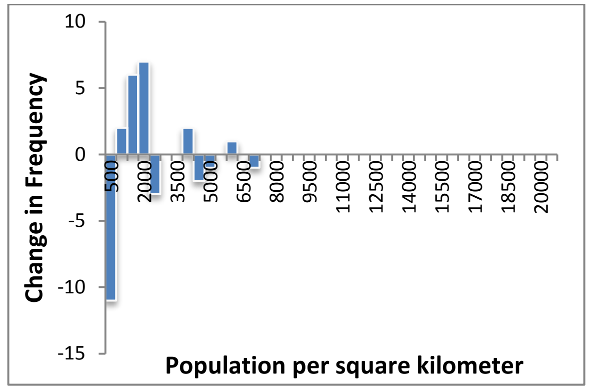


1990-2007

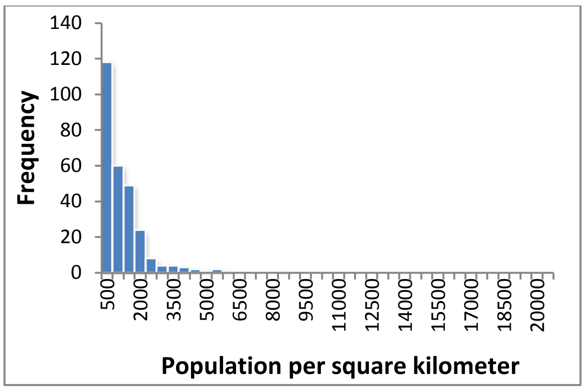
Richmond, VA



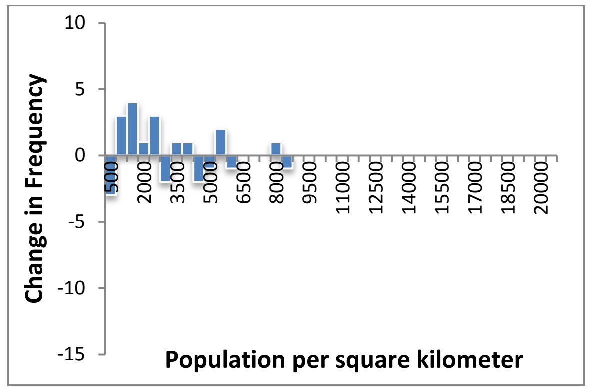
1990



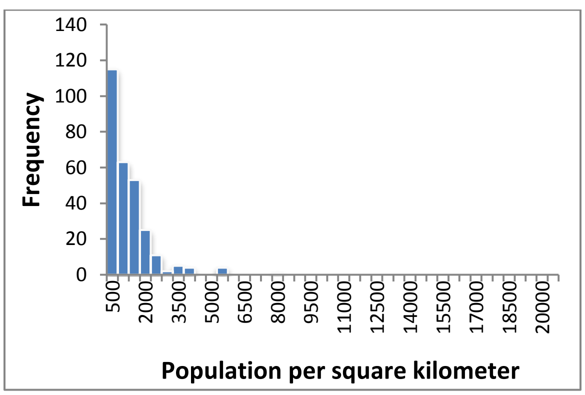
1990-2000



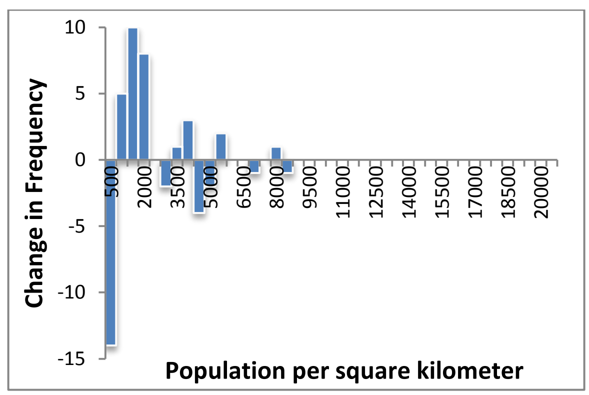
2000



2000-2007

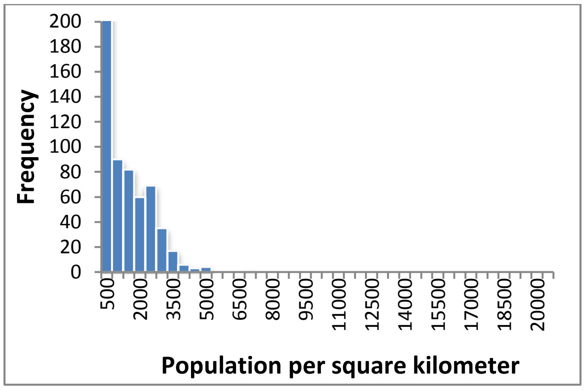


2007

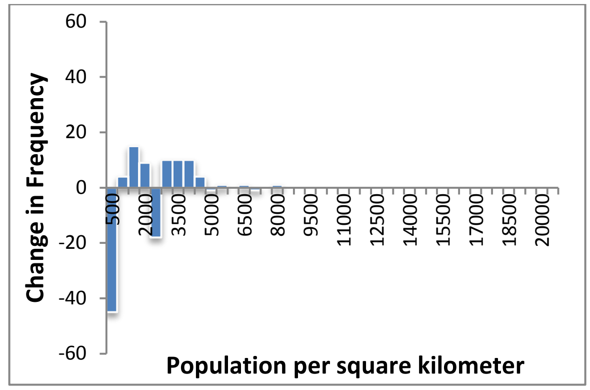


1990-2007

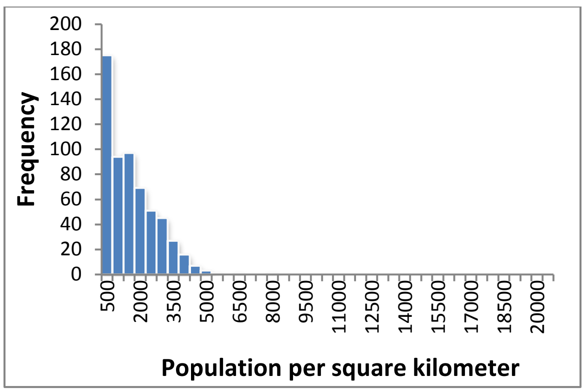
Riverside-San Bernardino-Ontario, CA



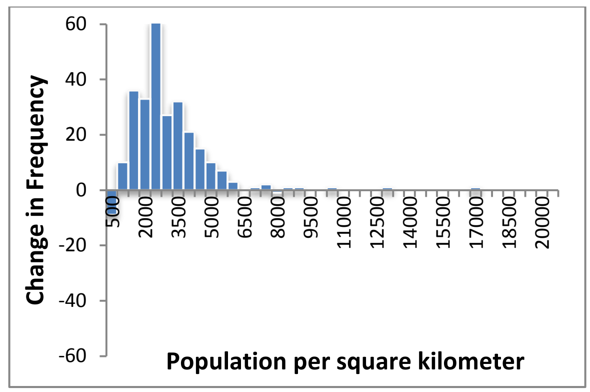
1990



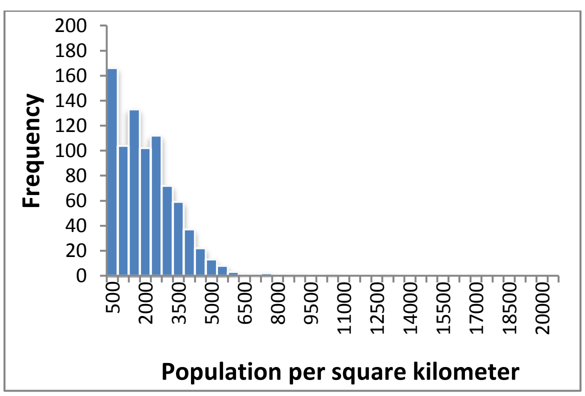
1990-2000



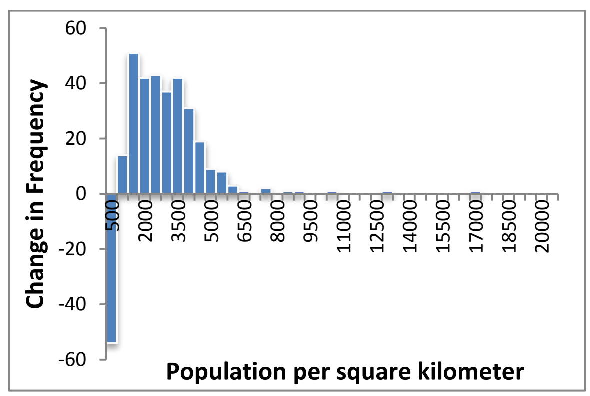
2000



2000-2007

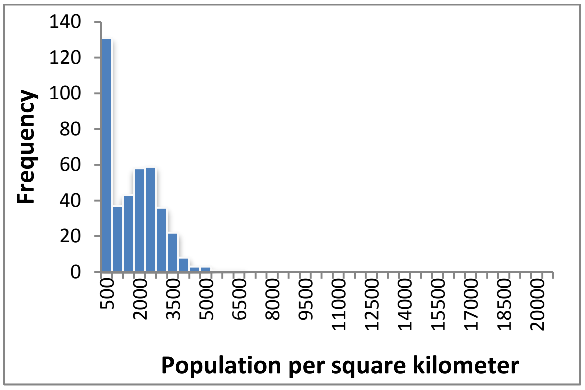


2007

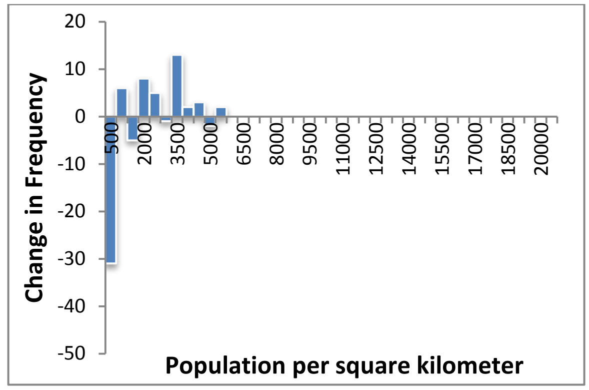


1990-2007

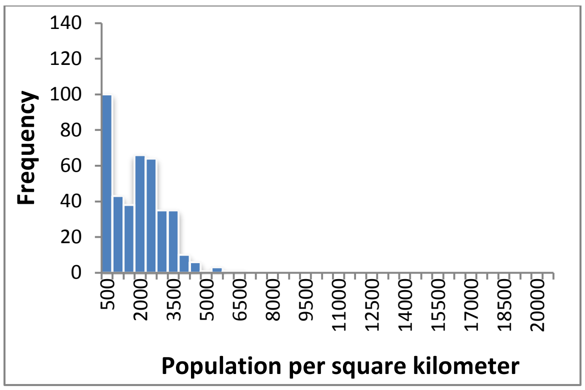
Sacramento-Arden-Arcade-Roseville, CA



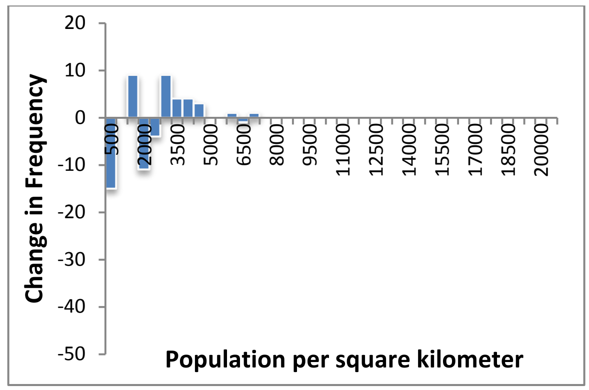
1990



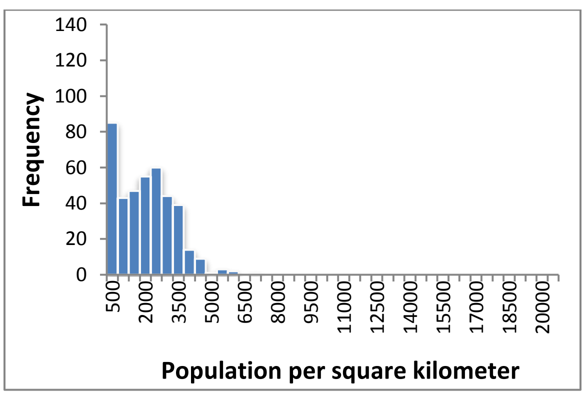
1990-2000



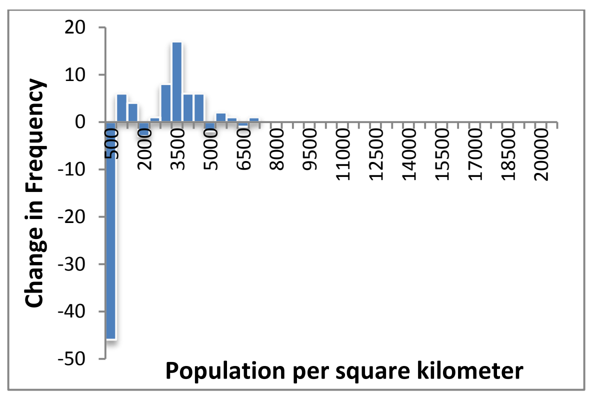
2000



2000-2007

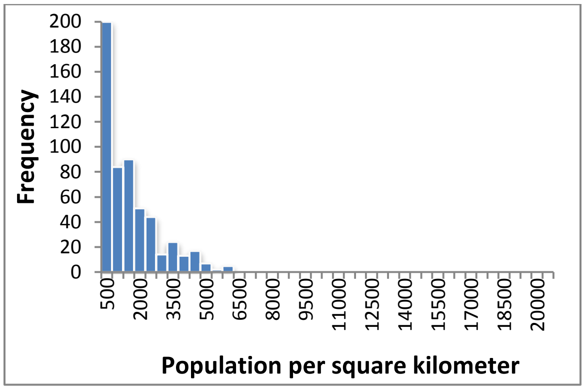


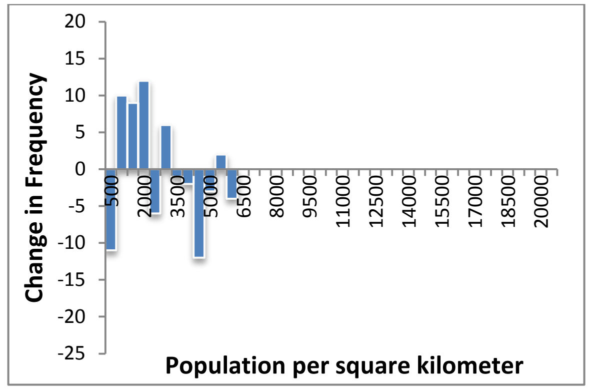
2007



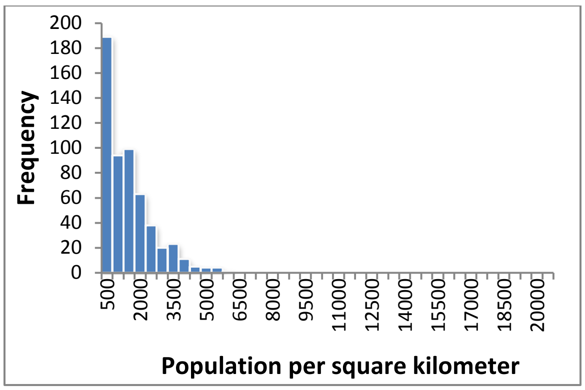
1990-2007

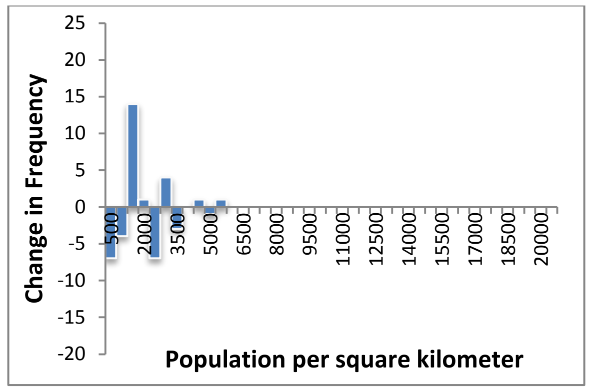
St. Louis, MO-IL



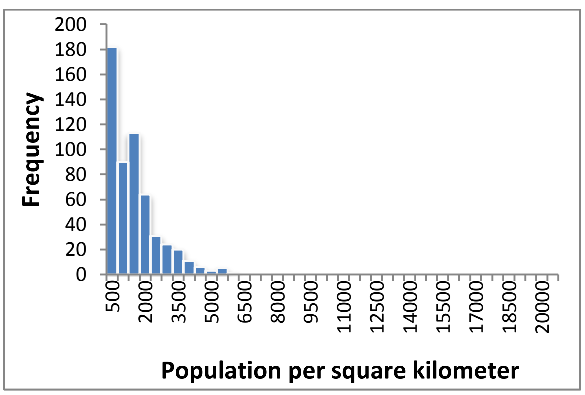


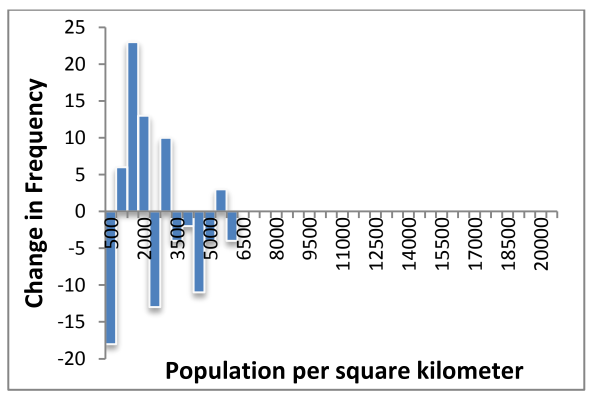
1990 **1990-2000**





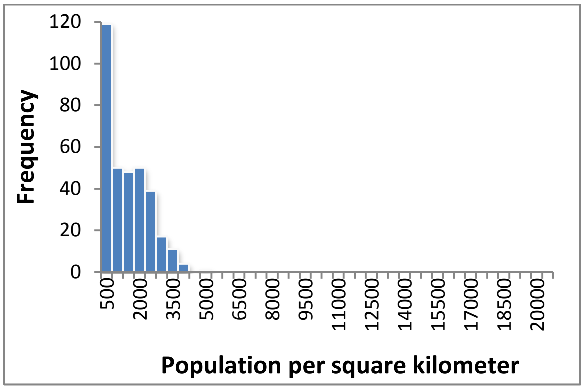
2000 **2000-2007**



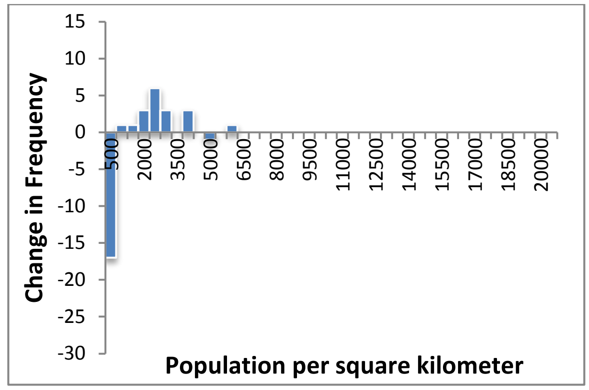


2007 **1990-2007**

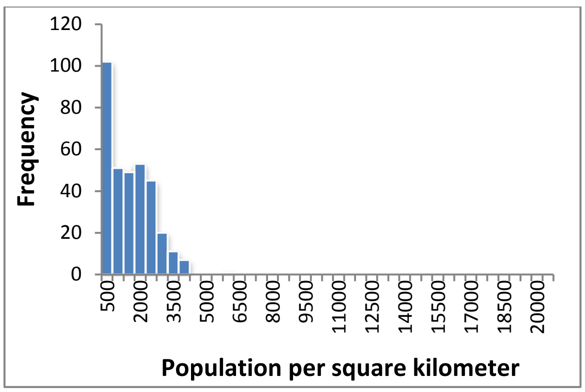
San Antonio, TX



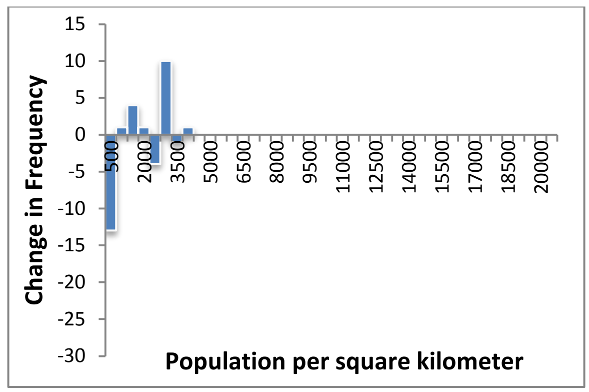
1990



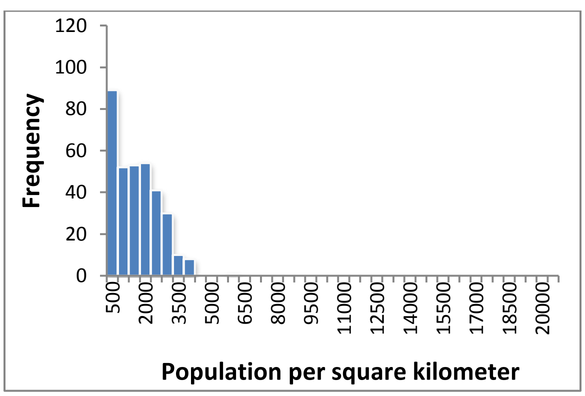
1990-2000



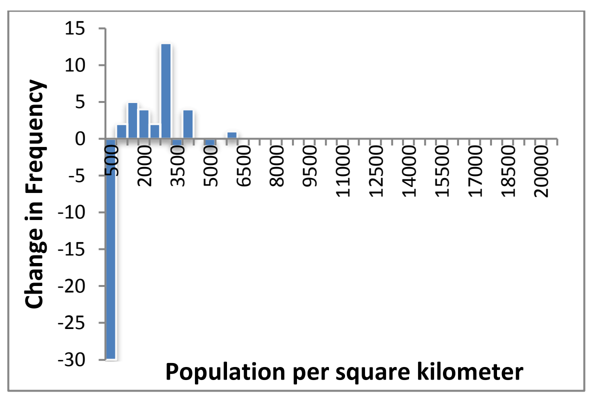
2000



2000-2007

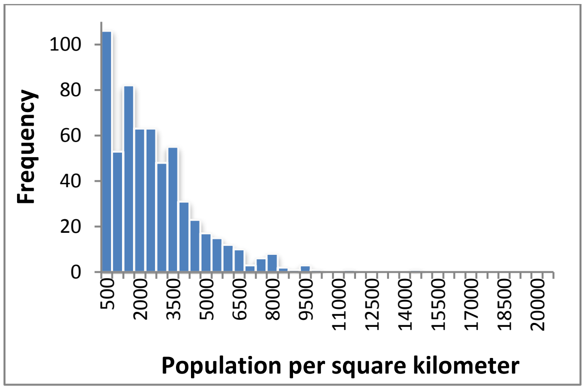


2007

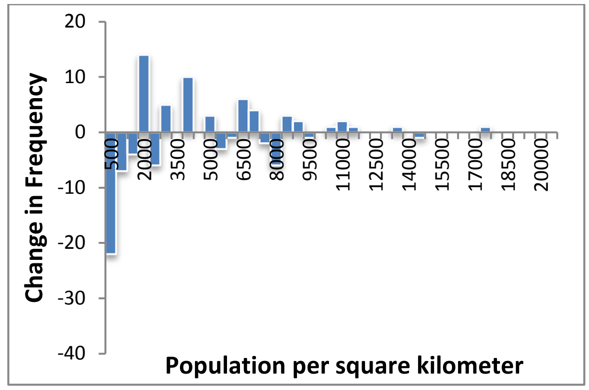


1990-2007

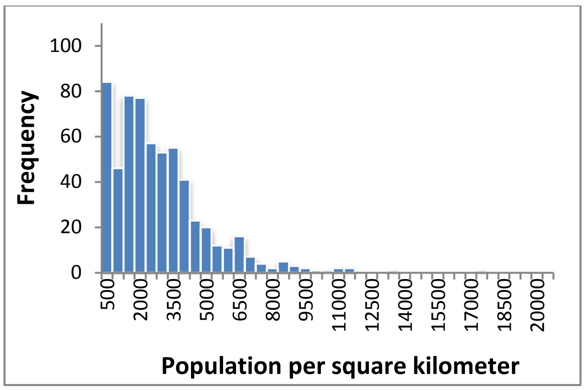
San Diego-Carlsbad-San Marcos, CA



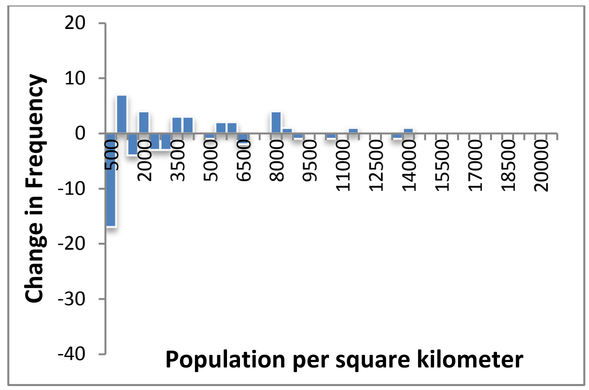
1990



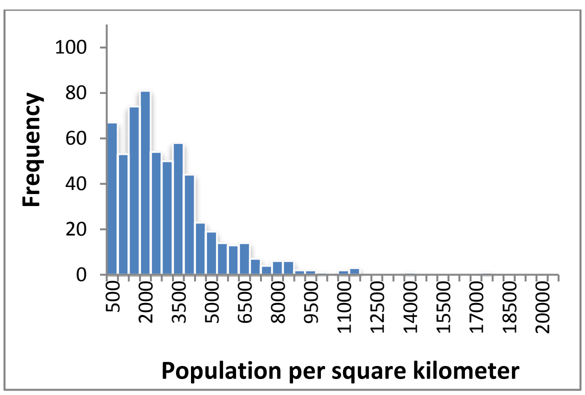
1990-2000



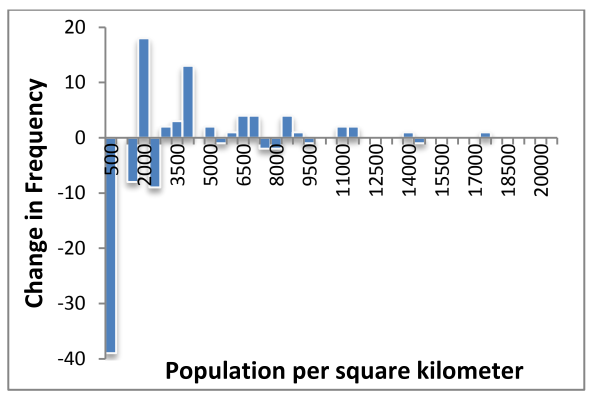
2000



2000-2007

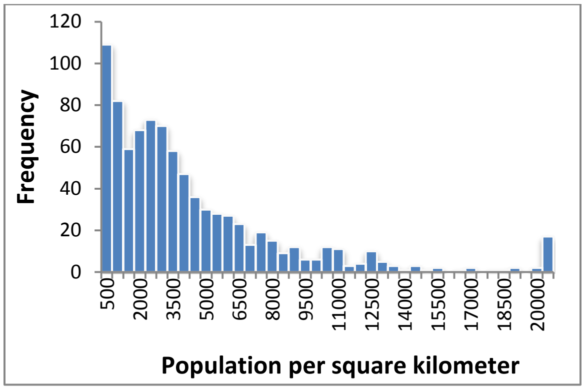


2007

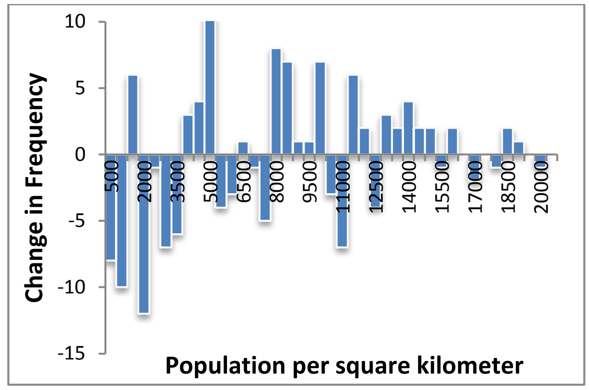


1990-2007

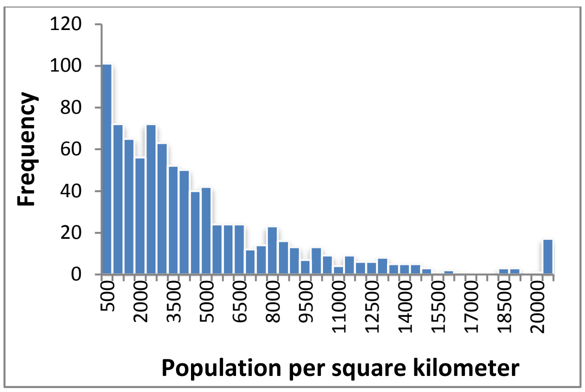
San Francisco-Oakland-Fremont, CA



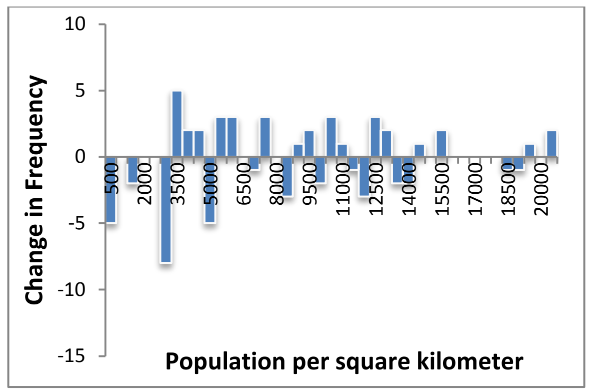
1990



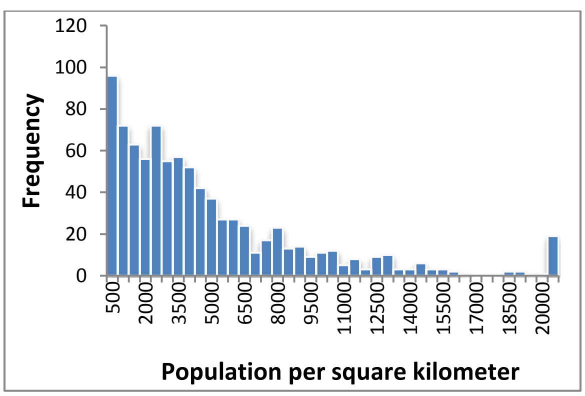
1990-2000



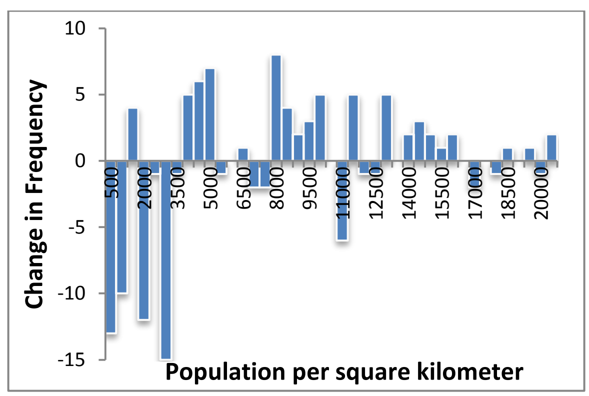
2000



2000-2007

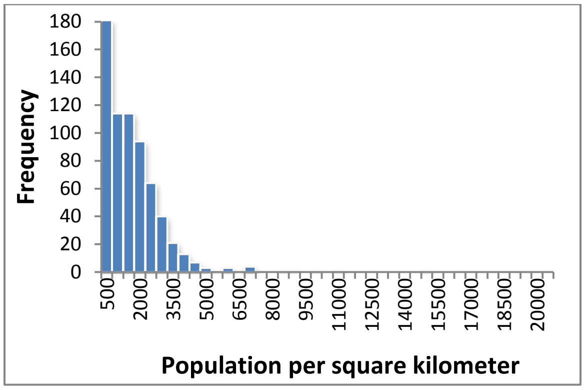


2007

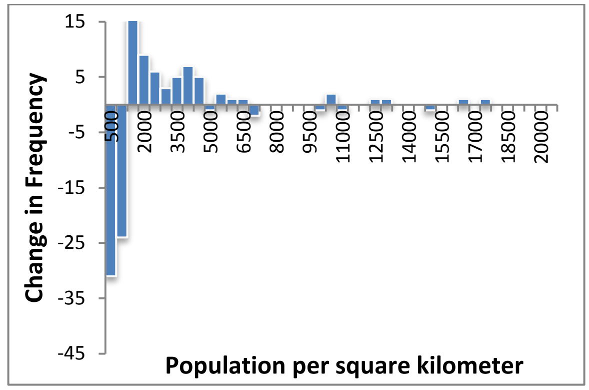


1990-2007

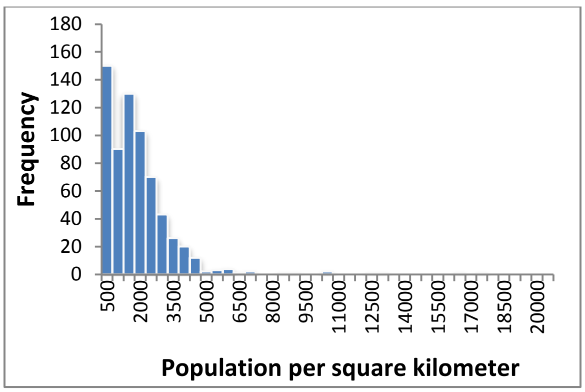
Seattle-Tacoma-Bellevue, WA



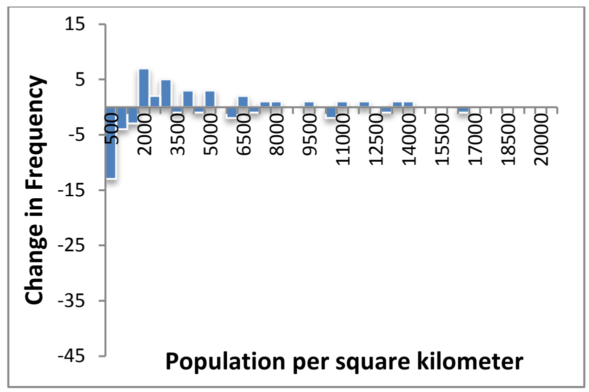
1990



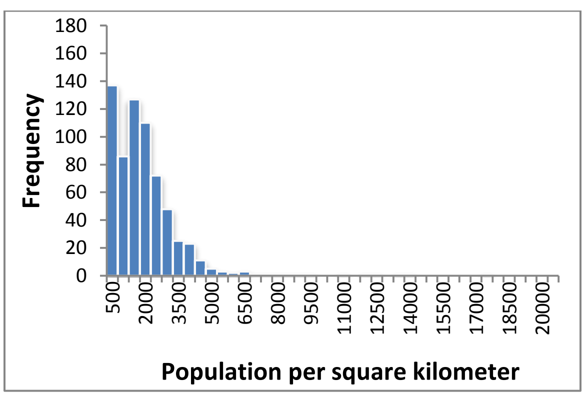
1990-2000



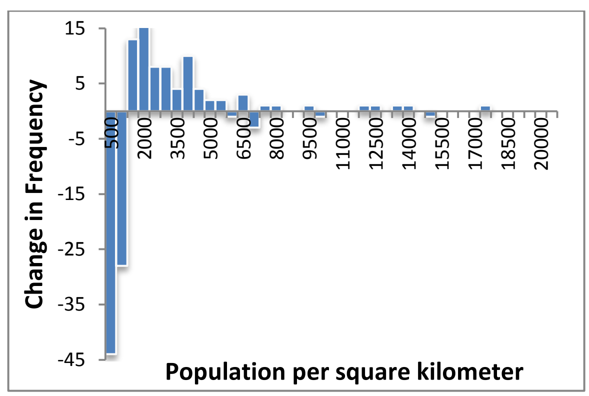
2000



2000-2007

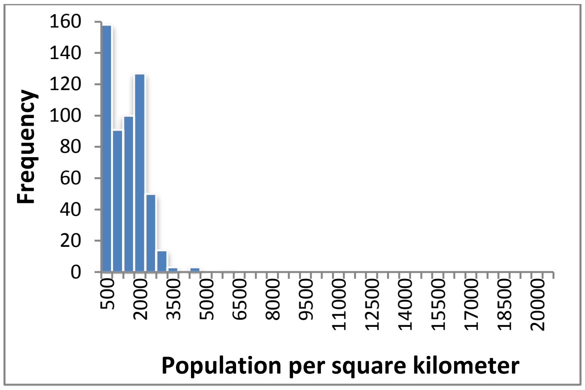


2007

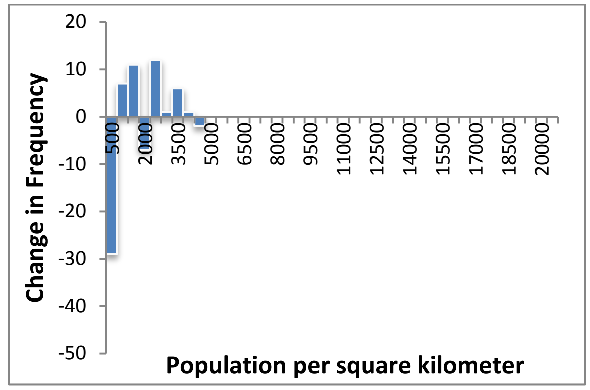


1990-2007

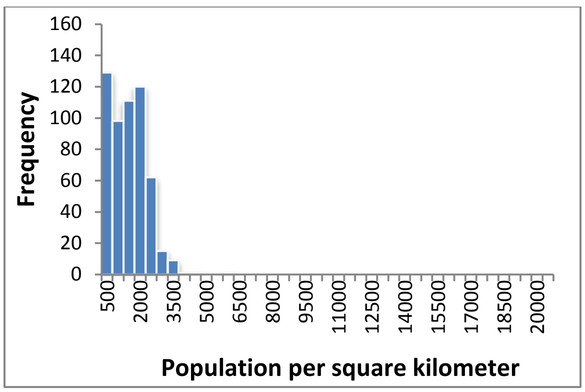
Tampa-St. Petersburg-Clearwater, FL



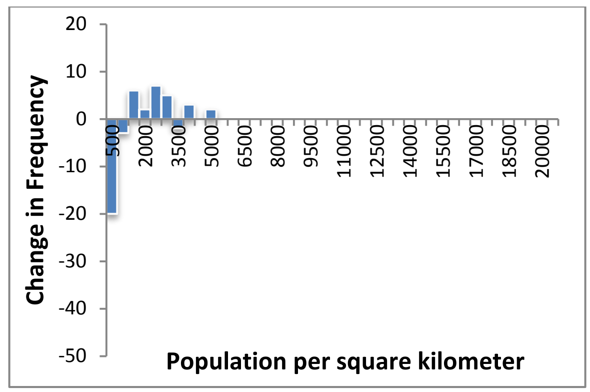
1990



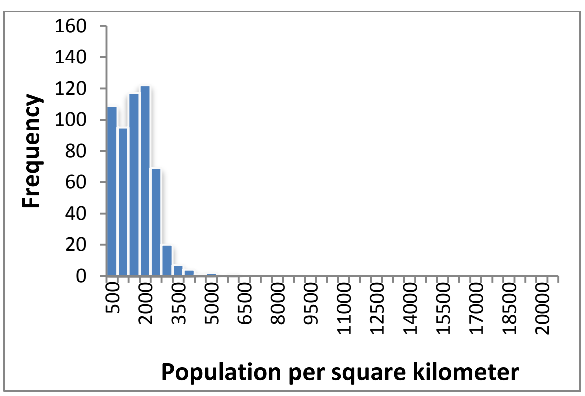
1990-2000



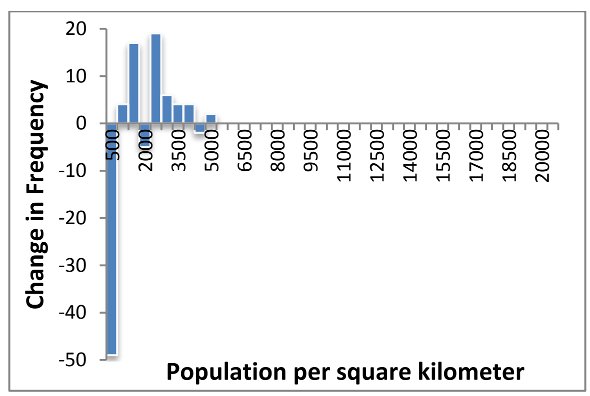
2000



2000-2007

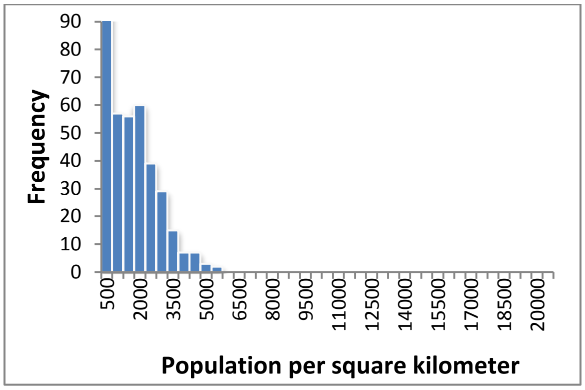


2007

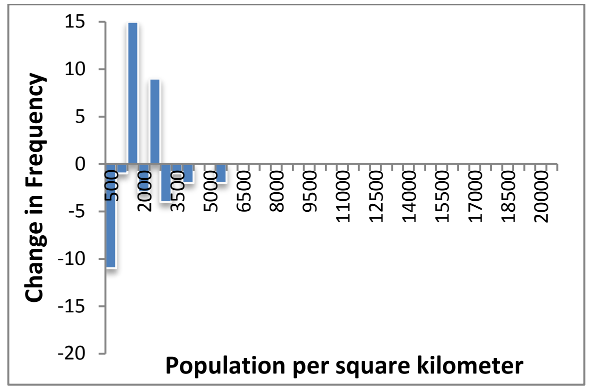


1990-2007

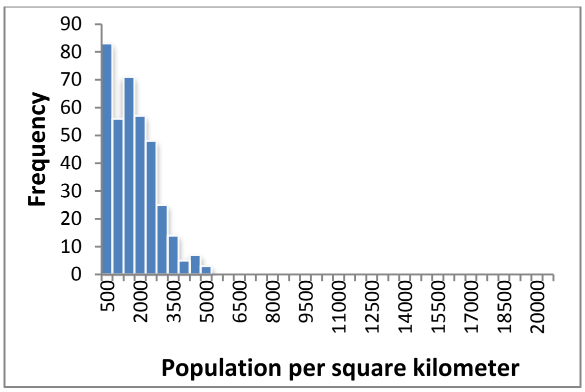
Virginia Beach-Norfolk-Newport News, VA-NC



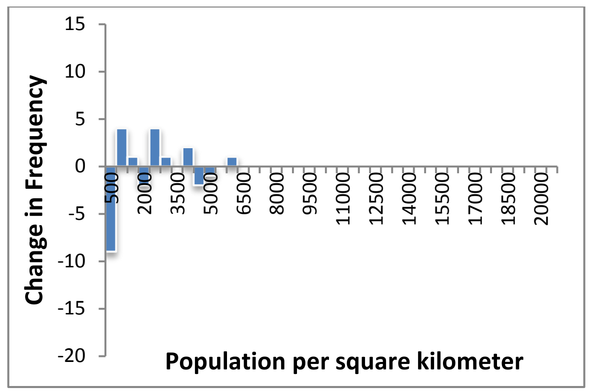
1990



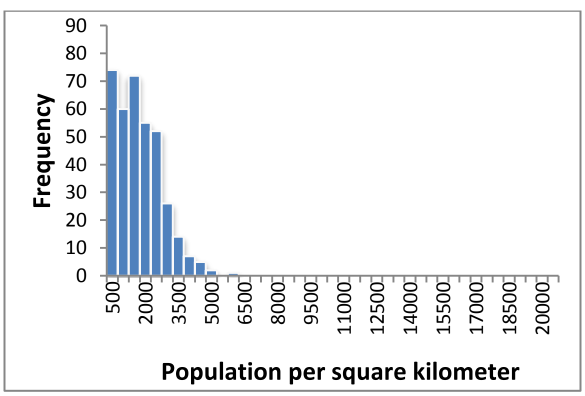
1990-2000



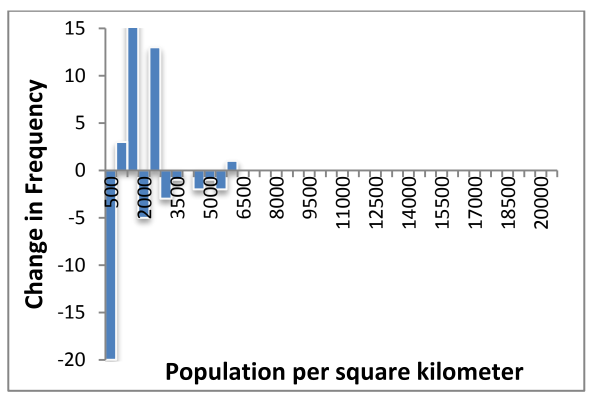
2000



2000-2007

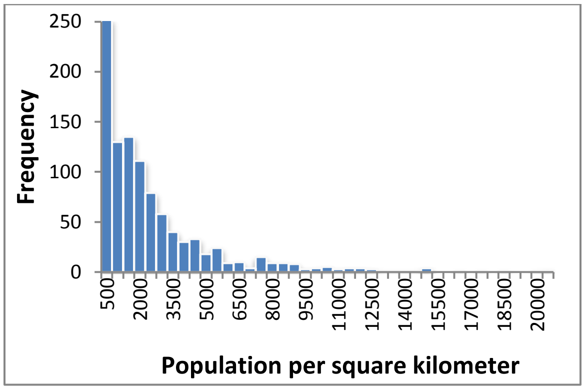


2007

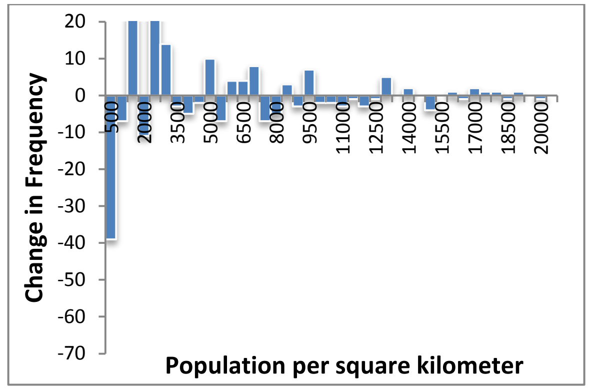


1990-2007

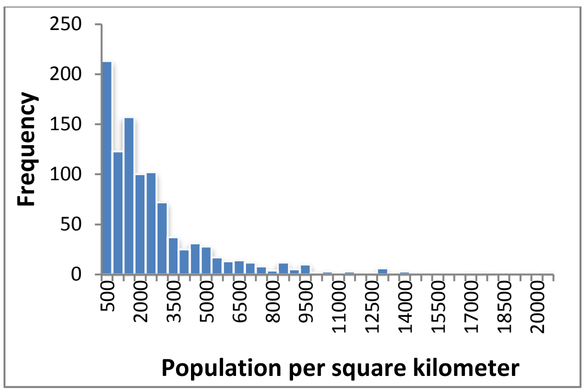
Washington-Arlington-Alexandria, DC-VA-MD-WV



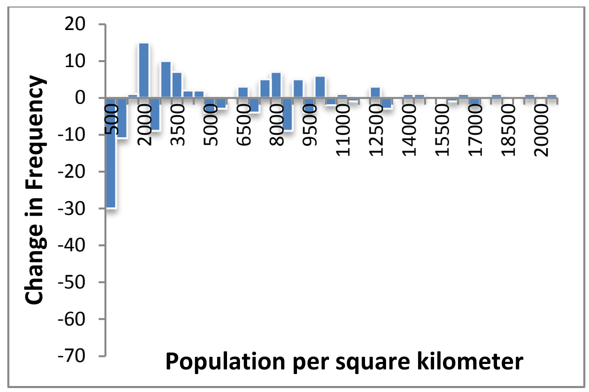
1990



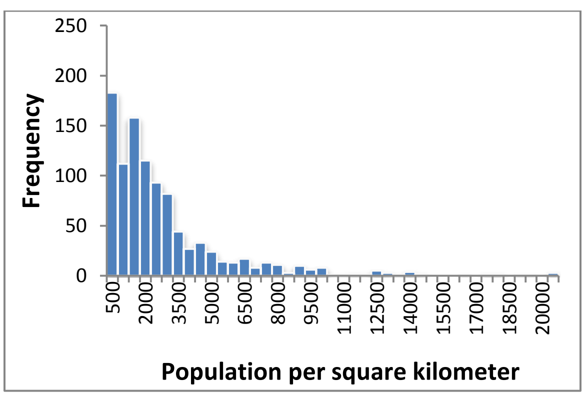
1990-2000



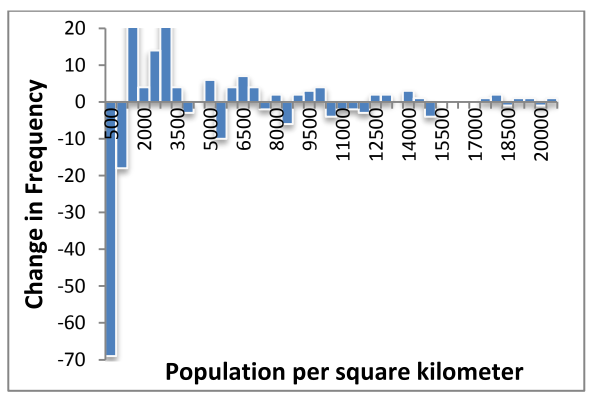
2000



2000-2007



2007



1990-2007

Urbanization

Table 16: Population Growth by Category – 1990-2007

CBSA	Urban 1990	Urbanized 1990-2000	Urbanized 2000-2007	Never Urban
Atlanta-Sandy Springs-Marietta, GA	29.11%	18.15%	14.27%	38.46%
Austin-Round Rock, TX	24.72%	21.77%	13.34%	40.17%
Baltimore-Towson, MD	22.40%	29.83%	6.13%	41.64%
Boston-Cambridge-Quincy, MA-NH	35.50%	4.88%	2.79%	56.83%
Charlotte-Gastonia-Concord, NC-SC	13.48%	17.94%	24.40%	44.19%
Chicago-Naperville-Joliet, IL-IN-WI	40.68%	27.00%	7.08%	25.23%
Cincinnati-Middletown, OH-KY-IN	-7.78%	30.33%	16.65%	60.80%
Cleveland-Elyria-Mentor, OH	-162.65%	45.90%	53.91%	162.84%
Dallas-Fort Worth-Arlington, TX	37.39%	20.96%	15.09%	26.56%
Denver-Aurora, CO	41.78%	24.87%	8.39%	24.96%
Detroit-Warren-Livonia, MI	-7.90%	28.14%	19.79%	59.97%
Houston-Sugar Land-Baytown, TX	45.37%	13.74%	16.21%	24.68%
Indianapolis, IN	9.58%	25.08%	15.04%	50.31%
Jacksonville, FL	20.78%	21.62%	12.00%	45.61%
Las Vegas-Paradise, NV	26.08%	40.12%	22.36%	11.44%
Los Angeles-Long Beach-Santa Ana, CA	88.29%	6.07%	1.14%	4.49%
Miami-Fort Lauderdale-Miami Beach, FL	57.39%	27.43%	5.22%	9.96%
Minneapolis-St. Paul-Bloomington, MN-WI	23.93%	22.18%	7.19%	46.70%
New York-Northern New Jersey-Long Island, NY-NJ-PA	80.84%	4.30%	1.84%	13.02%
Orlando-Kissimmee, FL	27.08%	25.98%	10.25%	36.69%
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	13.45%	20.53%	8.82%	57.19%
Phoenix-Mesa-Scottsdale, AZ	40.25%	18.43%	18.58%	22.74%
Pittsburgh, PA	-168.27%	3.10%	0.00%	65.17%
Portland-Vancouver-Beaverton, OR-WA	55.05%	20.54%	6.24%	18.17%
Richmond, VA	17.14%	14.75%	12.42%	55.69%
Riverside-San Bernardino-Ontario, CA	44.65%	21.07%	8.12%	26.16%
Sacramento-Arden-Arcade-Roseville, CA	31.30%	28.98%	13.92%	25.80%
St. Louis, MO-IL	-13.45%	28.74%	6.33%	78.38%
San Antonio, TX	33.44%	13.21%	14.92%	38.43%
San Diego-Carlsbad-San Marcos, CA	43.04%	24.06%	16.52%	16.37%
San Francisco-Oakland-Fremont, CA	70.49%	9.75%	3.16%	16.60%
Seattle-Tacoma-Bellevue, WA	56.96%	11.24%	8.04%	23.76%
Tampa-St. Petersburg-Clearwater, FL	35.09%	19.68%	14.10%	31.13%
Virginia Beach-Norfolk-Newport News, VA-NC	27.12%	16.91%	5.98%	49.99%
Washington-Arlington-Alexandria, DC-VA-MD-WV	46.22%	15.49%	8.61%	29.67%
All metropolitan areas	**41%**	**19%**	**11%**	**29%**

Table 17: Population Growth by Category – 1990-2000

CBSA	Urban 1990	Urbanized 1990-2000	Urbanized 2000-2007	Never Urban
Atlanta-Sandy Springs-Marietta, GA	34%	23%	11%	32%
Austin-Round Rock, TX	31%	23%	11%	35%
Baltimore-Towson, MD	18%	38%	5%	39%
Boston-Cambridge-Quincy, MA-NH	38%	6%	3%	53%
Charlotte-Gastonia-Concord, NC-SC	14%	22%	21%	43%
Chicago-Naperville-Joliet, IL-IN-WI	50%	27%	5%	18%
Cincinnati-Middletown, OH-KY-IN	-1%	30%	13%	58%
Cleveland-Elyria-Mentor, OH	-65%	38%	30%	97%
Dallas-Fort Worth-Arlington, TX	48%	24%	8%	20%
Denver-Aurora, CO	55%	24%	4%	17%
Detroit-Warren-Livonia, MI	-1%	28%	17%	56%
Houston-Sugar Land-Baytown, TX	51%	16%	10%	23%
Indianapolis, IN	18%	30%	9%	43%
Jacksonville, FL	17%	32%	12%	40%
Las Vegas-Paradise, NV	37%	52%	6%	5%
Los Angeles-Long Beach-Santa Ana, CA	89%	8%	0%	3%
Miami-Fort Lauderdale-Miami Beach, FL	60%	29%	3%	8%
Minneapolis-St. Paul-Bloomington, MN-WI	30%	28%	5%	38%
New York-Northern New Jersey-Long Island, NY-NJ-PA	83%	4%	1%	12%
Orlando-Kissimmee, FL	33%	32%	7%	28%
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	6%	28%	8%	58%
Phoenix-Mesa-Scottsdale, AZ	52%	25%	6%	16%
Pittsburgh, PA	-178%	4%	0%	74%
Portland-Vancouver-Beaverton, OR-WA	56%	22%	5%	17%
Richmond, VA	16%	19%	12%	54%
Riverside-San Bernardino-Ontario, CA	49%	28%	3%	21%
Sacramento-Arden-Arcade-Roseville, CA	40%	36%	9%	15%
St. Louis, MO-IL	-32%	33%	9%	89%
San Antonio, TX	38%	15%	8%	40%
San Diego-Carlsbad-San Marcos, CA	51%	30%	7%	12%
San Francisco-Oakland-Fremont, CA	77%	9%	1%	13%
Seattle-Tacoma-Bellevue, WA	61%	13%	4%	22%
Tampa-St. Petersburg-Clearwater, FL	39%	26%	11%	24%
Virginia Beach-Norfolk-Newport News, VA-NC	19%	26%	6%	49%
Washington-Arlington-Alexandria, DC-VA-MD-WV	51%	20%	5%	24%
All Metropolitan areas	**47%**	**23%**	**6%**	**24%**

Table 18: Population Growth by Category – 2000-2007

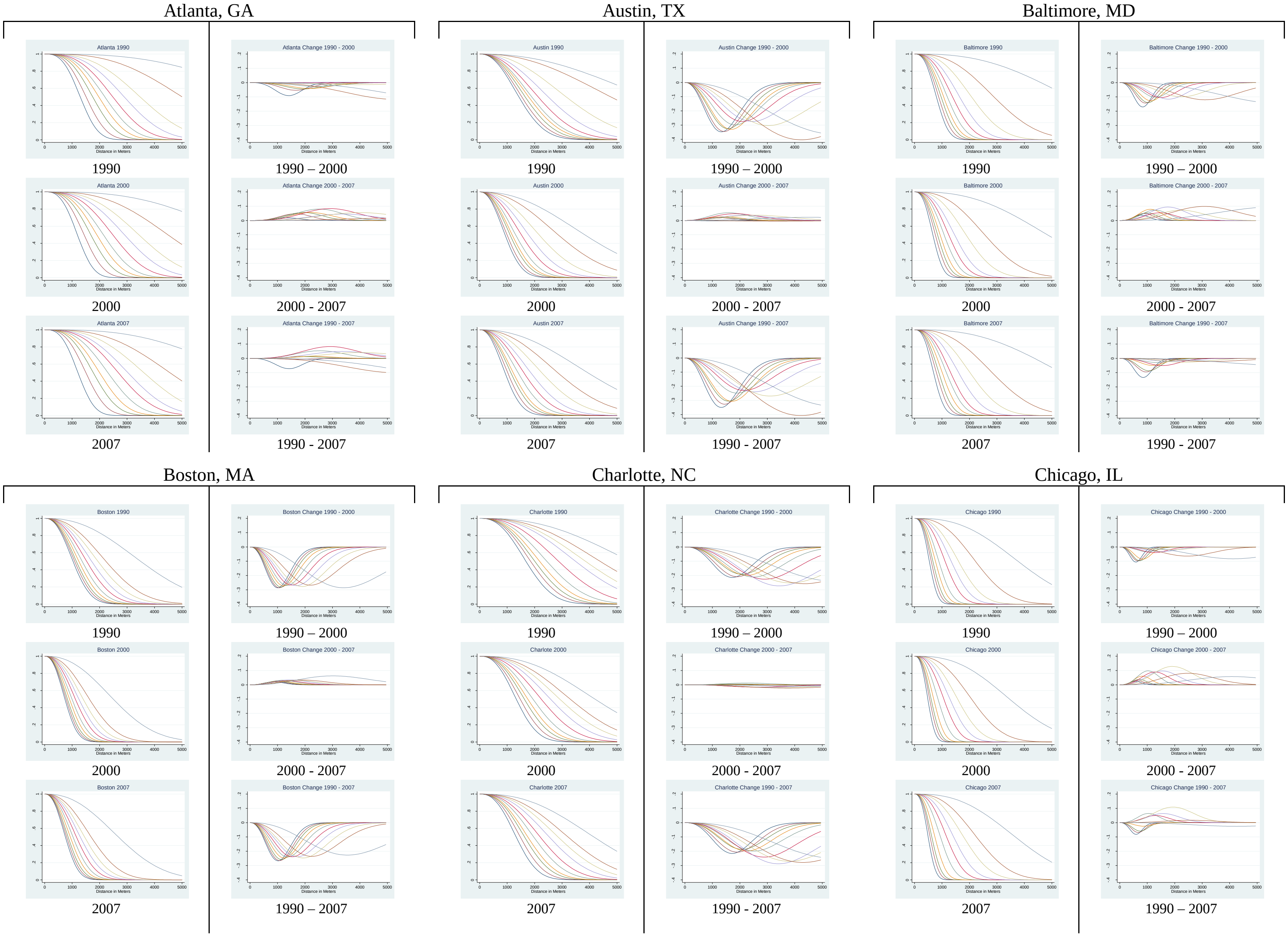
CBSA	Urban 1990	Urbanized 1990-2000	Urbanized 2000-2007	Never Urban
Atlanta-Sandy Springs-Marietta, GA	23.44%	13.26%	17.82%	45.48%
Austin-Round Rock, TX	16.43%	20.23%	16.52%	46.82%
Baltimore-Towson, MD	27.34%	20.52%	7.63%	44.52%
Boston-Cambridge-Quincy, MA-NH	30.50%	3.22%	2.50%	63.78%
Charlotte-Gastonia-Concord, NC-SC	13.35%	13.42%	27.63%	45.60%
Chicago-Naperville-Joliet, IL-IN-WI	27.26%	27.13%	10.45%	35.16%
Cincinnati-Middletown, OH-KY-IN	-18.79%	31.27%	22.03%	65.48%
Cleveland-Elyria-Mentor, OH	-2419.64%	235.75%	603.32%	1680.56%
Dallas-Fort Worth-Arlington, TX	24.57%	16.77%	23.90%	34.76%
Denver-Aurora, CO	18.16%	26.61%	16.27%	38.96%
Detroit-Warren-Livonia, MI	-20.22%	27.66%	25.54%	67.02%
Houston-Sugar Land-Baytown, TX	39.08%	11.18%	23.03%	26.71%
Indianapolis, IN	-0.92%	18.12%	22.48%	60.32%
Jacksonville, FL	24.16%	13.25%	12.14%	50.45%
Las Vegas-Paradise, NV	13.13%	25.36%	41.90%	19.61%
Los Angeles-Long Beach-Santa Ana, CA	87.59%	4.06%	2.00%	6.35%
Miami-Fort Lauderdale-Miami Beach, FL	53.46%	24.61%	8.94%	12.99%
Minneapolis-St. Paul-Bloomington, MN-WI	16.86%	15.49%	10.05%	57.60%
New York-Northern New Jersey-Long Island, NY-NJ-PA	77.43%	4.05%	3.31%	15.21%
Orlando-Kissimmee, FL	21.50%	20.38%	13.23%	44.89%
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	21.28%	12.52%	9.81%	56.39%
Phoenix-Mesa-Scottsdale, AZ	26.80%	10.84%	32.29%	30.06%
Pittsburgh, PA	-154.70%	2.19%	0.00%	52.51%
Portland-Vancouver-Beaverton, OR-WA	52.66%	18.69%	8.08%	20.56%
Richmond, VA	18.49%	10.25%	13.40%	57.86%
Riverside-San Bernardino-Ontario, CA	41.24%	14.92%	12.81%	31.03%
Sacramento-Arden-Arcade-Roseville, CA	23.31%	22.68%	18.50%	35.51%
St. Louis, MO-IL	2.36%	24.81%	3.85%	68.98%
San Antonio, TX	28.78%	11.04%	22.98%	37.19%
San Diego-Carlsbad-San Marcos, CA	33.22%	16.92%	28.14%	21.72%
San Francisco-Oakland-Fremont, CA	55.87%	12.58%	6.92%	24.62%
Seattle-Tacoma-Bellevue, WA	50.76%	8.62%	14.27%	26.35%
Tampa-St. Petersburg-Clearwater, FL	31.46%	14.22%	17.09%	37.23%
Virginia Beach-Norfolk-Newport News, VA-NC	36.18%	6.66%	5.68%	51.47%
Washington-Arlington-Alexandria, DC-VA-MD-WV	41.64%	10.80%	12.08%	35.47%
All metropolitan areas	**33%**	**15%**	**17%**	**35%**

APPENDIX B

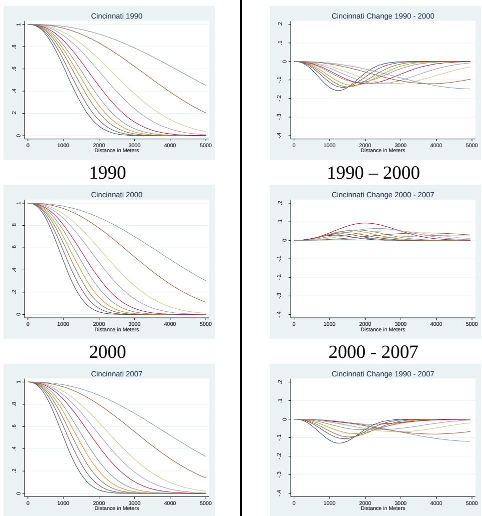
Table 19: Distance Measures Descriptive Statistics

		Dist. from Nearest Neighbor						Dist. From CBSA Center						Dist. From Local Center					
		1990		2000		2007		1990		2000		2007		1990		2000		2007	
	n	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
Atlanta-Sandy Springs-Marietta, GA	2,070	2,993	2,534	3,026	2,535	3,193	2,614	29,896	20,850	29,891	20,820	29,970	20,919	2,654	1,738	2,646	1,762	2,652	1,762
Austin-Round Rock, TX	771	2,485	2,914	2,532	2,864	2,733	2,953	18,947	15,206	18,958	15,196	19,084	15,296	3,191	2,039	3,221	2,009	3,242	2,011
Baltimore-Towson, MD	1,926	1,811	1,802	1,794	1,779	1,885	1,855	18,874	16,449	18,873	16,445	18,886	16,451	2,131	1,329	2,107	1,334	2,109	1,340
Boston-Cambridge-Quincy, MA-NH	2,760	1,717	1,469	1,731	1,472	1,817	1,536	27,071	22,568	27,074	22,565	27,089	22,601	1,868	1,126	1,879	1,143	1,875	1,147
Charlotte-Gastonia-Concord, NC-SC	801	2,872	2,348	2,884	2,347	3,061	2,406	21,867	15,541	21,879	15,518	21,904	15,587	3,246	2,106	3,239	2,074	3,267	2,058
Chicago-Naperville-Joliet, IL-IN-WI	6,156	1,262	1,429	1,275	1,428	1,357	1,508	29,552	22,628	29,552	22,622	29,561	22,650	2,995	2,657	3,000	2,655	3,005	2,653
Cincinnati-Middletown, OH-KY-IN	1,458	2,330	2,364	2,330	2,366	2,464	2,557	21,400	16,382	21,399	16,382	21,460	16,415	1,651	1,091	1,638	1,077	1,639	1,083
Cleveland-Elyria-Mentor, OH	2,079	1,418	1,345	1,422	1,348	1,493	1,388	19,225	14,515	19,229	14,513	19,230	14,538	1,978	1,152	1,981	1,160	1,979	1,163
Dallas-Fort Worth-Arlington, TX	3,138	1,986	2,236	2,005	2,231	2,153	2,294	31,767	20,792	31,745	20,781	31,786	20,816	3,245	1,957	3,258	1,955	3,261	1,958
Denver-Aurora, CO	1,560	1,711	2,243	1,704	2,184	1,922	2,788	16,048	11,221	16,041	11,199	16,157	11,525	2,630	1,977	2,628	1,941	2,652	1,988
Detroit-Warren-Livonia, MI	3,867	1,537	1,323	1,538	1,329	1,591	1,340	29,169	20,486	29,162	20,478	29,183	20,486	2,713	1,619	2,715	1,625	2,714	1,625
Houston-Sugar Land-Baytown, TX	2,685	2,255	2,524	2,275	2,531	2,484	2,797	29,082	20,912	29,086	20,915	29,136	20,987	4,075	2,669	4,088	2,679	4,108	2,692
Indianapolis, IN	945	2,499	2,600	2,507	2,590	2,644	2,703	17,570	14,327	17,566	14,310	17,623	14,360	2,551	1,427	2,553	1,417	2,562	1,419
Jacksonville, FL	603	2,624	2,358	2,647	2,310	3,038	2,793	19,426	16,662	19,434	16,694	19,518	16,662	6,535	4,142	6,528	4,147	6,576	4,178
Las Vegas-Paradise, NV	1,041	1,629	3,423	1,583	3,267	2,035	4,313	16,682	25,230	16,609	25,110	16,656	24,882	3,387	2,026	3,374	2,171	3,331	1,954
Los Angeles-Long Beach-Santa Ana, CA	7,113	949	826	942	811	1,010	941	24,928	15,006	24,926	14,996	24,944	15,032	3,183	2,615	3,178	2,615	3,181	2,623
Miami-Fort Lauderdale-Miami Beach, FL	2,670	1,331	737	1,325	588	1,469	1,073	46,871	34,528	46,877	34,539	46,937	34,552	1,674	1,252	1,650	1,096	1,655	1,100
Minneapolis-St. Paul-Bloomington, MN-WI	2,238	2,147	2,389	2,151	2,386	2,269	2,476	21,202	16,103	21,199	16,079	21,259	16,171	2,318	1,256	2,317	1,254	2,316	1,247
New York-Northern New Jersey-Long Island, NY-NJ-PA	13,188	864	938	862	937	905	1,039	26,405	22,162	26,408	22,163	26,397	22,170	2,636	1,920	2,632	1,922	2,633	1,924
Orlando-Kissimmee, FL	984	2,225	2,258	2,239	2,193	2,518	2,651	18,574	13,747	18,585	13,698	18,746	13,813	2,139	1,302	2,147	1,328	2,140	1,335
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	4,668	1,488	1,202	1,488	1,205	1,584	1,288	28,111	20,619	28,113	20,626	28,118	20,607	1,762	1,470	1,753	1,476	1,753	1,481
Phoenix-Mesa-Scottsdale, AZ	2,115	2,030	3,258	2,026	3,170	2,317	3,986	25,745	29,955	25,686	29,867	25,820	30,038	4,971	3,282	4,968	3,218	4,973	3,192
Pittsburgh, PA	2,163	2,017	2,065	2,026	2,067	2,128	2,133	24,207	18,189	24,203	18,180	24,237	18,241	1,288	1,078	1,277	1,086	1,279	1,099
Portland-Vancouver-Beaverton, OR-WA	1,278	2,137	2,448	2,124	2,411	2,424	3,205	17,457	13,719	17,464	13,748	17,592	14,062	2,444	1,888	2,419	1,845	2,395	1,746
Richmond, VA	849	3,328	3,760	3,329	3,731	3,558	3,990	22,063	19,218	22,087	19,207	22,185	19,249	2,662	1,740	2,660	1,743	2,668	1,748
Riverside-San Bernardino-Ontario, CA	2,520	1,896	2,790	1,882	2,597	2,258	4,764	41,344	27,347	41,330	27,401	41,372	27,322	2,440	2,038	2,432	1,861	2,428	1,840
Sacramento-Arden-Arcade-Roseville, CA	1,209	2,027	2,903	2,003	2,762	2,251	3,072	25,855	27,903	25,830	27,915	25,912	28,035	2,738	2,125	2,727	2,121	2,732	2,088
St. Louis, MO-IL	1,650	2,721	3,009	2,721	3,003	2,904	3,089	26,742	21,203	26,750	21,200	26,830	21,272	1,728	1,311	1,729	1,316	1,743	1,334
San Antonio, TX	1,014	2,726	3,179	2,723	3,170	2,983	3,412	19,015	16,010	18,986	15,957	19,164	16,213	3,951	2,506	3,965	2,508	3,953	2,509
San Diego-Carlsbad-San Marcos, CA	1,797	1,439	2,116	1,442	2,134	1,548	2,136	23,969	17,807	23,969	17,804	24,030	17,849	3,933	2,628	3,937	2,623	3,947	2,627
San Francisco-Oakland-Fremont, CA	2,613	1,083	981	1,079	960	1,181	1,115	24,070	15,365	24,076	15,358	24,076	15,369	2,174	1,357	2,166	1,351	2,169	1,351
Seattle-Tacoma-Bellevue, WA	1,992	1,844	1,760	1,832	1,720	1,956	2,028	29,220	17,607	29,213	17,592	29,277	17,702	2,329	1,703	2,318	1,691	2,317	1,723
Tampa-St. Petersburg-Clearwater, FL	1,641	1,839	1,269	1,847	1,252	1,983	1,396	26,547	14,392	26,533	14,371	26,581	14,446	2,594	1,907	2,575	1,908	2,591	1,911
Virginia Beach-Norfolk-Newport News, VA-NC	1,104	2,068	2,428	2,060	2,410	2,210	2,687	25,310	12,822	25,330	12,844	25,376	12,800	4,883	2,875	4,875	2,876	4,869	2,869
Washington-Arlington-Alexandria, DC-VA-MD-WV	3,012	1,704	1,746	1,700	1,727	1,813	1,825	23,332	19,734	23,337	19,742	23,355	19,757	1,740	1,281	1,736	1,289	1,736	1,287

Figure 36: Survival Functions and Changes

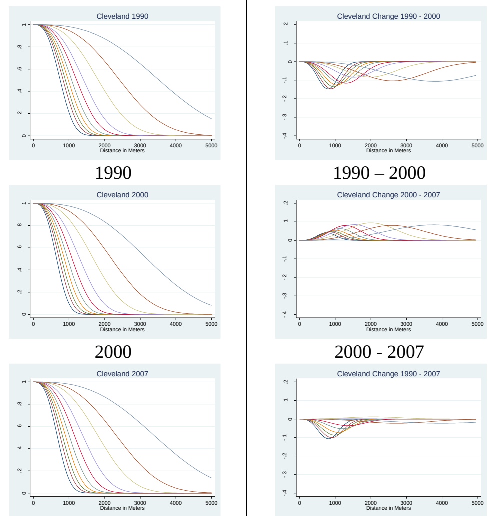


Cincinnati, OH



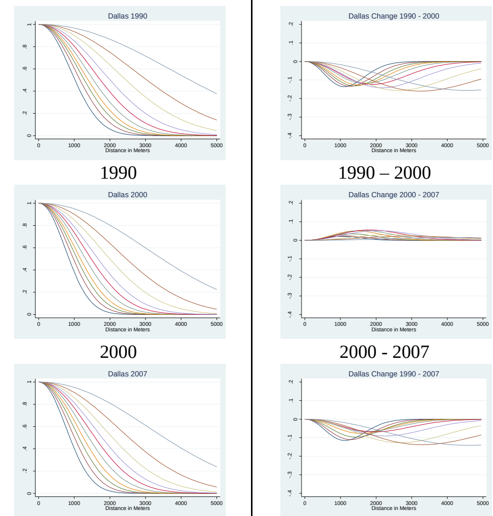
1990 | 1990 – 2000
2000 | 2000 - 2007
2007 | 1990 - 2007

Cleveland, OH



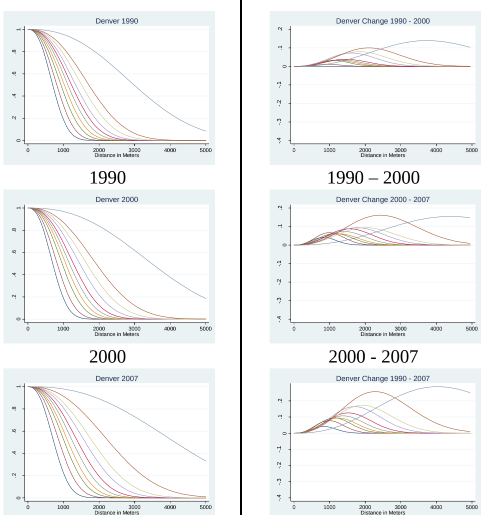
1990 | 1990 – 2000
2000 | 2000 - 2007
2007 | 1990 - 2007

Dallas, TX



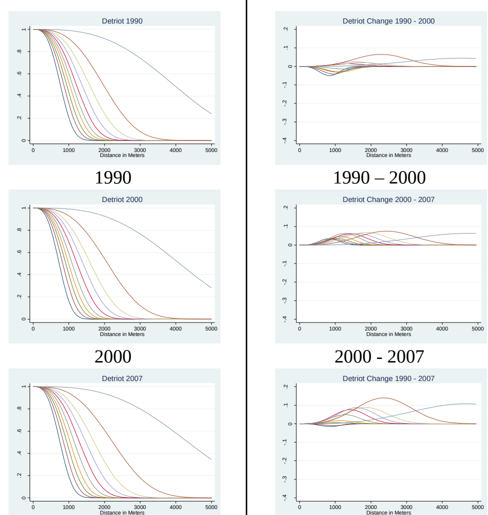
1990 | 1990 – 2000
2000 | 2000 - 2007
2007 | 1990 - 2007

Denver, CO



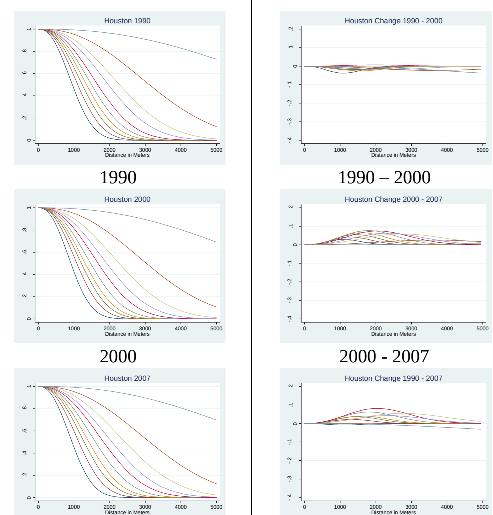
1990 | 1990 – 2000
2000 | 2000 - 2007
2007 | 1990 – 2007

Detroit, MI



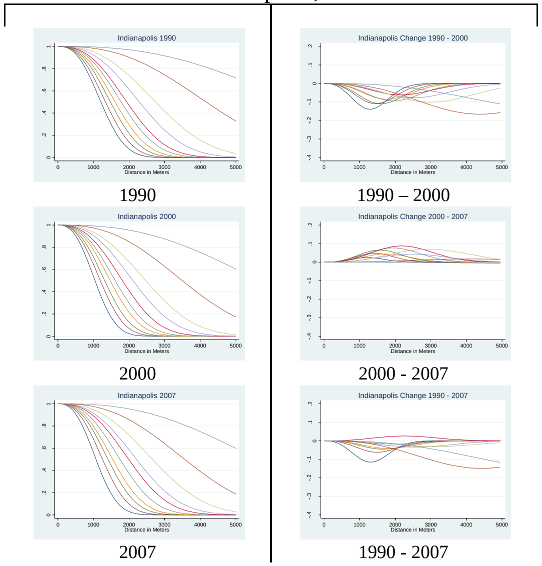
1990 | 1990 – 2000
2000 | 2000 - 2007
2007 | 1990 - 2007

Houston, TX

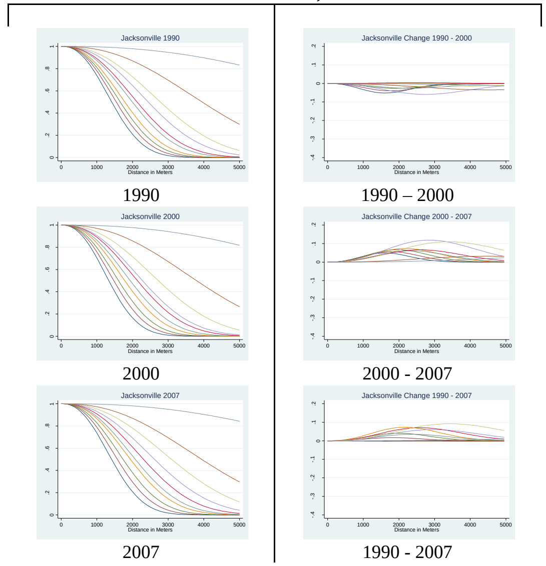


1990 | 1990 – 2000
2000 | 2000 - 2007
2007 | 1990 – 2007

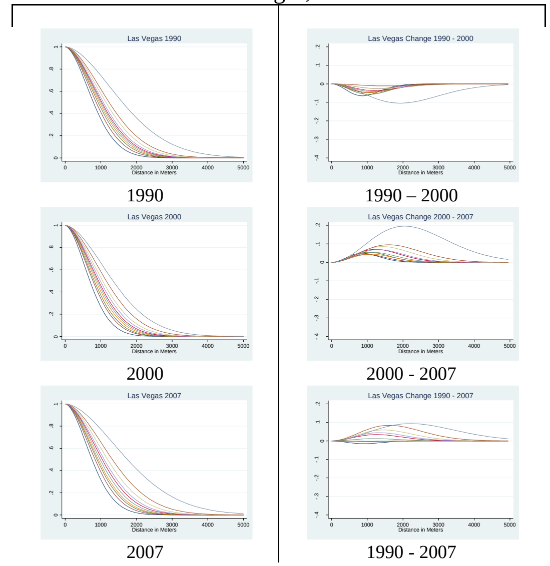
Indianapolis, IN



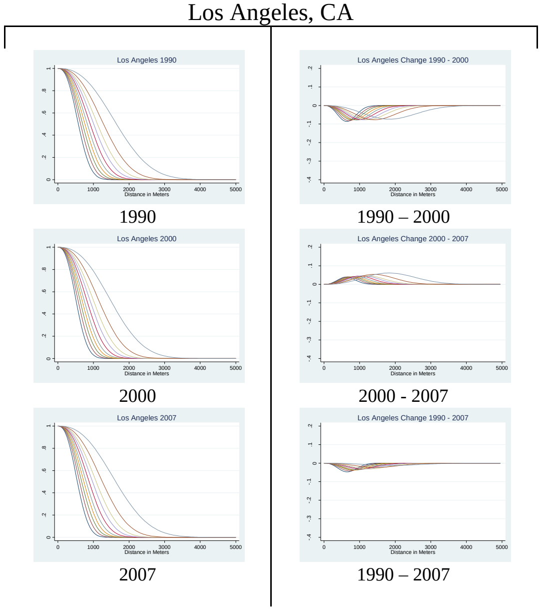
Jacksonville, FL



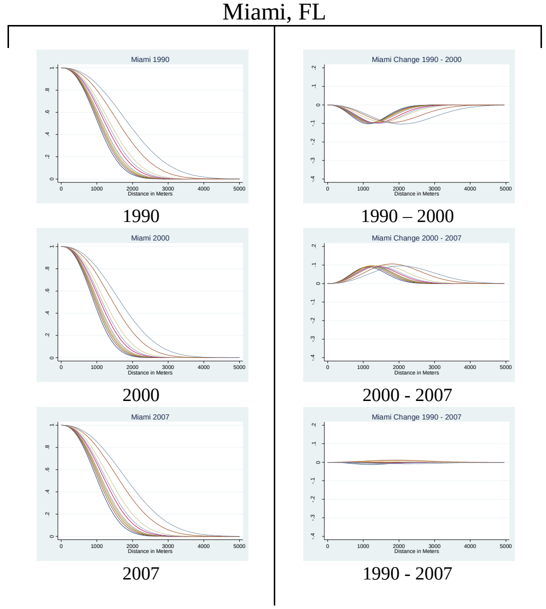
Las Vegas, NV



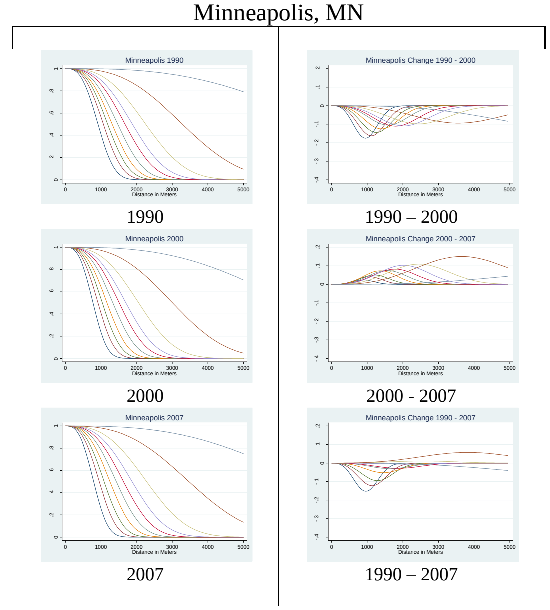
Los Angeles, CA



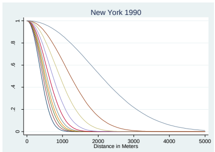
Miami, FL



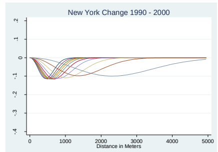
Minneapolis, MN



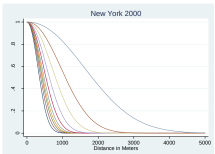
New York, NY



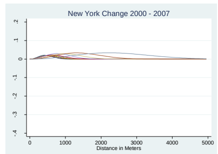
1990



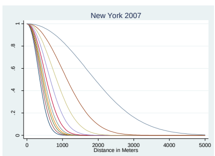
1990 – 2000



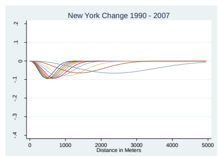
2000



2000 - 2007

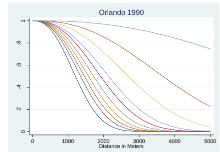


2007

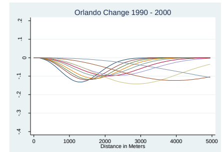


1990 - 2007

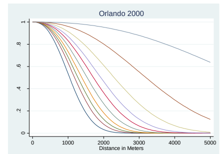
Orlando, FL



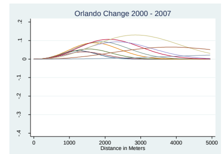
1990



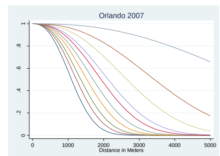
1990 – 2000



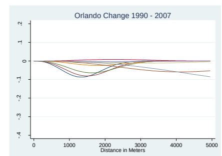
2000



2000 - 2007

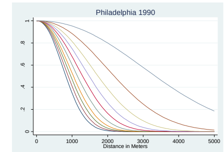


2007

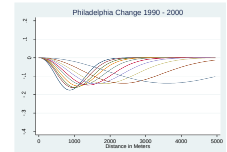


1990 - 2007

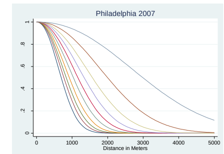
Philadelphia, PA



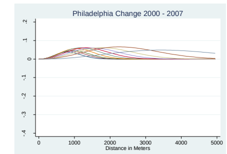
1990



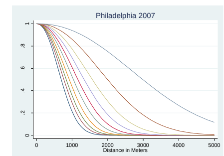
1990 – 2000



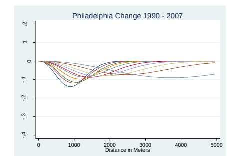
2000



2000 - 2007

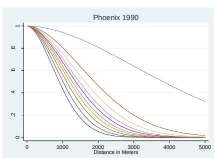


2007

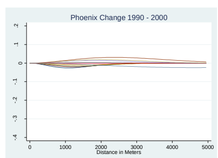


1990 - 2007

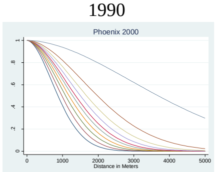
Phoenix, AX



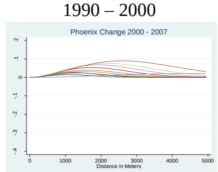
1990



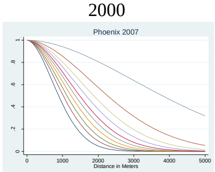
1990 – 2000



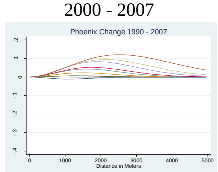
2000



2000 - 2007

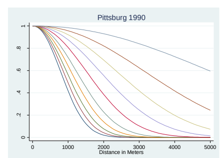


2007

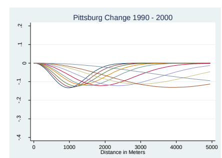


1990 - 2007

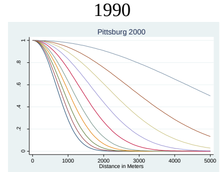
Pittsburg, PA



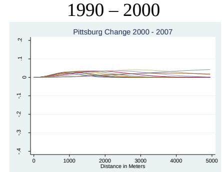
1990



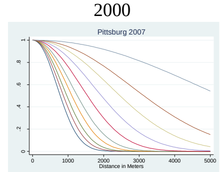
1990 – 2000



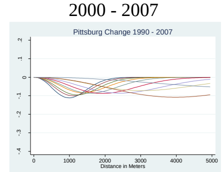
2000



2000 - 2007

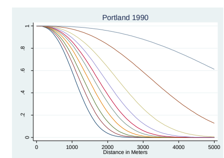


2007

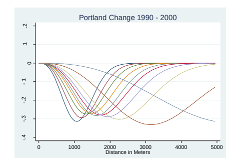


1990 - 2007

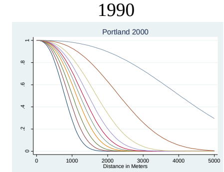
Portland, OR



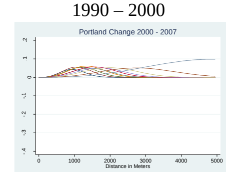
1990



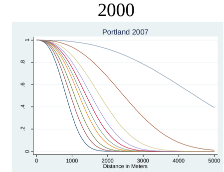
1990 – 2000



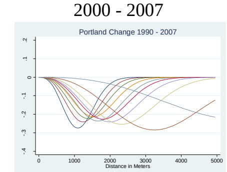
2000



2000 - 2007

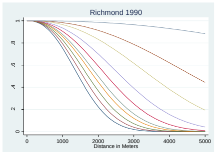


2007

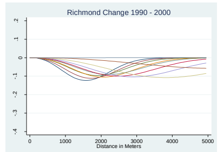


1990 – 2007

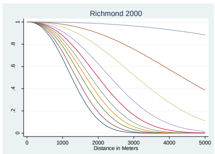
Richmond, VA



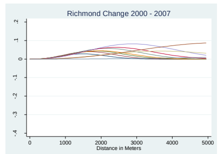
1990



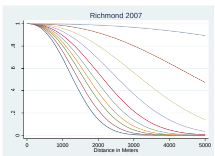
1990 – 2000



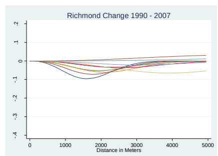
2000



2000 - 2007

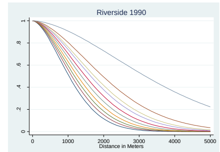


2007

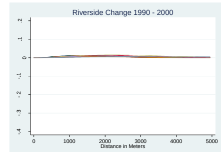


1990 - 2007

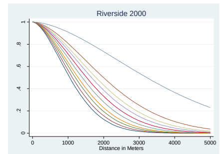
Riverside, CA



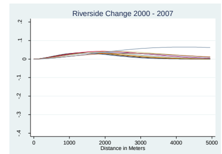
1990



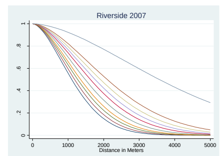
1990 – 2000



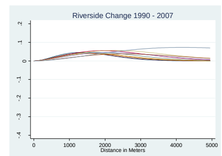
2000



2000 - 2007

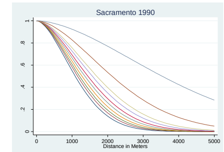


2007

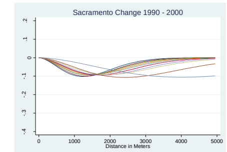


1990 - 2007

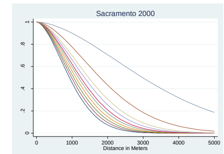
Sacramento, CA



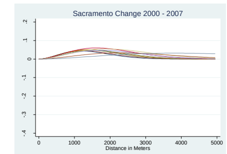
1990



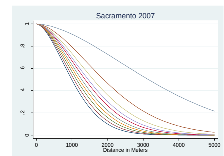
1990 – 2000



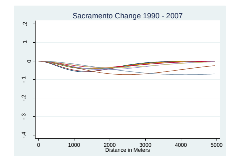
2000



2000 - 2007

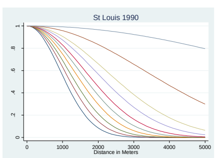


2007

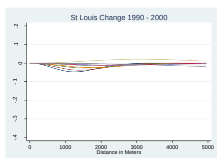


1990 - 2007

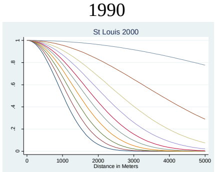
St Louis, MO



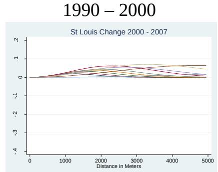
1990



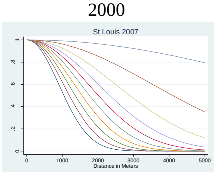
1990 – 2000



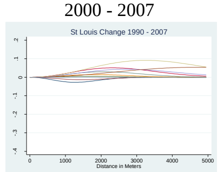
2000



2000 - 2007

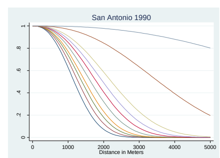


2007

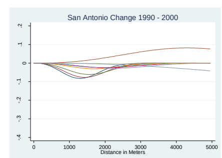


1990 - 2007

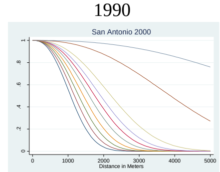
San Antonio, TX



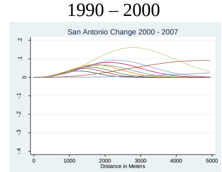
1990



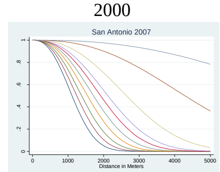
1990 – 2000



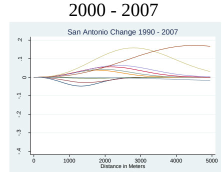
2000



2000 - 2007

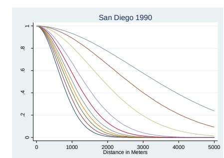


2007

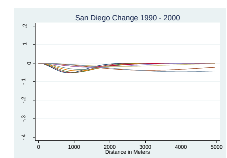


1990 - 2007

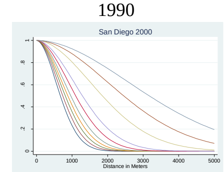
San Diego, CA



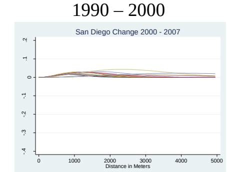
1990



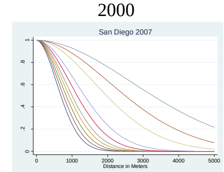
1990 – 2000



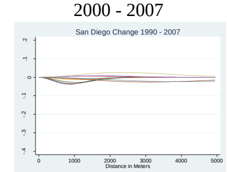
2000



2000 - 2007

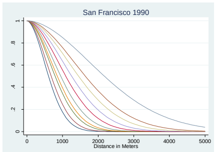


2007

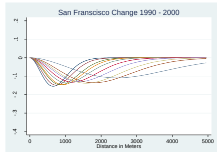


1990 – 2007

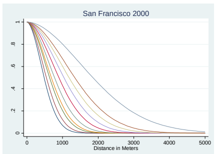
San Francisco, CA



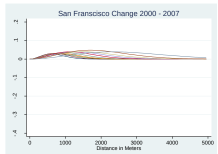
1990



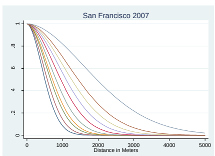
1990 – 2000



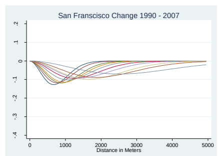
2000



2000 - 2007

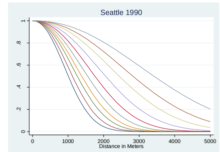


2007

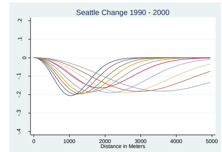


1990 - 2007

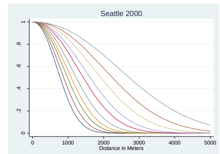
Seattle, WA



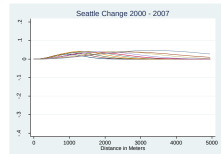
1990



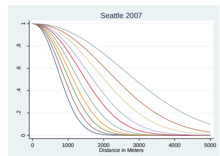
1990 – 2000



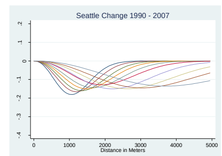
2000



2000 - 2007

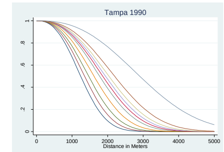


2007

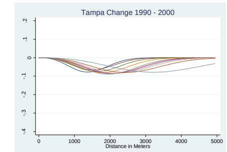


1990 - 2007

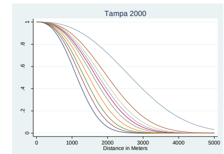
Tampa, FL



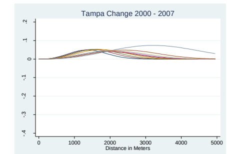
1990



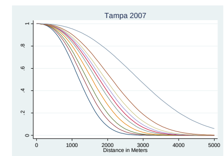
1990 – 2000



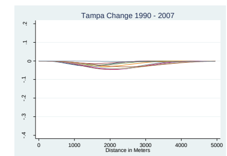
2000



2000 - 2007

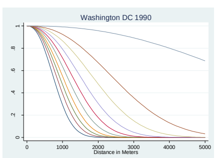


2007

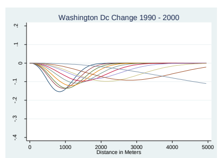


1990 - 2007

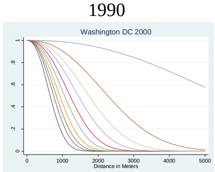
Washington, DC



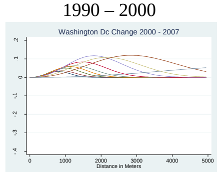
1990



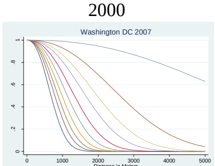
1990 – 2000



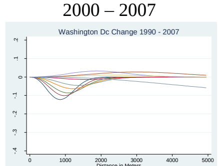
2000



2000 – 2007

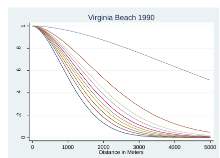


2007

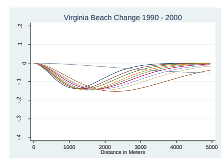


1990 – 2007

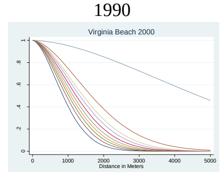
Virginia Beach, VA



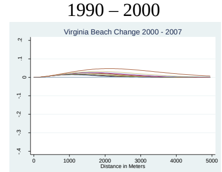
1990



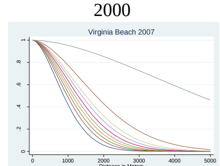
1990 – 2000



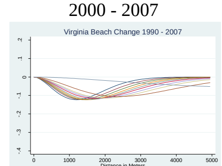
2000



2000 - 2007



2007



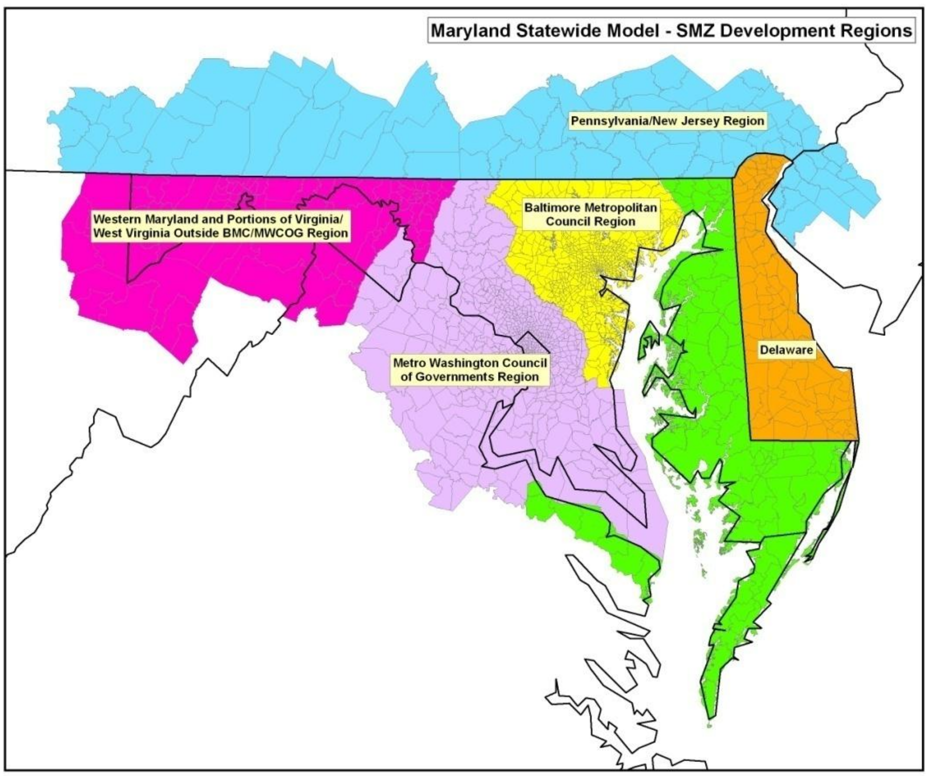
1990 - 2007

APPEDNIX C

Statewide Model Zones (SMZs)

SMZs are the polygon structures used in the statewide model and can be considered similar to TAZs in transportation planning. SMZs in the statewide model are equivalent to TAZs in high density development areas, or TAZs are nested under SMZs in the low density development areas. The MSTM SMZs were developed through an iterative process. The outer study area was identified from analysis of 2000 Census Transportation Package (CTPP) data on labor flows in/out of Maryland. Within this larger boundary, six regions were identified for SMZ formation, treating each region as a separate entity with its own datasets and issues. These regions are shown in Figure 1.

Figure 37: Regions used to develop SMZs



The remainder of this section discusses the process and assumptions made in developing SMZs for each of these sub-regions and overall. The goal was to respect the following major factors in the development of the SMZs.

- To the extent possible, SMZs conform to census geography to best utilize census data products in model development/updates and model calibration/validation. However, Washington MPO TAZs are retained, and do not follow census geography.
- SMZs must nest within Counties and conform to County boundaries.

- Aggregations of MPO zones, to facilitate linkages between MPO and statewide models.
    - Within Washington and Baltimore MPO areas, SMZs should be equal to or aggregations of MPO TAZs and nest within the MPO's TADs/RPDs.
    - SMZs should be more uniform in size than TAZs. In general, SMZ should be greater than 0.25 and less than 10 square miles. There should be greater aggregation in central areas where MPO TAZs are smaller (often individual street blocks) and little to no aggregation of larger MPO TAZs.
- SMZs should not straddle freeways, major rivers or other natural barriers.
- SMZs should separate the traffic sheds of major roads. MPO TAZs on opposite sides of a major road can be combined to define a traffic shed or corridor.
- SMZs should separate activity centers from surrounding areas and, where the activity center has been subdivided into multiple MPO TAZs, group adjacent TAZs into a single SMZ.

In each region, SMZs were developed with reference to various GIS overlays.

- MPO or other TAZ GIS shape file (where available) with activity density (ActDen) symbology (where TAZ data available) and Labels = TAZ number.

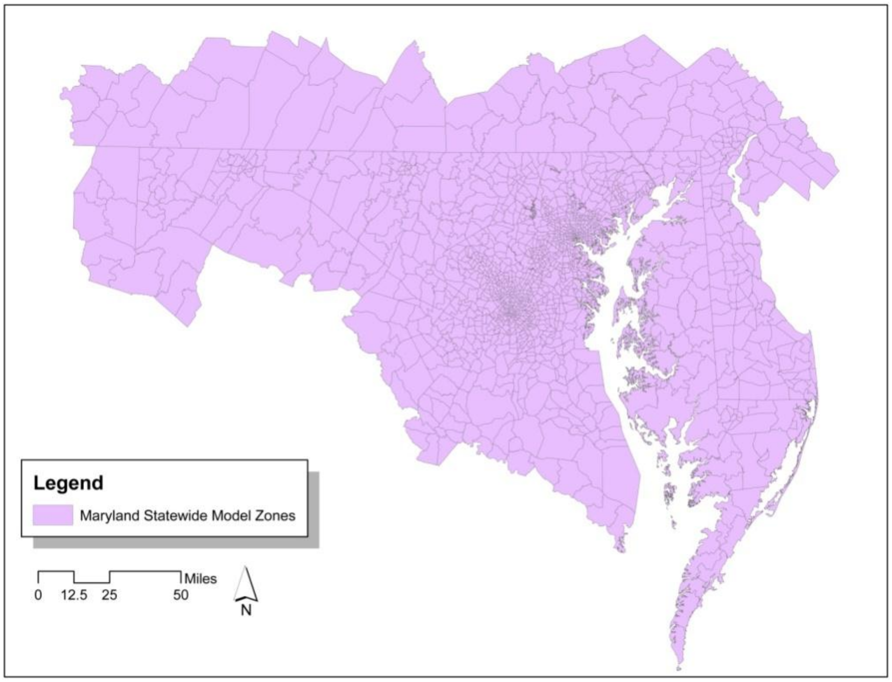
  Activity Density maps, calculated from historic/forecast demographic and acreage in areas of Maryland where TAZ demographic data is not available;
- Where TAZ shape files and related data are not available, use statewide land use or zoning coverage instead of Activity Density.
- Major roads coverage, from MPO networks where available, with Freeways and Major Arterials highlighted.
- MPO analysis districts (i.e., TAD or RPD) boundaries, where relevant.
- County boundaries.

The process for developing the zones consisted of a first cut based on the criteria above followed by review by SHA and other team members. Comments were addressed and conflicting comments resolved. During a final review the following additional changes were made:

- Isolate protected or restricted development lands for the land use model.
- Baltimore and District central business district aggregation to provide somewhat more uniform SMZ size and accentuate downtown activity levels on par with suburban centers.
- Distinctions were made to delineate areas with good accessibility to Metro rail stations.

To the extent possible, the SMZ boundaries outside the MPOs and Eastern Maryland were made to distinguish rural from urban/suburban development zoning boundaries, with zones centered upon activity/town centers and major crossroads.

Figure 38: Statewide Modeling Zones in MSTM



References

Alonso, W. (1964) Location and Land Use: Toward a General Theory of Land Rent. Harvard University Press: Cambridge, MA.

Alperovich, G. (1980). Determinants of population density gradient in Tel-Aviv metropolitan area. Urban Studies, 17, 185-192.

Alperovich, G. (1995). The Effectiveness of Spline Urban Density Functions: An Empirical Investigation. Urban Studies, Vol. 32, No. 9, 1537-1548.

Amromin, G. and A. L. Paulson (2009). Comparing Patterns of Default among Prime and Subprime Mortgages. Federal Reserve Bank of Chicago Economic Perspectives Q2, 18–37.

An, L., Brown, D.G. (2008). Survival Analysis in Land Use Change Science: Integrating with GIScience to Address Temporal Complexities. Annals of the American Association of Geographers, 98: 323 – 344.

Anas A., Arnott R., Small K.A. (1998). Urban Spatial Structure. Journal of Economic Literature, 36: 1426 -1464.

Anderson, J. (1982). Cubic spline urban density functions. Journal of Urban Economics, 12, 55-167.

Anderson, J. (1985a). Estimating generalized urban density functions. Journal of Urban Economics, 18, 1-10.

Anderson, J. (1985b). The changing structure of a city: Temporal changes in cubic spline urban density patters. Journal of Regional Science, 25, 413-425.

Amromin, G., and A. L. Paulson (2009). Comparing Patterns of Default Among Prime and Subprime Mortgages. Economic Perspectives, 33(2), 18–37.

Apgar, W.C., Duda, M (2005). Mortgage Foreclosures in Atlanta: Patterns and Policy Issues. NeighborWorks America, Washington, DC.

Batty, M., Kim, K.S. (1992). Form Follows Function: Reformulating Urban Population Density Functions. Urban Studies: 29: 1043 – 1070.

Batty M., Longley, P.A. (1987). Urban Shapes as Fractals. Area, 19: 215 – 221.

Batty M., Longley, P.A. (1994). Fractal Cities: A Geometry of Form and Function. London, UK: Academic Press.

Batty, M. and Xie, Y. (1996). Preliminary Evidence for a Theory of the Fractal City, Environment and Planning A, 28, 1745-1762.

Berkovec, J. A., Canner, G.B., Gabriel, S.A., and Hannan, T.H. (1994). Race, Redlining, and Residential Mortgage Loan Performance. Journal of Real Estate Finance and Economics, 9(3): 263-94.

Bertaud, A., and Malpezzi, S. (2003) The Spatial Distribution of Population in 48 World Cities: Implications for Economies in Transition. Wisconsin Real Estate Department Working Paper.

Bertaud, A., and Bertrand, R. (1997). Socialist Cities without Land Markets. Journal of Urban Economics, Elsevier, vol. 41(1), pages 137-151, January.

Bertaud, A. (2003). A Measure of Spatial Organization of 7 Large Cities. (unpublished paper)

Birch, E. (2009). Downtown in the "New American City". The ANNALS of the American Academy of Political and Social Science, vol. 626 no. 1

Bhutta, N., Dokko, J.K., and Shan, H. (2010). The Depth of Negative Equity and Mortgage Default Decisions. Federal Reserve Board of Governors Finance and Economics Discussion Series Working Paper 35.

Bogart, W.T. (2006). Don’t Call it Sprawl: Metropolitan Structure in the Twenty-first Century. New York, NY: Cambridge University Press.

Boots, B.N., Getis, A. (1988). Point Pattern Analysis. Newbury Park, CA: Sage.

Bradford, D., and Kelejian, H. (1973). An Econometric Model of the Flight to the Suburbs. Journal of Political Economy, May/June 1973, 81:3, 566-89.

Brinkman, J. (2008). An Examination of Mortgage Foreclosures, Modifications, Repayment Plans and Other Loss Mitigation Activities in the Third Quarter of 2007. Mortgage Bankers Association Working Paper.

Brueckner, J.K. (1982). A Note on the Sufficient Conditions for Negative Exponential Population Densities. Journal of Regional Science, 22: 353 – 359.

Brueckner, J.K. (1987). The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills Model. In Mills ES (ed.) Handbook of Regional and Urban Economics, Vol. II, pages 821 – 845. Amsterdam: North-Holland.

Brueckner, J.K.(2000a). Urban Sprawl: Diagnosis and Remedies. International Regional Science Review. 23:2, pp. 160–71.

Brueckner, J.K. (2000b). Urban growth models with durable housing: An overview. In: Hurion JM, Thisse JF (eds) Economics of cities: Theoretical perspectives. The Cambridge University Press, Cambridge, UK.

Burby, R.J., and May, P.J. (1997). Making governments plan: State experiments in managing land use. Baltimore: Johns Hopkins University Press.

Burchfield, M., Overman, H.G., Puga, D., and Turner, M.A. (2006). Cause of Sprawl: A Portrait from Space. Quarterly Journal of Economics (May): 587–633.

Capozza, D.R., Helsley, R. W. (1990). The stochastic city. Journal of Urban Economics, Volume 28, Issue 2.

Carruthers, J.I., Lewis, S., Knaap, G.J., Renner, R.N. (forthcoming) The American Way of Land Use: A Spatial Hazard Analysis of Changes Through Time. International Regional Science Review.

Carruthers, J.I., Lewis, S., Knaap, G.J., Renner, R.N. (2010). Coming Undone: A Spatial Hazard Analysis of Urban Form. Papers in Regional Science, 89: 65 – 88.

Carruthers, J.I .(2002a). The Impact of State Growth Management Programs: A Comparative Analysis. Urban Studies 39: 1959-1982.

Carruthers, J.I. (2002b). Evaluating the Effectiveness of Regulatory Growth Management Programs: An Analytical Framework. Journal of Planning Education and Research, 21: 406 – 420.

Casetti, E. (1969). Alternative population density models: An analytical comparison of their validity range. In Studies in Regional Science edited by Scott, A., pp105-116. Pion, London.

Chan, S., Gedal, M., Been, V., and Haughwout, A.F. (2010). The Role of Neighborhood Characteristics in Mortgage Default Risk: Evidence from New York City. Furman Center for Real Estate and Urban Policy, Working Paper 2010

Clark, C. (1951). Urban Population Densities. Journal of the Royal Statistical Society, 114: 490 –494.

Clark, W. A. V. (1992) Comparing cross-sectional and longitudinal analyses of residential mobility and migration. Environment and Planning A 24: 1291-1302.

Clifton, K., Ewing, R., Knaap, G.J., and Song, Y. (2008). Quantitative analysis of urban form: a multidisciplinary review. Journal of Urbanism: International Research on Placemaking and Urban Sustainability, 1:1, 17 - 45

Coulton, C., Chan, T., Schramm, M., Mikelbank, K. (2008). A Longitudinal Study of Mortgage Loans, Cleveland and Cuyahoga County, 2005-2008. Center on Urban Poverty and Community Development, Mandel School of Applied Social Sciences, Case Western Reserve University. Cleveland OH.

Cutsinger, J., Galster, G.,Wolman, H., Hanson, R., and Towns, D. (2005). Verifying the Multi-Dimensional Nature of Metropolitan Land Use: Advancing the Understanding and Measurement of Sprawl. Journal of Urban Affairs 27, 235–259.

Danis, M. A., and Pennington-Cross, A.N. (2005). The Delinquency of Subprime Mortgages. Federal Reserve Bank of St. Louis Working Paper 2005-022A.

Davies Withers, S. (1997). Methodological considerations in the analysis of residential mobility: A test of duration, state dependence, and associated events. Geographical Analysis 29: 354-72.

Dawkins, C.J., and Nelson, A.C. (2004). Urban Containment in the United States. Washington, DC : American Planning Association.

DeBorger, B. (1979). Urban population density functions: Some Belgian evidence. Annuals in Regional Science, 13, 15-24.

Demyanyk, Y. (2009). Ten Myths about Subprime Mortgages. Federal Reserve Bank of Cleveland Economic Commentary.

DHCD (Maryland Department of Housing and Community Development) (2009). Property Foreclosures in Maryland Third Quarter 2009. http://www.mdhope.org/Documents/PropertyForeclosureEventsinMaryland09Q3.pdf

Diggle, P.J. (2003). Statistical Analysis of Spatial Point Patterns. New York, NY: Arnold.

Ding, L. Quercia, R.G., Li, W., Ratcliffe, J. (2008). Risky Borrowers or Risky Mortgages: Disaggregating Effects Using Propensity Score Models. Working Paper. Center for Community Capital, University of North Carolina: Chapel Hill, NC.

Doms, M., Furlong, F., & Krainer, J. (2007). Subprime Mortgage Delinquency Rates. Federal Reserve Bank of San Francisco Working Paper.

Dowell, M. and Pitkin, J. (2009). Demographic Forces and Turning Points in the American City, 1950-2040. The ANNALS of the American Academy of Political and Social Science, vol. 626 no. 1.

Downs, A. (1997). The Challenge of Our Declining Big Cities. Housing Policy Debate 8, 359–408.

Downs, A. (1999). Some Realities about Sprawl and Urban Decline. Housing Policy Debate. 10:4, pp. 955–74.

Duany A, Plater-Zyberk E, Speck J. (2000). Suburban nation. The rise of sprawl and the decline of the American dream. New York: North Point Press.

Edmiston, K. D. (2009). Characteristics of High-Foreclosure Neighborhoods in the Tenth District. Federal Reserve Bank of Kansas City Economic Review (2nd Quarter):51–75.

Edmonston, B. (1975). Population Distribution in American Cities. Lexington, D.C.: Heath and Company.

Eldridge, N. (1984). Melbourne: Population density; distance; the relationship. Journal of Urban Economics 155, 310-316.

Elul, R., S. Chomsisengphet, D. Glennon, R. Hunt, and N. Souleles (2010). What 'Triggers' Mortgage Default? American Economic Review, Papers and Proceedings.

Esparza, A.X., Krmenec, A. (1994). Business Services in the Space Economy: A Model of Spatial Interaction. Papers in Regional Science, 73: 55 – 72.

Ewing, R. (1997) Is Los Angeles-style sprawl desirable? Journal of the American Planning Association 63(1): 107-126.

Ewing, R., Pendall, R., and Chen, D. (2002) Measuring Sprawl and Its Impact. Washington, DC: Smart Growth America.

Foote, C.L., Gerardi, K., and Willen, P.S. (2008). Negative Equity and Foreclosure: Theory and Evidence. Journal of Urban Economics, 64 (2): 234-245.

Foote, C. L., Gerardi, K.S., Goette, L.F, and Willen, P. (2009). Reducing Foreclosures. Federal Reserve Bank of Boston Public Policy Discussion Paper 2009-02.

Foster, C., and Van Order, R. (1984). An Option-based Model of Mortgage Default. Housing Finance Review, 3 (4), pp. 351-72.

Fotheringham , A.S., Batty, M., Longley, P.A. (1987). Diffusion-Limited Aggregation and the Fractal Nature of Urban Growth. Papers in Regional Science, 67:55 – 69.

Frankhauser, P. (1994). Urban structure fractality. Paris, France: Anthropos, Collection Villes.

Frey, W.H. (2009). Big City Populations Survive the Housing Crunch. The Brookings Institution: Metropolitan Policy Program, Washington, DC.

Fulton, W., Pendall, R., Nguyen, M., and Harrison, A. (2001). Who Sprawls Most? How Growth, Patterns Differ Across the U.S. Washington, DC: Fulton Center on Urban & Metropolitan Policy, The Brookings Institution.

Fujita, M., Krugman, P. And Venables, A. J. (1999). The Spatial Economy: Cities, Regions, and International Trade. Cambridge, MA: MIT Press.

Galster, G.,Hanson, R, Ratcliffe, M., Wolman, H., Coleman, S. and Freihage, J. (2001). Wrestling Sprawl to the Ground: Defining and Measuring an Elusive Concept. Housing Policy Debate 12:4. 681-717.

Getis, A. (1964). Temporal Land Use Pattern Analysis with the Use of Nearest Neighbor and Quadrat Methods. Annals of the American Association of Geographers, 54: 391 – 399.

Getis, A. (1983). Second-order analysis of point patterns: The case of Chicago as a multi-center urban region. Professional Geographer 35: 73–80.

Getis, A., and J. K. Ord. (1992). The Analysis of Spatial Association by Use of Distance Statistics. Geographical Analysis 24, no. 3.

Gerardi, K.S., Willen, P. S. (2008) Subprime mortgages, foreclosures, and urban neighborhoods. The B.E. Journal of Economic Analysis and Policy 9(3) Symposium Article 12.

Gerardi, K., Lehnert, A., Sherlund, S. M, and Willen, P. (2008). Making Sense of the Subprime Crisis. Brookings Papers on Economic Activity, Fall 2008, pp. 69-145.

Gini, C. (1936). On the Measure of Concentration with Special Reference to Income and Statistics. Colorado College Publication, General Series No. 208, 73-79.

Glaeser, E. L., Gottlieb, J. D. (2006). Urban Resurgence and the Consumer City. Urban Studies, 43: 1275-1299.

Glaeser, E.L., Kahn, M.E .(2004). Sprawl and Urban Growth. In Henderson JV, Thisse JF (eds.) Handbook of Urban and Regional Economics, 4. North-Holland, The Netherlands.

Glaeser, E., and Shapiro, J. (2001). Is There a New Urbanism? The Growth of U.S. Cities in the 1990s. NBER Working Papers 8357, National Bureau of Economic Research, Inc.

Glavac, S., and B. Waldorf. (1998). Segregation and residential mobility of Vietnamese immigrants in Brisbane, Australia. The Professional Geographer 50: 344-57.

Gramlich, E. M. (2007). Subprime Mortgages: America's latest boom and bust. Washington, DC: Urban Institute Press.

Haughwout, A., Peach, R.,and Tracy, J. (2008). Juvenile Delinquent Mortgages: Bad Credit or Bad Economy? Journal of Urban Economics, 64(2): 246-57.

Harris, Curtis C. Jr. (1968). A Stochastic Process Model of Residential Development. Journal of Regional Science, Volume 8, Issue 1.

Helms, A. (2003). Understanding gentrification: an empirical analysis of the determinants of urban housing renovation. Journal of Urban Economics 54: 474-498.

Howell-Moroney, M. (2007). Studying the Effects of the Intensity of US State Growth Management Approaches on Land Development Outcomes. Urban Studies, 44, 2163–2178.

HUD (2008) U.S. Housing Market Conditions, 2nd quarter 2008. U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

HUD (Department of Housing and Urban Development) (2010). Report to Congress on the Root Causes of the Foreclosure Crisis. Office of Policy and Development Research, Washington, DC.

Immergluck, D. (2010). The accumulation of lender-owned homes during the US mortgage crisis: examining metropolitan REO inventories. Housing Policy Debate, 619-645.

Immergluck, D. (2009a) Intrametropolitan Patterns of Foreclosed Homes: ZIP-Code-Level Distributions of Real-Estate-Owned (REO) Properties during the U.S. Mortgage Crisis. Federal Reserve Bank of Atlanta, Discussion Paper 01-09.

Immergluck, D. (2009b). Foreclosed: High-Risk Lending, Deregulation, and the Undermining of America's Mortgage. Ithaca, NY: Cornell University Press.

Immergluck, D., & Smith, G. (2005). Measuring the Effects of Subprime Lending on Neighborhood Foreclosures: Evidence from Chicago. Urban Affairs Review, 40, 362–389.

Immergluck, D., and Smith, G. (2006). The External Costs of Foreclosure: The Impact of Single-Family Mortgage Foreclosures on Property Values. Housing Policy Debate, 17, 57–79.

Immergluck, D., Smith, G. (2006). The Impact of Single-Family Mortgage Foreclosures on Neighborhood Crime. Housing Studies, vol. 21, no. 6, November 2006, pp. 851-66.

Ingram, G. K. , Carbonell, A. , Hong, Y. H. and Flint, A. (eds) (2009). Smart growth policies: An evaluation of programs and outcomes Lincoln Institute of Land Policy, Cambridge, MA.

Irwin, E.G., Bockstael, N.E. (2002). Interacting Agents, Spatial Externalities, and the Endogenous Evolution of Residential Land Use Patterns. Journal of Economic Geography, 2: 31 – 54.

Irwin, E.G., Bockstael, N.E. (2007). The Evolution of Urban Sprawl: Evidence of Spatial Heterogeneity and Increasing Land Fragmentation. Proceedings of the National Academy of Sciences, 104: 20672 – 20677.

James, I. R. (1989), Unemployment duration: Modeling and estimation with particular reference to the Australian Longitudinal Survey. Australian Journal of Statistics 31A: 197-212.

Jaret, C., Ghadge, R., Reid, L.W., and Adelman, R.M. (2009). The Measurement of Suburban Sprawl: An Evaluation. City & Community 8:65-84.

Ji, W., Ma, J., Twibell, R.W., and Underhill, K. (2006). Characterizing urban sprawl using multi-stage remote sensing images and landscape metrics. Computers, Environment and Urban Systems 2006; 30(6):861–879, November.

Jiang, W., Nelson, A.A., and Vytlacil, E. (2009). Liar's Loan? Effects of Origination Channel and Information Falsification on Mortgage Delinquency. Columbia University Business School Working Paper.

Johnson, S.R., Kau, J.B. (1980). Urban Spatial Structure: An Analysis with a Varying Coefficient Model. Journal of Urban Economics, 7: 141 – 154.

Jordan, S., Ross, J.P and Usowski. K.G. (1998). U.S. Suburbanization in the 1980s. Regional Science and Urban Economics, 28, pp. 611-27.

Katz, B. (2002). Smart Growth: The Future of the American Metropolis? London, UK: Centre for Analysis of Social Exclusion, London School of Economics. CASE paper 58.

Kau ,J.B., Lee, C.F. (1976). The Functional Form in Estimating the Density Gradient: An Empirical Investigation. Journal of the American Statistical Association, 71: 326 – 327.

Kau, J.B., Lee, C.F. (1977). A Random Coefficient Model To Estimate a Stochastic Density Gradient. Regional Science and Urban Economics, 7: 169 – 177.

Kau, J.B., Lee, C.F., and Chen, R.C. (1983). Structural Shifts in Urban Population Density Gradients: An Empirical Investigation. Journal of Urban Economics, 13: 364 – 377.

Kim, S. (2007). Changes In the nature of urban spatial structure in the United States, 1890-2000. Journal of Regional Science, 47, 2, 273-287.

LaCour-Little, M., Calhoun, C.A, Yu, W. (2011). What role did piggyback lending play in the housing bubble and mortgage collapse? Journal of Housing Economics, Volume 20, Issue 2, Pages 81-100.

Lang, R. E. (2003). Edgeless City. Washington, DC: Brookings Institution Press.

Lang, R.E. and Simmons, P.A. (2001). "Boomburbs": The Emergence of Large, Fast-Growing Suburban Cities in the United States. Washington, DC: Fannie Mae Foundation. Available: www.fanniemaefoundation.org/programs/census_notes_6.shtml

Latham, R. and Yeates, M. (1970). Population density in metropolitan Toronto. Geographical Analysis, 2:177 -. 185.

Long, J. S., and J. Freese. (2006). Regression models for categorical dependent variables using Stata. 2nd ed. College Station, TX: Stata Press.

Longley, P.A., and Mesev, V. (1997). Beyond Analogue Models: Space Filling and Density Measurement of an Urban Settlement. Papers in Regional Science, 76: 409 – 427.

Longley, P.A., and Mesev, V. (2000). On the Measurement and Generalization of Urban Form. Environment and Planning A, 32: 473 – 488.

Longley, P.A., and Mesev, V. (2002). Measurement of Density Gradients and Space Filling in Urban Systems. Papers in Regional Science, 81: 1 – 28.

Lopez, R., and Hynes, H. P. (2003). Sprawl in the 1990s: Measurement, Distribution, and Trends. Urban Affairs Review 38, 325–355.

Lucy, W. (2010). Foreclosing the Dream: How America's Housing Crisis Is Changing Our Cities and Suburbs. APA Planners Press.

Lucy, W., and D. Phillips. (2001). Suburbs and the Census: Patterns of Growth and Decline. Washington, DC: The Brookings Institution.

Macauley, M. (1985). Estimating recent behavior of urban population and employment density gradients. Journal of Urban Economics, 18, 301-310.

Malpezzi, S., and Guo, W. (2001). Measuring ''sprawl'': Alternative measures of urban form in U.S. metropolitan areas. Unpublished manuscript, Center for Urban Land Economics Research, University of Wisconsin, Madison.

Marshall, J. (2007). Urban Land Area and Population Growth: A New Scaling Relationship for Metropolitan Expansion. Urban Studies, Vol. 44, No. 10, 1889-1904.

Mayer, C., and Pence, K. (2009). Subprime Mortgages: What, Where, and to Whom?, in Edward Glaeser and John Quigley, eds. Housing Markets and the Economy: Risk, Regulation, and Policy. Cambridge, MA: Lincoln Land Institute of Land Policy.

Mayer, C., Pence, K., and Sherlund, S. (2009). The Rise in Mortgage Defaults. Journal of Economic Perspectives, 23(1): 27-50.

McDonald, J.F. (1989). Econometric Studies of Urban Population Density: A Survey. Journal of Urban Economics 26, 361-385.

McDonald, J.F. and Bowman, H. (1976). Some tests of alternative urban population density functions. Journal of Urban Economics, 3, 241-252.

McMillen D.P. and McDonald, John F. (1998). Population Density in Suburban Chicago: A Bid-Rent Approach. Urban Studies 35, 1119-1130

Mieszkowski, P., and Mills, E.S. (1993). The Causes of Metropolitan Suburbanization. Journal of Economic Perspectives 7(3), 135-147.

Mills, E. (1967) An aggregative model of resource allocation in a metropolitan area. American Economic Review, 57, 197-210.

Mills E. (1972). Studies in the Structure of the Urban Economy. Johns Hopkins University Press: Baltimore, MD.

Mills, E. S & Lubuele, L. (1994). Performance of Residential Mortgages in Low- and Moderate-Income Neighborhoods. The Journal of Real Estate Finance and Economics, Springer, vol. 9(3), pages 245-60.

Moran, P.A.P. (1950). Notes on Continuous Stochastic Phenomena. Biometrika, 37, 17–33.

Morello, C., Keating, D., Hendrix, S, (2011, May 5). Census: Young adults are responsible for most of D.C.'s growth in past decade. Washington Post. http://www.washingtonpost.com/local/census-young-adults-are-responsible-for-most-of-dcs-growth-in-past-decade/2011/05/04/AFJz5LtF_story.html

Mumford, L. (1961). The City in History: Its Origins, Its Transformations, and Its Prospects. Harcourt, Brace and World, New York.

Muniz, I., Galindo, A. and Garcia, M.A. (2003). Cubic spline population density functions and satellite city delimitation :The case of Barcelona. Urban Studies, 40, 7, 1303-1321.

Muth, R.F. (1969) Cities and Housing. University of Chicago Press: Chicago, IL.

Nassar, J. (2007). Foreclosure, predatory mortgage and payday lending in America's cities. Testimony before the U.S. House Committee on Oversight and Government Reform.

Nasser, H. E., and Overberg, P. (2001). What you don't know about sprawl. USA Today, 22 February, 1A, 6A–9A.

Narendranathan, W., and M. B. Stewart. (1993). Modeling the probability of leaving unemployment: Competing risks models with flexible base-line hazards. Applied Statistics 42: 63-83.

Nelson, A. C. (2004). Toward a new metropolis: The opportunity to rebuild America. Washington, DC: Brookings Institution.

Newling, B. (1969). The spatial variation of urban population densities. Geographical Review, 59, 242-252.

Oates, W., Howrey, W., and Baumol, W. (1971). The Analysis of Public Policy in Urban Models. Journal of Political Economy, January/February 1971, 79, 142-53.

Odland, J. (1997). Longitudinal approaches to analysing migration behaviour in the context of personal histories. In Recent developments of spatial analysis, edited by M. M. Fischerand A. Getis, 149-70. New York: Springer-Verlag.

Odland, J., Ellis, M. (1992). Variations in the spatial pattern of settlement locations: An analysis based on proportional hazards models. Geographical Analysis 24: 97–109.

Okah, E., and Orr, J. (2010). Subprime Mortgage Lending in New York City: Prevalence and Performance. Staff Reports, no. 432, February.

Ong, P., and D. Pfeiffer. (2008). Spatial variation in foreclosures in Los Angeles. Unpublished manuscript.

Pedersen, C., and Delgadillo, L. (2007). Residential Mortgage Default in Low- and High-Minority Census Tracts. Family and Consumer Sciences Research Journal, 35(4): 374-91.

Peiser, R. (1989). Density and Urban Sprawl. Land Economics, 65 (3), 193-294.

Perloff, H.S., Dunn, E.S., Lampard, E.E., Muth, R.F. (1960). Regions, Resources, and Economic Growth. Baltimore, MD: Johns Hopkins.

Pushkarev, B., and J. Zuban (1977). Public Transportation and Land use Policy. Indiana University Press, Bloomington, Indiana.

Quercia, R., and Stegman, M. (1992). Residential Mortgage Default: A Review of the Literature. Journal of Housing Research, 3, no. 2.

Raftery, A.E. (1996). Approximate Bayes factors and accounting for model uncertainty in generalized linear models. Biometrika (83: 251-266).

Rauterkus S.Y., Thall,G. and Hangen, E. (2010) Location Efficiency and Mortgage Default. The Journal of Sustainable Real Estate, Vol.2, No 1.

Roark, W. (2006). Concise Encyclopedia of Real Estate Business Terms. The Haworth Press, New York.

Schloemer, E., Li,W., Ernst, K., and Keest, K. (2006). Losing Ground: Foreclosures in the Subprime Market and Their Cost to Homeowners. Center for Responsible Lending.

Shen, G. (2002). Fractal dimension and fractal growth of urbanized area. International Journal of Geographical Information Science, 16(5), 419-437.

Schwartz, A., Susin, S., and Voicu, I. (2003). Has falling crime driven New York's real estate boom? Housing Policy Debate, 14(1), pp. 101–135.

Smart Growth Network. (2011). Smart Growth Online: http://www.smartgrowth.org/engine/index.php/principles/

Sierra Club (1998). The Dark Side of the American Dream: The Costs and Consequences of Suburban Sprawl, Challenge to Sprawl Campaign, College Park, MD.

Thomas, J. (2009). Residential Construction Trends in America's Metropolitan Regions Development, Community, and Environment Division. U.S. Environmental Protection Agency, Washington, DC.

Tiebout, Charles, M. (1956). A Pure Theory of Local Expenditure. Journal of Political Economy, October 1956, 64:5, 416-24.

Torrens, P.M. (2006). Simulating Sprawl. Annals of the American Association of Geographers, 96: 248 –275

Torrens, P.M. (2008). A Toolkit for Measuring Sprawl. Applied Spatial Analysis, 1: 5 – 36.

Tsai, Y.H. (2005). Quantifying urban form compactness vs sprawl. Urban Studies, 42, pp. 141—161.

Van Order, R., and Zorn, P. (2000). Income, Location and Default: Some Implications for Community Lending. Real Estate Economics, 28(3): 385-404.

von Thünen JH (1826) Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalekonomie. Hamburg, Germany.

Waldorf, B.S. (2003). Spatial Point patterns in a Longitudinal Framework. International Regional Science Review, 26: 269 – 288.

Wolman, H., Galster, G., Hanson, R., Ratcliffe, M., Furdell, K., and Sarzynski, A. (2005). The Fundamental Challenge in Measuring Sprawl: Which Land Should Be Considered? The Professional Geographer 57 (1): p. 94-104.

Wu, J. (2006). Environmental Amenities, Urban Sprawl, and Community Characteristics. Journal of Environmental Economics, and Management, 52, 527-547.

Zielinski, K. (1979). Experimental analysis of eleven models of urban population density. Environment and Planning A, 11, 629-641.