2013

Effects of Spatial Structure on Air Quality Level in U.S. Metropolitan Areas

Chang-shik Song
Cleveland State University

EFFECTS OF SPATIAL STRUCTURE ON AIR QUALITY LEVEL IN U.S. METROPOLITAN AREAS

CHANG-SHIK SONG

Bachelor of Arts in Public Administration
Konkuk University
February, 1994

Master of Arts in Public Administration
Yonsei University
February, 1998

Master of Public Administration
State University of New York at Albany
May, 2005

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN URBAN STUDIES AND PUBLIC AFFAIRS

at the

CLEVELAND STATE UNIVERSITY

MAY, 2013

This dissertation has been approved

for the Department of Urban Studies

and the College of Graduate Studies by

____

Dissertation Chairperson, William M. Bowen, Ph.D.

____

Department & Date

____

Brian A. Mikelbank, Ph.D.

____

Department & Date

____

Sugie Lee, Ph.D.

____

Department & Date

ACKNOWLEDGEMENT

Thank God for completing this dissertation with all of my heart and with all of my mind. God made it finally. I dare to confess that I have nothing without you, and that I cannot move forward without your intervention and encouragement.

My thanks to my advisor, Dr. William M. Bowen. His consistent encouragement and precious advice make me complete this dissertation very successfully. Without your advice I cannot stand on this moment. I would like to give my thanks to my committee, Dr. Brian A. Mikelbank. His great interests make me make a progress for my dissertation. I am eager to give my thanks to my committee, Dr. Sugie Lee. His invaluable assistance helps me get over a lot of difficult times. I would really like to thank a number of faculty members, staffs, Ph.D. students, Korean students and all other supporters in the Levin College of Urban Affairs, as well as all members at Korean Central Baptist Church of Cleveland. Their encouragement and prayers make me focus more on my doctoral studies.

I very deeply thank my parent and mom-in-law, my siblings and siblings-in-law in Korea for their endless support and love. I say to my wife, Hyo Jung Yu, and my son, JeongHwee William Song, “this dissertation is yours” and “your love and sacrifices make this dissertation completed.” May God bless all my family and supporters with his grace!

EFFECTS OF SPATIAL STRUCTURE ON AIR QUALITY LEVEL
IN U.S. METROPOLITAN AREAS

CHANG-SHIK SONG

ABSTRACT

The purpose of this dissertation is to investigate relationships between metropolitan spatial structure and air quality across U.S. metropolitan areas. Debates over compact city and sprawling development models as alternative patterns of metropolitan development and planning remain unsettled. This dissertation works from the hypothesis that compact regions with high-density, concentration, mixed land use, and better accessibility improve air quality.

To test the compact city hypothesis, this dissertation uses a combined spatial data of population, employment, government, land use, and air quality in 610 counties in U.S. metropolitan areas and their neighboring areas for 1990, 2000, and 2006. Indicators identified widely in literature are employed to measure compact city: land uses, density, concentration, accessibility, and centralization. This dissertation provides the empirical evidence on the basis of some stipulated causal relationships between compact regions

and air quality through multivariate regression models using spatial econometric analysis, that sheds light on the presence of spatial dependence between spatial variations in alternative spatial structures and changes in air quality level.

The empirical results show a number of interesting signs to the compact city hypothesis. Metropolitan areas with a higher percentage of developed open space or longer weighted average daily commute time bring out higher average air quality index values, leading to worsened air quality. On the contrary, metropolitan areas with a higher percentage of densely employed sub-areas produce lower average air quality index values, resulting in improved air quality.

The empirical findings contribute to the importance of compact development strategies, such as polycentric employment centers, on improved air quality over suburban sprawl in the United States towards successful sustainable metropolitan development and planning.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

1.1 Statement of the Problem

One of the most demanding principles of sustainable metropolitan development and planning is to improve environmental quality, which refers to the health of people and their natural environment (Berry et al., 1974; WCED, 1987; Wheeler, 2000; Paehlke, 2003). Improvements in environmental quality represent decreases in air pollution and water pollution and increases in protected and conserved land. Such improvements may lead to minimal public health threats associated with toxic chemicals or hazardous wastes, as well as less damage to agriculture, forestry, and natural ecosystems (Marquez & Smith, 1999). These improvements interact with the protected and conserved land providing environmental benefits such as water quality improvement and carbon sequestration

(Nowak, 2006; Kraft & Vig, 2003). Air quality improvements are greatly affected by the location decisions of people and firms. Households and firms tend to locate in areas with more benefits in social, economic, cultural, political, and ecological dimensions (Jacobs, 1961, 2001; Lynch, 1981; Capello, 2007). Increased settlement of residential and economic activities into specific areas, considered "urbanized" or "suburbanized," leads to significant changes in air quality in those areas.

Historical trends in suburbanization of metropolitan areas in America since the early 1800s have been expressed by two contrasting perspectives: concentration ("compactness") in an urban center and dispersion ("sprawl") from the central cities to the suburbs. A number of studies in the urban development and planning literature indicate that the consequences of suburban sprawl do more harm than good to public health and environmental health. The consequences are increased pollution, loss of open space in the landscape, and socioeconomic disparities that develop between urban areas and suburbs (Elkin, McLaren, & Hillman, 1991; Newman & Kenworthy, 1989, 1992, 1999; Ewing, 1994, 1997, 2000; Burchell et al., 1998). Those studies favor compact development patterns that utilize high-density, concentration, mixed land use, and better accessibility. Other studies posit that urban dispersion and low-density development outweigh the costs of sprawl leading to less congestion and pollution (Gordon, Richardson, & Jun 1991); more preferences and choices as to where to live and work (Gordon & Richardson, 1997; Glaeser & Kahn, 2003; Kahn, 2006); and a realization of the "American Dream" of homeownership, engagement with nature, and a livable community (Fishman, 1987; Hayden, 2004).

Debates over "compact city" and "sprawling development" models as alternative patterns of metropolitan development remain unsettled. A critical point of the debate is determining which of the patterns is more desirable for future metropolitan development with regard to the health of people and the environment. The compact city approach that holds itself as more sustainable than those sprawling patterns has gained wide acceptance (Wiersinga, 1997; Neuman, 2005). In this sense, the diagnosis and solutions for the negative impacts of suburban sprawl in America are addressed through "compact" or "smart" development strategies in metropolitan areas since the 1990s.

The urban literature has focused little attention on the link between alternative spatial structures and environmental quality as being embedded in a multi-dimensional context that comprises the interactions between people, firms, and governments in metropolitan space and over time, nor has it fully recognized the potential spatial dependence across neighboring areas (Anselin, 1988). Hence, a better understanding of the relationship between urban structures and environmental quality with respect to the health of people and the environment will be required for the future of metropolitan development and planning.

The critical problem concerns the extent to which the metropolitan structure influences environmental quality associated with public health in a multi-dimensional context over time. Which of the alternative development patterns is more desirable for future metropolitan development in consideration of the health of people and the environment? This dissertation works from the hypothesis that compact regions provide greater environmental quality improvements, considering air quality improvements as a proxy for environmental quality. Analyzing the empirical evidence and causal relationships between compact cities and air quality improvements shows the extent to which metropolitan structure influences air quality.

1.2 Purpose of the Dissertation

The purpose of this dissertation is three-fold: 1) to propose a comprehensive conceptual framework for the link between spatial configuration of metropolitan structures and air quality in a multi-dimensional context; 2) to investigate relationships between metropolitan spatial structures and air quality across U.S. metropolitan areas based on the proposed framework; and 3) to test for the presence of spatial dependence among neighboring areas, as well as for the magnitude and direction of spatial dependent effects on metropolitan structure and air quality.

This dissertation contributes to the knowledge surrounding the urban form debates between compact and sprawled cities. The research also adds insight into the relationships between alternative urban structures and environmental quality coupled with public health and environmental health. The effort contributes an advanced analytical framework that uses a combination of multidimensional measurements of metropolitan structure that are quantifiable in spatial terms referring to density, concentration, land-use diversity, accessibility, and centralization. The effort requires empirical evidence of spatial interaction effects for the alternative urban structures using spatial regression models. The practical policy contribution suggests implications for successful sustainable metropolitan development and planning, emphasizing the importance of compact development patterns over suburban sprawl in the United States.

1.3 Structure of This dissertation

This dissertation is composed of five chapters. Chapter 1 introduces a critical point of the competing debate over “compact city” and “sprawling development” models; its importance of the linkage between metropolitan spatial structure and air quality in a multi-dimensional context, the structure of this dissertation, and contributions to expected results.

Following this introductory chapter, Chapter 2 is divided into five major sections under the framework of a literature review. The first section defines metropolitan spatial structure rooted in location theory, regional development theory, and planning theory in the urban studies literature. Section 2 discusses the theoretical background of metropolitan spatial structure from the variety of disciplines that approach the subject. Section 3 provides descriptions of the three competing models of metropolitan spatial structures reflecting monocentric, polycentric, and sprawling patterns. Section 3 also describes the ways this dissertation applies each model to metropolitan structure, including variables reflecting the spatial distribution in population and employment in terms of land use changes, level of specialization in industrial structure, governmental structure, and other confounding factors. Section 4 investigates empirical evidence on the relations between alternative spatial structures and air quality. The final section of Chapter 2 provides some critique of the literature to testify the need for this dissertation research to address the limitations of the prior research on metropolitan spatial structures and air quality.

Chapter 3 describes the research methods of the dissertation, beginning with a conceptual framework to explain the relationships between metropolitan spatial structure, its intervening variables, and air quality. This dissertation employs 610 counties in metropolitan areas having air quality collected from air pollution monitoring sites. Based on the proposed conceptual framework, data sources and their measurements are discussed, particularly the multidimensional characteristics of metropolitan spatial structure – density, concentration, centralization, and accessibility in population or employment, and mixed land uses. With the hypotheses that drive the research, the last section of Chapter 3 outlines the empirical research design for the study, proposing an inter-regional analysis across 610 counties in the U.S. metropolitan areas for 1990, 2000, and 2006, through multivariate OLS regression models and spatial regression models.

Chapter 4 reports upon tests of the hypotheses. It identifies determinants of air quality level including interaction effects between spatial variation in alternative structures, characterized by density, concentration, accessibility, and centralization in terms of residential, economic, and land-use activities, and changes in air quality level as a proxy for environmental quality, as well as the presence of spatial dependence among neighboring counties in metropolitan areas.

Chapter 5 discusses major factors to improve air quality level, the presence of spatial dependence among metropolitan areas, and provides insight consistent with the empirical evidence regarding the causal links between compact regions and improved air quality level. This chapter then concludes with critical findings and policy implications towards compact development for urban planners, policy makers, and other stakeholders to tackle sprawling suburbs, relating to regional smart growth strategies. Lastly, this chapter describes limitations of this dissertation and the scope of further study.

CHAPTER II

LITERATURE REVIEW

2.1 Definitions of Metropolitan Spatial Structure (MSS)

Metropolitan spatial structure (MSS), as used in the urban research literature, is not easy to define because its structure has varied in space and over time (Gore, 1984). Some other terms found in the literature, such as "metropolitan spatial patterns (forms or shapes)," "metropolitan development patterns," "metropolitan suburbanization," or "urban spatial structure" will be used similarly throughout this dissertation. Several authors have formulated an understanding of metropolitan spatial structures as the spatial distribution of activities in terms of people, firms, and governments in space and over time in their location decisions (Lynch, 1981; Bourne, 1982; Anderson, Kanaroglou, & Miller, 1996; Tsai, 2005). Most notably, Kevin Lynch (1981) defined urban form as "the spatial pattern of the larger, inert, permanent physical objects in a city" (p. 47). Bourne (1982) referred to urban spatial structure in a comprehensive concept as a spatial system consisting of three elements: the urban form, urban interactions, and a set of organizing principles that determine the relationship between the first two. Anderson, Kanaroglou and Miller (1996) proposed that metropolitan development patterns as a process may represent spatial interactions (relations) among many significant elements and concepts that repeat and come together at the local and regional scale. They also explained that the results in changes in metropolitan development patterns may be characterized in terms of two simultaneous spatial trends: "the concentration of an increasing share of the population and economic activities into urban areas" (considered as concentric city) and the dispersion of population and economic activities within urban areas" (as dispersed city) (p.10). Tsai (2005) defined the spatial structure pattern of a metropolitan area as "the overall shape characterized by land use phenomena such as monocentric versus polycentric forms, centralized versus decentralized patterns, and continuous versus discontinuous developments" (p.142).

Such activities in a spatial setting were correlated and interdependent (Anselin, 1988; Irwin & Bockstael, 2002). The resultant changes in emerging metropolitan structures tended to be concentrated in some areas, centralized to the core area, or diversified in land uses in some areas. According to historic explanation of Anas, Arnott, and Small (1998), urban spatial structure is "the degree of spatial concentration of urban population and employment", along with not only the degree of "centralization" or "decentralization" of urban activities near the central business district (CBD) at the city-wide level, but also the degree of "clustering" or "dispersion" of the activities at a specific local level (p.1431).

2.1.1 Key Concepts of Metropolitan Spatial Structure

Metropolitan spatial structure (MSS) defined in this dissertation will be obviously explained by some key concepts. These concepts are rooted in location theory, regional development (growth) theory, and planning theory in the urban studies literature; they are particularly explained in geographic, social, economic, political, and ecological terms. First, one of the underlying concepts is the term "space." The conception of space refers to areas within a socioeconomic and ecological boundary beyond the level of administrative territories such as cities and townships. From an aspect of location theory and regional growth (development) theory, Capello (2007) highlighted space as "the source of advantages springing from the cumulative nature of productive processes in space" (p. 1) involved with spatial proximity, reduced transaction costs, agglomeration economies, and the spatial variations of activities, leading to geographical concentration and externalities in an urban context. From a seminal work in planning theory as to how greater cities operate in real life against traditional orthodox planning and rebuilding, [1] Jacobs (1961, 2001) conceived the term "cities" as the process of their death and life, which acted as "an immense laboratory of trial and error, failure and success" (p.6), arguing that cities in space were inherently embedded in diversity to give each other constant mutual support (p.14). In this sense, Guttenberg (1993) referred to "metropolitan" spaces as the use of tense in past, present, future natural (i.e. physiographic & biotic) and socio-cultural (i.e. beliefs, values, preferences, attitudes, rules, and habits) features of the spatial environment by human purposes (pp. 62-81). Also, the U.S. Census Bureau at the Office of Management and Budget (2003) defined "metropolitan" spaces as areas with greater population, larger jobs, and geographical expansion over time and in space. On the other hand, indicating urban form as a snapshot of process and an outcome of urbanization, Neuman (2005) envisioned the city (particularly the sustainable city) as

[1] The traditional urban planners proposed that the ideal forms of cities could be decentralized into individual cities or towns to enjoy individual freedom, prosperity, beauty, and lifestyle in a new urban order. The idealized planned cities between late 1800s and the 1930s were designed as follows: Howard's (1898) "the Garden City," Wright's (1932) "Broadacres," and Le Corbusier's (1935) "the Radiant City."

"the manifestation of many evolutionary processes between the city and its inhabitants and between the city and its environment" (p. 23).

Aligned with the concept of space, second, changes in emerging metropolitan structure over time will be closely related to "concentration," "spillover effects," and "externalities". Concentration of activities at the intra-metropolitan level can create lower production costs to the firms, the increasing size of the firms in the same industry sector, and the high density and variety of productive (i.e., innovative) and residential activities (Capello, 2007, pp. 17-20). Also, spillover effects since the importance of space conceptualized by Marshall (1920) can generate geographical clustering of innovative activities of different industries at the inter-metropolitan level (Capello, 2007, pp. 193-200; Maoh & Kanaroglou, 2007). In addition, while the spatial concentration promotes urban expansion and development, the change in the concentration of activities in the metropolitan context contributes to spatial decentralization and the emergence of metropolitan structures transforming from monocentric to polycentric or dispersed patterns. Such changes in emerging metropolitan structure over time bring out externalities associated with environmental costs (i.e. concentration of pollution, loss of open space).

In line with the concepts of space and metropolitan emergence, third, government mechanisms (Davis, 2002; Brown, 1999; DiGaetano & Klemanski, 1993; Stone, 1989; Stone & Sanders, 1987) will be one of the significant concepts in metropolitan spatial structure. Local jurisdictions are legal, institutional, and political entities with their own regulatory authority (i.e., zoning) to administer land use and land developments for households and developers, as well as planners and developers at a local level, used as locality (fragmentation) with a home rule (Tiebout, 1956). In this sense, political fragmentation may affect migration of people and firms. Regional, state and federal governments refer to hierarchies of legal and political systems over local jurisdictions through government spending and regulatory activities such as statewide growth management techniques.

Lastly, metropolitan spatial structure may be characterized by "tensions" between stakeholders, such as interests of households and governments or planners and developers, and homeowners' preferences and homebuilders' maximum profits. For example, there are tensions among households, homebuilders, and local or regional governments in location decisions (Hayden, 2004; Byun & Esparza, 2005; Vicino, 2008). A household's desire to achieve the American dream is likely to cause a move to suburban jurisdictions with a safer and cleaner environment and fewer growth controls. As well, homebuilders tend to move to suburban governments with less land use restrictions so that they may maximize their profits and meet households' preferences. Simultaneously, local governments are more competitively likely to attract households and homebuilders through non-controlled growth measures, while governments at the regional, state, and federal level tends to guide stringent land use planning measures. Such tensions between households, developers, governments, and some intervening variables will be taking place in a dynamic evolutionary process at a regional scale.

In summary, the term "metropolitan spatial structure" in the urban studies literature may be very difficult to define in a single and universal manner due to its abstractness. However, those elements and concepts that identify metropolitan structures - space, spillovers, and tensions between households' preferences, firms' profits, and government intervention - are required to better understand the metropolitan spatial structure, its emergence in space and over time, and its effects on air quality. The shape of metropolitan spatial structure (MSS) in this dissertation will be made up of the spatial distribution of people, firms, and governments in metropolitan space and the spatial variation of metropolitan structure over time in a continual and evolutionary process of those elements and concepts, representing from compact to polycentric or sprawling patterns.

2.2 History of Metropolitan Spatial Structure and its Theoretical Background

2.2.1 Transition to a Suburban Nation

Since World War II, advances in transportation (i.e., roads and vehicles) and communication (i.e., telegraph and telephone) have excelled, transforming the patterns of spatial development in metropolitan areas. Impacts of urban infrastructure and technological change enabled people and firms to move outward from the urban central area, leading to the creation of the suburban frontier (Anas, Arnott, & Small, 1998, pp.1428-1430). Suburb images after the post-World War II, noted "the decentralization of economic and residential life" at the farthest edges of metropolitan areas (Katz, 2002, p.4), remained the dominant growth patterns in the U.S. metropolitan areas, known as polycentric ("edge") city model with subcenters (Garreau, 1991) or sprawling ("edgeless") city model (Lang, 2003). The U.S. Census Bureau (2003) showed that more than 93 percentage of the U.S. residents live in metropolitan areas, and more than half of them with detached houses and automobile-dependent commuting live in the suburbs of metropolitan areas, to date.

The dominant suburban images are characterized by two different views in literature. Some explained suburban image in terms of the realization of the "American

Dream." Individuals living in the suburbs could seek affordable single-family housing, green nature, and neighborhood sociability for the pursuit of happiness (Fishman, 1987; Baxandall & Ewen, 2000; Hayden, 2004). Hayden (2004) highlighted the shapes of suburbia between 1820 and 2000 as the conflict between the triple dream of home, yard and community and the growth machines, which represented the complex process of "contestation between residents who wish to enjoy suburbia and developers who seek to profit from it" (p.9). Hayden contended that the "American dream" is intertwined with the seven historic suburban development patterns in the metropolitan landscape through borderlands around 1820, picturesque enclaves around 1850, streetcar buildouts around 1870, mail-order and self-built suburbs in 1900, mass-produced, urban-scale sitcom suburbs around 1940, edge nodes around 1960, and rural fringes around 1980. Emphasizing the historical patterns of suburbanization, Vicino (2008) described the making of a suburban nation as "the culmination of change in metropolitan residents' social characteristics, economic structure, desire for public services, and an improved built environment" (p. 378). On the other hand, some explained the suburban images as sprawl, which was conceived in terms of the unintended consequences of unbalanced, uncontrolled, or excessive growth. Its impacts could lead to loss of green space, aggravated environmental damages, and spatial disparities between urban and suburban areas (Real Estate Research Corporation, 1974; Ewing, 1994; Orfield, 1997, 2002; Burchell et al., 1998; Rusk, 1999; Katz, 2002).

2.2.2 Theoretical Background of Metropolitan Spatial Structure

From the post-World War II suburban image, five theoretical approaches to metropolitan structure may be identified. The most traditional approach is typically rooted in urban and regional economics from three primary schools. The bid-rent theory of von Thünen (1826) established the “monocentric city model”. The ecological models achieved by Park’s (1915) “the city,” McKenzie’s (1925) “human community,” and Hawley’s (1950) “human ecology,” helped explain urban and suburban decline after World War II. The third model was established by Hurd’s (1903) central and axial growth, Burgess’s (1925) concentric zones, Hoyt’s (1939) sectors, and Harris and Ullman’s (1945) multiple nuclei, which focused on the relationship between location and human activity in urban areas to account for emerging metropolitan structure.

The second approach describes emerging metropolitan structure may be focused in the urban and suburban decline theories in the development and planning literature, in which some scholars have tried to explain the creation of metropolitan suburbanization as a result of the tensions between the urban and older suburbs’ decline and rapid suburban growth. Aligned with the traditional ecological models, Bradbury, Downs, and Small (1982) and Mieszkowski and Mills (1993) commonly proposed suburbanization and urban decline in two main aspects: “natural-evolution” and “flight-from-blight.” The “natural-evolution” theory emphasized the migration of middle-class households and firms to suburban areas while leaving lower-income households behind in central and older suburbs, predicated on rising income levels. The “flight-from-blight” theory proposed that after World War II, non-Hispanic whites with increasing income levels were more likely to avoid the negative costs of the cumulative decline in central cities as well as in older suburbs, such as ethnic tensions, crime, tax, traffic congestion, environmental degradation, and other problems, to suburban or exurban areas, a phenomenon known as “white flight” (Frey, 1979; Massey & Denton, 1988, 1993).

Since the 1980s, urban research has emphasized that inner-ring suburban decline can contribute to increasing suburban expansion (growth) at the metropolitan fringe (Lee & Leigh, 2005; Hayden, 2004, see chapter 11; Lucy & Phillips, 1995, 2001, 2003; Orfield, 1997, 2002; Bollens, 1988; Jackson, 1985). Jackson (1985) pointed out that the inner-ring suburbs deteriorate, just as the central cities decline in terms of the filtering process of socio-economic features. Orfield (1997, 2002) identified that the inner-ring or at-risk suburbs decline more rapidly than the central cities and are less likely to suffer from decline, because of lack of the central cities' advantages, such as strong CBD, vital neighborhoods, amenities, and cultural resources. Also, Lucy and Phillips (2001) provided evidence of the decline of older suburbs for the 35 largest U.S. metropolitan areas between 1990 and 2000. They indicated that the slower-growing cities in the Midwest and Northeast are more likely to decline. Furthermore, Katz (2002) argued that the shape of metropolitan growth in America is characterized by "explosive sprawl where farmland once reigned, matched by decline or slower growth in the central cities and older suburbs. … The suburbs dominate employment growth … contain more people" (pp. 4-9). As argued above, the inner or older suburbs tended to decline rapidly due to lower income, poverty, population decline, and employment loss, thus leading to spatial decentralization of population and firms at the outer suburbs or exurbs.

The third approach relating to metropolitan suburbanization is the "market failure" explanation in urban economics and planning literature which considered sprawl as a process and spillovers. Byun and Esparza (2005) discussed how, since the 1970s, suburbanization grows and interacts with sprawl through a process-based conceptual model. They highlighted that local political fragmentation based on home rule (Tiebout, 1956) had an important role in the spatial shifts of households and homebuilders to the distant suburbs and urban fringe, leading to the "uncontrolled outward expansion of urban development" (p. 262).Their model identified market failures involving public goods that lead to sprawl, such as environmental quality, and externalities or spillover effects such as loss of open space, traffic congestion, air pollution, and social costs of inequality, racial segregation, and infrastructure (Ewing, 1997; Brueckner, 2000; Klosterman, 1985; Ridley & Low, 2001).

The fourth approach to understanding metropolitan structures is the urban regime theory from urban economics and politics literature (Davis, 2002; Brown, 1999; DiGaetano & Klemanski, 1993; Stone & Sanders, 1987). This approach focuses on the influence of public policies and government structures. The urban regime theory divides public policies and governmental structures into two views: polycentric or regional. The polycentric view began with Tiebout-style political fragmentation with a government's home rule (Tiebout, 1956; Ostrom, Tiebout, & Warren, 1961; Oates, 1972; Ostrom, 1974). Fragmented government measures such as local restrictive zoning might impact the location decisions of households and firms, thus contributing to changes in metropolitan structure. However, its measures ignored spillover effects between fragmented governments (Ward, 1987; Downs, 1994). Centering on effective solutions to such spillovers, the regional or central view focused on the role in metropolitan governmental structure beyond the Tiebout-based locality. This view also accounted for the influence of federal and state governments concerned with transportation and housing, such as the Federal Housing Acts and the Federal-Aid Highway Act after World War II, in transforming the shape of metropolitan spatial structures (Gelfand, 1982; Jackson, 1985; Aschauer, 1989; Holtz-Eakin, 1994; Kunstler, 1993; Boarnet, 1997; Transportation Research Board, 1995, 1997; Anas, Arnott, & Small, 1998; Voith, 1999, 2000; Gyourko & Voith, 2000; Fishman, 2000; Rusk, 2000; Peiser, 2001; Perky & Kurban, 2001; Katz, 2002; Byun & Esparza, 2005; Vicino, 2008).

Finally, the fifth approach to understanding metropolitan structure comes from theories and practices of sustainable metropolitan development patterns since the middle 1990s. The future patterns in metropolitan spatial structure are closely related with the definition of sustainable development, which refers to "development to enable to meet present generations' needs without compromising the ability of future generations to meet their needs" (WCED, 1987, p. 40), for the purposes of inter- and intra-generational equity, social justice, and environmental awareness (Haughton & Hunter, 1994).

In line with the conceptions of sustainable development, the future of metropolitan development presents a variety of forms: 1) statewide growth management strategies since the early 1970s to contain urban sprawl, preserve open space, farmlands and environmentally sensitive areas, and improve the quality of life, such as greenbelts, urban growth boundaries (UGB), urban service areas, urban containment policies, infill and redevelopment, zoning approaches, housing-related tools (Nelson & Duncan, 1995; Porter, 1997; Nelson, 2000; Nelson & Dawkins, 2004); 2) smart growth to contribute to save undeveloped land use, capital infrastructure consumption such as road and water/sewer, property development cost, and public service cost (Rutgers University, 1997; Maryland Office of Planning, 1997; Burchell, 2000; Nelson, 2001; Gillham, 2002; Katz, 2002); 3) new urbanism (or neotraditional development) since the early 1990s to curb suburban sprawl and inner-city decline, increase residential densities, enjoy neighborhood (community) lifestyles and encourage walking, mixed land use and fuel-efficiency, such as compact city, mixed-use development, transit-oriented development (TOD), pedestrian-oriented development, urban village, and walking urbanism (Duany & Plater-Zyberk, 1992; Calthorpe, 1993; Calthorpe & Fulton, 2001; Congress for New Urbanism (CNU), 1999; Ewing, 2000; Leinberger, 2007); 4) new regionalism since the 1990s to focus on the environment, equity, and efficiency under the interdependent approach between central cities and suburbs in a regional context, called a holistic approach, such as city-suburb cooperation, city-county consolidation, or joint city-suburb strategies (U.S. Housing and Urban Development, 1999; Benfield, Raimi, & Chen, 1999; Wachter, 2000; Frisken & Norris, 2001; Savitch & Vogel, 2001; Brenner, 2002; Wheeler, 2002; Fitzgerald & Leigh, 2002; Miller, 2002; Hamilton, 1999, 2013); and 5) eco-city to emphasize urban greening, ecological and cultural diversity, and passive solar design, such as eco-village, solar village, green city, sustainable housing, and sustainable community (Beatley, 2000; Beatley & Newman, 2009; Roseland, 1997, 2012).

For the practical applications of sustainable development, Jabareen (2006) proposed a sustainable urban form matrix in which form is desirable for sustainable and environmentally sound to contribute to practitioners and policy makers. He categorized the sustainable urban form matrix in terms of density, diversity, mixed land uses, compactness, sustainable transport, passive solar design, and greening. He asserted that "different urban forms contribute differently to sustainability" (p. 48), which accounts for that the ideal urban forms towards a sustainable city are closely involved in a high density and diversity, compact with mixed land uses, and less automobile dependency.

2.2.3. Summary Remarks

Urban scholars document that since the early 1800s the American metropolis has been characterized by spatial shifts of people and firms from urban centers to the suburbs and beyond, called metropolitan suburbanization. The spatial shifts can be understood as a process of tensions between residents' preferences to seek to live in a low-density and safe area and firms' (i.e., developers) profit maximization, as well as between private interest and public interest (Lincoln Institute of Land Policy, 2000; Kruse & Sugrue, 2006; Hanlon, Vicino, & Short, 2006; Vicino, 2008). As examined earlier, postwar suburbanization has been greatly promoted by the overwhelming impact of federal and state policies on the American metropolis such as transportation and housing policies, as well as the local political fragmentation. These government policies and systems dramatically affected the migration of people and firms to the suburbs, leading to rapidly sprawling development.

Impacts of suburban or sprawling development patterns can significantly spur geographic differentiation of decline in central cities and older and inner-ring suburbs, environmental degradation, excessive land consumption, loss of open space, racial segregation, and poverty concentration in blighted areas, whereas suburban sprawl as a realization of the American Dream can contribute to better lifestyles for those who can afford to live in dispersed suburbs. Since the 1990s, the aforementioned new forms of alternatives to conventional suburban sprawl under the term sustainable development emphasize an integrated regional approach to deal with social, economic, and environmental issues, which is concerned with both the new urbanism at the micro level and the new regionalism and growth management strategies at the macro level.

Taken together, postwar suburbanization in America tends to promote patterns of rapid sprawling development, due to improved transportation systems, technological advances such as automobiles, telephones and the Internet, and tensions between households' and firms' preferences and government strategies. Furthermore, the shape of future suburban development patterns will be transformed by a region's characteristics in space and time, which can reflect its complex social, economic, and political realities from CBD-oriented "compact" to "polycentric" or "sprawling" development patterns.

2.3 Theoretical Identifications of Metropolitan Spatial Structure

As explained previously, metropolitan spatial structure was emerged, formed, and transformed by spatial interaction (or distribution) of people and jobs in spatial, temporal, and political terms.

The modern theoretical foundations of metropolitan spatial structures (MSS) in urban and regional economics have evolved over the past four decades to empirically explain how metropolitan areas grew, following on the seminal work of a mathematical model of urban land use by Alonso (1964). IBI Group et al. (1990), Anderson et al. (1996), and Lang (2003) accounted for archetypal forms of metropolitan structure in terms of the distance from the central business district (CBD), representing “monocentric,” “polycentric,” and “sprawling” patterns.

Table 2-1 shows an evolutionary comparison of metropolitan spatial form, referring from some selected prior studies (Sharpe and Wallach, 1994; Anderson et al., 1996; Burchell et al., 1998; Lang, 2003).

Table 2-1 Competing Tensions of Metropolitan Spatial Structure

Form / Criteria	Monocentric	Polycentric	Sprawling
Key terms	CBD, concentric, centralized, single, high-density, downtown, core	Nodal, edge, concentric decentralized, suburbanized, village, subcenters, clustered, multinucleated, multicentered, specialized, clustered, fringe, high-density, countrified, disurb, outer, corridor	Edgeless, exurbanized, technoburbs, sprawling, low-density, post-polycentric
Time periods	post-war to present	mid- to late 1980s to present	late 1980s to present
Key figures	Alonso (1964), Mills (1967), Muth (1969)	McDonald (1987), Leinberger (1988), Garreau (1991)	Fishman (1990), Lang (2003)
Key forces	Agglomeration economies (services, high-tech)	Agglomeration economies (services, high-tech)	Agglomeration economies
Connection to CBD	stronger	strong	weak

2.3.1 The Monocentric Model

2.3.1.1 Theoretical Background

The monocentric models were originally based on two approaches: ecological models and traditional models. The two approaches were greatly influenced from von Thünen's (1826) land use theory for firms and households in urban areas, called *bid-rent* theory, depicting the relationships between location (defined as distance to the central city) and land rent (defined as a market price) at a given utility level. The ecological models with respect to human behavior in the city environment were developed by Robert E. Park's (1915) "the city", McKenzie's (1925) "human community", and Amos H. Hawley's (1950) "human ecology". These scholars regarded the growth of the city as a product of competition and cooperation, as well as a complex ecological process, leading to outward expansion, particularly by the size of the population, its concentration and distribution within the city area.

Under the influence of urban ecological approaches as a process of invasion and succession described above, the traditional models, established by Richard M. Hurd's (1903) central and axial growth, Ernest W. Burgess's (1925) concentric zones, Homer Hoyt's (1939) sectors, and Chauncy Harris and Edward Ullman's (1945) multiple nuclei (called the Chicago School), focused on spatial patterns of American large cities and suburbs in the first half of the twentieth century. These works contributed to theoretical explanations for the spatial pattern of urban growth resulting from roles in central zones such as proximity, accessibility to transportation systems, and internal characteristics of households. Subsequently, their efforts at specifying a generalized zonal pattern of urban growth influenced a wide range of thinking of contemporary urban scholars, such as urban economists, urban social and economic geographers, urban ecologists and environmentalists, urban planners and developers, and urban professional colleagues. However, many contemporary geographers, sociologists and urban economists have criticized those traditional models, alleging that they could cause an overly incomplete and inaccurate representation of the geography of American cities due to an abstract explanation of city growth patterns and processes in terms of succession and filtering (Harris, 1994; Dear & Flusty, 1998; Firey, 1947), failure of the city-suburban distinction that the commuters zone lie beyond city limits (Douglass, 1925; Ogburn, 1937; Queen & Thomas, 1939; Firey, 1946; Schnore, 1963; Rusk, 1993), ignorance of the sentimental and symbolic dimension of socioeconomic organization such as personal preference and motivation, the role of occupational status, culture (Claval, 2007; Firey, 1947), little attention to roles in political factors such as jurisdictions and policies, and no reflection of spatial impact related with population migration, employment activity, changes in housing market, and mutual relations between the cities.

2.3.1.2. Theoretical Assumptions and Empirical Applications

Alonso (1964) developed a mathematical model of urban land use based on von Thünen's (1826) bid rent theory. The developing works of Muth (1969), Mills (1967, 1972), and other scholars established an urban spatial model, known as a "monocentric city model", which emphasizes on the importance of the central business district (CBD) with respect to the degree of decentralization. The extensive work, such as a comparative static analysis of Wheaton (1974), Brueckner and Fansler's (1983) empirical study, and recently more evidence of McMillen (2006) and Spivey (2008), had a crucial role in identifying the spatial dimensions of urban and regional socio-economic activity grounded in urban economic theory.

Alonso-Mills-Muth models, as outlined empirically by Wheaton (1974), Brueckner & Fansler (1983) and Brueckner (1987), assumed that all residents (or consumers or employment) earned a common income at the CBD and had identical tastes over housing (or residential lot size) and a composite non-housing good. In urban equilibrium conditions, all residents reached the same utility level for the utility maximizing behavior with respect to a commuting cost from residence to the CBD and all producers maximize profit per unit of land associated with housing, in line with the indifference curve in neoclassical economics. The simple monocentric models were based on trade-offs between desire for housing space (or consumption) and perception of commuting cost. Higher income residents were likely to live farther from the CBD, because their increased utility from greater housing costs is larger than their decreased utility from increased commuting (or transport) costs.

After a William Wheaton's (1974) seminal work of the comparative static analysis using the traditional models, the fundamental parameters underlying spatial growth of cities were generalized by a function of population (or population density), household income, agricultural land rent, and commuting costs. McMillen (2006) indicated that the net effect of time costs of commuting and income on city size was ambiguous, because an increase in income enabled urban residents who prefer to live farther from the CBD to do so, as well as to increase their opportunity cost of commuting by selecting residential locations closer to the CBD, leading to a smaller city size, not a larger one. More recently, Spivey (2008) using the 2000 census data in the US urbanized areas[2] showed that the spatial size of the city grew as population or income level

---

[2] This dissertation extends the comparative statics predictions of the basic model tested by Brueckner and

increased, and as agricultural land rent or commuting costs[3] decreased, predicting that market forces drive urban spatial structure (or size), not uncontrolled urban sprawl, and that more densely populated urban areas had one or more employment centers.

More detailed functions to identify empirical regularities of post-war urban spatial structure were developed and estimated with respect to spatial patterns of population or employment. One approach was to examine “population decentralization” using an urban population density function, initiated by Colin Clark’s population densities (1951). Clark’s study and a number of extensive works defined population density as the number of people in the household divided by the land area, including all land uses, or residential land area, based on distance from the CBD. Their empirical results showed decentralization in U.S. cities which population density declined with distance from the existing central city along with increasing income and decreasing transport costs (Edmonston, 1975; Mills & Tan, 1980; McDonald, 1989; McDonald & McMillen, 2007, see Table 7.1.), referred to as a “negative exponential population function” (Papageorgiou & Pines, 1989). The concluding remarks of McDonald (1989)

Fanlser (1983) using the 1970 census data in the US urbanized areas,

[3] Even if the time cost of commuting is one of the crucial forces in shaping urban expansion, the statistical coefficients are consistent with the monocentric theory, accounting for an increase in urban spatial size and an decrease in commuting costs), but statistically insignificant because of negligence of geographical scale effect from smaller to larger urbanized areas (see Table 3, Spivey, 2008). In larger urbanized areas using the 2000 census data, increased demand for housing outweighs increased aversion to time cost of commuting, such as traveler measures, as income increase (see Tables A3 and A4 in Appendix 1, Spivey, 2008).

and Mills and Tan (1980) stated that an increase in population, particularly in larger urban population, was likely to correspond with a greater decentralization of employment as well as of population with a flatter gradient, meaning that population growth tends to be greater at the urban fringe. The other approach was to identify "employment decentralization' using employment density. With some empirical criticisms of population density gradients such as inaccurate measurement and lack of land use data, Mieszkowski and Mills (1993), using an employment density function, concluded that the density gradient was larger for employment than for households, even if the gradient dropped faster. Their empirical approach to interpreting spatial patterns of economic activity (i.e., manufacturing or services) by industry gave an important role in identifying decentralization of urban expansion, as depicted by McDonald (1987)'s definition of employment centers.

2.3.1.3. Theoretical Limitations and Extensions

The monocentric city models were generally considered as unrealistic. First, the basic models failed to predict that all jobs occur in the CBD in a location decision-making. That is, the models failed to capture the recent spatial evolution of U.S great cities, showing multiple subcenters or dispersed development patterns outside the central city, as identified by Mills's (2000) study that only 10 percent of employment in some metropolitan areas in the 1990s was located in the central city. Second, the assumptions that all urban consumers earn the same income and have same preferences were unrealistic, because each person had a different choice to where to live in or to how to commute. Third, housing needed to be viewed as a vector of its surrounding amenities and attributes, not a single composite good measured by floor space. Lastly, the monocentric models theoretically and empirically failed to capture the causes and consequences of environmental impact (or externalities) from spatial expansion of the city, assuming that such externalities as congestion, air pollution, noise, crime, and agglomerative effects disappears with distance from the city center. Subsequently, such externalities were likely to cause commercial and residential areas to fall farther away from the central city than the optimal boundary in the monocentric city model, leading to a larger, more decentralized urban area than before (McDonald & McMillen, 2007).

In spite of the limitations raised above, more extensive work with the monocentric models advocated that the models could still hold substantial predictive power vis-a-vis city spatial growth, as shown by Spivey's study (2008) that the Mills-Muth comparative statics predictions of urban growth in modern US cities remained valid. Those works included income heterogeneity (Hartwick et al., 1976; Wheaton, 1976), job decentralization (White, 1976; Thurston & Yezer, 1994; Spivey, 2008); multiple housing attributes (Büttler, 1981; Brueckner, 1983), public expenditures (Schuler, 1974; Yang & Fujita, 1983; Brueckner, 1997), and heterogeneous tastes (Anas, 1990; Beckmann & Papageorgiou, 1989). The monocentric city model and its extensive analytical predictions contributed to our understanding of the spatial expansion of the city over time, that is, the spatial variation in commuting costs, income, population, employment, agricultural land rents, and a home's price. The comparative statics predictions in the simplicity of urban spatial growth shed light on the dramatic changes in urban structure from the CBD to the polycentric or sprawling development occurring farther from the CBD.

2.3.1.4 Summary Remarks

The developing works of Muth (1969), Mills (1967, 1972), and other scholars established a "monocentric city model," emphasizing the importance of the CBD with respect to the degree of decentralization of population or employment. Some empirical evidence, following a pioneering work on population decentralization by Clark (1951), indicated that population density and transportation costs decline farther from the existing central city while incomes increase with greater distances from the CBD (Edmonston, 1975; Mills & Tan, 1980; McDonald, 1989; McDonald & McMillen, 2007).

Mieszkowski and Mills (1993), using an employment density function, concluded that the density gradient is larger for employment than for households, even if the gradient falls faster. An extensive body of work, including the comparative static analyses of Wheaton (1974), Brueckner and Fansler (1983), McMillen (2006), and Spivey (2008), concluded that the fundamental parameters underlying spatial growth of cities may be generalized by a function of population density, household income, agricultural land rent, and commuting costs. The comparative statics predictions tested by Brueckner and Fanlser (1983) and Spivey (2008), using the 1970 and 2000 census data in the US urbanized areas, pointed out that the spatial size of the city grew as population or income level increased and as agricultural land rent or commuting costs decreased. These relationships indicated that market forces drive and control urban spatial structure, rather than the notion of uncontrolled urban sprawl.

2.3.2 The Polycentric Model

2.3.2.1 Theoretical Background

The polycentric model was extended from the above-mentioned monocentric model. Theoretical and empirical background as to what determines subcenters formation was derived from theoretical limitations of the monocentric city models. It was characterized by the importance of suburban employment centers along with large specialized concentrations of office and retail space at the urban fringe, as well as nodes of major immediate accessible freeways (McDonald, 1987; Lockwood & Leinberger, 1988; Garreau, 1991; McMillen, 2001).

2.3.2.2 Theoretical Identifications

Prior numerous studies have tried to identify subcenters and their identity in large metropolitan areas[4]. A standard theoretical model by Fujita and Ogawa (1982) provided simple hypotheses of how changes in the population or changes in the commuting costs affect the subcenters formation, depending on spatial proximity. Their theoretical predictions showed that the equilibrium configuration of a polycentric area was likely to rise with population and the per-unit cost of commuting. McDonald (1987) seminally identified an employment center as a zone with a higher level of peak in gross employment density (measured by net employment density times the fraction of land devoted to employment use) than that of the employment density in the surrounding area, using 1970 Chicago area data. Giuliano and Small (1991) defined an employment center

[4] Prior studies have defined subcenters in various points of view: centers as defined by a regional planning agency (Greene, 1980; Griffith, 1981; Heikkila et al., 1989); subcenters as local municipalities (Erickson, 1986); historical growth nodes (Baerwald, 1982), and so on.

as a contiguous set of zones (or tracts) that each has both a density cutoff of at least 10 employees per acre and a minimum total 10,000 employees, using 1980 Census journey-to-work data for the Los Angeles region, regarding the peak of the center as the highest-density zone. Giuliano and Small identified employment centers as five clusters on a basis of agglomeration economies of industrial sectors: manufacturing-specialized; mixed industrial; mixed service; specialized entertainment; and service-oriented. Garreau (1991) also identified edge cities[5] with newer concentration of office-based employment associated with corporate headquarters, services, and FIRE (finance, insurance, and real estate) using 36 urban areas since the 1970s. He indicated that the New York area and the Los Angeles area showed a similar spatial pattern of urban areas with many edge cities, some traditional downtowns, and emerging additional edge cities, whereas the Chicago area had no emerging edge cites, some edge cities, and one traditional downtown. Reviewing the nature and role of subcenters in U.S. cities as the polycentric cities, Anas, Arnott, and Small (1998) tentatively generalized evidence on subcenters in large metropolitan areas[6], such as Los Angeles, Chicago, and San Francisco, into seven

[5] Garreau (1991) defined "edge cities" as places with at least 5 million square feet of office space, 600,000 square feet of retail space, more workers than residents, residents' perception as one place, and nothing like a recent city thirty years ago (pp.6-7).

[6] Some studies were referred to as the evidence on subcenters in U.S. large metropolitan areas. Los Angeles area emerges 32 subcenters and smaller outlying subcenters in 1980 (Giuliano and Small, 1991), as identified by Garreau's edge cities in Los Angeles (1991); Chicago area emerges 15 subcenters outside the city limits of Chicago for 1980 and 1990 (McMillen and McDonald, 1998); San Francisco area with 22 subcenters for 1990 (Cervero and Wu, 1998).

features (pp.1439-1444): 1) subcenters are prominent in both new and old cities; 2) the number of subcenters and their boundaries are quite sensitive to definition; 3) subcenters are sometimes arrayed in corridors; 4) employment centers help explain surrounding employment and population; 5) subcenters have not eliminated the importance of the main center; 6) most jobs are outside centers; and 7) commuting is not well explained by standard urban models, either monocentric or polycentric.

More extensive works by Craig and Ng (2001), McMillen (2001), Anderson and Bogart (2001), and McMillen and Smith (2003) found out the substantial regularities in multiple employment centers with highly specialized employment density across large metropolitan areas. They proposed that the size of the local peak for an employment center was higher in areas with commonly high levels of density. The empirical evidence presented by McMillen and Smith (2003), using 62 large US metropolitan areas in 1990, affirmed Fujita and Ogawa's (1982) theoretical model for subcenters formation. Their empirical results indicated that the two explanatory variables explained nearly 80% of the variation in the number of identified subcenters for Poisson regressions, which meant that higher levels of population and higher commuting costs (measured by traffic congestion levels) were more likely to increase the expected number of subcenters, along with some control variables such as median income, central city age, and median house age. The

Fujita-Ogawa theory and the McMillen-Smith empirical model implied that an urban area with a higher level of population and traffic congestion tended to form more subcenters.

More recently, McDonald and McMillen (2007) depicted that a polycentric urban area was an urban area with multiple employment centers, rather than the single economic center of monocentric city (p.144). An extensive research investigated by Marlay and Gardner (2010), using 50 most populous US metropolitan areas from census tracts-based Census 2000 data, identified the idea of Garreau (1991)'s edge cities that in large or small metropolitan areas employment-clustered sub-areas were apparently increasing, as of 2000, rather than only the CBDs were the dominant economic center across metropolitan areas.

2.3.3 The Sprawling Model

2.3.3.1 Multidimensional Definitions

Metropolitan sprawl is difficult to define as a single concept because of the nature of its formation; however, there are common underlying terms to define sprawling patterns. The dispersed sprawling development pattern in the U.S. metropolitan areas occurs at multiple dimensions of sprawl associated with space (or location), time, population, firms, the natural environment, and other internal or external compositions.

The urban studies literature utilizes such terms as: unplanned or chaotic (Fishman, 1987, 1990), edgeless (Lang, 2003), low-density, automobile-dependent, and isolated (i.e. strip, leapfrog, discontinuous) development far from the central areas and the polycentric cities (Real Estate Research Corporation, 1974; Downs, 1994; Nelson & Duncan, 1995; Ewing, 1997; U.S. Department of Housing and Urban Development, 1999; Burchell et al., 1998, 2002; Fulton, Pendall, Nguyen, & Harrison, 2001; Lee & Leigh, 2005). Reviewing as to how metropolitan areas have grown in the United States for the past 50 years, Downs (1998) and Johnson (2001) conceptualized sprawling development as multidimensional attributes – unlimited spatial expansion, low-density, automobile-dependent, segregated land uses, loss of open space, and fragmented governance system.

Robert Fishman (1987, 1990) viewed sprawling suburban form as a chaotic development pattern based on an individual's daily use of space (i.e. "household networks"), independent of the standards of the old metropolis associated with its geographical location from the center. Fishman described such structure as a "Technoburb" (1987, p. 190) with no clear boundaries and influenced by traffic access, population density, high-tech telecommunications, and income. Reviewing as to how metropolitan areas have grown in the United States for the past 50 years, Anthony Downs (1998) specified the form of such sprawling development with ten specific characteristics[7]: unlimited outward extension of new development, low-density residential and commercial settlements in new-growth areas, leapfrog development jumping out beyond established settlements, fragmented powers over land use among many local jurisdictions, automobile-dependent transportation system, no centralizing land use controls, strip commercial development, inter-regional fiscal disparities, and segregated types of land use zones, dependency on trickle down to provide housing to low-income households. Robert Lang (2003) identified dispersed development pattern as "edgeless cities" with a subset of non-CBD office space, non-cluster, non-edge city, and no well-defined boundary (p. 40), arguing that that an edgeless city was an urban geographic concept, but an elusive and hard-to-define one.

2.3.3.2 Some Evidence on Multidimensional Nature of Sprawling Patterns

Several important studies have tried to identify multidimensional characteristics of metropolitan structures. Galster et al. (2001) attempted to represent operational conceptualization of multidimensional nature of sprawl using 1990 census block housing data in 13 urbanized areas: density, continuity, concentration, clustering, centrality,

[7] Similarly, Johnson (2001) defined sprawl as a series of attributes: low-density, separation of land uses, leapfrog, strip retail, automobile-dependent, development of periphery area, employment decentralization, loss of rural and open space area, and fragmented governance system.

nuclearity, mixed uses, and proximity. They ranked an aggregated value of all six dimensions of sprawl (except continuity and mixed uses) to see overall housing sprawl scores for each area, based on equal weight of each dimension. Their results showed that New York sprawled the least and Atlanta sprawled the most, and that older urbanized areas such as New York (rank 1), Chicago (rank 3), and Boston (rank 4) were less likely to sprawl, while newer growing areas such as Denver (rank 10), Miami (rank 12), and Atlanta (rank 13) more likely to sprawl. Furthermore, two extensive works by Cutsinger et al. (2005) and Wolman et al. (2005) attempted to expand the operational conceptualization of the multidimensional nature of sprawl using housing, employment and land-use 1990 data in the U.S. 50 extended urban areas (EUAs) with consideration to measures of density, continuity, concentration, centrality, proximity, mixed uses, and nuclearity. They pointed out that in terms of multidimensional nature of land use patterns large populous EUAs had employment more concentrated and more housing centralized in the core, while older EUAs had housing and employment highly concentrated in the core. However, their combined metropolitan indices neglected to consider interactions with other complex metropolitan conditions, such as traffic behaviors, externalities, and initial regional characteristics.

The National Resources Inventory (NRI) density index from the U.S. Department of Agriculture (2001) were developed to see how dense the 50 most populous metropolitan areas in the United States were and how their density has changed from 1982 to 1997. To reduce the scale effects on large rural land to an urbanized area, this dissertation treated metropolitan density as two dimensions: population density per square acre in 1997 and percentage change in the population density from 1982 to 1997. This dissertation ranked each metropolitan area with a combined score of two density indexes defined above, ranging from 100 (most dense) to 2 (least dense). The results showed that most of the top 50 U.S. metropolitan areas were more likely to lose population density during the two decades. Regionally, almost all of the West (i.e. Los Angeles, San Francisco and Phoenix) tended to show much higher density scores (indicating positive percentage change in population density), while many of the South, such as Nashville, Richmond, Louisville, Memphis and Atlanta, were likely to have much lower density ones (meaning negative percentage change).[8]

Reviewing past efforts to define and measure sprawl, Lopez and Hynes (2003) developed a useful sprawl scale of the U.S. 330 metropolitan areas using the 1990 and

[8] Similarly, Lang's (2003) comparative study of sprawl and density provided evidence that edgeless cities were likely to have been grown at different development patterns: some (in case of the East) sprawled high, others (in case of the West) sprawled less, and some were balanced (i.e. medium or high sprawl – high or medium density) in Los Angeles, San Francisco, Denver, and Washington, D.C. (pp. 110-114).

2000 census data (using census tracts) and GIS tools. Their results pointed out that sprawl had more increased for the 1990-2000 period in many small (fewer than 250,000) and medium-sized (250,000 to 1,000,000 population) metropolitan areas, while larger metropolitan areas (greater than 1,000,000) appeared to be denser. Lopez and Hynes demonstrated that there were geographical variations of sprawl between regions: the great southern belt (i.e. Jacksonville, Charlotte and Atlanta), the Midwest, the North East regions, and some specific regions (i.e. Barnstable, MA) sprawled high, while the Pacific Coast (i.e. Los Angeles, San Francisco and Seattle), the Southern western parts of the country, and some particular regions (i.e. New York, Miami, Chicago and Boston) sprawled less (or tended to be denser). Furthermore, such variations among geography-based metropolitan areas in the level of sprawl could warrant further study as to how the regional effects interacted with other related factors, such as historical factors, geographic/climate features (i.e. coastal and temperature), socio-economic trends, land use policies, and other indirect factors.

Tsai (2005) quantitatively characterized metropolitan forms to distinguish compactness from sprawl. This dissertation provided a combination of four dimensions of metropolitan forms (metropolitan size, density, the degree of equal distribution, and the extent of clustering in sub-areas) and three degrees to distinguish compactness from sprawl (monocentric, polycentric, and sprawling), using the 1995 Census Transport Planning Package (CTPP) in the traffic analysis zones (TAZs) in the U.S. 219 metropolitan areas with less than 3 million populations. The empirical results indicated that employment was more concentrated (or less evenly distributed) than population across metropolitan areas, and that more than half of the metropolitan areas tended to show more compact development, even though a third of the metropolitan areas were likely to show a more sprawling pattern. The results also pointed out that large metropolitan areas were closely clustered among highly employed sub-areas. However, this dissertation indicated that the exact differences between compact and sprawling development patterns in the real world were hard to capture even with the Moran coefficients, because the levels of metropolitan areas partitioned, such as cities, census tracts, census blocks, were spatially different and inconsistent over time and the inclusion of undeveloped areas, such as rivers, mountains, or natural landscapes, could bring out measurement bias not to reflect only land use activities on the developed land.

Recently, Torrens (2008) portrayed the multidimensional nature of sprawl using a series of 42 measures in the fast-growing metropolitan area of Austin, Texas between 1990 and 2000, including aspects such as urban land development, population density, residential ownership, land use mix, decentralization, and accessibility index. The empirical results pointed out that Austin tended to have developed under successive waves of urbanization and urban growth over 10 years, which explained that sprawling and compact development patterns co-exist in the same geography and co-evolve in different urban systems due to ubiquitous accessibility region-wide. Torrens implied that Austin appeared to have both the central city with more jobs-oriented polycentric patterns and the suburbs with more fragmented and homogeneous land-use activities, like Los Angeles-style development patterns (Gordon & Richardon, 1997), rather than the sprawling development patterns from the central area to the periphery mainly noted in urban studies literature.

As reviewed by previous empirical works (Galster et al., 2001; Cutsinger et al., 2005; Wolman et al., 2005; U.S. Department of Agriculture, 2001; Lang, 2003; Lopez & Hynes, 2003; Tsai, 2005; Torrens, 2008), the literature on how to measure the shapes of metropolitan structure encompasses a variety of conceptual and operational dimensions such as density, concentration, clustering, centrality, proximity, mixed land uses, and so on. Table 2.2 represents the analytical terminology of multidimensional measurements of metropolitan development patterns to date in literature, which are operationalized as spatial distribution of population or employment or land uses. Feasible measurements of metropolitan structure at the micro or macro level will play important roles in diagnosing and managing sprawling metropolitan areas.

Table 2-2 Multidimensional Characteristics of Metropolitan Structure

Scholars	Definitions	Characteristics
Density		
Gordon & Richardson (1997), Burchell et al. (1998), Malpezzi & Guo (2001), Hess et al. (2001), Galster et al. (2001), Cutsinger et al. (2005), Tsai (2005). Torrens (2008)	Defined as overall activity intensity of population or employment in land area in a metropolitan area, referring to density per capita in a certain sub-area according to land cover and land use; Operationalized as total number of population or employment in land area in a metropolitan area	A high value may mean compactness; A low value can characterize sprawl
Unequal distribution (or inequality), Dissimilarity, Concentration		
Lorenz (1905), Galster et al. (2001), Hess et al. (2001), Cutsinger et al. (2005), Wolman et al. (2005), Tsai (2005), Torrens (2008)	Defined as the degree to which human activities are equally or unequally distributed (concentrated) in a few sub-areas in a metropolitan area; Operationalized as the Gini coefficient measured by unequal distribution of population or employment by spatial sub-areas among metropolitan areas, borrowing from inequality of income distribution	A higher coefficient (close to 1) means that population or employment is unevenly concentrated in some sub-areas; A lower coefficient (close to 0) means that population or employment is evenly distributed in a metropolitan area
Clustering versus Scattering (Spread)		
Galster et al. (2001), Cutsinger et al. (2005), Tsai (2005), Torrens (2008)	Defined as the degree to which high-density sub-areas (or development) are clustered or randomly distributed; Operationalized as the global Moran coefficient and adjusted Geary coefficients using an inverse-distance-based weighting between sub-areas	A high positive coefficient means that high-density sub-areas are closely clustered; A medium value for polycentric; A coefficient close to 0 means random scattering; A -1 value indicates a chessboard pattern of development (decentralized sprawling)
Centrality versus Decentrality		
Galster et al. (2001), Malpezzi & Guo (2001), Hess et al. (2001)	Defined as the degree to which a land use (i.e. residential or nonresidential) is located close to the CBD, weighed by the number of population or jobs in each sub-area in a metropolitan area; Operationalized as the ratio of the average distance to the CBD of centroids of all the sub-areas relative to the average distance to the CBD of employment in each sub-area in a metropolitan area	A high value means that a land use of population or employment is located near the CBD; A low value (close to 0) indicates that a land use of population or employment is located farther from the center leading to more sprawl
Continuity		
Galster et al. (2001), Malpezzi & Guo (2001)	Defined as the degree to which developable land has been developed in an unbroken fashion throughout the metropolitan area; Operationalized as the share of all the sub-areas in the metropolitan area that are developed (i.e., more than 50% or more land)	A high value (R^2) means a high level of continuity; A low value (R^2) indicates the extent of leapfrog (discontinuous) development pattern

Mixed Land Uses		
Galster et al. (2001), Rajamani et al. (2003), Cutsinger et al. (2005), Torrens (2008)	Defined as the degree to which substantial number of different land uses (i.e., residents or jobs) exist within the same sub-area in a metropolitan area and this pattern is common across the metropolitan area; Operationalized as the average density of a certain land use in another land use in a certain sub-area using Massey and Denton's exposure index (1988)	A high level means an equal proportion of population and employment in a metropolitan area, leading to increase in land use mix diversity affecting a greater preference for walking, biking, and transit modes to travel; A low level means patterns of an exclusive land use which represents more sprawl-like development pattern, leading to separation of homes and workplaces, more trip length and times and its resulting consequence of traffic congestion
Accessibility		
Ewing (1997), Ewing et al. (2002), Rajamani et al. (2003), Torrens (2008)	Defined as the degree to which households or jobs are accessible to a range of the destinations according to travel modes-related variables; Operationalized as straight-line and road network distance to a range of urban opportunities, such as the CBD, sub-centers, and major educational opportunities (universities, libraries, museums); Measured by average trip length, average commute time, vehicle miles travelled per person, percentage of households to commute by private automobiles or public transit	A greater value of accessibility to the CBD is less sprawled; A lower value of accessibility to the CBD is more sprawl-like
Proximity		
Galster et al (2001), Cutsinger et al (2005), Wolman et al. (2005)	Defined as the degree to which residents, jobs, or residents/jobs pairs are close to each other, relative to the distribution of all land composing of the study area; Operationalized as the ratio of the average distance among centroids of square-mile cells in a certain area to the weighted average distance among jobs (or residents or jobs/residents) across all cells in the same area	A high level of proximity is less sprawled; A low level of proximity is more sprawl-like
Nuclearity		
Griffith (1981), Gordon, Kumar, & Richardson (1989), Small & Song (1992), Cervero & Wu (1998), Malpezzi & Guo (2001), Galster et al. (2001), McMillen (2001)	Defined as the degree to which jobs within a metropolitan area disproportionately located in the nuclei, either at the CBD or sub-centers outside the CBD; Operationalized as the ratio of jobs in the CBD to jobs in all other nuclei; CBD measured by the highest-density nucleus and its adjacent nodes within one standard deviation of the highest-density nucleus	A high value means that development is intensely located close to the CBD or maximized around the CBD; A metropolitan area with mononuclear or polynuclear pattern of development may contain an agglomeration of activities and shorter journey-to-work

2.3.3.3 Consequences of Sprawling Development Patterns

A clearer understanding of sprawl may be possible after a review of the debates on its impact vis-à-vis benefits and costs, as seen in Table 2-3.[9] One approach considers it as a desirable urban form that provides: safe and cheap places (Bank of America, 1996); higher consumer satisfaction and benefits based on free-market merits of continued suburbanization (Gordon & Richardson, 1997); housing affordability and equal housing opportunities (Kahn, 2001); significant improvements in quality of living (Glaeser & Kahn, 2003); and lower municipal spending per capita (Cox & Utt, 2004).

An alternative approach sees negative outcomes involving urban and environmental problems (Berry et al., 1974; Real Estate Research Corporation, 1974; Kunstler, 1993; Ewing, 1997; Burchell et al., 1998, 2002; Sierra Club, 1998; Fulton, Pendall, Nguyen, & Harrison, 2001; Ewing, Pendall, & Chen, 2003; American Farmland Trust, 2007). Such problems are identified as follows: racial/social segregation, income inequality, regional disparities for concentrated poverty, land consumption (i.e., loss of open space, loss of agricultural farmland), poor health, increased crime, more public/private expenditure, energy cost, travel and transportation impacts (i.e., traffic

[9] The debate over sprawl went on (Burchell et al., 1998, see table 6). The debates included Ewing (1997) versus Gordon and Richardson (1997), and the Urban Lawyer versus the Housing Policy Debate. The former was anti-sprawl, and the latter was pro-sprawl.

congestion and travel time), environmental pollution (i.e., air, water, and land), and other intangible costs.

A third approach attributes sprawling development with both negative and positive impacts. Lang (2003) explained that there is a "trade-off" of sprawl with impacts that vary by issue. The expansion of an edgeless city is likely to produce more negative impacts on the environment, open space, and transportation, while it also tends to increase market preferences such as residential ownership, commuting cost and job positions. Lang's explanation was consistent with two prior surveys, "Fannie Mae" survey of likeness of Americans about sprawl (Lang & Hornburg, 1997) and "visual preferencing" survey (Nelessen, 1994). The two survey results pointed out that people prefer to live in their own housing and its suburban location, but that people think that sprawl looks ugly and yields increasingly congested suburbs.

Table 2-3 Unsettled Debates over Sprawling Development Patterns

	Sprawling	Key Figures
Positive	Less congestion & pollution; More preferences & choices; Safe places; American dream; Lower municipal spending	Fishman (1987), Bank of America (1996), Gordon & Richardson (1997), Glaeser & Kahn (2003), Hayden (2004), Cox & Utt (2004), Kahn (2006)
Negative	Traffic congestion; Higher fuel consumption; Increased pollution; Loss of open space	Real Estate Research Corporation (1974), Newman & Kenworthy (1989, 1999), Elkin McLaren, & Hillman (1991), Kunstler (1993), Burchell et al. (1998), Fulton et al. (2001), Ewing, Pendall, & Chen (2003)
Mixed	People prefer to live in their own housing and its suburban location, but people think that sprawl yields increasingly congested suburbs.	Nelessen (1994), Lang & Hornburg (1997), Lang (2003)

2.3.4 Determinants of Metropolitan Spatial Structure

2.3.4.1 Geographic Distribution of Population or Employment

An emerging metropolitan structure may be determined by the distributional configuration of location decisions which households can settle to areas outside of and farther from the central areas. Such shapes cannot be fully explained by urban decentralization measured by population density in the assumption of monocentric city models. Some research argues that there is little evidence of strong convergence (or compactness) or divergence (or sprawl) regarding spatial distribution in population levels on urban growth at the metropolitan area level. Glaeser, Scheinkman, and Shleifer (1995) pointed out that while some cities with higher population densities are likely to converge, almost all of the larger U.S. cities with high population levels between 1960 and 1990 show less convergence in metropolitan areas. Beeson, DeJong, and Troesken (2001) also showed that there has been little evidence for either concentration or deconcentration in U.S. counties between 1840 and 1990, even if there has been population deconcentration only in all but the most-densely-populated counties in 1840.

Regarding the link between population and employment, job-housing imbalance and spatial mismatch tends to encourage the distributional consequences of suburban deconcentration. After seminal works of spatial mismatch hypothesis by Kain (1964, 1968), the spatial dispersal of urban and suburban employment due to new development and fragmented land use controls such as exclusionary zoning and other policies took place far from the central zones. It brought out such mismatch that lower-income and minority households remain in blighted central areas and in less affluent (i.e. inner-ring and outer-ring) suburbs called “concentrated poverty” (Downs, 1998), leading to the effects of housing market segregation/discrimination, unequal educational opportunities, and environmental degradation (Cervero, 1989, 1996; Mieszkowski & Mills, 1993; Glaeser & Kahn, 2001; Orfield, 1997, 2002; Kain, 2004).

2.3.4.2 Agglomeration Economies and Human Capital

A change in metropolitan structure may be attributed to scale-dependent processes from the agglomerative forces which cause job clusters in a certain region. After a seminal work of Alfred Marshall (1890) on localization economies with geographical proximity and a extensive work of Jane Jacobs (1969) on technological innovations, many studies have identified spatial variations of knowledge spillovers not only from endogenous technical progress (Romer, 1986; Henderson, 1986; Lucas, 1988; Porter, 1990; Krugman, 1991; Jaffe, Trajtenberg, & Henderson, 1993; Fujita & Thisse, 1996;

Ellison & Glaeser, 1997; Hanson, 2000) but also from human capital associated with education (Chinitz, 1961; Rauch, 1993; Glaeser, 1994; Glaeser et al., 1995; Henderson, Kurono, & Turner, 1995; Rappaport, 1998; Simon & Nardinelli, 1996, 1998; Glaeser & Kahn, 2001; Fujita & Thisse, 2002), and spatial distribution of employment in industrial sectors as to which it is concentrated or diversified in a region (Cooke, 1983; Noyelle & Stanback,1983; Carlino, 1985; Henderson, 1986, 2003; Glaeser, Kallal, Scheinkman, & Shleifer, 1992; Glaeser and Kahn (2001); Felsenstein, 2002; Stanback, 2002; Burchfield et al., 2005) for urban and regional growth.

Glaeser et al. (1992) examined growth in employment from 1956 to 1987 in the six largest industries at the two-digit SIC code in the U.S. 170 largest urban areas. They measured four relevant variables of dynamic agglomeration economies: employment in the industry in the urban area in 1956 was used as a proxy for the size of the local industry; the average size of establishments in the local industry relative to the nation; the urban area's other top five industries' share of total employment in the urban area was used as a proxy for the degree of diversity in the local economy; and the location quotient for the industry in the urban area was used as a proxy for a combination of a dynamic localization effect and a dynamic urbanization effect.

Glaeser and Kahn (2001) investigated the decentralization of employment using zip code data within the U.S. 335 metropolitan areas on employment across 3-digit SIC industries between 1969 and 1997. They gave empirical evidence that regions with more suburbanized populations in 1969 have faster decentralization of employment by 1997. Their findings pointed out that metropolitan areas with more specialized employment in manufacturing industries appeared to sprawl more, while those specializing in services and idea-intensive industries tended to be more dense and centralized (Cooke, 1983; Carlino, 1985).

Noyelle and Stanback (1983) used a functional classification system to define the types of goods and services produced in urban areas, which are grouped into eight basic functional sectors[10]. Subsequently, Stanback (2002) updated a classification system of economic activity based on high-tech element: high-tech manufacturing (drug, computer, communication, electronic, aircraft, space, surgery instruments, detection); high-tech services (telephone communication, computer programming, data, motion picture, engineering, and R&D). In addition, Henderson (1986, 2003) examined the existence of localization and urbanization economies in sixteen manufacturing industries. The

[10] Their eight basic industrial sectors are follows: manufacturing; agriculture, extractive, construction (not necessary); distribution services (transportation, communication, utilities, wholesale); corporate activities (finance, insurance, and real estate, headquarters), nonprofit services (health, education); retailing; consumer services (hotels, auto repair, motion pictures, recreation, private households); and government enterprises.

empirical results indicated that localization economies occur in several manufacturing industries, but that urbanization economies in manufacturing industries are absent.

Aligned with agglomeration advantages, Felsenstein (2002) analyzed the relationship between high technology employment concentrations and urban sprawl (measured by the magnitude of land conversion in the outer suburbs) using two counter-factual simulated situations in the city of Chicago, its inner suburbs, and its outer suburbs. Felsenstein pointed out that an increase in high technology industries in the outer suburbs of the Chicago metropolitan area was associated with higher costs of sprawl, such as congestion, pollution, loss of open space, and public health risks.

Using the sprawl index for 1976-1992 undeveloped land surrounding residential development in 275 metropolitan areas, Burchfield et al (2005) pointed out that sprawl tends to increase in metropolitan areas with more decentralized employment sectors such as restaurants and bars, but that it tends to decrease in regions with more centralized employment sectors such as business services.

2.3.4.3 Roles in Governments

The impact of political forces on emerging metropolitan structures is likely to be significant; different forms of governmental structure and the role of public policies may affect the spatial distribution of population or employment and the consequences of urban sprawl.

2.3.4.3.1 The role of Governmental Structure

Different forms of governmental structure from the polycentric view to the centralist view (including a regionalist view) may impact location decisions of households and firms. One major influence is the extent to which governmental structure as conducive to a vote-within-its-foot principle (Tiebout, 1956) works in shaping urban spatial patterns. Aligned with the Tiebout hypothesis, since late 1960s to the present, the polycentric view has focused on decentralized, fragmented systems of metropolitan government in accord with voters' preferences and locally service-related problems (Ostrom, Tiebout, & Warren, 1961; Oates, 1972; Bish & Ostrom, 1973; Ostrom, 1974; Parks & Oakerson, 1993; McGinnis, 1999; Thurmaier & Wood, 2002; Wood, 2006). From the 1990s to the present, the centralist or regionalist view started from beyond Tiebout's major assumption that there are no spillover issues between communities such as socio-economic disparities, traffic congestion, environmental pollution, and loss of green space, which are associated with adverse impacts of sprawling development patterns (Ward, 1987; Frisken, 1991; Rusk, 1993; Downs, 1994; Wallis, 1995; Dodge, 1996; Adams, 1997; Bollens, 1997; Orfield, 1997, 2002; Nelson & Foster, 1999; Stephens & Wikstrom, 2000; Dreier, Mollenkopf, & Swanstrom, 2001; Squires, 2002; Hamilton, Miller, & Paytas, 2004; Howell-Moroney, 2008).

Prior empirical studies on the role of governmental structure are not conclusive. Many advocates of the polycentric model have provided evidence that the fragmented local governments in a metropolitan area can contribute to economic performance (Boyne, 1992; McGinnis, 1999; Thurmaier & Wood, 2002; Stansel, 2005; Hammond & Tosun, 2009). In contrast, many regionalists have provided evidence for a more centralized metropolitan governmental structure having an impact on economic development (Nelson & Foster, 1999; Hamilton, Miller, & Paytas, 2004; Jeong & Feiock, 2006) and social equity (Rusk, 1993; Pierce, Johnson & Hall, 1993; Bollens, 1997; Orfield, 1997, 2002), while some pointed out that such structure may not have any impact on solutions to such spillover issues (Blair & Zhang, 1994; Carr & Feiock, 1999; Savitch & Vogel, 2004). Challenging works by Carruthers and Ulfarsson (2002) and Carruthers (2003) illuminated evidence that the fragmented governmental structure can contribute to the growth of outlying areas in the U.S. metropolitan areas leading to sprawl. Urban scholars have not yet devoted much attention to the impact of metropolitan governmental structure on environmental performance.

2.3.4.3.2 The Role of Public Policies

As explained in previous sections 2.2.2 and 2.2.3, those governmental measures at a local/regional and state/national scale have an effective role in shaping metropolitan spatial structure. Along with the housing policies and the transportation projects since World War II, the scope of public policies aiming to reverse excessive decentralization (viewed as sprawl), to reduce automobile use, and to revitalize the central cities, such as statewide growth management strategies, smart growth programs, new urbanism, new regionalism, and eco-city approaches, can contribute to economic development and environmental quality.

Studies on the role of public policies are still not conclusive. Recent works supporting growth management programs within metropolitan areas suggest that they offer a beneficial impact on economic growth (Nelson & Peterman, 2000), public finance (Carruthers & Ulfarsson, 2008), and urban sprawl (Nelson & Duncan, 1995; Carruthers, 2002). Yet again, evidence of the effect of government intervention on environmental concerns remains inconclusive. Johnson (2001) presented evidence of the mixed impacts of Portland's urban growth boundary in that there was increased density and growth containment, but rising housing prices. Brueckner (2001) also suggested that a remedy for attacking urban sprawl should consider not only the potential market failures such as the amenity value of open space, social costs of congestion, and infrastructure costs of new development, but also the pitfalls of growth management policy.

The debate over which public policies will be beneficial for compact or sprawling development patterns still lasts between planners and market-oriented advocates, as typified by the debate of Gordon & Richardson (1997) and Ewing (1997).

2.3.4.4 Other Confounding Forces

There are other confounding forces that can have an important role in shaping metropolitan spatial structure, such as income level, race, education, and regional amenities. The debate over whether income level matters to the evolution of urban structure is still being argued. Barro and Sala-i-Martin (1992) pointed out that there is more convergence (deconcentration) in per-capita income since 1840 across the U.S. states. Empirical results by Carruthers and Ulfarsson (2002), Carruthers (2003) and Faggian, Olfert, & Partridge (2011) supported the contention that rising income can affect the spatial distribution of population growth occurring at the urban fringe in U.S metropolitan areas. In agreement with the white flight hypothesis (Frey, 1979; Massey & Denton, 1993; Kruse, 2005), they suggested that the role of racial composition associated with income level can contribute to metropolitan spatial structure. However, some recent studies argued that there is little evidence of convergence in income levels across the U.S. cities and counties (Baumol, 1986; DeLong, 1988; Barro, 1991; Glaeser et al., 1995).

For education, a growing number of studies pointed out that education, defined as the role of human capital, had an important role in shaping metropolitan structure, because regions with more highly-skilled (or highly-educated) workers brought out greater economic growth (Chinitz, 1961; Glaeser et al., 1995; Beeson et al., 2001; Berliant & Wang, 2004) and paid more attention to green policies (Glaeser & Kahn, 2003; Portney, 2003; Kahn, 2006). Chinitz (1961) and many scholars provided evidence that the role of human capital (measured by years of schooling, high school graduation rate, or college graduates) is more likely to support the role of intellectual spillovers leading to driving urban growth. Glaeser et al. (1995) and Glaeser and Kahn (2001) empirically identified the level of human capital (measured by degree of intellectual intensity) as a key force of urban growth for population, employment, and income growth, leading to productive externalities of growth that can reinforce circular causation between agglomerative knowledge effects and growth.

Many research studies have emphasized that regional amenities, such as temperature and geographical location, can affect the shapes of a region. A lot of studies indicated that air quality improvements tended to be sensitive to temperature (Robson, 1977; Rao, Zalewsky, & Zurbenko, 1995; Marquez & Smith, 1999; Beeson et al., 2001; Stone, 2005, 2008; Rappaport, 2007; Clark, Millet & Marshall, 2011; Faggian et al., 2011) and physical locations of Census regions (U.S. Department of Agriculture, 2001; Glaeser & Kahn, 2001; Lang, 2003; Lopez & Hynes, 2003; Kahn, 2006; Lee & Gordon, 2007; Clark et al., 2011). Regions with more geographical advantages (i.e., mild temperature and development-friendly place) can influence more changes in air quality produced by spatial distribution of residential and economic activities than those with less geographical advantages. Across a geographical location and central cities matrix, Glaeser and Kahn (2001) pointed out that geographical location effects vary according to regions in U.S. They found that the central cities of the Northeast and the West are more likely to be anti-business, causing employment to go farther from the central areas, while the central cities of the South and the Midwest seem to pro-business, leading to high employment density in the central cities

2.3.4.5 Summary Remarks

The shapes of metropolitan structure in space and over time have been represented from monocentric to polycentric or sprawling patterns. Its formation tended to emerge or grow faster in the presence of strong agglomerative effects of firms, well-educated people, durable infrastructure and transport systems, productive public policies, and other underlying forces. As Anas et al. (1998) explained, such changes in emerging metropolitan structure can be influenced by "positive and negative externalities, all acting with different strengths, among different agents, at different distances" (p. 1459), which are interrelated with spatial distribution of population or employment spreading out from central cities to suburbs or exurbs.

2.4 Empirical Evidence for Air Quality

2.4.1. Multidimensional Nature of Environmental Quality

The urban studies literature documents that environmental quality will be included in a broader concept of the quality of life, which implies overall increase or decrease of both the welfare (or well-being) of people and that of the environment in which people live. In this sense, Berry et al. (1974) referred to environmental quality as "a product of the joint influences of human processes and dynamics of the biosphere" (p.14). In a sustainable view, Paehlke (2003) referred to environmental quality as "the capacity to continuously produce the necessities of a quality human existence within the bounds of a natural world of undiminished quality" (p.57).

From a policy perspective, the Council on Environmental Quality (CEQ) in the National Environmental Protection Act (NEPA) of 1969 established the environmental quality standards associated with the health and welfare effects of different pollution intensities, such as air quality, water quality, toxic and hazardous wastes, noise, pesticides and radiation reported by the Environmental Protection Agency (EPA). Kraft and Vig (2003) evaluated that impressive progress on environmental quality between 1970 and 2000 has been made in controlling conventional pollutants and in expanding green space. However, they indicated that substantial improvements in environmental quality will be more difficult, costly, and controversial because of the interaction of changes in short-term and long-term social, economic, technological, political, and ecological forces over time. In short, the conception of environmental quality involves more complicated and sometimes intractable interactions with the shapes of metropolitan structure.

2.4.2. Empirical Evidence for Air Quality

Some studies have been conducted at the metropolitan scale in America. The scope of the empirical evidence has mainly dealt with the interaction between metropolitan structure, land development, transportation, and air quality. A critical overview of the interactions is of importance to a better understanding of the significance of metropolitan spatial structure to air quality for further research. As stated in the introductory section¸ we review empirical evidence in terms of the two major arguments over the relationship between urban structure and air quality.

Some research has argued that the compact and large metropolitan areas can contribute to improved air quality. Through a comprehensive approach for the 76 urban regions between 1950 and 1970, Berry et al. (1974) analyzed that which of the different alternative urban forms and urban land use pattern most improved environmental quality in 76 largest metropolitan areas. They defined environmental quality as the level, intensity, and spatial distribution of environmental pollution (air, water, solid wastes, noise, pesticides and radiation), as well as urban form as a process of urban expansion in terms of population dispersion and economic concentration, and their relationships (seen as linkages and interactions). From a pollution-sensitive typology for the 76 urban regions, they clustered the seventy-six urban regions into the eleven groups based on similarities in environmental pollution through a Q-mode factor analysis. The findings pointed out that worse air quality tends to appear in regions with larger, dispersed, manufacturing-concentrated patterns (i.e. Indianapolis, Washington, DC), while the better environmental quality tends to appear in those with small, more affluent, non-manufacturing, and core-oriented patterns (i.e. Salt Lake City, Phoenix, and Tulsa). Berry et al. (1974) concluded that the pace of metropolitan suburbanization can lead to increasing city size, increasing dispersion, increasing automobile use, changing urban forms and land use patterns (i.e. less open space), and the resulting increase of environmental pollution. Identifying a significant role of urban form in contributing to the concentration of environmental pollution, Berry et al. (1974) suggested that changes in the direction of current dispersed development patterns will be required to reduce the current trend of environmental pollution.

More notably, Newman and Kenworthy (1989[11], 1999) examined the influence of urban form (measured by density) on automobile dependence and on air quality in 37 large cities in the world in 1990 using multiple regression analysis. They pointed out that the regions with low-density level (i.e. Houston, Phoenix, and Detroit) tend to have increased automobile use, but that more dense regions such as Chicago and New York tend to have more compact and more public transit use. Also, they emphasized that Portland, Oregon, with a more compact pattern is most effective for reduced automobile use (1000 Friends of Oregon, 1997). Newman and Kenworthy suggested that some

[11] An original analysis of Newman & Kenworthy in 1989 provided implications of auto dependence in 32 international cities (including five cities in the US) from 1960, 1970, and 1980. This analysis found that high density areas tend to have less automobile use, leading to shorter travel distance and decreased gasoline use.

fundamental policies to overcome car dependence towards a sustainable urban form will be required, such as a multi-nodal city model with high-density development patterns, mixed land-use zoning, and an extensive public transit system to connect to "urban village" sub-centers in the suburbs (Newton, 1997, 2000, Masnavi, 2000).

Recently, Moore (2001) conducted a comparative analysis between Atlanta, Georgia, and Portland, Oregon about the effects of development choices on air quality. The findings pointed out that Portland has had a greater reduction rate in carbon monoxide and ozone levels over the period 1988-1997 than Atlanta. Moore suggested that the land use strategies, such as compact and mixed-use development, and transportation policies in Portland have been effective for reductions in air pollution. Looking at the multidimensional nature of metropolitan sprawl and its impact, Ewing, Pendall and Chen (2002, 2003) created an overall metropolitan sprawl index (called Smart Growth America (SGA) Index) associated with residential density, land use mix, centeredness, and accessibility of the street network. Based on the multidimensional dimensions of sprawl and a composite measure using principal components analysis, they examined the relationship between metropolitan expansion and its impacts on travel and transportation outcomes (i.e. vehicle ownership, fatal accidents, commute mode and time, and maximum 8-hour average ozone level) for the 83 U.S. large metropolitan areas for

1990 and 2000 through multiple regression analysis. The findings pointed out that residents living in more sprawling regions tend to drive longer distances, own more cars, face a greater risk of fatal accidents, walk and use public transit less, and breathe more polluted air. They suggested compact development strategies to improve quality of life, such as urban infill, mixed use development, and smart growth management.

More recently, Stone (2008) explored the impact of urban spatial structure on air quality (measured by the number of annual ozone exceedances) in the 45 largest U.S. metropolitan areas between 1990 and 2002. Indicating the lack of Ewing et al. (2002, 2003) sprawl index analysis,[12] he conducted an integrated multiple regression analysis of the links of urban sprawl to air quality in large metropolitan areas, while controlling for population size, average annual precursor emissions (i.e. nitrogen oxides and volatile organic compounds), average annual temperature, and average ozone season temperature (May to September). The empirical results supported the hypothesis that urban form drives ozone formation, which accounts for urbanized regions with high levels of sprawl (i.e. density and connectivity) having significantly higher levels of mean annual ozone. He suggested the importance of region-scale land use planning strategies such as urban growth boundaries (Nelson, 1994; Song & Knaap, 2004) and form-based codes related to

[12] Stone (2008) indicated ignorance of important variables in Ewing et al. (2002, 2003) analysis, such as meteorological factors (i.e. temperature), ozone precursors, and the occurrence of high ozone days.

street network connectivity against traditional zoning ordinances in improving benefits of regional air quality. In order to address air quality outcomes (i.e. human exposures to criteria air pollutants), Schweitzer and Zhou (2010) examined the relationships between urban form,[13] criteria air pollutants (i.e. concentration in ozone & fine particulates ($PM_{2.5}$) and neighborhood-level population[14] exposures in monitors-installed in 80 U.S. metropolitan areas using two-scale (i.e. neighborhood and regional) multiple linear regression models for 2000. The two-level regression models pointed out that urban form, particularly in more compact regions, has an important role in lowering ozone concentrations at the regional level, while population exposures to both ozone and fine particulates, particularly in poor and minority residential areas, are higher in more compact regions than in more sprawled regions at the neighborhood level. The findings suggested that urban and regional planners should consider opposite directions between air quality concentrations and population exposures when putting infill and new compact development into practices. Clark et al. (2011) investigated the link between air quality (measured by long-term population-weighted ozone and particulate matter ($PM_{2.5}$) concentrations) and urban form (particularly measured by density and centrality) in 111

[13] To measure sprawl at the regional scale, Schweitzer and Zhou (2010) re-used the Smart Growth America (SGA) index scores developed by Ewing et al. (2002, 2003): residential density, street connectedness, regional centeredness, and land use mix.
[14] Neighborhood-level population was composed of two groups: children under 5 years old and people aged 65 and older.

U.S. urban areas. Through standardized coefficients with interquartile range (IQR) changes of dependent and independent variables using cross-sectional stepwise linear regression analysis, they found that population density positively impacted population-weighted $PM_{2.5}$ concentrations at the 99% significance level, while population centrality negatively impacted population-weighted ozone and $PM_{2.5}$ concentrations at the 99% significance level. Their findings pointed out that spatial distributions of population are statistically significant predictors of air quality, which shed light on the necessity for effective regional planning to improve air quality. However, this dissertation failed to consider important factors that may affect changes in air quality, such as spatial changes in the built-in environment (i.e., land uses over time), industrial concentration, and public policies, even if it contributed to statistical power in predicting long-term air quality with interquartile range (IQR) changes in urban form at the urban scale.

The other side of the debate posits that the dispersed and large metropolitan areas can enjoy reduced air pollution levels. Robson (1977) provided challenging evidence regarding whether increased dispersion of both residences and destinations (i.e. temperature, dispersion of population, population growth rate, and the fraction of the workforce) can affect the concentration of pollution (i.e. particulates and nitrogen dioxide) from transportation using statistical equations in the 44 larger SMSAs between 1920 and

1950. Robson concluded that 44 larger metros with increased dispersion of population and firms can lead to reduce air pollution concentration (i.e. particulates and nitrogen dioxide) because of more public transit use. Rodson suggested that more public transit use will have a role in lowering air pollution.

Most recently, Kahn and Schwartz (2008) found evidence that there is a positive relationship between sprawling development patterns and urban air pollution in major California cities. Using the zip-coded California random road-side emissions tests from 1997 to 2002 and log-linear OLS regressions, this dissertation investigated estimates of average vehicle emissions (measured by hydrocarbons, carbon monoxide, and nitrogen oxides) at monitoring stations by vehicle model year and by calendar year to measure overall technological emissions progress from 1982 to 2000, controlling for per capita income and population at a county level. The empirical results pointed out that, due to technological advances, a greater decline in the average vehicle's emissions (i.e. carbon monoxide) can offset an increase in population growth and per-capita income, leading to improvements in ambient air quality. They suggested that the technological progress for emissions control can play an important role in reductions in the costs of sprawl such as air pollution (Kahn, 2006; Glaeser, Kolko, & Saiz, 2001).

Meanwhile, Emison (2001) examined the relationships between sprawl (measured by change in population density or in urbanized area) and air quality improvements (measured by ozone exceedances) in 52 metropolitan areas that exceeded air quality standards for ozone over the time period 1982-1996 using OLS regression models. This dissertation particularly considered impacts of policies and environmental expenditures at the state level on air quality improvements in the metropolitan areas. The findings pointed out that the 52 metropolitan areas with ozone exceedances have tended to grow in a sprawling pattern, while population density decreases, urbanized area expands, and vehicle miles travelled per capita for automobile use increase. However, this dissertation showed that changes in population density and higher environmental protection expenditures had no impact on improvements in ozone air quality. Emison suggested a necessity for further examination to identify the sprawl-policy-air quality relationships.

Taken together, the empirical evidence for larger U.S. metropolitan areas shows mixed outcomes, positive or negative. Such evidence pointed out that either the compact region or the sprawling region may be desirable for environmental quality improvements and vice versa. The evidence is summarized in Table 2-4.

Table 2-4 Summary of Empirical Evidence for Air Quality

	Authors	Unit/Time/Model	Key Variables	Findings
Compact	Berry et al. (1974)	76 urban regions; 1950-1970; Q-mode factor analysis	Environmental pollution (air, water, solid wastes, noise, pesticides and radiation); Urban expansion in terms of population dispersion and economic concentration.	Larger, dispersed & manufacturing-concentrated areas have worse air quality; Small, non-manufacturing & core-oriented areas have better air quality.
	Newman & Kenworthy (1989, 1999)	37 large cities in the world (including 13 large cities in US); 1990; Multiple Regression analysis	Automobile dependence & air quality; Urban form (measured as density)	Regions with low-density level (i.e. Houston, Phoenix, & Detroit) have more automobile use, while regions with high-density level (i.e. Chicago, New York, & Portland) have more compact & more public transit use.
	Moore (2001)	Comparison between Atlanta & Portland; 1988-1997	Carbon monoxide & ozone levels; Vehicle miles travelled (VMT)	Portland has greater reduction rate in carbon monoxide & ozone levels than Atlanta.
	Ewing et al. (2002, 2003)	83 large metros; 1990-2000; Multiple regression analysis	Quality of life (i.e. 8-hour average ozone level); SGA index (i.e. residential density, land use mix, centeredness, & Accessibility	More sprawling regions drive longer distance, own more cars, walk & use public transit less, and breathe more polluted air.
	Stone (2008)	45 large metros; 1990-2002; Integrated multiple regression analysis	Air quality (measured by the number of annual ozone exceedances); Urban sprawl (i.e. density and connectivity)	High sprawling regions have high levels of mean annual ozone.
	Schweitzer & Zhou (2010)	80 metros; 2000 Census; Two-scale (i.e. neighborhood and regional) linear regression models	Air pollutants (i.e. concentration in ozone & fine particulates ($PM_{2.5}$); Urban form using SGA index; Neighborhood-level human exposures	Compact regions lower ozone concentrations, whereas ozone exposures in neighborhoods in compact regions are higher.
	Clark et al. (2011)	111 urban areas; 1990 Census & 2000 air quality; Cross-sectional stepwise linear regression analysis	Air quality (measured by long-term population-weighted ozone and particulates ($PM_{2.5}$) concentrations; Urban form (measured by density and centrality)	Spatial distributions of population are statistically significant predictors of air quality
Sprawling	Robson (1977)	44 large metros; 1920-1950, 1970; Statistical equations	Concentration in particulates &nitrogen dioxide; Dispersion of population & workforce and temperature	Large metros with increased dispersion of population and firms lead to reduce air pollution concentration (i.e. particulates and nitrogen dioxide) because of more public transit use
	Kahn & Schwartz	California cities; 1997-2002;	Zip-coded California emissions (measured by	Sprawling patterns reduce urban air pollution due to technological

	(2008)	Log-linear OLS regressions	hydrocarbons, carbon monoxide, and nitrogen oxides); Vehicle model year	advance to vehicle's emissions
No link	Emison (2001)	52 metros; 1982-1996; OLS regression models	Sprawl (measured by change in population density or in urbanized area); Environmental quality (measured by improvements in ozone air pollution)	Changes in population density and higher environmental protection expenditures had no impact on improvements in ozone air quality

2.5 Limitations of Prior Literature

A growing body of knowledge in urban development and planning studies has been paying attention to the future of alternative development patterns in a sustainable aspect. Its nexus is about how the shapes of a region contribute to the health and quality of life for people and the environment. As explained earlier, postwar suburbanization in metropolitan areas refers to spatial transition as a process of tensions between households, firms, and governments, which can shift from compact to sprawling development. Its impacts can drive different changes in environmental quality within a region. However, the existing theories surrounding the relationship between urban form and environmental quality lack full understanding of the multidimensional nature of metropolitan spatial structure, which accompanies the determinants mentioned in Section 2.3.4. Empirically, not much of the prior evidence has provided any definitive answers to the future of metropolitan patterns in a multidimensional context, which is particularly associated with the effect of compact pattern to improvement in environmental quality since the 1990s. Furthermore, little attention has been drawn to comparison studies on the links between different urban forms (i.e. compact cites versus sprawling cities) and environmental quality using the presence of spatial dependence among neighboring areas.

2.5.1 Necessity for a Comprehensive Framework

Both the nature of metropolitan structure and its intervening variables should be reflected in a better understanding of the relationship with environmental quality. A comprehensive framework will be needed to explore the nexus between metropolitan spatial structure and environmental quality in a multidimensional context which reflects spatial interactions between different strengths, different agents, and different distances in order to provide better information for urban policy decision-making processes. Such a framework must be developed on the basis of the strongest theoretical models (Berry et al., 1974; Newton, 1997; Stone, 2008) and use an integrated approach of metropolitan structure, its intervening variables, and environmental quality.

2.5.2 Recognition of Compact Development Patterns

Since the early 1800s, the trends in metropolitan suburbanization in America have reached spatial shifts of households and firms from central cities to suburbs and beyond. Since the 1990s, alternatives to overcome negative externalities of suburban sprawl (i.e. environmental degradation) have emphasized the future of metropolitan development patterns in a sustainable view. Furthermore, many works suggested empirical predictions that a more compact city can contribute to environmental quality improvements (Newman & Kenworthy, 1999; Newton, 1997, 2000; Masnavi, 2000; Williams, 2000; Neuman, 2005) compared to a more sprawling city. However, little attention has been paid to empirical applications to test the compact city hypothesis that regions with more compact patterns are more desirable for environmental aims than those with more dispersed patterns (Neuman, 2005).

2.5.3 Development of Empirical Models in a Multidimensional Context

Empirically, previous studies have shown some limitations regarding the links of metropolitan structure and air quality. First, the use of larger metropolitan areas does not allow for more detailed characteristics of smaller regions such as counties and census tracks, which can fail to identify patterns of concentration or deconcentration happening at smaller regions The analysis of a county-level unit can better reflect the nature and characteristics of a geographical distribution of population and economic activity in space and over time (Beeson et al., 2001; Rappaport, 2007). Secondly, the measurement of density in population and employment over the entire area, including natural areas such as lake, river, or open space, can lead to misleading predictions to policy makers. To overcome geographical boundary issues related to density or size, we will use the nature of metropolitan structure developed by Galster et al. (2001) and their colleagues, considered “developable” area except for natural space. Thirdly, the key determinants of urban spatial structure and their interactions have not been considered simultaneously and comprehensively in an effort to identify changes in metropolitan shapes. An exploratory analysis of a three-way interaction between land use, socioeconomic characteristics and travel patterns in the UK by Stead, Williams and Titheridge (2000) suggested the importance of interrelationships between different intervening factors. They suggested that the success of compact cities may result as much from the socio-economic characteristics of the residents as from the land use characteristics. Yet the urban studies literature indicated that there are still uncertainties associated with interdependence between significant intervening variables which identifies the complexity of urban systems in solving local and regional problems, particularly in environmental quality.

Further research will be needed to examine some unproven issues on geographic concentration of economic activity by industry (Hanson, 2000), vehicle attributes (Beeson et al., 2001; Rappaport, 2007; Stone, 2008), and interdependence between intervening variables (Torrens, 2008). Lastly, failure to identify the presence of spatial interaction among neighboring areas and the time periods can lead to model misspecification related to unobserved differences of contiguous spatial units (Anselin, 1988, 2003; Ward & Gleditsch, 2008). The use of spatial dependence model will be required to provide better information for the future of alternative development patterns at the regional level (counties or metros).

CHAPTER III

RESEARCH DESIGN AND METHODS

3.1 Conceptual Framework

The conceptual framework for this dissertation is built from main elements of metropolitan spatial structure reviewed in the previous chapter. This includes the complex interactions of the process of location decisions of residents, firms, and governments; as well as the intended and unintended impacts on human health and the environment as a result of those decisions. In the complex urban system, the change in metropolitan spatial structure (MSS) (as measured by land use patterns) across regions and over time contributes to the change in air quality level (as measured by average air quality index values), the proxy for environmental conditions.

Depending on the complex linkages between regional basic components and political characteristics, the interaction between spatial structure and its intermediate factors (i.e. travel behaviors and regional amenities) plays a crucial role in the changes in air quality level. Each element of the framework in the link between spatial structure and air quality determines the extent to which the level of air quality changes. Figure 3-1 lays out the analytical framework that relates metropolitan spatial structure (MSS) to its intervening variables.

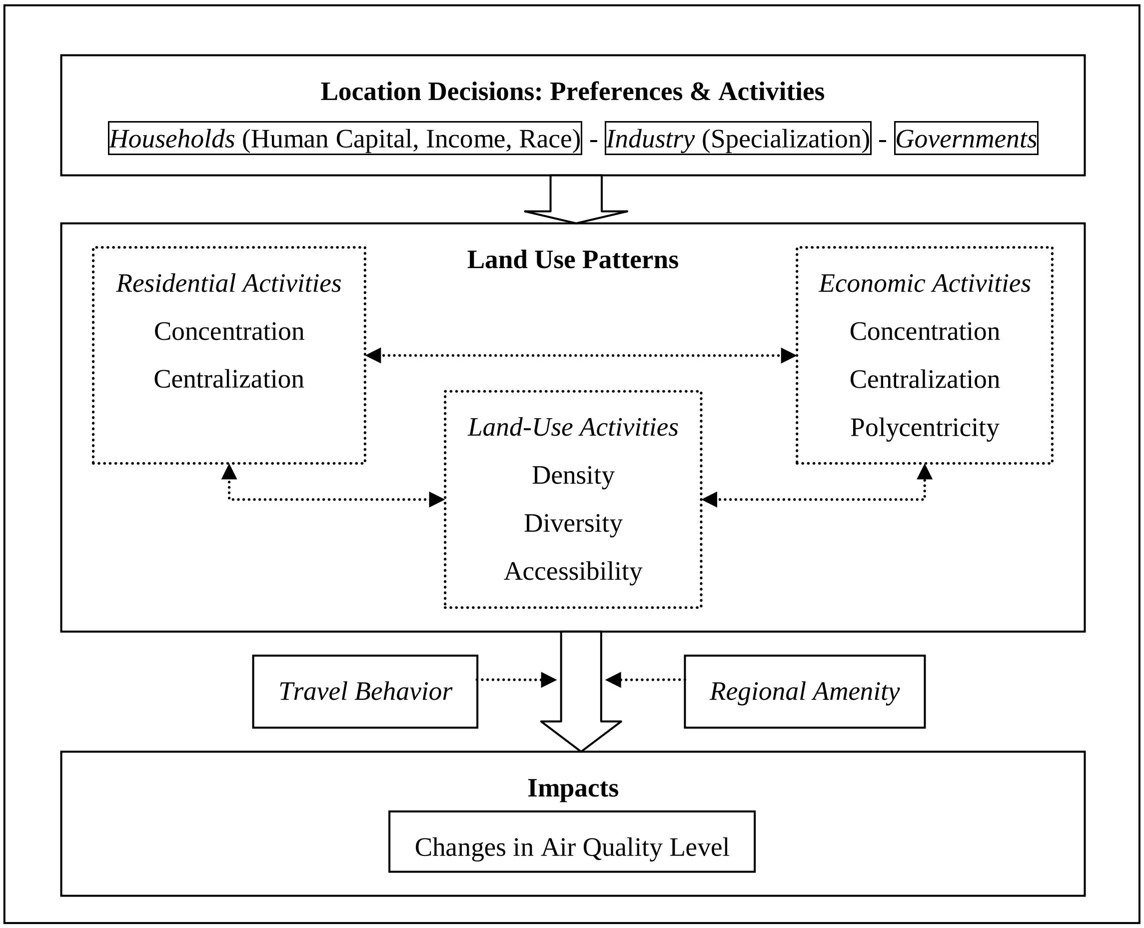


Figure 3-1 Spatial Interaction between MSS, Its Intervening Variables and Air Quality Level

The analytical framework is created by reduced forms between those relating variables:

Air Quality Level = f (MSS, Basic Components, Intermediate Effects, Initial Effect, *u*)

Basic Components

= f (human capital, income, race, specialization, political features) (1-a)

Intermediate Effect = f (travel behaviors, regional amenities) (1-b)

Initial Effect = f (initial conditions of population, jobs, and land in 1990) (1-c)

Metropolitan spatial structure (MSS) may be examined through the spatial variations of population or employment density that interacts with land-use activities in a specific area in that region. The spatial interactions between residential activities and economic activities result in land use activities that can greatly contribute to changes in air quality level. Spatial interaction will be identified as follows: the degree of change in residential or non-residential land uses, the concentration of residents or jobs in some specific areas, the centralization of residents or jobs in some specific areas located closer to each other, the polycentricity of jobs in some specific areas, and the accessibility of households or jobs to a range of urban opportunities (i.e. the central area).[15] The steps to measure changes in land use patterns as a proxy for metropolitan spatial structure are discussed in Section 3.2.2.

This analysis considers main elements affecting changes in metropolitan spatial structure (MSS). First, the spatial pattern of residential or economic activities in a region (1-a) is shaped by the characteristics of the region's human capital, income level, level of specialization, and racial composition. Secondly, local jurisdictions that exercise "home rule" (1-a) influence changes in metropolitan patterns. In addition to the resulting

[15] This dissertation uses and extends precedents in measuring metropolitan spatial structure, as developed by Galster et al. (2001), Ewing et al. (2002), Cutsinger et al. (2005), Tsai (2005), and Torrens (2008).

political fragmentation of "home rule", state- and regional-level land use programs interact to elicit changes in the spatial, economic, and demographic patterns of that region. Thirdly, spatial effects of different development patterns to changes in air quality level can vary by intermediate impacts, such as commuter travel behaviors and regional amenities including average temperature and Census locations (1-b). Local travel behaviors, influenced by federal and state-level environmental policies, can yield spatial differences in ozone precursor emissions (i.e. carbon monoxide, nitrogen oxides, and volatile organic compounds resulting from internal combustion engines). Lastly, the initial conditions of population, employment, and of natural open space in that region (1-c) can also play an important role in changes in metropolitan structure and air quality level (Beeson et al., 2001).

3.2 Unit of Analysis, Data Sources and Variables

3.2.1 Unit of Analysis

This dissertation uses 610 counties in the level of metropolitan statistical area (MSA).[16] The county is considered as an aggregate of sub-areas (census tracts).[17] The

[16] The terms "counties in the level of metropolitan statistical area (MSA)" or "metropolitan areas" will be used similarly for the unit of analysis in this dissertation.
[17] A sub-area refers to a census tract of a county that is mapped as a point feature (centroid) representing the mean value of density in population, employment, and land-use activities for that tract.

county represents the spatial distribution of population, employment, land uses, governments, and other major confounding variables for 1990, 2000, and 2006. The county also reflects the detailed characteristics of small areas (i.e., census tracts) within a county, because it better reflects patterns of concentration or deconcentration happening among smaller local economies than do larger economies, such as states, regions, or nations (Beeson et al., 2001; Desmet & Fafchamps, 2005). The county represents the potential significance of proximity among neighboring counties related to spatial spillovers (Anselin, 1988; Desmet & Fafchamps, 2006; Ward & Gleditsch, 2008). The county stands as a more consistent spatial boundary, having experienced less change over the period 1990-2006 than the boundaries of cities or metropolitan areas (Beeson et al., 2001; Rappaport, 2007).

For the consistency of the data for 1990, 2000, and 2006, excluded from this dissertation are counties in micropolitan statistical areas, counties in Alaska and Hawaii, counties with missing values,[18] and counties not having air pollutant monitoring stations. Figure 3-2 shows counties in the U.S. metropolitan areas with air quality index values used in this dissertation.

---

[18] The 4 counties in the New York-Northern New Jersey-Long Island, NY-NJ-PA (MSA code 35620) - Bronx, Queens, Richmond and Kings, and the Broomfield county (county FIPS, 08014) in the Denver-Aurora, CO (MSA code 19740) are excluded in this dissertation.

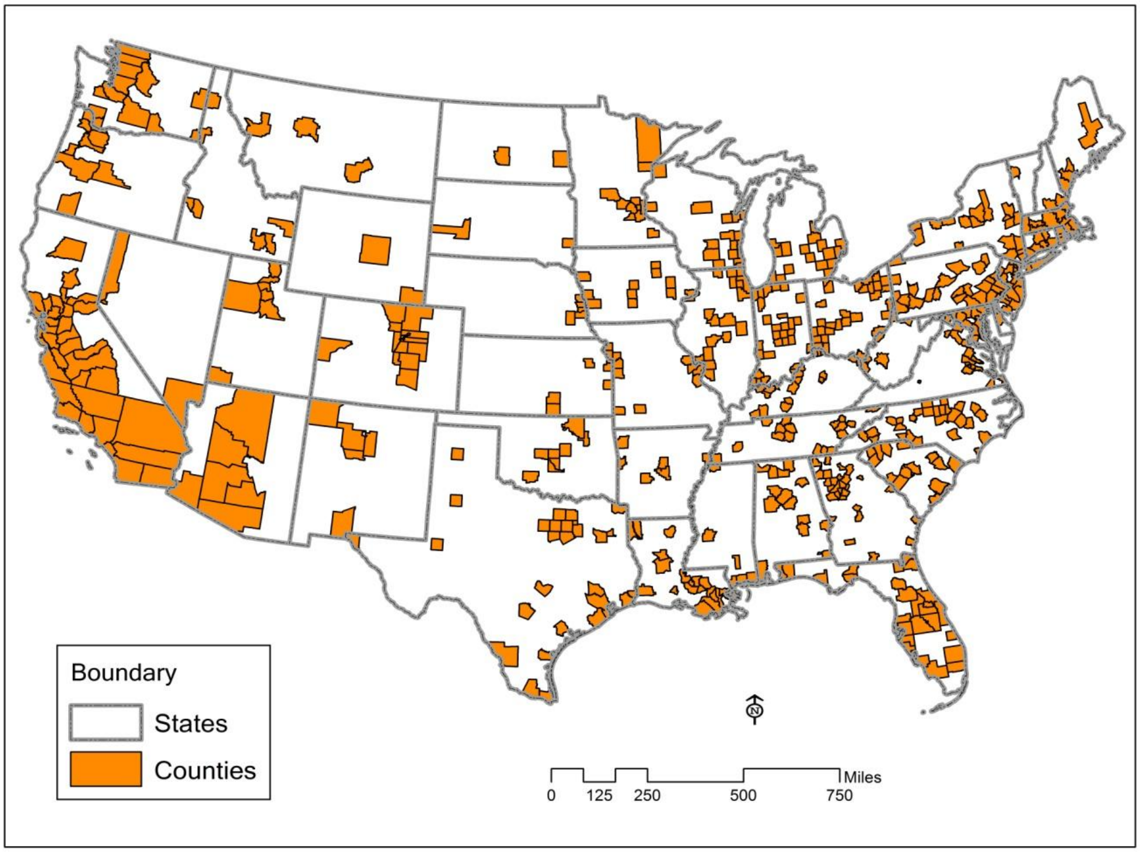


Figure 3-2 Counties in the U.S. Metropolitan Areas with Air Quality Index Values

3.2.2 Data Sources and Independent Variables

This dissertation creates a combined database of population, employment, government, land use, travel behavior, and average air quality index (AQI) values for 610 counties in the metropolitan areas using geographic information system (GIS) tools. The integrated spatial database is compiled from different sources.

3.2.2.1 Population

Population data is from the 1990, 2000, and 2005-2009 5-year period estimates of U.S. Census Bureau. The 2005-2009 ACS data are used as a proxy for the year 2006 to achieve a consistent source for 2006 data at the census tract level. Differences in total population between the 1-year period estimates in 2006 and the 5-year period estimates in 2005-2009 at the county level appeared to be very small or less than 5 % margin of error. This indicates that a use of 2005-2009 5-year period estimates is reliable to calculate a change in population at the census tract level for 2006.

3.2.2.2 Employment

Employment data is from the 2000 Census Transportation Planning Package (CTPP) by the Bureau of Transportation Statistics (BTS) at the traffic analysis zones (TAZs) for place of work data.[19]

For data on level of specialization in industries, we utilize the Moody's economy.com data on industrial activities at a county level. The Moody's economy.com employment data are derived from the annual employment data adjusted historically to

[19] TAZs are geographic boundary to delineate traffic-related data at the census tract level, particularly for place-of-work (see geographic area description, http://www.census.gov/geo/www/cob/tz_metadata.html#gad).

reconcile the Bureau of Labor Statistics (BLS) Current Employment Statistics for employment as well as the Regional Economic Information System (REIS) by the Bureau of Economic Analysis (BEA) at a county level. We categorize the industrial activities into four industrial sectors based on the literature review: manufacturing (MNFG), services, environmentally-friendly industry (ENV), and research and development (R&D), using the SAS software package. A list of environment-friendly industries is based on the Bureau of Labor Statistics (BLS) definition. The BLS defines environmentally-friendly industry as ones that "produce goods and provide services that benefit the environment." This dissertation focuses on jobs associated with air pollution, particularly on "pollution reduction and removal, greenhouse gas reduction, and recycling and reuse," as categorized by the BLS.[20] Table 3-1 shows definitions of the four industrial sectors to represent level of specialization in industries.

[20] See the BLS green jobs definition (http://www.bls.gov/green/green_definition.pdf).

Table 3-1 Definitions of Four Industrial Sectors

Variables	Sector--Title	Definition
MNFG	31-33 Manufacturing	Comprises establishments engaged in the mechanical, physical, or chemical transformation of materials, substances, or components into new products
Services	44-45 Retail Trade	Comprises establishments engaged in retailing merchandise and rendering services incidental to the sale of merchandise.
	61 Education Services	Comprises establishments that provide instruction and training in a wide variety of subjects
	62 Health Care & Social Assistance	Comprises establishments providing health care and social assistance for individuals
	71 Arts, Entertainment, & Recreation	Includes a wide range of establishments that operate facilities or provide services to meet varied cultural, entertainment, and recreational interests of their patrons
ENV	4851 Urban Transit Systems	Comprises establishments primarily engaged in operating local and suburban passenger transit systems within a metropolitan area and its adjacent nonurban areas, such as light rail, commuter rail, subways, streetcars, buses, & other motor vehicles
	4852 Interurban & Rural Bus Transportation	Comprises establishments primarily engaged in providing bus passenger transportation , principally outside a single metropolitan area and its adjacent nonurban areas
	4854 School & Employee Bus Transportation	Comprises establishments primarily engaged in providing buses and other motor vehicles to transport pupils to and from school or employees to and from work
	4855 Charter Bus Industry	Comprises establishments primarily engaged in providing buses for charter; Associated with multi-passenger commuter services
	5112 Software Publishers	Comprises establishments primarily engaged in computer software publishing or publishing and reproduction; Associated with software used to reduce or monitor energy usage
	562 Waste Management & Remediation Services	Waste collection, waste treatment and disposal, and remediation and other waste management
R&D	5417 Scientific Research & Development Services	Comprises establishments engaged in conducting original investigation undertaken on a systematic basis to gain new knowledge (research) and/or the application of research findings or other scientific knowledge; Associated with pollution reduction via research on biofuels and organisms

Note: Industrial sectors and titles are identified according to the 2007 NAICS definition (http://www.census.gov/eos/www/naics/).

Methodologically, the level of specialization for the four sectors at a county level is measured by the location quotient index of their industrial activities. Location quotient (LQ) is measured as a ratio of a region's share of jobs in an industry relative to the nation's share of jobs in that industry over time. We assume that highly concentrated industries (those with an LQ greater than 1.0)--considered a region's economic base--will be export-oriented industries which can contribute more to potential regional employment growth over the given period relative to other industries, leading to attract more people or jobs to move into a given region. Regions with manufacturing-dominated industrial activities (or high-LQ manufacturing industry) tend to be associated with worsened air quality, whereas regions with more environmentally-friendly industrial activities (or high-LQ environmental industry) tend to be associated with improved air quality.

$$LQ_i = \frac{[e_i/e]}{[E_i/E]}$$

, where

LQ_i is the location quotient for industry *i* in metropolitan area A.

e_i is local employment in industry *i;*

e is total local employment;

E_i is national employment in industry *i;*

E is total national employment.

3.2.2.3 Government

Government-related data comes from the Census of Governments by the U.S. Census Bureau and the National Association of Counties (NACo). The government-related data are divided into two categories. One is for general purpose municipalities, including cities and townships in incorporated places. The other is for special purpose governments, those that provide goods and services relating to public health, sewer, fire, police, parks, library, and school but do so independently of general purpose municipalities. They may or may not share resources or spatial boundaries. Political fragmentation is measured as the number of local governments per 1000 population in that county. Regions with more fragmented local governments in the county--based on the local governments ‘vote with their feet’ principle (Tiebout, 1956)--can contribute to the development of the suburbs or urban fringe in the county, leading to dispersed development patterns.

The level of environmental policy innovation, based on Resource Renewal Institute’s (RRI’s) the State of the States (Siy, Koziol, & Rollins, 2001), is used to assess the capacity for achieving sustainable development of the states, which accounts for “the degree to which a state seeks continuous improvement of its environmental programs” (Siy et al., 2001, p. 13). The level of environmental policy innovation was scaled from 0 (lowest) to 40 (highest) points; measured by the 11 policy-based indicators, including air quality standards, pollution prevention programs, energy policy supportive of renewable, existence of National Environmental Performance Partnership System (NEPPS) program, existence of environmental leadership program, existence of state climate change action plan, state authored inventories of greenhouse gas emissions, existence of state-level "Right-to-Know" act, existence of "bottle bill" legislation, existence of environmental assessment requirements, and innovation in comprehensive plan requirements. The states with higher policy innovation scores are considered the greater continuous progress in improving environmental performance for a sustainable future, particularly on air quality and land use. We assign all 50 states' policy innovation scores (Siy et al., 2001, p. 61) to the counties corresponding to each state, respectively.

In addition, we measure the importance of statewide growth management programs (SGMPs). We specify a dummy variable to point out if a state adopted statewide management programs up to 2006. Yin & Sun (2007) considered counties within 15 states that have adopted state growth management programs as the presence of statewide planning measures. The states with SGMPs are as follows: Hawaii in 1961, California in 1965, Vermont in 1970, Oregon in 1973, Florida in 1985, New Jersey in 1986, Maine in 1988, Rhode Island in 1988, Georgia in 1989, Washington in 1990, Maryland in 1992, Arizona in 1998, Tennessee in 1998, Colorado in 2000, and Wisconsin in 2000. Assuming that counties with state growth management programs have a relatively greater magnitude in improving environmental quality, the counties with SGMPs are more likely to have great change in the pollution index and their natural footprint over time than those without SGMPs.

3.2.2.4 Land Use Activities and Their Interrelating Variables

Land-use activities data comes from the National Land Cover Database (NLCD) 1992/2001/2006 created by the Multi-Resolution Land Characteristics Consortium (MRLC) at the U.S. Geological Survey (USGS). The NLCD 1992/2001/2006 contains the land cover classification scheme, based on a 30-meter spatial-pixel resolution Landsat Thematic Mapper (TM) satellite data, to provide spatial reference and components of the land cover, such as water, developed, barren, forest, scrubland, herbaceous, planted/cultivated, and wetlands.

Table 3-2 illustrates differences in land use classification codes and descriptions between NLCD 1992 and NLCD 2001, which will bring out actual misleading results in land cover change when a direct comparison between the two land cover databases is made. For more accurate and reliable estimates of land cover change between NLCD

1992 and NLCD 2001, as in Table 3-2, the two NLCD class codes were cross-walked to the modified Anderson Level I land cover classification codes and descriptions (Fry, Coan, Homer, Meyer, & Wickham, 2009) derived from the Anderson Level I and II classification system (Anderson, Hardy, Roach, & Witmer, 1976). Using the modified Anderson Level I land cover classification codes and descriptions, as highlighted in Table 3-2, the NLCD 1992 identifies "urban" land (class code 2) including residential, commercial, industrial, and transportation land uses, corresponding to "developed" land in NLCD 2001. The total "urban" land in NLCD 1992 is considered the sum of urban/recreation grasses (code 85), low intensity residential (code 21), high intensity residential (code 22), and commercial/industrial/transportation (code 23). Land cover change between 1992 and 2001 in the United States metropolitan areas will be comparable in terms of the "urban" (or "developed") land in a modified Anderson Level I class codes and description.[21]

[21] The very cautious point is to compare and interpret the variable land use mixes between 1992 and 2001 directly. Direct comparison between NLCD 1992 and NLCD 2001 is not advisable, because the two land cover products were independently created in terms of substantial differences in imagery, legends, and methods (Fry, Coan, Homer, Meyer, & Wickham, 2009, pp. 1-2).

Table 3-2 Crosswalk of 1992-2001 NLCD Class Code to Anderson Level I Class Code

NLCD 1992		NLCD 2001		Modified Anderson Level I	
Class Code	Description	Class Code	Description	Class Code	Description
11	Open water	11	Open water	1	Open water
85	Urban, recreational grasses	*21*	**Developed**, Open Space	*2*	**Urban**
21	Low intensity residential	*22*	**Developed**, Low Intensity	*2*	**Urban**
22	High intensity residential	*23*	**Developed**, Medium Intensity	*2*	**Urban**
23	Commercial, industrial, roads	*24*	**Developed**, High Intensity	*2*	**Urban**
31	Bare rock, sand	31	Barren Land, Rock, Sand, Clay	3	Barren
32	Quarry, strip mine, gravel pit	31	Barren Land, Rock, Sand, Clay	3	Barren
33	Transitional barren	31	Barren Land, Rock, Sand, Clay	3	Barren
41	Deciduous forest	41	Deciduous forest	4	Forest
42	Evergreen forest	42	Evergreen forest	4	Forest
43	Mixed forest	43	Mixed forest	4	Forest
51	Scrubland	52	Shrub, Scrub	5	Grass/shrub
71	Grasslands, herbaceous	71	Grasslands, herbaceous	5	Grass/shrub
61	Orchards, vineyards, other	82	Cultivated Crops	6	Agriculture
81	Pasture, hay	81	Pasture, hay	6	Agriculture
82	Row crops	82	Cultivated Crops	6	Agriculture
83	Small grains	82	Cultivated Crops	6	Agriculture
84	Fallow	82	Cultivated Crops	6	Agriculture
91	Woody wetlands	90	Woody wetlands	7	Wetland
92	Emergent/herbaceous wetland	95	Emergent/herbaceous wetland	7	Wetland
12	Perennial ice, snow	12	Perennial ice, snow	8	Ice/snow

Note. Referenced from Completion of the National Land Cover Database (NLCD) 1992–2001 Land Cover Change Retrofit Product (see Table 1 Modified Anderson Level I and II land cover classification codes and brief descriptions, p.4). (http://pubs.usgs.gov/of/2008/1379/pdf/ofr2008-1379.pdf)

Table 3-3 illustrates land use classification codes and descriptions for NLCD 2001 and NLCD 2006. The 2001/2006 NLCD identifies open space, residential or nonresidential (i.e. commercial, industrial, recreational) land use activities, particularly in developed land. This dissertation focuses on land use changes in the "developed" land category; which is considered the sum of the *open space* (code 21), *low intensity* (code

22), *medium intensity* (code 23), and *high intensity* (code 24) subcategories. The "developed" land cover changes between 2001 and 2006 in the United States metropolitan areas will be quantifiably comparable, because the National Land Cover Database 2006 (NLCD2006) is produced following the same protocols as NLCD2001 products. This dissertation uses the definition of "net density" as the proportion of one land-use activity (i.e. highly intensive developed land) to the total "developed" land, not all land[22] lying with the administrative municipality boundaries.

[22] Gross density as the proportion of one land-use activity to all land is not used in this dissertation, because all land includes undeveloped areas like water, wetlands or forests in which people did not reside. The inclusion of such undeveloped areas will lead to measurement bias which can not reflect only land-use activities on the area of "developed" land.

Table 3-3 NLCD 2001/2006 Land Cover Class Descriptions

Land Cover Category	Classification Description
Water	All areas of open water or permanent ice/snow cover.
Developed, Open Space (21)	Areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20 percent of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes.
Developed, Low Intensity (22)	Areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20-49 percent of total cover. These areas most commonly include single-family housing units.
Developed, Medium Intensity (23)	Areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 50-79 percent of the total cover. These areas most commonly include single-family housing units.
Developed, High Intensity (24)	Highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses and commercial/industrial. Impervious surfaces account for 80-100 percent of the total cover.
Barren	Areas characterized by bare rock, gravel, sand, silt, clay, or other earthen material, with little or no "green" vegetation present regardless of its inherent ability to support life. Vegetation, if present, is more widely spaced and scrubby than that in the green vegetated categories; lichen cover may be extensive.
Forest	Areas characterized by tree cover (natural or semi-natural woody vegetation, generally greater than 6 meters tall); tree canopy accounts for 25-100 percent of the cover.
Scrubland	Areas characterized by natural or semi-natural woody vegetation with aerial stems, generally less than 6 meters tall, with individuals or clumps not touching to interlocking. Both evergreen and deciduous species of true shrubs, young trees, and trees or shrubs that are small or stunted because of environmental conditions are included.
Herbaceous	Areas characterized by natural or semi-natural herbaceous vegetation; herbaceous vegetation accounts for 75-100 percent of the cover.
Planted /Cultivated	Areas characterized by herbaceous vegetation that has been planted or is intensively managed for the production of food, feed, or fiber; or is maintained in developed settings for specific purposes. Herbaceous vegetation accounts for 75-100 percent of the cover.
Wetlands	Areas where the soil or substrate is periodically saturated with or covered with water

Note. It was referred from the Multi-Resolution Land Characteristics Consortium (MRLC) NLCD2001/NLCD2006 Product Legend (www.mrlc.gov/nlcd06_leg.php).

3.2.2.4.1 Land Use Mix Index

This dissertation examines the spatial distribution of land use activity in five land categories: four land categories in the “developed” land (i.e., open space, low-intensity, medium-intensity, and high-intensity) and the undeveloped land category including forests, wetlands, barren, cultivated, and scrubland, but not including the area of water. A land use mix (LUM) index[23] is used to quantify the evenness of developed land-use activities across five land use categories, based on an index developed by Frank et al. (2006). The land use mix index scores equal zero when one land use is maximally dominated, whereas the scores equal one when a variety of land uses are maximally mixed. Regions with high mixed land use values are likely to decrease travel time and travel distance. The land use mix index is calculated as follows:

Land use mix (LUMix) =

$$-\frac{\text{Area}}{\ln(\text{N})}$$

, where

Area = $(b_1/a)*\ln(b_1/a) + (b_2/a)*\ln(b_2/a) + (b_3/a)*\ln(b_3/a) + (b_4/a)*\ln(b_4/a) + (b_5/a)*\ln(b_5/a)$;

a = Total land area in square miles for all five land categories in the county A;

b_1 = Open space area in square miles;

b_2 = Low-intensity developed area in square miles;

b_3 = Medium-intensity developed area in square miles;

b_4 = High-intensity developed area in square miles;

b_5 = Undeveloped area in square miles;

N = Number of land use categories in the county A.

3.2.2.4.2 Density

[23] It was derived from Shannon's entropy index (Shannon, 1948).

Density-based measurements of land use characteristics are commonly used, even as the debate on density and its proper role in the multidimensional aspects of land use patterns continue (Mees, 2010; Ewing, 1997; Gordon & Richardson, 1997; McLoughlin, 1991). This dissertation examines the dynamic characteristics of land-use activities associated with population or employment. Net population (or employment) density is defined as the number of total population (or employment) per square mile of total "developed" land at the county level. In a similar way, a census tract's (the sub-area) population (or employment) density is defined as the number of total population (or employment) per square mile of the total "developed" land at that sub-area within that county. As reviewed in the literature (Burchell et al. 1998; Ewing et al. 2002; Galster et al. 2001; Lang, 2003), high net population (or employment) density has association with compact patterns while low population (or employment) density has association with sprawling patterns.

However, net density alone cannot account for spatial patterns of sub-area proximity or patterning within that county, or whether some highly populated (or employed) sub-areas are located closer to the central business district (CBD) within that county. We consider the central business district (CBD)[24] as an area containing primary

[24] The dominance of the CBD for residential and economic activities has an effect on the spatial variation

cities with highest dense population (or employment) designated by the U.S. Census Bureau.

To reflect dimensions-based land use characteristics to tackle limits to net density by itself, as developed by Tsai (2005) and Torrens (2008), measurement of net density will be made at multiple scales - at the county level or at the census tract level (the sub-area). To better identify land-use characteristics at the sub-area level, we specify the sub-area as spatial variation of the net population (or employment) density, as developed by previous works (Galster et al. 2001; Cutsinger et al. 2005; Wolman et al. 2005; Marlay & Gardner, 2010). As seen in Table 3-4, highly populated (or employed) sub-areas are considered as sub-areas with the highest (> 95 percentile or top decile), very high (> 90 percentile) or high (> 75 percentile) net population (or employment) density relative to the rest of the metropolitan county.

Table 3-4 shows the net population density quartile thresholds for the high-population sub-areas across all census tracts within the metropolitan counties used for this dissertation, which are at approximately 17,000, 11,300, and 6,500 residents per square mile in 2000, and 16,700, 11,100, and 6,400 residents per square mile in 2006.

---

of population or employment density (Savitch, Collins, Sanders, & Markham, 1993; Rusk, 1993, 2003; Hill, Wolman, & Ford, 1995; Ihlanfeldt, 1995).

To identify a high employment sub-area, we use an employment-population ratio developed by Garreau (1991), calculated as the ratio of the number of employment to the number of population in the sub-area; and the net employment density used in this dissertation, measured as the number of employment per square mile of developed land area in that sub-area. We consider a sub-area to be high employment when a sub-area has more workers than residents, or an employment-population ratio of greater than 1.0, and the highest (> 95 percentile or top decile) or very high (> 90 percentile) or high (> 75 percentile) employment density. Table 3-4 shows the net employment density quartile thresholds for the high employment sub-areas across all census tracts within the metropolitan counties used for this dissertation, which are at approximately 7,500, 4,700, and 2,300 workers per square mile. The minimum employment density used in this dissertation is greater than 500 workers per square mile (> 25 percentile). The higher net employment density sub-areas are more likely related with high employment sub-areas, particularly in the central business districts (CBDs), while the lower net employment density sub-areas are more likely related with low employment sub-areas, particularly in outlying employment sub-areas.

Table 3-4 Density Quartile Thresholds for the High Population (or Employment) Sub-areas

Density Category	Quartiles	Net Population Density, 2000	Net Population Density, 2006	Net Employment Density, 2000
Highest (top decile)	95	16,985.5	16,686.6	7,475.7
Very High	90	11,296.2	11,127.0	4,667.5
High	75	6,476.6	6,427.5	2,314.5
Medium	50	3,820.0	3,838.2	1090.6
Low (bottom)	25	2,157.1	2,257.2	497.1
Outlying	< 25			

Source: U.S. Census; CTPP 2000
Note: The number of observations used is 45,091 census tracts for 610 counties.

3.2.2.4.3 Concentration

A concentration index is used to identify the extent to which populated (or employed) sub-areas are equally distributed within that county, or which highly populated (or employed) sub-areas are concentrated in some sub-areas within that county. Based on the identification of spatial variation of net population (or employment) density in all sub-areas, we estimate the Gini index in order to estimate the evenness of distribution across all sub-areas within the county. As developed by prior works (Galster et al. 2001; Cutsinger et al. 2005; Wolman et al. 2005) building on Lorenz (1905)'s curve of income concentration and the Gini coefficient of unequal distribution, the Gini index is calculated as the proportion of the total number of population (or employment) in highly populated (or employed) sub-areas to that of population (or employment) in all sub-areas within the county (see formula below), called "population concentration index" or "employment concentrations index." The index is scaled from 0 (lowest) to 1 (highest). Higher population (or employment) concentration index (e.g., close to 1) suggests that some highly populated or employed sub-areas are disproportionately located in the county, whereas lower population (or employment) concentrated index (e.g., close to 0) indicates that populated (or employed) sub-areas are more evenly located (or sprawl-like) in the county.

Population (or employment) concentration index =

$$\frac{\sum_{i=1}^{m} \text{Population (or Employment)}}{\sum_{i=1}^{N} \text{Population (or Employment)}}$$

, where

$\sum_{i=1}^{m} \text{Population (or Employment)}$

= The total number of population (or employment) in 1 to m very highly populated (or employed) sub-areas in the county;

$\sum_{i=1}^{N} \text{Population (or Employment)}$

= The total number of population (or employment) of all sub-areas in the county;

m = The number of highly populated (or employed) sub-areas in the county;

N = The number of all sub-areas in the county.

3.2.2.4.4 Accessibility

The accessibility index is associated with travel time and distance for workers from each sub-area to commute the CBD or very highly employed sub-areas in the county. We use two indicators relating to travel behavior, the average commute time and the weighted average drive-alone commute time, as used by previous works (Ewing, 1997; Sierra Club, 1998; HUD, 1999; Ewing et al. 2002). The average commute time is defined as “travel time to work for workers 16 years and over who did not work at home” (U.S. Census Bureau, 2000), called “commuters.” The weighted average drive-alone commute time is calculated as the average commute time weighted by total drive-alone commuters in the county. Regions with shorter average commute times or shorter weighted average drive-alone commute times can bring about less gasoline consumption, leading to lower levels of air quality index than those with longer commute time.

3.2.2.4.5 Centralization

Along with the degree of concentration in population (or employment) distribution, the centralization index represents the spatial distribution of highly populated sub-areas. The centralization index measures whether highly populated (or employed) sub-areas are geographically clustered, dispersed, or random in the county. The centralization index is useful to supplement the limitation of the concentration index as to which populated (or employed) sub-areas are equally distributed within that county, or which highly populated (or employed) sub-areas are concentrated in some sub-areas within that county.

We use the Moran's *I* statistic for spatial autocorrelation metrics (Moran, 1950; Fotheringham et al., 2000; Anselin, 2003; Ward & Gleditsch, 2008) of the ordinary least squares' (OLS) residuals in order to estimate the level of clustering among the sub-areas in the county. The Moran's *I* coefficient is calculated by the inverse-distance-based weighting between the centroids of two sub-areas. The Moran's *I* coefficient ranges from -1 to +1. Theoretically, the coefficient scores equal -1 when highly populated (or employed) sub-areas are distributed in a chessboard (or decentralized sprawling) pattern, the scores 0 when highly populated (or employed) sub-areas are randomly scattered, and the scores +1 when highly populated (or employed) sub-areas are geographically clustered. Highly clustered regions produce shorter travel distances and time to reduce air pollution emissions, whereas less populated (or employed) clustered regions bring out longer travel distances and time to increase air pollution emissions.

Moran's I =

$$\left(\frac{\mathrm{N}}{\sum_{\mathrm{i}=1}^{\mathrm{N}}\sum_{\mathrm{j}=1}^{\mathrm{N}}\mathrm{W_{ij}}}\right)\left(\frac{\sum_{\mathrm{i}=1}^{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}\mathrm{W_{ij}}\ (\mathrm{x_i}-\bar{\mathrm{x}})(\mathrm{x_j}-\bar{\mathrm{x}})}{\sum_{\mathrm{i}=1}^{\mathrm{N}}(\mathrm{x_i}-\bar{\mathrm{x}})^2}\right)$$

, where $-1 \leq I \leq +1$;
$\mathrm{x_i}$ is the number of population (or employment) in sub-area *i*;
$\mathrm{x_j}$ is the number of population (or employment) in sub-area j;
$\bar{\mathrm{x}}$ is the mean of population (or employment) of all sub-areas N;
$\mathrm{W_{ij}}$ is the row-standardized inverse-distance-based weights matrix between sub-areas *i* and j;
N is the number of all sub-areas.

Using a normalized factor as a z score with the mean and variance component, Moran's *I* is simplified as follows:

Moran's *I* =

$$\frac{1}{2}\sum_{ij} W_{ij} z_i z_j$$

, where $\forall_{i \neq j}$;
W_{ij} is the row-standardized inverse-distance-based weights matrix between sub-areas *i* and j;
z_i is the z score (with a mean of 0 and a standard deviation of 1) in sub-area *i;*
z_j is the z score in sub-area j.

3.2.2.5 Other Confounding Variables

The three other confounding forces of race, income, education, and travel behavior can contribute to a change in metropolitan structure, as examined by previous works (Glaeser et al., 1995; Ewing, 1997; Carruthers & Ulfarsson, 2002; Carruthers, 2003; Glaeser & Kahn, 2003; Portney, 2003; Berliant & Wang, 2004; Kahn, 2006; Faggian & McCann, 2009). Racial composition is measured as proportion of black or Hispanic residents in the county. Income level is defined as average median household income in the county for 1990, 2000, and 2006. Education as a proxy for human capital is calculated as proportion of college graduates or higher (i.e., bachelors, masters, and doctorates) for 1990, 2000, and 2006. Car-dependent commuters as a proxy for travel behavior are measured as proportion of drive-alone commuters for workers 16 years and over by means of transportation in the county for 1990, 2000, and 2006.

In relating to the white flight hypothesis (Frey, 1979; Massey & Denton, 1993), migration of high-income white people to the suburbs can excel suburban development, and consequence of suburban growth can bring out more increased commute time and distance leading to more air pollution emissions and more land consumption. In view of the demand for green governance (Kahn, 2006), regions with more highly educated people can support pro-environment policies including environmental regulation to reduce air pollution threats. Relating to accessibility index, regions with higher proportion of drive-alone workers can produce more increased commute time and distance, more gasoline consumption, and less public transit use, leading to more air pollution emissions than those with lower proportion of drive-alone workers.

Regional amenities such as climate and geographical location can have an impact on the shape of a region, leading to changes in air quality across the counties in the metropolitan areas. We use climate scores from 0 (worst) to 100 (best) percentiles for the U.S. metropolitan areas judged by *Places Rated* rating in terms of four factors[25]: winter

[25] The four factors are defined: winter mildness (i.e., wind-chill defined as air temperature reduced by

mildness, summer mildness, hazardousness, and seasonal effect (Savageau, 2007, pp. 497-499). The higher the average score for each metropolitan area, the ‘better’ the metropolitan area is considered to be, with respect to weather. We assign the average score for each metropolitan area as those for all counties located within the metropolitan area. Also, we use a binary variable in terms of the census divisions (dummy = 1) from the U.S. Census Bureau to control for the Pacific division (dummy = 0). Regions with better climate and location advantage (i.e., Southwestern division) are more likely to grow the number of population (or employment) than those with worse climate and locational disadvantage,[26] which can bring out increased congestion and environmental degradation (Carruthers & Ulfarsson, 2002; Carruthers, 2003; Chen, Irwin & Jayaprakash, 2009).

3.2.3 Dependent Variable: Air Quality Index (AQI)

Air Quality Index (AQI) data at a county level were obtained from the Air Quality System (AQS) of the U.S. Environmental Protection Agency. The Air Quality Index

---

wind), summer mildness (i.e., humidity, the average 24-hour temperature of the hottest month, and the number of months the thermometer tops 90˚F), hazardousness (i.e., winter snowfall and the frequencies of strong winds and thunderstorms), and seasonal affect (i.e., the number of cloudy days, wet days and fog days) (Savageau, 2007, pp. 497-499).

[26] According to the residents’ choices of locations (Tiebout, 1956; Ferguson, Ali, Olfert, & Partridge, 2007), more people tend to move into the region to have better benefits (or amenities).

(AQI)[27] is an indicator of overall air quality measured from any monitoring sites in the county for one year and reported to Air Quality System (AQS) database on a daily base. The AQI represents the six ambient air pollutants regulated by the Clean Air Act (CAA) amended in 1990: carbon monoxide (CO), nitrogen dioxide (NO_2), ground-level ozone (O_3), sulfur dioxide (SO_2), and particulate matters ($PM_{2.5}$ or PM_{10}), which may be harmful to public health of exposed sensitive groups such as children and older persons (American Lung Association, 2012).

We use the EPA's definition that the highest reported AQI value of the air pollutants for the county for each day is considered the "defining" AQI value for that date, called "main pollutant." The annual summary values of AQI for one year and for the county were downloaded from EPA's Air Quality System (AQS) database for the period 1990-2006.[28] The AQI values for the maximum air pollutants concentrations were

[27] According to the EPA's definition of AQI(http://www.epa.gov/ttn/airs/airsaqs/), the AQI formula is as follows:

$$\mathrm{AQI} = \left(\left(\frac{\mathrm{I}_{High} - \mathrm{I}_{Low}}{\mathrm{BP}_{High} - \mathrm{BP}_{Low}}\right) \times \left(Conc - \mathrm{BP}_{Low}\right)\right) + \mathrm{I}_{Low}$$

Where:

Conc is the concentration of the pollutant

BP is the upper and lower bounds of each AQI level classification, called "breakpoints" for the level

IHigh is the AQI value of the upper breakpoint of the level

ILow is the AQI value of the lower breakpoint of the level

BPHigh is the concentration associated with the upper breakpoint of the level

BPLow is the concentration associated with the upper breakpoint of the level

[28] See detailed information on air AQS, http://www.epa.gov/ttn/airsaqs/detaildata/AQIindex.htm.

calculated between the lowest AQI value of 0 and the highest of 500, and classified into the six AQI categories within the defined ranges on the basis of the National Ambient Air Quality Standards (NAAQS) for major air pollutant concentrations identified by the U.S. EPA. The higher of the AQI values corresponding to the greater level of air pollution is considered to be a greater concern to public health for the county, while the lower of the AQI values considered being a lesser concern to public health for the county. For example, the AQI value above 100 might be unhealthy and hazardous for people living in the county, particularly children and the elderly. Table 3-5 shows the AQI category corresponding to the major air pollutant concentrations.

Table 3-5 Concordance of the Air Pollutants Concentrations to the AQI Categories

Air Pollutants	Concentrations	AQI Values	AQI Levels
8-hour Ozone	0.000 - 0.059 ppm	0 - 50	good
	0.060 - 0.075 ppm	51 - 100	moderate
	0.076 - 0.095 ppm	101 - 150	unhealthy for sensitive groups
	0.096 - 0.115 ppm	151 - 200	unhealthy
	0.116 - 0.374 ppm	201 - 300	very unhealthy
	> 0.375 ppm	301 - 500	hazardous or very hazardous
24-hour Particle Matter ($PM_{2.5}$)	0.0 - 15.4 μg/m^3	0 - 50	good
	15.5 - 35.0 μg/m^3	51 - 100	moderate
	35.1 - 65.4 μg/m^3	101 - 150	unhealthy for sensitive groups
	65.5 - 150.4 μg/m^3	151 - 200	unhealthy
	150.5 - 250.4 μg/m^3	201 - 300	very unhealthy
	> 250.5 μg/m^3	301 - 500	hazardous or very hazardous
8-hour Carbon Monoxide	> 9 ppm	> 101	> unhealthy for sensitive groups
1-hour Nitrogen Dioxide	> 100 ppb	> 101	> unhealthy for sensitive groups
24-hour Particle Matter (PM_{10})	> 150 μg/m^3	> 101	> unhealthy for sensitive groups
1-hour Sulfur Dioxide	> 75 ppb	> 101	> unhealthy for sensitive groups

Source: The EPA's National Ambient Air Quality Standards (NAAQS) and Air Quality Index Dictionary; American Lung Association's *the State of the Air 2012*.

Note: Unit of measures are parts per million (ppm) by volume, parts per billion (ppb) by volume, and micrograms per cubic meter of air (μg/m3).

The AQI indicator was calculated to identify county-level AQI values. Using the county-level average AQI calculation developed by the American Lung Association (pp. 40-42), we compute the average AQI value over the 3-years for 1990-1992, 2000-2002, and 2004-2006 for the county. The AQI value for 2006 is used to calculate the sum of AQI values for 2004-2006 divided by 3 for the period 2004-2006. For example, if a county had an AQI value of 89 for 2004, 101 for 2005, and 78 for 2006, the average AQI value over 3 years for 2006 for the county would be 89.3, or (89 + 101 + 78)/3. The reason to use the AQI values averaged over 3 years is to be consistent with the EPA's use of 3-year averages to prevent abnormal conditions in any single year from adversely impacting the interpretation of ambient air quality standards. The AQI indicator is relevant to capture the effects of the overall air quality trend on human health (Olewiler, 2006; Stone, 2008).

Table 3-6 provides a brief description of the dependent variables, independent variables, and their respective data sources at the county level and at the census tract level.

Table 3-6 Description of Data Sources and Variables

Variables	Description		Sources	Unit
AQI	Average 3-year AQI, 1990-1992, 2000-2002, 2004-2006		US. EPA	County
Land-Use	Amount in developed land; proportion in different land-use activities; Land use mix, 1992, 2001, 2006		NLCD, MRLC	County; Tract
Density (square mile)	Number of total population (employment) per square mile in developed land, 1990, 2000, 2006		U.S. Census; CTPP; NLCD	County; Tract
Concentration	Proportion of high population (employment) density sub-areas to all sub-areas, 2000, 2006		US. Census; CTPP; NLCD	County; Tract
Accessibility	Average commute time; Weighted average drive-alone commute time, 1990, 2000, 2006		US. Census; CTPP; NLCD	County; Tract
Centralization	Degree of closeness between high population (employment) density sub-areas and the CBD or very highly populated (employed) sub-areas, 2000, 2006		US. Census; CTPP; NLCD	County; Tract
Industrial Specialization	Manufacture	LQ of the manufacturing industry (NAICS 31-33)	Moody's economy.com	County
	Services	LQ of the service industry (NAICS 44-45, 61, 62, 71)		
	R&D	LQ of the R&D industry (NAICS 5417)		
	Environmental Industry	LQ of the environmental industry (NAICS 485, 5112, 562)		
Political Properties	Fragmented	Log of the number of general-purpose local governments per 1,000 persons, 1992, 2002	Census of Governments	County
		Log of the number of special districts local governments per 1,000 persons, 1992, 2002		
	Policy Effects	A state's environmental policy innovation score, 2001	RRI	County
		Regions with statewide growth management programs (SGMP) (dummy, 0, 1)	Yin & Sun (2007)	County
		Age of statewide growth management programs		
Socio-demographic Features	Racial Composition	Proportion of Black or Hispanic (Black + Hispanic) residents, 1990, 2000, 2006	US. Census	County
	Median Household Income	Annual median household income, 1990, 2000, 2006	US. Census	County
	Human Capital	Proportion of college graduates or higher, 1990, 2000, 2006	US. Census	County
Intermediate Effect	Travel Behavior	Proportion of drive-alone commuters by means of transportation, 1990, 2000, 2006	US. Census	County
	Regional Amenity	Climate, 2006	PlacesRated Almanac	County
		9 Census Regions (dummy, 0,1)	US. Census	County
		Log of undeveloped land, 1992	NLCD	County
Initial Effect	Population	Log of total population, 1990	US. Census	County
	Employment	Log of total employment, 1990	Moody's economy.com	County
	Land	Log of total or developed land, 1992	NLCD	County

3.3 Hypotheses

Hypothesis 1: Metropolitan areas with a lower net population density produce lower average air quality index values than those with a higher net population density.

Hypothesis 2: Metropolitan areas with a lower percentage of developed land produce lower average air quality index values than those with a higher percentage of developed land.

Hypothesis 3: Metropolitan areas with more highly diverse mix of land-use activities produce lower average air quality index values than those with less diverse mix of land-use activities.

Hypothesis 4: Metropolitan areas with a higher percentage of densely populated sub-areas produce lower average air quality index values than those with a lower percentage of densely populated sub-areas.

Hypothesis 5: Metropolitan areas with a higher percentage of densely employed sub-areas produce lower average air quality index values than those with a lower percentage of densely employed sub-areas.

Hypothesis 6: Metropolitan areas with shorter average daily commute time produce lower average air quality index values than those with longer commute time.

Hypothesis 7: Metropolitan areas with shorter weighted average drive-alone commute time produce lower average air quality index values than those with longer weighted average drive-alone commute time.

Hypothesis 8: Metropolitan areas with a higher clustering of densely populated sub-areas produce lower average air quality index values than those with a lower clustering of densely populated sub-areas among all sub-areas.

Hypothesis 9: Metropolitan areas with a higher clustering of densely employed sub-areas produce lower average air quality index values than those with a lower clustering of densely employment sub-areas among all sub-areas.

Hypothesis 10-a: Metropolitan areas with a higher level of specialization in the manufacturing industry produce lower average air quality index values than those with a lower level of specialization in the manufacturing industry.

Hypothesis 10-b: Metropolitan areas with a higher level of specialization in the service industry produce lower average air quality index values than those with a lower level of specialization in the service industry.

Hypothesis 10-c: Metropolitan areas with a higher level of specialization in the research & development (R&D) industry produce lower average air quality index values than those with a lower level of specialization in the research & development (R&D) industry.

Hypothesis 10-d: Metropolitan areas with a higher level of specialization in the environmental industry produce lower average air quality index values than those with a lower level of specialization in the environmental industry.

Hypothesis 11-a: Metropolitan areas with more numbers of general-purpose local governments per 1,000 persons produce lower average air quality index values than those with smaller numbers of general-purpose local governments per 1,000 persons.

Hypothesis 11-b: Metropolitan areas with more numbers of special-purpose local governments per 1,000 persons produce lower average air quality index values than those with smaller numbers of special-purpose local governments per 1,000 persons.

Hypothesis 12: Metropolitan areas with highly innovative pro-environment policies produce lower average air quality index values than those with lowly innovative pro-environment policies.

Hypothesis 13: Metropolitan areas with statewide growth management programs produce lower average air quality index values than those without statewide growth management programs.

Hypothesis 14: Metropolitan areas with a lower percentage of Black or Hispanic residents produce lower average air quality index values than those with a higher percentage of Black or Hispanic residents.

Hypothesis 15: Metropolitan areas with a higher level of median household income produce lower average air quality index values than those with a lower percentage of median household income.

Hypothesis 16: Metropolitan areas with a higher percentage of college graduates or higher produce lower average air quality index values than those with a lower percentage of college graduates or higher.

Hypothesis 17: Metropolitan areas with a lower percentage of drive-alone commuters produce lower average air quality index values than those with a higher percentage of drive-alone commuters.

Hypothesis 18: Metropolitan areas with better climate produce higher average air quality index values than those with worse climate.

Hypothesis 19: Metropolitan areas with larger size in total population in 1990 produce higher average air quality index values than those with lower size.

Hypothesis 20: There is no spatial dependence among neighboring regions for the changes in average air quality index values.

3.4 Statistical Methods

Using multiple ordinary least squares (OLS) and spatial regression models, this dissertation seeks to identify the relationships between spatial variations in population, employment, governments, and land-use activities and changes in air quality, as well as the presence of spatial dependence among neighboring regions in metropolitan areas.

3.4.1 Multiple Ordinary Least Squares (OLS) Regression Models

Multivariate regression analysis is used to identify the determinants that significantly influence the air quality improvements in the 610 counties in metropolitan areas for 1990, 2000, and 2006. OLS models are used to estimate the equation-by-equation functions, which are assumed to be linear in parameters and have zero mean and no covariance in the disturbance terms (Gujarati, 2003). The statistical specification is as follows:

$$\text{Air Quality Index}_{j,t} = \alpha + \beta_1 * X_{j,t} + \eta * (\text{MSS})_{j,t} + \varepsilon_{j,t}$$

$$\alpha + \beta_1 * X_{j,t} + \eta * (\text{MSS})_{j,t} + \beta_2 * \log(\text{initial condition})_{j,t-1} + \varepsilon_{j,t}$$

$$\alpha + \beta_1 * X_{j,t} + \eta * (\text{MSS})_{j,t} + \beta_2 * \log(\text{initial condition})_{j,t-1} + \theta * (\text{location})_{j,t} + \varepsilon_{j,t}$$

where j ranges across metropolitan counties; t ranges from the period of 1990 (t-1) through 2006 (t); α represents the overall constant; β_1 represents a $k \times 1$ vector of regression coefficients estimates on the explanatory variables (X); X represents major intervening variables to air quality index, such as human capital, income, race, agglomeration, political properties, residential travel behaviors, and regional amenities; β_2 represents a $k \times 1$ vector of regression coefficients estimates on the initial conditions according to size in population, employment, or developed land in 1990; θ represents the location-fixed effects on different geographical locations; η represents a $k \times 1$ vector of regression coefficients estimates on metropolitan spatial structure measures; and ε is the $n \times 1$ vector of error terms.

3.4.2 Spatial Regression Models

The spatial interaction between neighboring regions may play a significant role in changes in air quality. The spatial effects of 610 counties in metropolitan areas and their neighboring counties are characterized by spatial dependence in the dependent and explanatory variables to influence changes in air quality. OLS regression models ignoring the presence of spatially correlated observations trigger three motivations for including the presence of spatial dependence among neighboring regions in the standard OLS regression models (Anselin, 2002) – theoretical, data-driven, and analytical. OLS regression models ignoring the presence of spatially correlated observations motivate theoretical specifications for including spatial dependence in dependent and explanatory variables in OLS regression models (Anselin, 1988, 2002; LeSage, 1997; LeSage & Pace, 2009). Also, OLS regression models ignoring the presence of spatially correlated observations motivate data-driven specifications[29] for including spatial dependence in omitted variables in the OLS estimates and the spatial regression models (Dubin, 1988; Anselin, 2002; Brasington & Hite, 2005; Pace & LeSage, 2010). Lastly, OLS estimates ignoring the presence of spatially neighboring regions motivate analytical specifications for including the spatial weights matrix (*W*) to reflect the connections between each region and neighboring regions. The analytical specifications can be formed through two different spatial regression models in the spatial econometric literature (Anselin, 1988, 2002, 2003; Anselin & Bera, 1998; LeSage, 1997; LeSage & Pace, 2009; Fotheringham, Brunsdon & Charlton, 2000; Ward & Gleditsch, 2008): spatial error model of explanatory variables and spatial lag model of the dependent variable.

[29] In a similar way, Lesage and Pace (2009) suggests omitted variable or uncertainty motivations correlated or not correlated with the explanatory variables among neighboring regions.

3.4.2.1 Spatial Lag Models

Spatial lag models seek to account for the spatial dependence between spatially lagged values of dependent variable as an extra independent variable. The specified spatial lag model is as follows:

$$Y = \beta X^{*} + \rho WY + \varepsilon$$

$$= (\mathrm{I}_n - \rho W)^{-1} \beta X^{*} + (\mathrm{I}_n - \rho W)^{-1} \varepsilon$$

$$\varepsilon \sim N(0, \sigma^2 I_n)$$

where the terms β and ε are defined in the previous equation. Y is a $n \times 1$ vector of dependent variable (Y) containing logged average air quality index for each region, X^{*} is a $n \times k$ matrix containing explanatory variables including metropolitan spatial structures (*MSS*), initial condition, and location, I is the $n \times n$ identity matrix, W is a $n \times n$ spatial weights matrix for a row-standardized form where the row elements sum to 1, WY is a $n \times 1$ spatial lag vector reflecting a spatially weighted neighborhood average value of the dependent variables (Y) accounted for by continuous inverse distance between neighboring regions specified by the spatial weights matrix W, ρ is a scalar spatial autoregressive coefficient reflecting the strength of spatial dependence in spatially lagged dependent variables, and ε is a $n \times 1$ vector of independent and normally distributed error terms with a vector mean zero (0) and constant variance (σ^2).

When we focus on the reduced form of the spatial lag model in terms of the associated data generating process, we need to consider the spatial multiplier $(\mathrm{I}_n - \rho W)^{-1}$ (Anselin, 2002; Ward &Gleditsch, 2008), which reflects how much the dependent variable *Y* in each region is determined by the spatially lagged dependent variables (*WY*) from neighboring regions or by the error terms in the explanatory variables in that region. This simultaneous feedback in the spatial autoregressive data generating process makes the spatially lagged dependent variable (*WY*) endogenous (Anselin, 2002; LeSage & Pace, 2009; Fingleton & Le Gallo, 2010), which means that the changes in dependent variables from neighboring regions, on average, influence a change in the dependent variable itself in the region. When the spatial autoregressive coefficient $\rho = 0$, we can interpret that there is no spatial dependence in the spatially lagged dependent variables from neighboring regions. However, when the spatial autoregressive coefficient $\rho \neq 0$, we can interpret that there is spatial dependence in the spatially lagged dependent variables from neighboring regions, indicating that the expected value of the dependent variable itself is influenced by the spatially weighted average value in dependent variables from neighboring regions. A higher positive value of spatial autoregressive parameter (ρ)

reflects the presence of the higher strength of spatial dependence in the spatially lagged dependent variables from neighboring regions in the OLS regression models.

3.4.2.2 Spatial Error Models

Due to the data-driven misspecification of functional forms (Anselin, 2002), or omitted variable, or uncertainty motivations (LeSage & Pace, 2009; Pace & LeSage, 2010; Fingleton & Le Gallo, 2010), spatial error models seeks to account for the spatial dependence between the spatially correlated error terms in the explanatory variables in the OLS regression models. The specified spatial error model is as follows:

$$Y = \beta X^* + \varepsilon$$

$$\varepsilon = \lambda W\varepsilon + \mu \text{ and } \varepsilon = (\mathrm{I}_n - \lambda W)^{-1}\mu$$

$$\mu \sim N(0, \sigma^2 I_n)$$

where the terms β and X^* is defined in the previous equations. ε is the unobserved errors in explanatory variables from neighboring regions containing a spatially correlated error term, μ is a $n \times 1$ vector of independent and normally distributed error terms with a vector mean zero (0) and constant variance (σ^2), W is a $n \times n$ spatial weights matrix for a row-standardized form where the row elements sum to 1, $W\varepsilon$ is a $n \times 1$ spatial error vector reflecting a spatially weighted neighborhood average value of the unobserved errors (ε) in explanatory variables accounted for by continuous inverse distance between neighboring regions specified by the spatial weights matrix W, λ is a scalar spatial autoregressive coefficient in terms of a spatially weighted average of the errors explained by continuous inverse distance measure among neighboring regions ($W\varepsilon$).

When we focus on the reduced form of the spatial error model, we need to consider the spatial multiplier $(\mathrm{I}_n - \lambda W)^{-1}$ (Anselin, 2002; Ward & Gleditsch, 2008), which reflects how much the errors in the explanatory variables from neighboring regions are spatially correlated. When the spatial autoregressive coefficient $\lambda = 0$, we can interpret that there is no spatial dependence between the error terms in the explanatory variables from neighboring regions. However, when the spatial autoregressive coefficient $\lambda \neq 0$, we can interpret that there is spatial dependence between the error terms in the explanatory variables from neighboring regions, indicating the strength of the spatial correlation of the residuals among neighboring regions.

3.4.3 Diagnostic Tests of Regression Models

In order to choose a good model, model diagnostics tests detect multicollinearity, normality, heteroscedasticity, or autocorrelation based on the OLS residuals for model specification errors, such as omitting relevant variables, including irrelevant variables, adopting the incorrect functional form, errors of measurement, and incorrect specification of the stochastic error term (Gujarati, 2003).[30]

3.4.3.1 Diagnostic Tests of OLS Regression Models

The diagnostics tests of the OLS regression models consist of three measures: multicollinearity, normality, and heteroscedasticity. Multicollinearity is needed to measure when a linear relationship among some or all independent variables of the OLS regression model exists, such as high correlations between two independent variables. To detect multicollinearity, the first measure of multicollinearity is the variance-inflation factor (VIF), the VIF reflects how the variance of a OLS regression coefficient is inflated by the presence of multicollinearity. The inverse of VIF is called tolerance (TOL) (Gujarati, 2003, p. 350-353). The higher the VIF of an independent variable is, the more collinear the variable is with other independent variables. On the other hand, the closer the TOL of an independent variable is to 0, the higher collinear the variable is with other independent variables, which means that the OLS regression coefficient of the variable can be difficult to precisely estimate with high multicollinearity among other variables. In

[30] For methodological details on these specification and diagnostics testing, see chapter 13 in Gujarati (2003)'s Basic Econometrics.

order to remedy high multicollinearity among independent variables, we drop one of the collinear variables, or transform a variable as a ratio or a natural logarithm value.

$$\text{TOL} = \frac{1}{\text{VIF}}$$

The second detection of multicollinearity is the condition number *k* using eigenvalues defined as

$$k = \frac{\text{Maximum eigenvalue}}{\text{Minimum eigenvalue}}$$

Based on a rule of thumb, if the value of k of the OLS regression model is lower, we consider that there is low multicollinearity of the regression model. Typically, there is moderate to high multicollinearity if *k* is between 100 and 1000 (Gujarati, 2003, p.362).

The Jarque-Bera (JB) test (Jarque & Bera, 1987) is employed to detect normality. The JB test statistic is defined as the chi-square distribution with 2 degrees of freedom (S and K). Using the OLS regression residuals, if the value of the JB statistic is close to zero, which mean the *p* value of the JB statistic is high, we cannot reject the null hypothesis that the residuals are normally distributed.

$$\text{JB} = \text{n}\left[\frac{\text{S}^2}{6} + \frac{(\text{K}-3)^2}{24}\right]$$

where n is sample size, S is skewness coefficient, and K is kurtosis coefficient.

The next three diagnostic tests are employed to detect heteroscedasticity, defined as unequal spread (or variance) of errors: the Breusch-Pagan (BP) test, the Koenker-Bassett (KB) test, and the White test.[31] On the basis of the OLS regression residuals, or squared residuals, the three test statistics are defined as the chi-square distribution with degrees of freedom. The BP and KB tests assume a specific functional form for heteroscedasticity as tests on random coefficients. On the other hand, the White test does not depend on the normality assumption, but introduces the powers and cross-products of the independent variables in the regression model. If the chi-square value with degrees of freedom in the model exceeds the critical chi-square value at the chosen level of significance, we can reject the hypothesis of homoscedasticity. In other words, there is heteroscedasticity in the error variance; otherwise there is homoscedasticity in it.

3.4.3.2 Diagnostic Tests of Spatial Regression Models

3.4.3.2.1 Spatial Autocorrelation Tests

As described previously, Moran's *I* statistic is used to test diagnostics for spatial autocorrelation of OLS regression residuals among independent variables. Autocorrelation is considered as spatial error dependence (Moran, 1950; Cliff & Ord,

[31] For methodological details on these three measures for heteroscedasticity, see Chapter 11 in Gujarati (2003)'s Basic Econometrics.

1972; Anselin, 2003, 2009; Ward & Gleditsch, 2008), which is defined as the presence of correlation between the error terms in space (Gujarati, 2003). Using the spatial weights matrix, the Moran's *I* statistic inference is based on a normal standardized value. If the Moran's *I* value is highly significant, we can reject the null hypothesis that there is no autocorrelation between the error terms in space. In order words, spatial autocorrelation among the OLS regression residuals is present.

3.4.3.2.2 Maximum Likelihood Based Tests

The maximum likelihood (ML) based diagnostic tests are utilized to detect the significance of coefficients (λ & ρ) of spatial regression models, such as a spatial error or a spatial lag model. The ML tests[32] are more appropriate for larger data sets to detect the presence of spatial autocorrelation in the residuals from the OLS regression models.

To obtain the ML estimation, the Lagrange multiplier (LM) test statistic calculated by the chi-square with degrees of freedom is used for detection of the presence of a spatial lag or a spatial error autocorrelation. Based on the *GeoDa* program,[33] LM-Error (i.e., presence of spatial error specification) or LM-Lag (i.e., presence of spatial lag

[32] For methodological details on maximum likelihood (ML) based tests, particularly on the Lagrange Multiplier (LM) statistic to obtain maximum likelihood (ML) estimation, see chapter 14 in the SAGE Handbook Spatial Analysis (Anselin, 2009).
[33] According to spatial regression decision process in *GeoDa*, when both LM-Error and LM-Lag test statistics are significant, we can consider the Robust LM diagnostics, Robust LM-Error and LM-Lag (Anselin, 2005, pp. 197-200).

specification) test statistics are considered to detect the presence of spatial autocorrelation to the OLS regression residuals. If the LM-Error or LM-Lag statistics are significant, we can reject the null hypothesis that there is no spatial dependence. The statistical significance of the spatial autoregressive coefficients (λ & ρ) implies that strong spatial effects are interrelated among its neighboring regions. If the LM-Error or LM-Lag statistics are not significant, we do not reject the null hypothesis that there is no spatial dependence, keeping the OLS regression models.

CHAPTER IV

STATISTICAL RESULTS

4.1 Descriptive Statistics

Descriptive statistics examine the central tendency (i.e., mean), variability around the mean (i.e., standard deviation), deviation from normality (i.e., skewness and kurtosis), and spread of the distribution (i.e., minimum and maximum) for each variable in this dissertation. Tables 4-1 and 4-2 provide descriptive statistics for the distributions of all of the variables, including the size of sample (N), Minimum (lowest), Maximum (highest), Mean (or average), Std. Deviation (standard deviation), Skewness, and Std. error of the Skewness.

4.1.1 Descriptive Statistics for AQI and MSS

Table 4-1 shows descriptive statistics for air quality index values (AQI) and metropolitan spatial structure (MSS), such as land use change, population or employment density, level of concentration in population or employment, accessibility, and level of centralization in population or employment at the county level for the years 1990, 2000, and 2006. All the 610 counties have complete data for each variable, except for air quality index values (AQI) for 1990 (*A_mean90*). For AQI in 1990, 133 of the total 610 counties are excluded because there were no monitoring sites, or inadequate sites in those counties between 1990 and 1992.

We can consider the variables for AQI and MSS to be approximately normally distributed, because the variables for AQI and MSS have their skewness values either between -1.0 and 1.0 or between -2.0 and 2.0. Two variables have skewness; Centralization in employment (*CentE00*) below -2.0, Concentration in population (*conPop00*) above 2.0.

4.1.1.1 Changes in AQI and MSS

As in Table 4-1, the mean value for the air quality index (AQI) has decreased consistently from 1990 to 2006, by 7.24-percent overall; or from 45.9 in 1990 (*A_mean90*) to 42.6 in 2006 (*A_mean06*). This indicates that the average air quality across metropolitan areas in U.S. has been improving during this period. The mean value for land-use activities has greatly increased by 64.0-percent from 1992 to 2006, or from 12.9-percent of "developed" land to the total developed and undeveloped land in 1992 (*pct_urb92*) to 21.2 percentage in 2006 (*pct_urb06*). This represents that more land areas across U.S metropolitan areas has been developed for residential, commercial, industrial, or recreational uses.

As in Table 4-1, the mean values for mixed land uses (LUM) between 1992 and 2006 have been greatly increased by 57.1-percent, or from 0.26 (*LUMix92*) to 0.41 (*LUMix06*). This indicates that mixed land development in U.S. metropolitan areas has strengthened over time.

Table 4-1 Descriptive Statistics for AQI and MSS

Variables	Description	N	Minimum	Maximum	Mean	Std. Deviation
AQI	A_mean90	477	11.467	128.188	45.902	17.346
	A_mean00	610	5.176	100.874	45.809	12.974
	A_mean06	610	13.297	90.996	42.579	10.566
MSS	pct_urb92	610	0.128	93.457	12.943	17.075
	pct_urb01	610	0.663	97.883	20.381	18.963
	pct_urb06	610	0.676	97.991	21.226	19.368
	pct_open01	610	0.276	36.997	8.303	5.533
	pct_open06	610	0.276	37.650	8.599	5.694
	pct_low01	610	0.076	44.949	6.891	7.174
	pct_low06	610	0.077	45.245	7.186	7.341
	pct_med01	610	0.008	39.877	3.558	5.740
	pct_med06	610	0.008	39.923	3.758	5.817
	pct_high01	610	0.001	53.659	1.629	4.143
	pct_high06	610	0.001	53.630	1.683	4.159
	LUMix92	610	0.000	0.961	0.259	0.220
	LUMix01	610	0.029	0.980	0.396	0.226
	LUMix06	610	0.030	0.980	0.408	0.229
	lnNetE90	610	5.488	11.699	7.344	0.532
	lnNetE00	610	4.229	11.614	6.786	0.733
	lnNetE06	610	4.220	11.577	6.797	0.711
	lnNetP90	610	7.103	11.218	8.222	0.390
	lnNetP00	610	5.318	11.139	7.592	0.540
	lnNetP06	610	5.337	11.192	7.631	0.530
	conPop00	610	0.000	0.992	0.093	0.165
	conPop06	610	0.000	0.988	0.089	0.164
	conEmp00	610	0.000	0.938	0.369	0.239
	AveCom90	610	12.335	36.106	21.383	4.074
	AveCom00	610	15.123	41.089	24.413	4.713
	AveCom06	610	15.100	40.700	24.014	4.617
	lnTotCom90	610	10.404	18.151	13.931	1.142
	lnTotCom00	610	10.635	18.194	14.247	1.097
	lnTotCom06	610	10.630	18.337	14.312	1.111
	CentP00	610	-0.985	0.522	0.078	0.178
	CentP06	610	-0.999	0.747	0.075	0.175
	CentE00	610	-0.996	0.437	0.002	0.153

4.1.1.2 Net Population or Employment Density

As displayed in Table 3-4, we identify the net population or employment density quartile thresholds for the high-population or high-employment sub-areas across all the 45,091 census tracts within the 610 metropolitan counties for this dissertation. Table 4-1 shows that the mean value for net population density in natural logarithms has decreased by 7.67-percent between 1990 and 2000, or from 8.222 (*lnNetP90*) to 7.592 (*lnNetP00*), but has displayed slight growth by 0.52-percent between 2000 and 2006, or from 7.592 (*lnNetP00*) to 7.631(*lnNetP06*).

The net employment density shows a similar trend to the net population density during this period. The mean value for net employment density in natural logarithms has dropped by 7.6-percent between 1990 and 2000, or from 7.344 (*lnNetE90*) to 6.786 (*lnNetE00*), but has risen by 0.16-percent between 2000 and 2006, or from 6.786 (*lnNetE00*) to 6.797 (*lnNetE06*), a nearly imperceptible change in the transformed values.

4.1.1.3 Concentration Index

We specify the population or employment concentration index as the ratio of the total population in highly populated sub-areas to that of population in all sub-areas with the county, as stated in Section 3.2.2.4.3. An average 0.093 (or 0.089) population concentration index in 2000 (or in 2006), as in Table 4-1, indicates that 9.3-percent (or 8.9-percent) of the total population is distributed in highly populated sub-areas in the county, while most of total population are distributed in medium or lowly populated sub-areas. Table 4-1 shows that the mean value for population concentration index between 2000 and 2006 has dropped by 3.67-percent, or from 0.093 (*conPop00*) to 0.089 (*conPop06*), which indicates that more people have tended to move to medium or lowly populated sub-areas from highly populated sub-areas in the county, particularly to suburbs or outlying sub-areas. On the other hand, an average 0.369 employment concentration index in 2000 points out that a 36.9-percent of the total employment is disproportionately located in the CBD or highly employed sub-areas in the county.

4.1.1.4 Accessibility Index

This dissertation specifies the accessibility index relating to commuters' travel behavior, such as the commute time to their workplaces, either the CBD or the very highly employed sub-areas in the county, as described in Section 3.2.2.4.4. Table 4-1 shows that the average commute time and the total average commute time weighted by total drive-alone commuters have increased between 1990 and 2006. During this period the average commute time was raised by 12.3-percent from 21.38 minutes (*AveCom90*) to 24.01 minutes (*AveCom06*) and the weighted total commute time rose by 2.7-percent from 13.93 in natural logarithms (*lnTotCom90*) to 14.32 in natural logarithms (*lnTotCom06*). Overall, more workers 16 years and over not working at home tend to drive alone longer in commuting to work in the county during this time period.

4.1.1.5 Centralization Index

As with the population or employment concentration index, we specify the level of centralization as to which highly populated or employed sub-areas are located closer to one another based on the Moran's *I* coefficients ranging from -1 (distributed in a sprawling pattern) to 0 (scattered in a polycentric pattern) or +1 (clustered in a monocentric pattern), as stated in Section 3.2.2.4.5. Table 4-1 shows that the mean values for population or employment centralization index across all metropolitan areas are 0.078 for population in 2000 (*CentP00*), 0.075 for population in 2006 (*CentP06*), and 0.002 for employment in 2000 (CentE00). The Moran's I scores close to 0 suggest that highly populated or employed sub-areas are randomly scattered in a polycentric pattern, not in a sprawling pattern nor geographically clustered in a monocentric pattern.

4.1.1.5.1 Centralization Index in Cuyahoga County, Ohio

We take an example of the extent to which highly populated or employed sub-areas are clustered closer to one another, seeking spatial patterns of net population or employment density in Cuyahoga county in Ohio in 2000. The concentration index is specified by the Moran's *I* coefficients and the Local Indicator of Spatial Association (LISA) cluster map measured by the degree of spatial autocorrelation based only on the local neighborhood (Anselin, 1995).

Figure 4-1 displays spatial distribution of net employment density at the census tract level in Cuyahoga County in Ohio, in 2000. As identified in Section 3.2.2.4.2 (see Table 3-4), highly employed sub-areas in Figure 4-1 are those sub-areas with greater than a job-resident ratio of 1.0 and above the high (> 75 percentile) employment density, or with more than 2,300 workers per square mile. As in Figure 4-1 below, the spatial distribution of net employment density in Cuyahoga County in Ohio shows spatial clustering among highly employed census tracts. That is, the highly employed sub-areas (in dark blue) are located closer to highly employed sub-areas (in dark blue), particularly in the central business districts (CBDs) or in some higher employment sub-areas (in dark blue) in the suburbs, while the lowly employed sub-areas (in light or medium blue) are located closer to lowly employed sub-areas, particularly on the urban fringe.

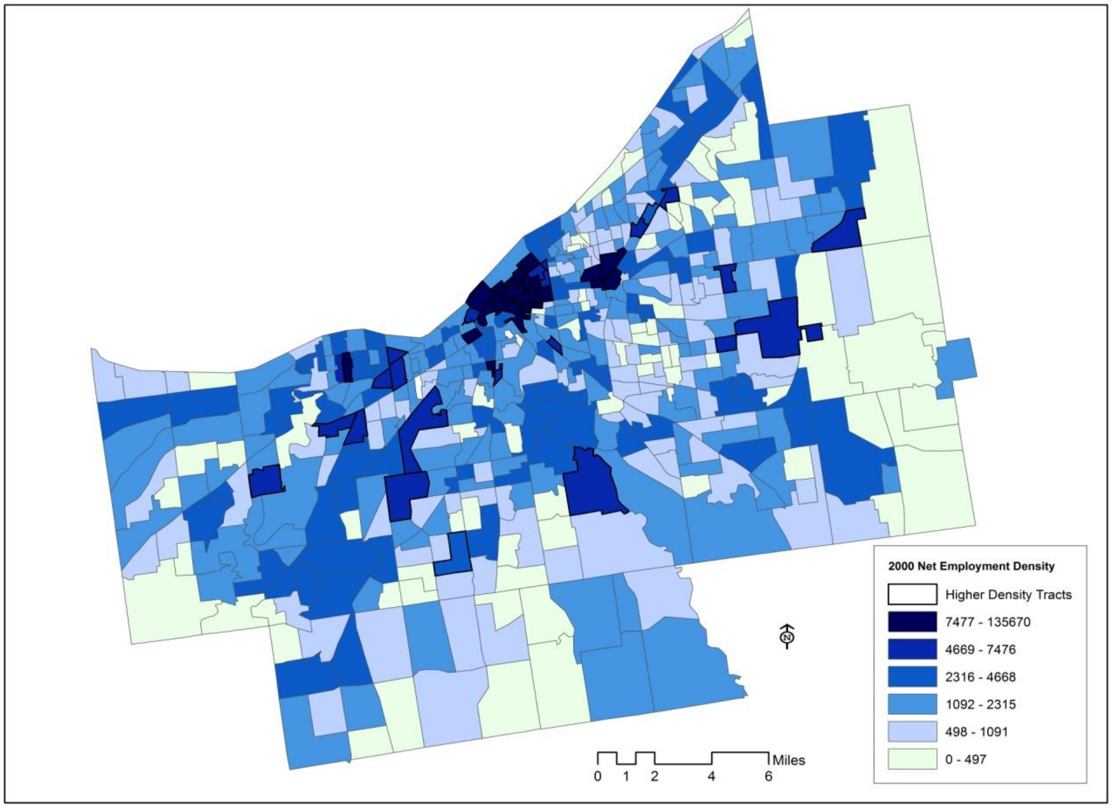


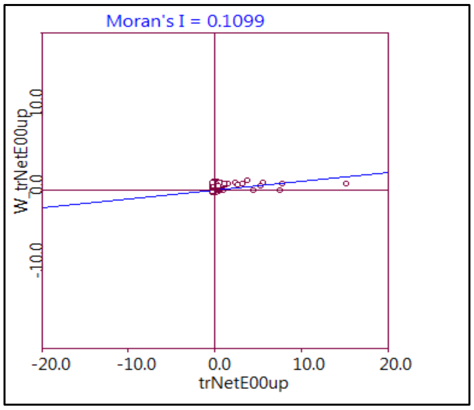
Figure 4-1 Net Employment Density Distribution in Cuyahoga County, OH, 2000

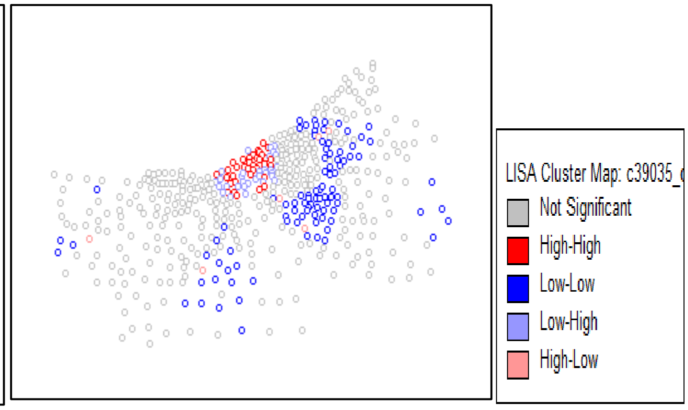
Figures 4-2, 4-3, and 4-4 below represent a level of centralization in Cuyahoga County in Ohio, using Moran's I coefficients. A-1 in Figure 4-2, B-1 in Figure 4-3, and C-1 in Figure 4-4 show the slope of the regression line (or Moran *I* coefficient) for net employment density in 2000 (0.1099), net population density in 2000 (0.3193), and net population density in 2006 (0.2899), respectively, using the Euclidean distance-band weights defined by the distance between the points (or census-tract polygon centroids). The Moran's *I* coefficients for the centralization index of net population or employment density in Cuyahoga County in Ohio at a 0.001 significance level represent the presence of spatial clustering or association across geographically neighboring sub-areas in Cuyahoga County, Ohio. For example, higher values of Moran's *I* for net population density (0.3193 in 2000 and 0.2899 in 2006), as in B-1 in Figure 4-3 and C-1 in Figure 4-4, point out that there is stronger positive clustering among neighboring sub-areas, meaning that highly populated sub-areas are located closer to one another.

In addition to the Moran scatter plots, as in A-1 in Figure 4-2, B-1 in Figure 4-3, and C-1 in Figure 4-4, the Local Indicator of Spatial Association (LISA) cluster maps below display the significant locations for spatial autocorrelation, color coded in the legend into four categories. Legend categories indicate the levels within the area in question, and the levels of surrounding areas. For example, high-high indicates an area with high concentration, surrounded by areas with high concentration; surrounding areas may be high-low, indicating high internal levels but low levels in surrounding areas., as in A-2 in Figure 4-2, B-2 in Figure 4-3, and C-2 in Figure 4-4.

The LISA cluster map for level of centralization of net employment density in 2000 in Cuyahoga County in Ohio, as in A-2 in Figure 4-2, shows the spatial clusters in the high-high (in downtown areas, see areas in dark red in Figure above) and low-low (in eastern suburban areas) locations, using 499 permutations and a 0.001 significance level.

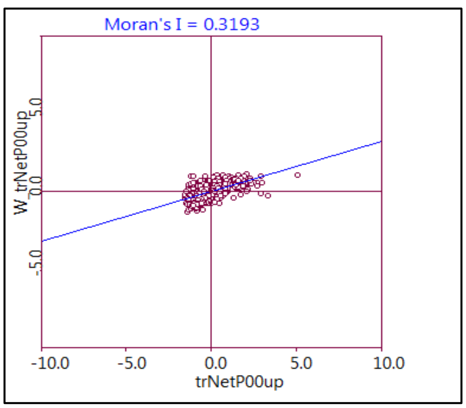
That is, there is positive local spatial autocorrelation that the highly employed sub-area is located closer to highly employed sub-areas (or the lowly employed sub-area closer to lowly employed ones). On the other hand, the LISA cluster map, as in A-2 in Figure 4-2, does not illustrate signicant high-low and low-high locations in that county which reflects negative local spatial autocorrelation, even if such locations are scattered in that county.

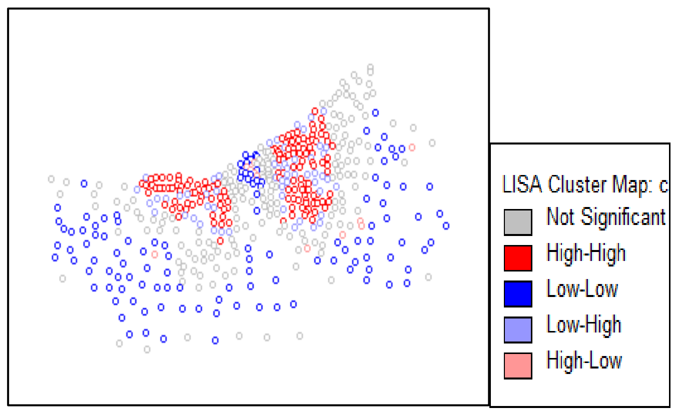




A-1 Moran Scatter Plot A-2 LISA Cluster Map

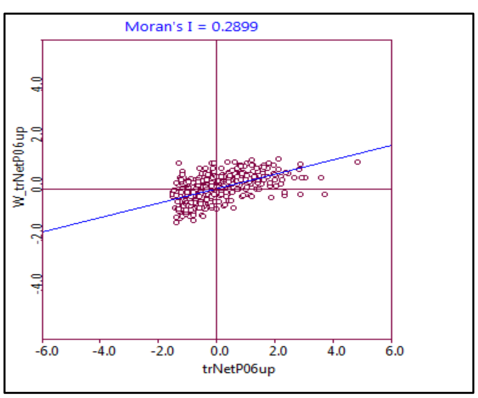
Figure 4-2 Level of Clustering of 2000 Net Employment Density in Cuyohaga County, OH

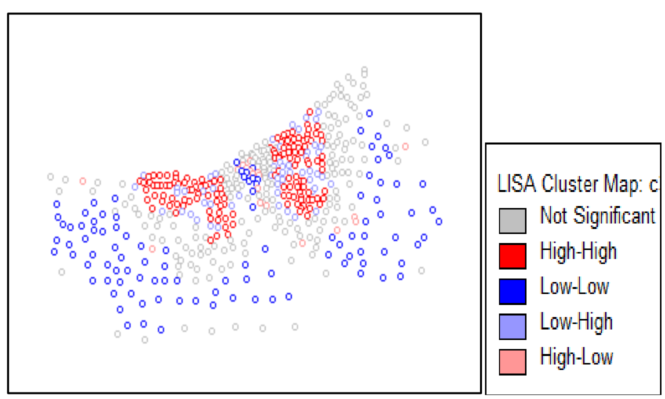




B-1 Moran Scatter Plot B-2 LISA Cluster Map

Figure 4-3 Level of Clustering of 2000 Net Population Density in Cuyohaga County, OH





C-1 Moran Scatter Plot C-2 LISA Cluster Map

Figure 4-4 Level of Clustering of 2006 Net Populaton Density in Cuyohaga County, OH

4.1.2 Descriptive Statistics for Major Control Variables

Table 4-2 below provides descriptive statistics for major confounding variables including agglomeration economies in industrial sectors, governmental structures, environmental policy, socio-demographic features (i.e., racial composition, income level and human capital), travel behavior, and regional amenities, as defined in Section 3.2.2, which can contribute to changes in average air quality index (AQI) and in metropolitan spatial structure (MSS) at the county level for the years 1990, 2000, and 2006. All the 610 counties have complete data for each variable.

We can consider the major confounding variables to be approximately normally distributed, because the variables have their skewness values either between -1.0 and 1.0 or between -2.0 and 2.0, but a few variables have skewness values above -2.0 (i.e., travel behavior (*PctDriA*)) or 2.0 (i.e., agglomeration economies in R&D industry (*M_rd*) and environmental industry (*M_env*)).

4.1.2.1 Level of Specialization in Industrial Sectors

Level of specialization in industrial sectors are identified by location quotients (LQ) reflecting how concentrated an industry is in a given metropolitan area relative to the nation, as defined in Section 3.2.2.2. As in Table 4-2, the mean value of the location quotient (LQ) for level of specialization in manufacturing industry has increased by 4.29 percent from 1990 to 2006, or from an average LQ of 1.07 (*m_mnf90*) to an average LQ of 1.12 (*m_mnf06*), which reflects that the relatively high employment concentration in manufacturing industry above LQs of 1.0 across metropolitan areas in U.S. has increased during this period. For service industries, the mean value of the location quotient (LQ) has only increased by 0.16 percent between 1990 and 2006, or from 1.013 (*m_ser90*) to 1.014 (*m_ser06*), which reflects that the relatively high employment concentration in services industry above LQs of 1.0 across metropolitan areas in U.S. has a little increased during this period.

The mean value of the location quotient (LQ) in research & development (R&D) industry has increased by 3.0-percent from 1990 to 2006, or from an average LQ of 0.71 (*m_rd90*) to an average LQ of 0.73 (*m_rd06*), which reflects that the relative employment concentration in R&D industry across metropolitan areas in U.S. has consistently increased during this period even if LQs in this industry (0.71 in 1990 and 0.73 in 2006) are well below 1.0. The mean value of the location quotient (LQ) in environmental industry has increased by 0.98-percent from 1990 to 2006, or from an average LQ of 0.94 (*m_env90*) to an average LQ of 0.95 (*m_env06*), which reflects that the relative employment concentration in R&D industry across metropolitan areas in U.S. has increased during this period. Both research and development (R&D) industry and environmental industry which not yet concentrated in the areas are becoming more concentrated during this period, leading to more potential to contribute to the regional growth.

4.1.2.2 Governmental Structure

Governmental structure at the county level is identified by the degree of fragmented local governments based on the governments 'vote-with-their-feet' principle (Tiebout, 1956), as stated in Section 3.2.2.3. Table 4-2 provides that both the average number of general-purpose local governments and of special-purpose governments across metropolitan areas have increased from 1992 to 2002, which points out that the local governmental structure in the county has been more fragmented, leading to more development in the suburbs or on the urban fringe. During this period the average number of local general-purpose governments has increased by 5.1-percent from 2.60 in natural logarithms (*LnGenGt92*) to 2.73 in natural logarithms (*LnGenGt02*) and the average number of special-purpose governments in the county has increased by 3.62-percent from 2.20 in natural logarithms (*LnSpeGt92*) to 2.28 in natural logarithms (*LnSpeGt02*).

4.1.2.3 Socio-demographic Features and Travel Behavior

As reviewed in Section 2.3.4.4, racial composition, income level, education, and commuters' preferences to travel modes can have an important role in shaping metropolitan structure. Table 4-2 provides that the proportion of Black or Hispanic residents in the county has sharply increased by 48.3-percent from 1990 to 2006, or from 15.08-percent (*PctHisB90*) to 22.37-percent (*PctHisB06*), which reflects that the number of minority communities in the county has grown to have a great impact on a shape of metropolitan structure. During this period, the proportion of persons with bachelors or higher degrees in the county has sharply increased by 36.9-percent from 19.07-percent (*PctBA90*) to 26.1-percent (*PctBA06*), which reflects that high-skilled or high-educated workers to affect economic growth and environmental policies have greatly grown in the county.

As in Table 4-2, the median household income level has decreased by 1.65-percent from 1990 to 2000, or from 11.0 in natural logarithms (*LnMHHI90*) to 10.81 in natural logarithms (*LnMHHI00*), but increased by 0.32-percent from 2000 to 2006, or from 10.81 in natural logarithms (*LnMHHI00*) to 10.85 in natural logarithms (*LnMHHI06*). This reflects that high-income residents tend to migrate to the suburbs, leading to excessive suburban development according to the white flight hypothesis, and to spend a higher proportion of their time as commute time, as evidenced in Section 4.1.1.5 accessibility index (indicating a 12.3-percent increase in average commute time (*AveCom*)). Furthermore, the average proportion of drive-alone commuters across metropolitan areas has consistently increased by 4.75-percent from 1990 to 2006, or from 78.8-percent (*PctDriA90*) to 82.6-percent (*PctDriA06*). A gradual increase in median household income, commute time, and car-dependent commuters during this period can bring out a consistent change in metropolitan shape.

Table 4-2 Descriptive Statistics for Major Control Variables

Variables	Description	N	Minimum	Maximum	Mean	Std. Deviation
Industrial Specialization	M_mnf90	610	0.06	3.88	1.07	0.61
	M_mnf00	610	0.04	4.64	1.09	0.65
	M_mnf06	610	0.03	5.55	1.12	0.70
	M_ser90	610	0.16	1.84	1.01	0.23
	M_ser00	610	0.23	1.95	1.03	0.22
	M_ser06	610	0.31	2.03	1.01	0.21
	M_rd90	610	0.00	20.32	0.71	1.55
	M_rd00	610	0.00	13.72	0.73	1.39
	M_rd06	610	0.00	13.60	0.73	1.35
	M_env90	610	0.00	30.35	0.94	1.43
	M_env00	610	0.00	17.59	0.93	1.19
	M_env06	610	0.06	15.19	0.95	1.18
Government /Policy	LnGenGt92	610	-0.78	7.37	2.60	1.22
	LnGenGt02	610	-0.69	7.34	2.73	1.22
	LnSpeGt92	610	-0.72	6.61	2.20	1.17
	LnSpeGt02	610	-0.87	6.64	2.28	1.15
	EnvPolicy	610	2.00	30.00	13.53	7.09
Socio -demographic	PctHisB90	610	0.30	93.99	15.08	14.93
	PctHisB00	610	0.89	94.35	19.70	16.54
	PctHisB06	610	1.41	94.87	22.37	17.07
	LnMHHI90	610	10.32	11.64	11.00	0.22
	LnMHHI00	610	10.28	11.48	10.81	0.22
	LnMHHI06	610	10.31	11.63	10.85	0.23
	PctBA90	610	4.45	52.30	19.07	7.96
	PctBA00	610	4.92	60.22	23.18	9.26
	pctBA06	610	6.70	68.80	26.10	9.80
Intermediate	PctDriA90	610	8.29	89.43	78.81	6.67
	PctDriA00	610	8.06	92.15	81.68	6.66
	PctDriA06	610	7.50	90.55	82.56	6.67
	LnClimate	610	0.00	4.61	3.60	0.89
	LnUnde92	610	15.80	24.60	21.01	1.07
Initial Effect	LnTotPop90	610	8.75	16.00	11.93	1.11
	LnTotEmp90	610	-0.03	8.35	4.14	1.30

4.2 Ordinary Least Squares (OLS) Regression and Spatial Regression Estimation

Ordinary least squares (OLS) multiple regressions are used to estimate the relationships between changes in air quality index values and changes in land use activities, population, employment, and governmental structure. OLS multiple regression models are used to identify the predictive parameters that have a significant impact on air quality improvements in 610 counties in the metropolitan areas for 1990, 2000, and 2006. The predictive regression coefficients (β) of air quality improvements for the time period are tested with *GeoDa* and SPSS statistical packages.

The spatial regression models for 1990, 2000, and 2006 are used to identify the presence of spatially correlated observations, considered spatial dependence or spatial autocorrelation between observations. The analytical specifications associated with spatial dependence will be formed through two spatial regression models: spatial lag estimation with a spatially lagged dependent variable (air quality index) and spatial error estimation with spatial autoregressive effects of independent variables (i.e., metropolitan spatial structure) to affect changes in air quality.

In order to form a good fit model, we run model diagnostic tests for OLS regression estimation which consist of multicollinearity, normality, heteroscedasticity, and spatial autocorrelation (spatial lag model or spatial error model) to conduct Lagrange Multiplier (LM) test statistics.

4.2.1 The 1990 OLS Estimation and Spatial Regression Results

We first estimate the OLS regression coefficients with /or without metropolitan spatial structure (MSS) and conduct the diagnostics tests for the OLS estimation in terms of three measures: multicollinearity, normality, and heteroscedasticity. Simultaneously, we detect for spatial autocorrelation or dependence based on the Lagrange Multiplier (LM) test statistics (i.e., LM-Lag or LM-Error). These spatial dependence diagnostics are useful in choosing an alternative spatial regression model specification, either spatial lag or spatial error model. We estimate all the regression coefficients with spatial autoregressive coefficients (ρ and λ) based on the maximum likelihood estimation to the OLS regression model. Finally, we compare not only the alternative spatial regression results to the OLS regression estimation, but also the results between the spatial lag and error model.

4.2.1.1 The 1990 OLS Regression Estimation with Metropolitan Spatial Structure (MSS)

Table 4-3 below shows the ordinary least squares (OLS) multiple regression estimates for independent variables, including metropolitan spatial structures (MSS). Table 4-3 displays the 1990 regression estimation for Models 1 and 2, with no consideration for spatial dependence.

Table 4-3 illustrates the summary characteristics for Model 1 showing all of the regression standardized coefficients for the dependent variable (average air quality index in 1990), not including metropolitan structures (MSS) and regression diagnostics. Table 4-3 shows the number of observations (477 counties), the number of variables including the constant term (17), and the degrees of freedom (460) for Model 1.

Table 4-3 shows that the R-squared value and the adjusted R-squared value for Model 1 are about 0.231 and 0.204, respectively. This means that 23.1% (or 20.4%) of the variance in changes in air quality are predicted from the combination of agglomeration effect, governmental structures, socio-demographic features, travel behavior, and regional amenities. Table 4-3 also shows that the analysis of variance (ANOVA) F-statistic with 17 and 460 degrees of freedom for all of the regression coefficients is 8.61at less than the 1-percent significance level, indicating that the combination of all of the independent variables significantly predicts changes in air quality.

Table 4-3 displays a number of interesting patterns for Model 1. The negative regression coefficients of special-purpose local governments (*lnSpeGt92*, $\beta = -0.134$) and college graduates or higher (*PctBA90*, $\beta = -0.324$) indicate that these variables are statistically significant for improved air quality at the 5-percent and 1-percent significance levels, respectively. The positive regression coefficients of general-purpose local governments (*lnGenGt92,* $\beta = 0.136$) and total population in 1990 (*lnTotPop90,* $\beta = 0.520$) indicate that these variables are statistically significant predictors for worsened air quality at the 10-percent and 1-percent significance level, respectively.

Table 4-3 also displays a number of interesting patterns for Model 2[34] with metropolitan spatial structure (MSS). The negative regression coefficients of special-purpose local governments (*lnSpeGt92*, $\beta = -0.141$) and college graduates or higher (*PctBA90,* $\beta = -0.252$) indicate that these variables are statistically significant as predictors for improved air quality at the 5-percent and 1-percent significance level, respectively. The positive coefficients of general-purpose local governments (*lnGenGt92,* $\beta = 0.163$), net population density per square mile (*lnnetP90*, $\beta = 0.175$) and weighted

[34] Model 2 results with the variable metropolitan spatial structure (MSS) are shown in parentheses.

average drive-alone commute time (*lnTotCom90*, $\beta = 0.438$) indicate that these variables are statistically significant as predictors for worsened air quality at the 5-percent or 1-percent significance level.

However, other confounding variables for Model 2 are together considered to obtain this result because the null hypothesis that all of the regression coefficients are simultaneously equal to zero is rejected. Table 4-3 displays that the ANOVA F-statistic with 19 and 458 degrees of freedom ($F = 8.20$) for Model 2 is statistically significant at less than the 1-percent level of significance (or the $p < 0.0000000$). This indicates that all of the independent variables significantly combine together to predict changes in air quality.

Table 4-3 OLS Estimation with MSS, 1990

N	477 (477)	F-statistic	*8.6146 (8.1986)*
# Variables	17 (19)	Prob (F-statistic)	*0.0000 (0.0000)*
DF	460 (458)	Log likelihood	-1974.87 (-1970.76)
R-squared	*0.231 (0.244)*	Akaike info criterion	*3983.73 (3979.52)*
Adjusted R-squared	0.204 (0.214)	Schwarz criterion	4054.58 (4058.7)

Component	Model 1 (2)	Standardized Coefficients (β)	Sig.
	CONSTANT	-57.442 (-48.622)	0.305 (0.440)
Specialization	m_mnf90	0.007 (0.007)	0.900 (0.899)
	m_ser90	-0.041 (-0.018)	0.405 (0.715)
	m_rd90	0.017 (0.010)	0.734 (0.843)
	m_env90	-0.062 (-0.043)	0.197 (0.378)
Government	*lnGenGt92*	0.136 (0.163)	*0.077 (0.030)*
	lnSpeGt92	-0.134 (-0.141)	*0.032 (0.023)*
	EnvPolicy	-0.086 (-0.022)	0.128 (0.702)
	SGMP	-0.073 (-0.081)	0.155 (0.108)
Socio -demographic	PctHisB90	-0.085 (-0.077)	0.138 (0.177)
	lnMHHI90	0.078 (-0.064)	0.228 (0.435)
	PctBA90	-0.324 (-0.252)	*0.000 (0.000)*
Intermediate	PctDriA90	-0.091 (-0.002)	0.124 (0.983)
	lnClimate	0.035 (0.010)	0.482 (0.844)
	LnUnde92	-0.073	0.192
	region_dum	-0.096 (-0.081)	0.104 (0.181)
Initial	*lnTotPop90*	0.520	*0.000*
MSS	lnNetP90	(0.175)	(0.002)
	LUMix92	(0.063)	(0.358)
	AveCom90	(0.041)	(0.479)
	lnTotCom90	(0.438)	(0.000)

Note: the values in parentheses are the OLS estimation with metropolitan spatial structures (MSS) for Model 2.

4.2.1.2 Diagnostic Tests of the1990 OLS Estimation with MSS

As stated in Section 3.4.3.1, we test three measures for the diagnostics of OLS regression estimation: multicollinearity, normality, and heteroscedasticity. The first detection of multicollinearity is about the impact between one or more highly correlated independent variables using the tolerance (TOL) and the variance inflation factor (VIF).

Due to high multicollinearity among independent variables in Models 1 and 2, we remove variables with high VIF: total employment (*LnTotEmp90*), net employment density (*LnNetE*90), and total developed land (*pct_urb92*). Their VIFs are higher than VIF of 4.06 or lower than TOL of 0.247 between these variables.

To reduce the impact of multicollinearity, we use the cutoff threshold for the tolerance (TOL) of below 0.247 (above a VIF of 4.06). To remedy high multicollinearity among independent variables, we transform the following variables using the natural logarithm: number of local governments (*lnGenGt92* and *lnSpeGt92*), net population density (*lnNetP90*), median household income (*lnMHHI90*), undeveloped land size (*lnUnde92*), and weighted drive-alone commute time (*lnTotCom90*), and climate (*lnClimate*). The natural logarithm of aforementioned independent variables is approximately normally distributed between a skewness statistic of -2.0 and 2.0. The semi-log regression estimation using the natural logarithm of aforementioned independent variables, called a lin-log model (Gujarati, 2003, pp.178-183), is a better fit for the estimate OLS regression coefficients.

The value of multicollinearity condition number (k) between 100 and 1000 is also used to detect multicollinearity. Table 4-4 shows that the multicollinearity condition number (k) between 100 and 1000 and are 403.725 for Model 1 and 503.027 for Model 2. We can consider that there is moderately high multicollinearity of the 1990 OLS regression models based on a rule of thumb.

The Jarque-Bera (JB) test to detect normality on the OLS regression residuals is used. Table 4-4 displays that the value of the JB statistic with 2 degrees of freedom is 230.382 for Model 1 and 221.771 for Model 2 at the 1-percent significance level. We can reject the null hypothesis that the residuals are normally distributed, meaning that there is non-normality of the errors.

The three diagnostic tests to detect heteroscedasticity on the OLS regression residuals are used for Models 1 and 2. Table 4-4 shows that both the Breusch-Pagan (BP) test and Koenker-Bassett (KB) test on random coefficients for the Models 1 and 2 are statistically significant at the 1-prcent and 5-percent significance level, respectively.[35] We can reject the null hypothesis that there is an equal variance of errors, which indicates that the OLS regression errors are unequally spread.

[35] The White test on specification-robust test for heteroscedasticity in *GeoDa* reports N/A, because *GeoDa* is not able to correct for this (Anselin, 2005, p.195), as in Table 4-4.

In summary, Table 4-4 illustrates that the diagnostic tests of the 1990 OLS estimation with metropolitan spatial structure (MSS) detect non-normality and heteroscedasticity among the regression residuals.

Table 4-4 Diagnostic Tests of the1990 OLS Estimation with MSS

REGRESSION DIAGNOSTICS			
MULTICOLLINEARITY CONDITION NUMBER			403.725 (503.027)
TEST ON NORMALITY OF ERRORS			
TEST	DF	VALUE	Sig.
Jarque-Bera	2 (2)	230.382 (221.771)	0.0000 (0.0000)
DIAGNOSTICS FOR HETEROSKEDASTICITY			
RANDOM COEFFICIENTS			
TEST	DF	VALUE	Sig.
Breusch-Pagan test	16 (18)	58.385 (68.038)	0.0000 (0.0000)
Koenker-Bassett test	16 (18)	25.499 (29.966)	0.0615 (0.0378)
SPECIFICATION ROBUST TEST			
TEST	DF	VALUE	Sig.
White	152 (189)	N/A	N/A

Note: the values in parentheses are the OLS estimation with metropolitan spatial structures (MSS) for Model 2.

4.2.1.3 Diagnostic Tests for Spatial Dependence in the1990 OLS Estimation with MSS

Diagnostic tests for spatial dependence in the 1990 OLS estimation with metropolitan spatial structure (MSS) in Model 2 are computed for the row-standardized weights matrix (*cnty477_y90_point_w38.gwt*), as in Table 4-5.

Table 4-5 shows that both LM-Lag and LM-Error statistics are highly significant. Of the robust forms, both the Robust LM-Lag and the Robust LM-Error statistics are significant at a 0.0048 and a 0.0042 significance level, respectively. We can reject the null hypothesis that there is no spatial autocorrelation or dependence to the 1990 OLS regression residuals. This indicates that strong spatial effects are interrelated among its neighboring regions. We consider the alternative spatial regression estimation in terms of a spatial lag model and a spatial error model.

Table 4-5 Diagnostic Tests for Spatial Dependence in the1990 OLS Estimation with MSS

DIAGNOSTICS FOR SPATIAL DEPENDENCE			
FOR WEIGHT MATRIX : cnty477_y90_point_w38.gwt	(row-standardized weights)		
TEST	MI/DF	VALUE	PROB
Moran's I (error)	0.1498	N/A	N/A
Lagrange Multiplier (lag)	1	107.3161	0.0000000
Robust LM (lag)	1	7.9168	0.0048979
Lagrange Multiplier (error)	1	107.5768	0.0000000
Robust LM (error)	1	8.1774	0.0042416
Lagrange Multiplier (SARMA)	2	115.4935	0.0000000

4.2.1.4 The 1990 Maximum Likelihood Spatial Regression Estimation Results

Table 4-6 below provides the spatial regression estimation results when considering spatial dependence between dependent or independent variables for Models 3 and 4 against the 1990 OLS regression residuals, as seen in Section 3.4.2.

4.2.1.4.1 The 1990 Maximum Likelihood Spatial Lag Estimation

Table 4-6 displays that the maximum likelihood (ML) spatial lag estimation with all the OLS regression standardized coefficients for the dependent variable, air quality index (AQI) for Model 3. Table 4-6 shows the number of observations (477 counties), the number of variables including the constant term (19), and the degrees of freedom (458) for Model 3. Table 4-6 illustrates that the R-squared value for Model 3 is approximately 0.3403. [36] This points out that 34% of the variance in changes in air quality can be predicted from the combination of agglomeration effects, governmental structures, socio-demographic features, travel behavior, regional amenities, and a spatially lagged dependent variable (*W_A_mean90* (ρ)) for distance-based weights matrix (*cnty477_y90_point_w38.gwt*).

The three measures in Table 4-6 are used to select a good fitting spatial regression model. They are the log likelihood (-1943.56), the Akaike information criterion (3927.12), and the Schwarz information criterion (4010.47). To compare the values in the 1990 spatial lag model in the right column in Table 4-6 to those for the 1990 OLS estimation in Table 4-3, we identify an increase in the log likelihood from -1970.76 for OLS to -

[36] The R-squared value in the Model 3 spatial lag model, 0.3403, is not a real R-squared, but a *pseudo* R-squared. With the pseudo R-squared, we need to be cautious of a direct comparison of all the spatial regression coefficient estimates to the OLS regression results (Anselin, 2005, p. 207).

1943.56, a decrease in the AIC from 3979.52 for OLS to 3927.12, and a decrease in the SIC from 4058.7 for OLS to 4010.47. The lower the values of both the Akaike information criterion (AIC) and the Schwarz information criterion (SIC) are, the better the spatial lag model is fitted. The higher the log likelihood is, the better the spatial lag model is fitted.

Table 4-6 shows that the spatial autoregressive coefficient in the spatial lag model (*W_A_mean90*, $\rho = 0.592$) are statistically highly significant ($p < 0.0000001$). This indicates that there is spatial dependence of spatially lagged dependent variable to the 1990 OLS regression estimation.

To compare all the regression estimates between the spatial lag model in Table 4-6 and the OLS estimation in Table 4-3, the magnitude of all the regression coefficients is affected by the coefficient of the spatially lagged dependent variable (*W_A_mean90*, ρ = 0.592, $p < 0.0000001$). Some of the coefficients, such as college graduates or higher (*PctBA90*), special-purpose local governments (*lnSpeGt92*), net population density per square mile (*lnNetP90*), and weighted average drive-alone commute time (*lnTotCom90*), show a decrease in absolute value. Some of the coefficients, such as general-purpose local governments (*lnGenGt92*) and median household income (*lnMHHI90*), show an increase in absolute value. The significance of other regression coefficients is also affected by the coefficient of the spatially lagged dependent variable used to reflect the spatial effects on neighboring regions. As a result of neighboring region effects, the significance of a number of other regression coefficients is also changed. The significance of college graduates or higher (*PctBA90*) changes from $p < 0.000498$ to $p < 0.022$. The median household income (*lnMHHI90*) changes from $p < 0.435$ to $p < 0.065$, or from insignificant to significant, indicating that the median household income level is statistically significant as a predictor for changes in air quality at the 10-percent significance level.

Table 4-6 the 1990 Maximum Likelihood Spatial Regression Estimation Results

Spatial Weight	*cnty477_y90_point_w38.gwt*		
N	477	Log likelihood	-1943.56 (-1939.70)
# Variables	19 (18)	Akaike info criterion	*3927.12 (3917.4)*
DF	458 (459)	Schwarz criterion	4010.47 (3996.58)
R-squared	*0.3403(0.3593)*		

Component	Model 3 (4)	Standardized Coefficients	Sig.
	CONSTANT	15.519 (62.440)	0.788 (0.357)
Specialization	m_mnf90	0.733 (1.210)	0.645 (0.447)
	m_ser90	-1.970 (-3.072)	0.584 (0.402)
	m_rd90	-0.389 (-0.726)	0.398 (0.106)
	m_env90	-0.036 (0.167)	0.942 (0.720)
Government	*lnGenGt90*	2.342 (2.710)	*0.016 (0.008)*
	lnSpeGt90	-1.688 (-1.625)	*0.046 (0.077)*
	EnvPolicy	-0.017 (0.093)	0.889 (0.533)
	SGMP	-2.157 (-0.726)	0.198 (0.715)
Socio-demographic	PctHisB90	-0.101 (-0.088)	0.088 (0.205)
	lnMHHI90	-10.746 (-12.256)	0.065 (0.068)
	PctBA90	-0.337 (-0.270)	*0.022 (0.080)*
Intermediate	PctDriA90	0.006 (0.016)	0.967 (0.925)
	lnClimate	-0.322 (-0.004)	0.723 (0.998)
	LnUnde92		
	region_dum	-2.151 (-7.454)	0.4715 (0.154)
Initial	*lnTotPop90*		
MSS	lnNetP90	5.059 (4.034)	0.037 (0.131)
	LUMix92	6.115 (4.956)	0.196 (0.333)
	AveCom90	0.270 (0.330)	0.255 (0.204)
	lnTotCom	5.874 (6.102)	0.000 (0.000)
Spatial Lag	W_A_mean90 (ρ)	0.592	0.0000000
Spatial Error	LAMBDA (λ)	(0.704)	(0.0000000)

Note: the values in parentheses are the spatial error regression estimation with metropolitan spatial structures (MSS) for Model 4.

Table 4-7 displays the diagnostic tests for the 1990 maximum likelihood spatial lag estimation for Model 3 for the row-standardized weights matrix (*cnty477_y90_point_w38.gwt*). The Breusch-Pagan statistic on random coefficients in the error terms for Model 3 to detect heteroscedasticity is 55.509 at a 0.0001 significance level. We can reject the null hypothesis that there is an equal variance of errors (homoscedasticity) indicating that the spatial lag errors are unequally spread. The Likelihood Ratio Test is used to compare the OLS regression estimation to the alternative spatial lag model. Table 4-7 also shows that the Likelihood Ratio Test value of 54.3997 is statistically significant ($p < 0.0000001$). This indicates that the spatial autoregressive coefficient (ρ) of spatial lag model for Model 3 is strongly significant.

Table 4-7 Diagnostic Tests for the 1990 Spatial Regression Models

REGRESSION DIAGNOSTICS			
DIAGNOSTICS FOR HETEROSKEDASTICITY			
RANDOM COEFFICIENTS			
TEST	DF	VALUE	PROB
Breusch-Pagan test	18 (18)	55.5090 (50.2456)	0.000010 (0.000069)
DIAGNOSTICS FOR SPATIAL LAG MODEL			
SPATIAL *LAG* DEPENDENCE FOR WEIGHT MATRIX : *cnty477_y90_point_w38.gwt*			
TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	54.3997	0.0000000
DIAGNOSTICS FOR SPATIAL ERROR MODEL			
SPATIAL *ERROR* DEPENDENCE FOR WEIGHT MATRIX : *cnty477_y90_point_w38.gwt*			
TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	62.1214	0.0000000

Note: the values in parentheses are the spatial error regression estimation with metropolitan spatial structures (MSS) for Model 4.

4.2.1.4.2 The 1990 Maximum Likelihood Spatial Error Estimation

Table 4-6 illustrates the maximum likelihood (ML) spatial error estimation for Model 4 with all the OLS regression standardized coefficients for the dependent variable, air quality index (AQI). We can assume that the errors for neighboring observations of a spatial regression model are spatially correlated, as in Section 3.4.2.2. Table 4-6 shows the number of observations (477 counties), the number of variables including the constant term (18), and the degrees of freedom (459) for Model 4. Table 4-6 shows that the R-squared value for Model 4 is approximately 0.3593.[37] This indicates that about 36% of the variance in changes in air quality can be predicted from the combination of specialization in industries, governmental structures, socio-demographic features, travel behavior, regional amenities, and a spatially weighted average of the errors (*LAMBDA* (λ)) for distance-based weights matrix among neighboring regions (*cnty477_y90_point_w38.gwt*).

The three measures in Table 4-6 are used to select a good fitting spatial regression model. They are the log likelihood (-1939.70), the Akaike information criterion (3917.4), and the Schwarz information criterion (3996.58). To compare the values in the 1990

[37] The R-squared value in the Model 4 spatial error model, 0.3593, is not a real R-squared, but a *pseudo* R-squared. With the pseudo R-squared, we need to be cautious of a direct comparison of all the spatial regression coefficient estimates to the OLS regression results (Anselin, 2005, p. 207).

spatial error model in Table 4-6 to those for the 1990 OLS estimation in Table 4-3 (or the 1990 spatial lag model in Table 4-6), we identified an increase in the log likelihood from -1970.76 for OLS (or -1943.56 for spatial lag) to -1939.70, a decrease in the AIC from 3979.52 for OLS (or 3927.12 for spatial lag) to 3917.4, and a decrease in the SIC from 4058.7 for OLS (or 4010.47 for spatial lag) to 3996.58. The lower the values of both the Akaike information criterion (AIC) and the Schwarz information criterion (SIC) are, the better the spatial error model is fitted. The higher the log likelihood is, the better the spatial error model is fitted.

As in Table 4-6 above, the spatial autoregressive coefficient in the spatial error model (*LAMBDA*, $\lambda = 0.704$) are statistically highly significant ($p < 0.0000001$). This indicates that there is spatial dependence of spatially weighted average of error terms between neighboring regions to the 1990 OLS regression estimation.

To compare all the regression estimates between the spatial error model in Table 4-6 and the OLS estimation in Table 4-3, the magnitude of all the regression coefficients is affected by the coefficient of the spatially weighted average of errors (*LAMBDA*, $\lambda = 0.704$, $p < 0.0000001$). The results between the spatial error model and the OLS estimation are similar to those between the spatial lag model and the OLS estimation in terms of the sign of the regression coefficients, but different in terms of the magnitude and significance of the coefficients. For example, some of the coefficients, such as college graduates or higher (*PctBA90*), special-purpose local governments (*lnSpeGt92*), net population density per square mile (*lnNetP90*), and weighted average drive-alone commute time (*lnTotCom90*), show a decrease in absolute value. Some of the coefficients, such as general-purpose local governments (*lnGenGt92*) and median household income (*lnMHHI90*), show an increase in absolute value. The significance of college graduates or higher (*PctBA90*) changes from $p < 0.000498$ to $p < 0.0796$. The median household income (*lnMHHI90*) changes from $p < 0.4347$ to $p < 0.067$, or from insignificant to significant, indicating that the median household income level is statistically significant to predict changes in air quality at the 10-percent significance level.

Table 4-7 also shows the diagnostic tests for the 1990 maximum likelihood spatial error estimation in Model 4 for the row-standardized weights matrix (*cnty477_y90_point_w38.gwt*). The Breusch-Pagan statistic on random coefficients in the error terms for Model 4 to detect heteroscedasticity is 50.246 at a 0.0001 significance level. We can reject the null hypothesis that there is an equal variance of errors for neighboring observations (homoscedasticity). Table 4-7 also displays that the Likelihood Ratio Test value of 62.1214 are statistically significant ($p < 0.0000001$). This indicates that the spatial autoregressive coefficient (λ) of spatial error model for Model 4 is strongly significant.

To compare the results between the spatial lag and error model, as in Table 4-6, the spatial error model is a better fit than the spatial lag model because the spatial error model can interpret a higher value of the Log Likelihood and lower values of the AIC and SIC.[38] Additionally, the estimation results of the spatial error model are similar to those of the spatial lag model in terms of the sign of the regression coefficients, but different in terms of the magnitude and significance of the coefficients. For example, the magnitude of some of the coefficients, such as general-purpose local governments (*lnGenGt92*), medium household income (*lnMHHI90*), and weighted average drive-alone commute time (*lnTotCom00*), show an increase in absolute value. The significance of most of the other regression coefficients also changed. The significance of general-purpose local governments (*lnGenGt92*) changes from $p < 0.0155$ to $p < 0.0085$. The significance of the net population density (*lnNetP90*) changes from $p < 0.037$ to $p < 0.13$,1 or from significant to insignificant.

[38] Along with an increase in the log likelihood, a decrease in the AIC, and a decrease in the SIC, the spatial regression model specifications using an error and lag model are well fitted to the OLS regression estimation.

4.2.2 The 2000 OLS Estimation and Spatial Regression Results

As conducted in the 1990 OLS estimation and spatial regression results, we estimate the 2000 OLS regression coefficients with metropolitan spatial structure (MSS) and conduct the diagnostics tests for the 2000 OLS estimation in terms of three measures: multicollinearity, normality, and heteroscedasticity. Continuously throughout the process, we detect the diagnostics for spatial autocorrelation or dependence based on Lagrange Multiplier (LM) test statistics (i.e., LM-Lag or LM-Error). We estimate all the regression coefficients with spatial autoregressive coefficients (ρ and λ) relating to the spatial dependence based on the maximum likelihood estimation to the 2000 OLS regression model. Finally, we compare not only the alternative spatial regression results to the OLS regression estimation, but also the results between the spatial lag and error model.

4.2.2.1 The 2000 OLS Regression Estimation with Metropolitan Spatial Structure (MSS)

Table 4-8 below shows that the summary characteristics of the Model 5 shows all regression standardized coefficients for the air quality index (AQI) in 2000, not including metropolitan spatial structures (MSS) and regression diagnostics. Table 4-8 displays the number of observations (610 counties), the number of variables including the constant term (17), and the degrees of freedom (593) for Model 5.

Table 4-8 shows that R-squared value and adjusted R-squared value are about 0.237 and 0.217, respectively. This indicates that 23.7% (or 21.7%) of the variance in changes in air quality is predicted from the combination of agglomeration effects, governmental structures, socio-demographic features, travel behavior, and regional amenities. The analysis of variance (ANOVA) F-statistic with 17 and 593 degrees of freedom for all of the regression coefficients is 11.52 at less than the 1-percent significance level, indicating that the combination of all of the independent variables significantly predicts changes in air quality.

Table 4-8 illustrates a number of interesting patterns for Model 5. The negative regression coefficient of college graduates or higher (*PctBA00*, $\beta = -0.170$), undeveloped land (*LnUnde92*, $\beta = -121$), a state's pro-environmental policy (*EnvPolicy*, $\beta = -0.178$), agglomerative effects in the service industry (*M_ser00*, $\beta = -0.142$) and agglomerative effects in the environmental industry (*M_env00*, $\beta = -0.087$) indicate that these variables are statistically significant predictors of improved air quality at the 5-percent or 1-percent significance level. The positive regression coefficient of total population in 1990 (*lnTotPop90*, $\beta = 0.508$) indicates that this variable is a statistically significant predictor of worsened air quality at the 1-percent significance level.

Table 4-8 also shows a number of interesting signs for Model 6[39] with MSS. The negative regression coefficients of a state's pro-environmental policy (*EnvPolicy*, β = -0.128), level of specialization in the service industry (*M_ser00*, β = -0.141), level of specialization in the environmental industry (*M_env00*, β = -0.087), medium household income (*LnMHHI00*, β = -0.268`), and regional locations except for the Pacific division (*region_dum*, β = -0.104), mixed land use (*LUMix01*, β = -0.239), and proportion of high employment density sub-areas (*ConEmp00*, β = -0.124) indicate that these variables are statistically significant as predictors for improved air quality at the 5-percent or 1-percent significance level, except for the variable proportion of high employment density sub-areas (*ConEmp00*) at the 10-percent significance level. The positive coefficients of proportion of drive-alone commuters (*PctDriA00*, β = 0.199), proportion of developed open space (P*ct_open01*, β = 0.266), net population density per square mile (*lnNetP00*, β = 0.262), proportion of high population density sub-areas (C*onPop00*, β = 0.168), average commute time (*AveCom00*, β = 0.135), and weighted average drive-alone commute time (*lnTotCom00*, β = 0.419) indicate that these variables are statistically significant as predictors for worsened air quality at the 5-percent or 1-percent significance level.

[39]Model 6 results with the variable metropolitan spatial structure (MSS) are shown in parentheses.

However, other confounding variables for Model 6 are together considered to obtain this result because the null hypothesis that all regression coefficients are simultaneously equal to zero is rejected. Table 4-8 shows that the ANOVA F-statistic with 24 and 586 degrees of freedom (=10.09) for Model 6 is statistically significant at less than the 1-percent significance level (or $p < 0.0000001$). This indicates that all of the independent variables significantly combine together to predict changes in air quality.

Table 4-8 OLS Estimation with MSS, 2000

N	610 (610)	F-statistic	*11.5242 (10.0939)*
# Variables	17 (24)	Prob(F-statistic)	*0.0000 (0.0000)*
DF	593 (586)	Log likelihood	-2345.85 (-2326.64)
R-squared	*0.237 (0.284)*	Akaike info criterion	*4725.7 (4701.27)*
Adjusted R-squared	0.217 (0.256)	Schwarz criterion	4800.73 (4807.2)

Component	Model 5 (6)	Standardized Coefficients (β)	Sig.
	CONSTANT	-6.386 (79.709)	0.860 (0.045)
Specialization	m_mnf00	0.023 (0.065)	0.614 (0.168)
	m_ser00	-0.142 (-0.141)	0.002 (0.002)
	m_rd00	-0.030 (-0.031)	0.466 (0.437)
	m_env00	-0.087 (-0.087)	0.028 (0.025)
Government	*lnGenGt02*	-0.002 (-0.002)	*0.981 (0.978)*
	lnSpeGt02	-0.033 (-0.049)	*0.559 (0.372)*
	EnvPolicy	-0.178 (-0.128)	0.000 (0.011)
	SGMP	0.012 (-0.015)	0.793 (0.731)
Socio -demographic	PctHisB00	0.062 (0.078)	0.132 (0.131)
	lnMHHI00	0.027 (-0.268)	0.628 (0.000)
	PctBA00	-0.170 (-0.018)	*0.003 (0.798)*
Intermediate	PctDriA00	0.080 (0.199)	0.125 (0.006)
	lnClimate	0.058 (-0.055)	0.179 (0.224)
	LnUnde92	-0.121	0.008
	region_dum	-0.057 (-0.104)	0.250 (0.036)
Initial	*lnTotPop90*	0.508	*0.000*
MSS	lnNetP00	(0.262)	(0.001)
	PctOpen01	(0.266)	(0.000)
	LUMix01	(-0.239)	(0.009)
	ConPop00	(0.168)	(0.021)
	ConEmp00	(-0.124)	(0.099)
	AveCom00	(0.135)	(0.024)
	lnTotCom00	(0.419)	(0.000)
	CenPop00	(-0.041)	(0.411)
	CenEmp00	(0.047)	(0.353)

Note: the values in parentheses are the OLS estimation with metropolitan spatial structures (MSS) for Model 6.

4.2.2.2 Diagnostic Tests of the 2000 OLS Estimation with MSS

Because of high multicollinearity among independent variables in Models 5 and 6, we remove variables with high VIF: total employment (*LnTotEmp00*), net employment density (*LnNetE0*0), and total developed land (*pct_urb01*). Their VIFs are higher than VIF of 6.80 or lower than TOL of 0.147 between these variables. To reduce the impact of multicollinearity, we use the cutoff threshold for the tolerance (TOL) of below 0.147 (above a VIF of 6.80). We also use a natural log transformation for the following variables: the number of local governments (*lnGenGt02* and *lnSpeGt02*), net population density (*lnNetP00*), median household income (*lnMHHI00*), undeveloped land size (*LnUnde92*), and weighted average drive-alone commute (*lnTotCom00*). The natural logarithm of aforementioned independent variables is approximately normally distributed between a skewness statistic of -2.0 and 2.0 based on the semilog regression estimation. Table 4-9 shows that the value of multicollinearity condition number (*k)* is 396.593 for Model 5 and 557.081 for Model 6, respectively. We can consider that there is moderately high multicollinearity of the OLS regression models, based on a rule of thumb.

The Jarque-Bera (JB) test to detect normality on the 2000 OLS regression residuals is used. Table 4-9 displays that the value of the Jarque-Bera (JB) statistic with 2 degrees of freedom is 40.208 for Model 5 and 41.581 for Model 6 at the 1-percent significance level, respectively. We can reject the null hypothesis that the residuals are normally distributed, meaning that there is non-normality of the errors.

The three diagnostic tests to detect heteroscedasticity are used for Models 5 and 6. Table 4-9 displays that both the Breusch-Pagan test and Koenker-Bassett test on random coefficients for the Models 5 and 6 are statistically significant at the 1-percent significance level.[40] We can reject the null hypothesis that there is an equal variance of errors, homoscedasticity, meaning that the 2000 OLS regression errors are unequally spread.

In summary, as in Table 4-9, the diagnostic tests of the 2000 OLS regression estimation with metropolitan spatial structure (MSS) detect non-normality and heteroscedasticity among the residuals.

[40] The White test on specification-robust test for heteroscedasticity in *GeoDa* reports N/A, because *GeoDa* is not able to correct for this (Anselin, 2005, p.195), as in Table 4-9.

Table 4-9 Diagnostic Tests of the 2000 OLS Estimation with MSS

REGRESSION DIAGNOSTICS			
MULTICOLLINEARITY CONDITION NUMBER			396.593 (557.081)
TEST ON NORMALITY OF ERRORS			
TEST	DF	VALUE	Sig.
Jarque-Bera	2 (2)	40.208 (41.581)	0.0000 (0.0000)
DIAGNOSTICS FOR HETEROSKEDASTICITY			
RANDOM COEFFICIENTS			
TEST	DF	VALUE	Sig.
Breusch-Pagan test	16 (23)	116.186 (94.473)	0.0000 (0.0000)
Koenker-Bassett test	16 (23)	74.204 (59.874)	0.0000 (0.00004)
SPECIFICATION ROBUST TEST			
TEST	DF	VALUE	Sig.
White	152 (299)	N/A	N/A

Note: the values in parentheses are the 2000 OLS estimation with metropolitan spatial structures (MSS) for Model 6.

4.2.2.3 Diagnostic Tests for Spatial Dependence in the 2000 OLS Estimation with MSS

Diagnostic tests for spatial dependence in the 2000 OLS estimation with MSS in Model 6 are computed for the row-standardized weights matrix (*cnty610_point_w3.3.gwt*), as in Table 4-10.

Table 4-10 shows that both LM-Lag and LM-Error statistics are highly significant. Of the robust forms, both the Robust LM-Lag and the Robust LM-Error statistics are significant at a 0.00001 and a 0.0022 significance level, respectively. We can reject the null hypothesis that there is no spatial autocorrelation (or dependence) to the 2000 OLS regression residuals, meaning that strong spatial effects are interrelated among neighboring regions. We consider the alternative spatial regression estimation in terms of a spatial lag model and a spatial error model.

Table 4-10 Diagnostic Tests for Spatial Dependence in the 2000 OLS Estimation with MSS

DIAGNOSTICS FOR SPATIAL DEPENDENCE			
FOR WEIGHT MATRIX : *cnty610_point_w3.3.gwt* (row-standardized weights)			
TEST	MI/DF	VALUE	PROB
Moran's I (error)	0.1825	N/A	N/A
Lagrange Multiplier (lag)	1	243.6061	0.0000000
Robust LM (lag)	1	45.9836	0.0000000
Lagrange Multiplier (error)	1	207.0094	0.0000000
Robust LM (error)	1	9.3869	0.0021854
Lagrange Multiplier (SARMA)	2	252.9930	0.0000000

4.2.2.4 The 2000 Spatial Regression Estimation Results for AQI

Table 4-11 below provides the spatial regression estimation results with considering spatial dependence between dependent or independent variables in Models 7 and 8 to the 2000 OLS regression residuals, as in the formulas in Section 3.4.2.

4.2.2.4.1 The 2000 Maximum Likelihood Spatial Lag Estimation

Table 4-11 shows the maximum likelihood (ML) spatial lag estimation with all the 2000 OLS regression standardized coefficients for the dependent variable, air quality index (*A_mean00*) for Model 7. Table 4-11 displays the number of observations (610 counties), the number of variables including the constant term (25), and the degrees of freedom (585) for Model 7. Table 4-11 shows that the R-squared value is approximately 0.4181.[41] This indicates that 41.81% of the variance in changes in air quality is predicted from the combination of agglomeration effect, governmental structures, socio-demographic features, travel behavior, regional amenities, and a spatially lagged dependent variable (*W_A_mean00* (ρ)) for distance-based weights matrix (*cnty610_point_w3.3.gwt*).

The three measures are used to choose select a good fitting spatial lag model. They are the log likelihood (-2271.6), the Akaike information criterion (4593.19), and the Schwarz information criterion (4703.53), as in the right hand column in Table 4-11. To compare the values in the 2000 spatial lag model in the right column in Table 4-11 to those for the 2000 OLS estimation in Table 4-8, we identified an increase in the log likelihood from -2326.64 for OLS to -2271.6, a decrease in the AIC from 4701.27 for OLS to 4593.19, and a decrease in the SIC from 4807.2 for OLS to 4703.53. The lower the values of both the Akaike information criterion (AIC) and the Schwarz information

[41] The R-squared value in the Model 7 spatial lag model, 0.4181, is not a real R-squared, but a *pseudo* R-squared. With the pseudo R-squared, we need to be cautious of a direct comparison of all the spatial regression coefficient estimates to the 2000 OLS regression results (Anselin, 2005, p. 207).

criterion (SIC) are, the better the spatial lag model is fitted. The higher the log likelihood is, the better the fit of the spatial lag model is.

As in Table 4-11, the spatial autoregressive coefficient in the spatial lag model (*W_A_mean00*, $\rho = 0.644$) are statistically highly significant ($p < 0.0000001$), meaning that there is spatial dependence of the spatially lagged dependent variable to the 2000 OLS regression estimation.

To compare all the regression estimates between the spatial lag model in Table 4-11 and the OLS estimation in Table 4-8, the magnitude of all the regression coefficients is affected by the coefficient of the spatially lagged dependent variable (*W_A_mean00*, $\rho = 0.644$, $p < 0.0000001$). Most of the coefficients, such as pro-environment policy (*EnvPolicy*), median household income (*lnMHHI00*), services industry (*M_ser00*), employment concentration (*ConEmp00*), net population density per square mile (*lnNetP00*), and weighted average drive-alone commute time (*lnTotCom00*), show a decrease in absolute value. A few of the coefficients, such as special-purpose local governments (*lnSpeGt02*), show an increase in absolute value. The significance of most of the other regression coefficients also changed. The significance of the net population density per square mile (*lnNetP00*) and the services industry (*M_ser00*) change from $p < 0.00114$ to $p < 0.019$ and from $p < 0.0022$ to $p < 0.018$, respectively. The significance of special-purpose local governments (*lnSpeGt02*) changes from $p < 0.3723$ to $p < 0.091$, or from insignificant to significant. This indicates that the change in special-purpose local governments is statistically significant as a predictor for changes in air quality at the 10-percent significance level. Some of the coefficients change from significant to insignificant. The mixed land use (*LUMix01*) changes from $p < 0.009$ to $p < 0.516$, or from significant to insignificant. This points out that the mixed land use is not statistically significant as a predictor for changes in air quality index values, even if the mixed land use is a significant predictor for air quality index values at the 1-percent significance level without a spatial lagged dependent variable.

Table 4-11 the 2000 Maximum Likelihood Spatial Regression Estimation for AQI

Spatial Weight	cnty610_point_w3.3.gwt		
N	610	Log likelihood	-2271.6 (-2267.26)
# Variables	25 (24)	Akaike info criterion	*4593.19 (4582.51)*
DF	585 (586)	Schwarz criterion	4703.53 (4688.44)
R-squared	*0.4181 (0.4359)*		

Component	Model 7 (8)	Standardized Coefficients	Sig.
	CONSTANT	59.213 (97.617)	0.093 (0.011)
Specialization	m_mnf00	0.716 (1.011)	0.389 (0.237)
	m_ser00	-5.567 (-3.478)	0.018 (0.146)
	m_rd00	-0.513 (-0.591)	0.125 (0.066)
	m_env00	-0.413 (-0.130)	0.272 (0.715)
Government	*lnGenGt02*	0.403 (1.095)	*0.548 (0.129)*
	lnSpeGt02	-0.932 (-1.095)	*0.091 (0.079)*
	EnvPolicy	-0.192 (-0.175)	0.017 (0.096)
	SGMP	-0.134 (1.635)	0.900 (0.231)
Socio-demographic	PctHisB00	0.013 (0.046)	0.725 (0.283)
	lnMHHI00	-13.229 (-13.500)	0.0005 (0.0014)
	PctBA00	0.040 (0.092)	*0.638 (0.296)*
Intermediate	PctDriA00	0.151 (0.021)	0.222 (0.874)
	lnClimate	-0.686 (-0.104)	0.241 (0.893)
	LnUnde92		
	region_dum	-2.638 (-3.807)	0.170 (0.336)
Initial	*lnTotPop90*		
MSS	lnNetP00	3.988 (1.980)	0.019 (0.274)
	PctOpen01	0.270 (0.174)	0.067 (0.279)
	LUMix01	-3.022 (-1.910)	0.516 (0.709)
	ConPop00	3.848 (-0.661)	0.444 (0.898)
	ConEmp00	-6.424 (-9.182)	0.074 (0.011)
	AveCom00	0.140 (0.083)	0.336 (0.595)
	lnTotCom00	4.783 (5.831)	0.000 (0.000)
	CenPop00	-0.746 (-0.339)	0.814 (0.913)
	CenEmp00	0.920 (-0.237)	0.808 (0.949)
Spatial Lag	W_A_mean00 (ρ)	0.644	0.000
Spatial Error	LAMBDA (λ)	(0.765)	(0.000)

Note: the values in parentheses are the spatial error regression estimation with metropolitan spatial structures (MSS) for Model 8.

Table 4-12 illustrates the diagnostic tests for the year 2000 maximum likelihood spatial lag estimation in Model 7 for the row-standardized weights matrix (*cnty610_point_w3.3.gwt*). The Breusch-Pagan (BP) statistic on random coefficients in the error terms for the Model 7 to detect heteroscedasticity is 50.647 at a 0.0008 significance level. We can reject the null hypothesis that there is an equal variance of errors (homoscedasticity), meaning that the spatial lag errors are unequally spread. The Likelihood Ratio Test is used to compare the 2000 OLS regression estimation to the alternative spatial lag model. The Likelihood Ratio Test value of 110.0826 is statistically significant ($p < 0.0000001$), indicating that the spatial autoregressive coefficient (ρ) of spatial lag model in Model 7 is strongly significant.

Table 4-12 Diagnostic Tests for the 2000 Spatial Regression Models

REGRESSION DIAGNOSTICS			
DIAGNOSTICS FOR HETEROSKEDASTICITY			
RANDOM COEFFICIENTS			
TEST	DF	VALUE	PROB
Breusch-Pagan test	23 (23)	50.64684 (46.97684)	0.0007573 (0.0022567)
DIAGNOSTICS FOR SPATIAL LAG MODEL			
SPATIAL *LAG* DEPENDENCE FOR WEIGHT MATRIX : *cnty610_point_w3.3.gwt*			
TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	110.0826	0.0000000
DIAGNOSTICS FOR SPATIAL ERROR MODEL			
SPATIAL *ERROR* DEPENDENCE FOR WEIGHT MATRIX : *cnty610_point_w3.3.gwt*			
TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	118.761	0.0000000

Note: the values in parentheses are the spatial error regression estimation with metropolitan spatial structures (MSS) for Model 8.

4.2.2.4.2 The 2000 Maximum Likelihood Spatial Error Estimation

Table 4-11 shows the maximum likelihood (ML) spatial error estimation for Model 8 with all the 2000 OLS regression standardized coefficients for the dependent variable, air quality index (*A_mean00*). We can assume that the errors for neighboring observations of a spatial regression model are spatially correlated, as in Section 3.4.2.2. Table 4-11 shows the number of observations (610 counties), the number of variables including the constant term (24), and the degrees of freedom (586) for Model 8. Table 4-11 also shows that the R-squared value for Model 8 is about 0.4359.[42] This indicates that about 43.59% of the variance in changes in air quality is predicted from the combination of agglomeration effect, governmental structures, socio-demographic features, travel behavior, regional amenities, and a spatially weighted average of the errors (*LAMBDA* (λ)) for distance-based weights matrix among neighboring regions (*cnty610_point_w3.3.gwt*).

The three measures in Table 4-11 are used to select a good fitting spatial lag model. They are the log likelihood (-2267.256), the Akaike information criterion (4582.51), and the Schwarz information criterion (4688.44). Comparing the values in the 2000 spatial error model in Table 4-11 to those for the 2000 OLS estimation in Table 4-8

[42] With the pseudo R-squared value of 0.4359 in the Model 8 in the spatial error model, we need to be cautious of a direct comparison of all the spatial regression coefficient estimates to the 2000 OLS regression results.

(or the 2000 spatial lag model in Table 4-11), we identify an increase in the log likelihood from -2326.64 for OLS (or -2271.6 for spatial lag) to -2267.26, a decrease in the AIC from 4701.27 for OLS (or 4593.19 for spatial lag) to 4582.51, and a decrease in the SIC from 4807.2 for OLS (or 4703.53 for spatial lag) to 4688.44. The lower the values of both the Akaike information criterion (AIC) and the Schwarz information criterion (SIC) are, the better the spatial lag model is fitted. The higher the log likelihood is, the better the fit of the spatial lag model is.

Table 4-11 displays that the spatial autoregressive coefficient in the error terms in the spatial error model (*LAMBDA*, $\lambda = 0.765$) is statistically highly significant ($p < 0.0000001$). This indicates that there is spatial dependence of spatially weighted average of errors among neighboring regions to the 2000 OLS regression estimation.

When comparing all of the regression estimates between the spatial error model in Table 4-11 and the 2000 OLS estimation in Table 4-8, the magnitude of all of the regression coefficients is affected by the coefficient of the spatially weighted average of errors (*LAMBDA*, $\lambda = 0.7647$, $p < 0.0000001$). The results between the spatial error model and the OLS estimation in 2000 are similar to those between the spatial lag model and the OLS estimation in terms of the sign of the regression coefficients, but different in terms of the magnitude and significance of the coefficients. For example, most of the coefficients, such as pro-environment policy (*EnvPolicy*), median household income (*lnMHHI00*), the services industry (*M_ser00*), net population density per square mile (*lnNetP00*), and weighted average drive-alone commute time (*lnTotCom00*), show a decrease in absolute value. Some of the coefficients, such as special-purpose local governments (*lnSpeGt02*), the research and development (R&D) industry (*M_rd00*), concentration in high employment density sub-areas (*ConEmp00*), and weighted average drive-alone commute time (*lnTotCom00*), show an increase in absolute value.

The significance of most of the other regression coefficients also changed. The significance of the medium household income level (*lnMHHI00*) and the pro-environment policy (*EnvPolicy*) change from $p < 0.00026$ to $p < 0.0014$ and from $p < 0.0106$ to $p < 0.096$, respectively. The significance of special-purpose local governments (*lnSpeGt02*) and the R&D industry (*M_rd00*) change from $p < 0.3723$ to $p < 0.079$ and from $p < 0.4374$ to $p < 0.066$, respectively, or from insignificant to significant. This means that both special-purpose local governments and the R&D industry are statistically significant predictors of change in air quality at the 10-percent significance level. Some of the coefficients change from significant to insignificant. The significance of the land use mix (*LUMix01*) and the net population density per square mile (*lnNetP00*) change from $p < 0.009$ to $p < 0.709$ and from $p < 0.0011$ to $p < 0.274$, respectively, or from significant to insignificant. This means that both the mixed land use and the net population density are statistically insignificant as predictors for change in air quality, even if they are significant as predictors for air quality at the 1-percent significance level without a spatial lagged dependent variable and spatially correlated error terms.

Table 4-12 above displays the diagnostic tests for the 2000 maximum likelihood spatial error estimation for Model 8 for the row-standardized weights matrix (*cnty610_point_w3.3.gwt*). The Breusch-Pagan (BP) statistic on random coefficients in the error terms for Model 8 to detect heteroscedasticity is 46.977 at a 0.002 significance level. We can reject the null hypothesis that there is an equal variance of errors for neighboring observations (homoscedasticity). The Likelihood Ratio Test value of 118.761 is statistically significant ($p < 0.0000001$). This indicates that the spatial autoregressive coefficient (λ) of spatial error model for Model 8 is strongly significant.

To compare the results between the spatial lag and error model, as in Table 4-11, the spatial error model is a better fit than the spatial lag model because the spatial error model has a higher value of the Log Likelihood and lower values of the AIC and SIC. Additionally, the estimation results of the spatial error model are similar to those of the spatial lag model in terms of the sign of the regression coefficients, but different in terms of the magnitude and significance of the coefficients. For example, the magnitude of some of the coefficients, such as special-purpose local governments (*lnSpeGt02*), medium household income (*lnMHHI00*), the R&D industry (*M_rd00*), concentration in high employment density sub-areas (*ConEmp00*), and weighted average drive-alone commute time (*lnTotCom00*), show an increase in absolute value. A few of the coefficients, such as pro-environment policy (*EnvPolicy*), the services industry (*M_ser00*), and net population density (*lnNetP00*), show a decrease in absolute value. The significance of most of the other regression coefficients also changed. The significance of medium household income (*lnMHHI00*) changes from $p < 0.0006$ to $p < 0.0014$. The significance of the R&D industry (*M_rd00*) changes from $p < 0.125$ to $p < 0.066$, or from insignificant to significant. The significance of the services industry (*M_ser00*) and the net population density (*lnNetP00*) change from $p < 0.018$ to $p < 0.146$ and from $p < 0.019$ to $p < 0.274$, respectively, or from significant to insignificant.

4.2.3 The 2006 OLS Estimation and Spatial Regression Results

As conducted in the 1990 and 2000 OLS estimation and spatial regression results, we estimate the 2006 OLS regression coefficients with MSS and conduct the diagnostics tests for the 2006 OLS estimation in terms of three measures: multicollinearity, normality, and heteroscedasticity. Continuously, we detect the diagnostics for spatial autocorrelation or dependence based on Lagrange Multiplier (LM) test statistics (i.e., LM-Lag or LM-Error). We estimate all the regression coefficients with spatial autoregressive coefficients (ρ and λ) relating to the spatial dependence based on the maximum likelihood (ML) estimation to the 2006 OLS regression model. Finally, we compare not only the alternative spatial regression results to the 2006 OLS regression estimation, but also the results between the spatial lag and error model.

4.2.3.1 The 2006 OLS Estimation with MSS

Table 4-13 below illustrates the summary characteristics of Model 9 showing all regression standardized coefficients for the air quality index (AQI) in 2006 (*A_mean06*), not including metropolitan spatial structures (MSS) and regression diagnostics. Table 4-13 shows the number of observations (610 counties), the number of variables including the constant term (17), and the degrees of freedom (593) for Model 9.

Table 4-13 shows that R-squared value and adjusted R-squared value for Model 9 are about 0.251 and 0.231, respectively. This indicates that 25.1% (or 23.1%) of the variance in changes in air quality are predicted from the combination of agglomeration effect, governmental structures, socio-demographic features, travel behavior, and regional amenities. Table 4-13 also displays that the ANOVA F-statistic with 17 and 593 degrees of freedom for all of the regression coefficients is 12.43 at less than the 1-percent significance level. This indicates that the combination of all of the independent variables is statistically significant as a predictor for changes in air quality.

Table 4-13 shows a number of interesting patterns for Model 9. The negative regression coefficients of agglomerative effects in the services industry (*M_ser06*, β = -0.110), college graduates or higher (*PctBA06*, β = -0.219), and a state's pro-environmental policy (*EnvPolicy*, β = -0.169) indicate that these variables are statistically significant predictors to improved air quality at the 5-percent or 1-percent significance level. The positive regression coefficients of medium household income (*lnMHHI06*, β = 0.102), proportion of Black and Hispanic population (*PctHisB06*, β = 0.134), and total population in 1990 (*lnTotPop90*, β = 0.514) indicate that these variables are statistically significant predictors to worsened air quality at the 10-percent, 5-percent, and 1-percent significance level, respectively.

Table 4-13 also shows a number of interesting patterns for Model 10[43] with metropolitan spatial structure (MSS). The negative regression coefficients of a state's pro-environmental policy (*EnvPolicy*, β = -0.123), agglomerative effects in the services industry (*M_ser06*, β = -0.190), medium household income (*LnMHHI06*, β = -0.190),

[43] Model 10 results with the variable MSS are shown in parentheses.

concentration in high employment density (*ConEmp00*, β = -0.160), and land use mix (*LUMix06*, β = -0.292) indicate that these variables are statistically significant as predictors for improved air quality at the 5-percent or 1-percent significance level. The positive regression coefficients of proportion of Black or Hispanic population (*PctHisB06*, β = 0.092), concentration in high population density (*ConPop06*, β = 0.177), weighted average drive-alone commute time (*lnTotCom06*, β = 0.588), net population density per square mile (*lnNetP06*, β = 0.165), and proportion of developed open space (*Pct_open06*, β = 0.231) indicate that these variables are statistically significant as predictors for worsened air quality at the 10-percent or 5-percent or 1-percent significance level.

However, other confounding variables for Model 10 are together considered to obtain this result because the null hypothesis that all regression coefficients are simultaneously equal to zero is rejected. Table 4-13 shows that the ANOVA F-statistic with 24 and 586 degrees of freedom (F=10.80) for Model 10 is statistically significant below the 1-percent significance level (or $p < 0.0000001$). This indicates that all of the independent variables significantly combine together to predict changes in air quality.

Table 4-13 OLS Estimation with MSS, 2006

N	610 (610)	F-statistic	*12.4289 (10.8029)*
# Variables	17 (24)	Prob(F-statistic)	*0.0000 (0.0000)*
DF	593 (586)	Log likelihood	-2215 (-2195.4)
R-squared	*0.251 (0.298)*	Akaike info criterion	*4464.01 (4438.8)*
Adjusted R-squared	0.231 (0.270)	Schwarz criterion	4539.04 (4544.72)

Component	Model 9 (10)	Standardized Coefficients (β)	Sig.
	CONSTANT	-51.509 (40.743)	0.063 (0.202)
Specialization	m_mnf06	0.044 (0.053)	0.325 (0.240)
	m_ser06	-0.110 (-0.099)	0.013 (0.025)
	m_rd06	-0.021 (-0.009)	0.610 (0.824)
	m_env06	-0.051 (-0.047)	0.183 (0.209)
Government	lnGenGt02	0.025 (0.018)	*0.714 (0.799)*
	lnSpeGt02	-0.037 (-0.051)	*0.502 (0.346)*
	EnvPolicy	-0.169 (-0.123)	0.0008 (0.013)
	SGMP	0.011 (-0.028)	0.795 (0.517)
Socio -demographic	*PctHisB06*	0.134 (0.092)	0.010 (0.076)
	lnMHHI06	0.102 (-0.190)	0.059 (0.011)
	PctBA06	-0.219 (-0.097)	*0.000 (0.171)*
Intermediate	PctDriA06	-0.001 (0.052)	0.991 (0.006)
	lnClimate	0.006 (-0.071)	0.889 (0.114)
	LnUnde92	-0.010	0.829
	region_dum	-0.008 (-0.050)	0.874 (0.313)
Initial	*lnTotPop90*	0.514	*0.000*
MSS	lnNetP06	(0.165)	(0.038)
	PctOpen06	(0.231)	(0.001)
	LUMix06	(-0.292)	(0.001)
	ConPop06	(0.177)	(0.015)
	ConEmp00	(-0.160)	(0.025)
	AveCom06	(0.040)	(0.509)
	lnTotCom06	(0.588)	(0.000)
	CenPop06	(-0.050)	(0.262)
	CenEmp00	(0.016)	(0.733)

Note: the values in parentheses are the OLS estimation with metropolitan spatial structures (MSS) for Model 10.

4.2.3.2 Diagnostic Tests of the 2006 OLS Estimation with MSS

Due to high multicollinearity among independent variables in Models 9 and 10, as in Section 3.4.3.1, we remove variables with high variance inflation factors (VIF): total employment (*LnTotEmp06*), net employment density (*LnNetE06*), and total developed land (*pct_urb06*). Their VIFs are higher than VIF of 6.91 or lower than the tolerance (TOL) of 0.145 between these variables. To reduce the impact of multicollinearity, we use the cutoff threshold for the tolerance (TOL) of below 0.145 (above a VIF of 6.91). To remedy high multicollinearity among independent variables, we transform the following variables using natural logarithm: number of local governments (*lnGenGt* & *lnSpeGt*), net population density (*lnNetP06*), median household income (*lnMHHI06*), undeveloped land size (*LnUnde92*), and weighted average drive-alone commute (*lnTotCom06*). The natural logarithm of aforementioned independent variables is approximately normally distributed between a skewness statistic of -2.0 and 2.0 using the semilog regression estimation. Table 4-14 shows that the value of multicollinearity condition number (*k)* is 373.281 for Model 9 and 549.149 for Model 10, respectively. We can consider that there is moderately high multicollinearity of the 2006 OLS regression models, based on a rule of thumb.

The Jarque-Bera (JB) test to detect normality on the 2006 OLS regression residuals is used. Table 4-14 shows that the value of the Jarque-Bera (JB) statistic with 2 degrees of freedom is 102.102 for Model 9 and 67.707 for Model 10 at less than a 1-percent significance level, respectively. We can reject the null hypothesis that the residuals are normally distributed, meaning that there is non-normality of the errors.

The three diagnostic tests to detect heteroscedasticity are used for Models 9 and 10. Table 4-14 illustrates that both Breusch-Pagan (BP) test and Koenker-Bassett (KB) test on random coefficients for Models 9 and 10 are statistically significant at the 1-percent significance level. We can reject the null hypothesis that there is an equal variance of errors (homoscedasticity), indicating that the 2006 OLS regression errors are unequally spread.

In summary, as in Table 4-14, the diagnostic tests of the 2006 OLS regression estimation with metropolitan spatial structure (MSS) detect non-normality and heteroscedasticity among the residuals.

Table 4-14 Diagnostic Tests of the 2006 OLS Estimation with MSS

REGRESSION DIAGNOSTICS			
MULTICOLLINEARITY CONDITION NUMBER			373.281 (549.149)
TEST ON NORMALITY OF ERRORS			
TEST	DF	VALUE	Sig.
Jarque-Bera	2 (2)	102.102 (67.707)	0.0000 (0.0000)
DIAGNOSTICS FOR HETEROSKEDASTICITY			
RANDOM COEFFICIENTS			
TEST	DF	VALUE	Sig.
Breusch-Pagan test	16 (23)	107.605 (90.679)	0.0000 (0.0000)
Koenker-Bassett test	16 (23)	54.791 (50.475)	0.0000 (0.0000)
SPECIFICATION ROBUST TEST			
TEST	DF	VALUE	Sig.
White	152 (299)	N/A	N/A

Note: the values in parentheses are the OLS estimation with metropolitan spatial structures (MSS) for Model 10.

4.2.3.3 Diagnostic Tests for Spatial Dependence in the 2006 OLS Estimation with MSS

Diagnostic tests for spatial dependence in the 2006 OLS estimation with metropolitan spatial structure (MSS) in Model 10 are computed for the row-standardized weights matrix (*cnty610_point_y06_w3.3.gwt*), as in Table 4-15.

Table 4-15 illustrates that both LM-Lag and LM-Error statistics are highly significant. Of the robust forms, both the Robust LM-Lag and the Robust LM-Error statistics are significant at a 0.000008 and a 0.000031 significance level, respectively. We can reject the null hypothesis that there is no spatial autocorrelation or dependence to the 2006 OLS regression residuals. This means that strong spatial effects are interrelated among neighboring regions. We consider the alternative spatial regression estimation in terms of a spatial lag model and a spatial error model.

Table 4-15 Diagnostic Tests for Spatial Dependence in the 2006 OLS Estimation with MSS

DIAGNOSTICS FOR SPATIAL DEPENDENCE			
FOR WEIGHT MATRIX : *cnty610_point_y06_w3.3.gwt* (row-standardized weights)			
TEST	MI/DF	VALUE	PROB
Moran's I (error)	0.1668	N/A	N/A
Lagrange Multiplier (lag)	1	175.31332	0.0000000
Robust LM (lag)	1	19.9276	0.0000080
Lagrange Multiplier (error)	1	172.7543	0.0000000
Robust LM (error)	1	17.3685	0.0000308
Lagrange Multiplier (SARMA)	2	192.6818	0.0000000

4.2.3.4 The 2006 Spatial Regression Estimation Results for AQI

Table 4-16 below displays the spatial regression estimation results for Models 11 and 12 against the 2006 OLS regression residuals when considering spatial dependence between dependent or independent variables, as in Section 3.4.2.

4.2.3.4.1 The 2006 Maximum Likelihood Spatial Lag Estimation

Table 4-16 shows the maximum likelihood (ML) spatial lag estimation with all the 2006 OLS regression standardized coefficients for the dependent variable, air quality index (*A_mean06*) for Model 11. Table 4-16 shows the number of observations (610 counties), the number of variables including the constant term (25), and the degrees of freedom (585) for Model 11. Table 4-16 shows that the R-squared value for Model 11 is approximately 0.3973.[44] This indicates that 39.73% of the variance in changes in air quality can be predicted from the combination of agglomeration effect, governmental structures, socio-demographic features, travel behavior, regional amenities, and a spatially lagged dependent variable (*W_A_mean06* (ρ)) for distance-based weights matrix (*cnty610_point_y06_w3.3.gwt*).

As shown in the right column in Table 4-16, the three measures to choose a good fitting spatial lag model are the log likelihood (-2155.03), the Akaike information criterion (4360.05), and the Schwarz information criterion (4470.39). Comparing the values in the 2006 spatial lag model in the right column in Table 4-16 to those for the 2006 OLS estimation in Table 4-13, we identify an increase in the log likelihood from -2195.4 for OLS to -2155.03, a decrease in the AIC from 4438.8 for OLS to 4360.05, and a decrease in the SIC from 4544.72 for OLS to 4470.39. The lower the values of both the Akaike information criterion (AIC) and the Schwarz information criterion (SIC) are, the

[44] With the pseudo R-squared value of 0.3973 in the Model 11 spatial lag model, we need to be cautious of a direct comparison of all the spatial regression coefficient estimates to the 2006 OLS regression results (Anselin, 2005, p. 207).

better the 2006 spatial lag model is fitted. The higher the log likelihood is, the better the 2006 spatial lag model is fitted.

As in Table 4-16, the spatial autoregressive coefficient in the spatial lag model (*W_A_mean06*, $\rho = 0.576$) is statistically highly significant ($p < 0.0000001$). This means that there is spatial dependence of the spatially lagged dependent variable to the 2006 OLS regression estimation.

To compare all the regression estimates between the 2006 spatial lag model in Table 4-16 and the 2006 OLS estimation in Table 4-13, the magnitude of all of the regression coefficients is affected by the coefficient of the spatially lagged dependent variable (*W_A_mean06*, $\rho = 0.576$, $p < 0.0000001$). Most of the coefficients, such as median household income (*lnMHHI06*), the services industry (*M_ser06*), concentration in high employment density sub-areas (*ConEmp00*), net population density per square mile (*lnNetP06*), and weighted average drive-alone commute time (*lnTotCom00*), and land use mix (*LUMix06*), show a decrease in absolute value. A few of the coefficients, such as pro-environment policy (*EnvPolicy*) and special-purpose local governments (*lnSpeGt02*), show an increase in absolute value. The significance of the other regression coefficients also changed. The significance of pro-environment policy (*EnvPolicy*) and land use mix (*LUMix06*) change from $p < 0.013$ to $p < 0.006$ and from $p < 0.0014$ to $p < 0.082$, respectively. The significance of special-purpose local governments (*lnSpeGt02*) and the climate amenity (*lnClimate*) change from $p < 0.346$ to $p < 0.077$ and from $p < 0.114$ to $p < 0.050$, respectively, or from insignificant to significant. This means that the special-purpose local governments and the regional climate are statistically significant predictors of changes in air quality at the 10-percent and 5-percent significance level, respectively. The significance of the services industry (*M_ser06*) and the concentration in high employment density sub-areas (*ConPop06*) change from $p < 0.025$ to $p < 0.108$ and from $p < 0.015$ to $p < 0.542$, respectively, or from significant to insignificant. This means that agglomerative effects in the services industry and the concentration in high population density sub-areas are statistically insignificant as predictor for changes in air quality, even if the level of specialization in the services industry and the concentration in high population density sub-areas are significant predictors to air quality at the 5-percent significance level without a spatial lagged dependent variable.

Table 4-16 the 2006 Maximum Likelihood Spatial Regression Estimation for AQI

Spatial Weight	*cnty610_point_y06_w3.3.gwt*		
N	610	Log likelihood	-2155.03 (-2148.70)
# Variables	25 (24)	Akaike info criterion	4360.05 (4345.39)
DF	585 (586)	Schwarz criterion	4470.39 (4451.31)
R-squared	*0.3973 (0.4183)*		

Component	Model 11 (12)	Standardized Coefficients	Sig.
	CONSTANT	29.308 (70.686)	0.313 (0.026)
Specialization	m_mnf06	0.817 (1.134)	0.185 (0.071)
	m_ser06	-3.202 (-2.056)	0.108 (0.313)
	m_rd06	-0.122 (-0.043)	0.672 (0.875)
	m_env06	-0.143 (-0.082)	0.641 (0.781)
Government	*lnGenGt02*	0.318 (0.578)	*0.573 (0.333)*
	lnSpeGt02	-0.801 (-1.122)	*0.077 (0.028)*
	EnvPolicy	-0.183 (-0.224)	0.006 (0.008)
	SGMP	-0.286 (1.409)	0.747 (0.205)
Socio -demographic	PctHisB06	0.022 (0.044)	0.458 (0.191)
	lnMHHI06	-8.072 (-9.637)	0.009 (0.004)
	PctBA06	-0.059 (-0.012)	*0.394 (0.866)*
Intermediate	PctDriA06	-0.070 (-0.154)	0.477 (0.128)
	lnClimate	-0.945 (-0.745)	0.050 (0.234)
	LnUnde92		
	region_dum	-1.387 (-3.354)	0.386 (0.269)
Initial	*lnTotPop90*		
MSS	lnNetP06	2.872 (2.376)	0.046 (0.111)
	PctOpen06	0.267 (0.237)	0.027 (0.069)
	LUMix06	-6.693 (-5.971)	0.082 (0.150)
	ConPop06	2.600 (-1.879)	0.542 (0.667)
	ConEmp00	-6.416 (-8.392)	0.026 (0.004)
	AveCom06	0.019 (0.016)	0.879 (0.902)
	lnTotCom06	5.189 (5.877)	0.000 (0.000)
	CenPop06	-1.398 (-0.782)	0.567 (0.744)
	CenEmp00	-0.726 (-1.619)	0.808 (0.580)
Spatial Lag	W_A_mean06 (ρ)	0.576	0.000
Spatial Error	LAMBDA (λ)	(0.708)	(0.000)

Note: the values in parentheses are the spatial error regression estimation with metropolitan spatial structures (MSS) for Model 12.

Table 4-17 shows the diagnostic tests for the 2006 maximum likelihood spatial lag estimation for Model 11 for the row-standardized weights matrix (*cnty610_point_y06_w3.3.gwt*). The Breusch-Pagan (BP) statistic on random coefficients in the error terms for Model 11 to detect heteroscedasticity is 56.996 at a 0.0001 significance level. We can reject the null hypothesis that there is an equal variance of errors (homoscedasticity), meaning that the spatial lag errors are unequally spread. Table 4-17 also shows that the Likelihood Ratio Test value of 80.746 is statistically significant ($p < 0.0000001$). This indicates that the spatial autoregressive coefficient (ρ) of spatial lag model for Model 11 is strongly significant.

Table 4-17 Diagnostic Tests for the 2006 Spatial Regression Models

REGRESSION DIAGNOSTICS			
DIAGNOSTICS FOR HETEROSKEDASTICITY			
RANDOM COEFFICIENTS			
TEST	DF	VALUE	PROB
Breusch-Pagan test	23 (23)	56.99597 (55.16569)	0.0001026 (0.0001849)

DIAGNOSTICS FOR SPATIAL LAG MODEL			
SPATIAL *LAG* DEPENDENCE FOR WEIGHT MATRIX : *cnty610_point_y06_w3.3.gwt*			
TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	80.74575	0.0000000

DIAGNOSTICS FOR SPATIAL ERROR MODEL			
SPATIAL *ERROR* DEPENDENCE FOR WEIGHT MATRIX : *cnty610_point_y06_w3.3.gwt*			
TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	93.40775	0.0000000

Note: the values in parentheses are the spatial error regression estimation with metropolitan spatial structures (MSS) for Model 12.

4.2.3.4.2 The 2006 Maximum Likelihood Spatial Error Estimation

Table 4-16 displays the maximum likelihood (ML) spatial error estimation for Model 12 with all the 2006 OLS regression standardized coefficients for AQI (*A_mean06*). This spatial error model employs the spatial weights matrix for the independent variables explained by continuous inverse distance between neighboring regions (*cnty610_point_y06_w3.3.gwt*). Table 4-16 shows the number of observations (610 counties), the number of variables including the constant term (24), and the degrees of freedom (586) for Model 12. Table 4-16 shows that the R-squared value for Model 12 is approximately 0.4183.[45] This indicates that 41.83% of the variance in changes in air quality can be predicted from the combination of agglomeration effect, governmental structures, socio-demographic features, travel behavior, regional amenities, and a spatially weighted average of the errors (*LAMBDA* (λ)) for distance-based weights matrix among neighboring regions (*cnty610_point_y06_w3.3.gwt*).

The three measures in the right hand column in Table 4-16 are used to choose a good fitting spatial regression model. They are the log likelihood (-2148.696), the Akaike information criterion (4345.39), and the Schwarz information criterion (4451.31), as in the right hand column in Table 4-16 above. To compare the values in the 2006 spatial

[45] With the pseudo R-squared value of 0.4183, we need to be cautious of a direct comparison of all the spatial regression coefficient estimates to the 2006 OLS regression results.

error model in Table 4-16 to those for the 2006 OLS estimation in Table 4-13 (or the 2006 spatial lag model in Table 4-16), we identify an increase in the log likelihood from -2195.4 for OLS (or -2155.03 for spatial lag) to -2148.70, a decrease in the AIC from 4438.8 for OLS (or 4360.05 for spatial lag) to 4345.39, and a decrease in the SIC from 4544.72 for OLS (or 4470.39 for spatial lag) to 4451.31. The lower the values of both the Akaike information criterion (AIC) and the Schwarz information criterion (SIC) are, the better the spatial error model is fitted. The higher the log likelihood is, the better the spatial error model is fitted.

As in Table 4-16 above, the spatial autoregressive coefficient in the spatial error model (*LAMBDA*, $\lambda = 0.7081$) is statistically highly significant ($p < 0.0000001$). This means that there is spatial dependence of the spatially weighted average of error terms between neighboring regions to the 2006 OLS regression estimation.

To compare all the regression estimates between the spatial error model in Table 4-16 and the 2006 OLS estimation in Table 4-13, the magnitude of all of the regression coefficients is affected by the coefficient of spatially weighted average of errors (*LAMBDA*, $\lambda = 0.7081$, $p < 0.0000001$). The results between the spatial error model and the OLS estimation in 2006 are similar to those between the spatial lag model and the OLS estimation in terms of the sign of the regression coefficients, but different in terms of the magnitude and significance of the coefficients. For example, some of the coefficients, such as pro-environment policy (*EnvPolicy*), median household income (*lnMHHI06*), the manufacturing industry (*M_mnf06*), special-purpose local governments (*lnSpeGt02*), concentration in high employment density sub-areas (*ConEmp00*), and weighted average drive-alone commute time (*lnTotCom00*), show an increase in absolute value.

The significance of most of the other regression coefficients also changed. The significance of medium household income level (*lnMHHI06*) and concentration in high employment density sub-areas (*ConEmp00*) change from $p < 0.010$ to $p < 0.0044$ and from $p < 0.0251$ to $p < 0.0039$, respectively. The significance of special-purpose local governments (*lnSpeGt02*) and level of specialization in the manufacturing industry (*M_mnf06*) change from $p < 0.346$ to $p < 0.028$ and from $p < 0.240$ to $p < 0.070$, respectively, or from insignificant to significant. This means that both special-purpose local governments and level of specialization in the manufacturing industry are statistically significant predictors of changes in air quality at the 5-percent and 10-percent significance level, respectively. The significance of mixed land use (*LUMix06*) and net population density per square mile (*lnNetP06*) change from $p < 0.0014$ to $p < 0.150$ and from $p < 0.038$ to $p < 0.111$, respectively, or from significant to insignificant. This indicates that both mixed land use and net population density are statistically insignificant as predictors for air quality, even if they are statistically significant predictors of changes in air quality at the 1-percent and 5-percent significance level, respectively, without spatial lagged dependent variable.

Table 4-17 above displays the diagnostic tests for the 2006 maximum likelihood spatial error estimation in Model 12 for the row-standardized weights matrix (*cnty610_point_y06_w3.3.gwt*). The Breusch-Pagan (BP) statistic on random coefficients in the error terms for Model 12 to detect heteroscedasticity is 55.166 at the 0.00018 significance level. We can reject the null hypothesis that there is an equal variance of error terms for neighboring observations (homoscedasticity). Table 4-17 also displays that the Likelihood Ratio Test value of 93.408 is statistically significant ($p < 0.0000001$). This indicates that the spatial autoregressive coefficient (λ) of spatial error model for Model 12 is strongly significant.

Comparing the results between the spatial lag and error models in Table 4-16, as indicated previously, the spatial error model are a better fit than the spatial lag model because the spatial error model can interpret a higher value of the Log Likelihood and lower values of the AIC and SIC. Additionally, the estimation results of the spatial error model are similar to those of the spatial lag model in terms of the sign of the regression coefficients, but different in terms of the magnitude and significance of the coefficients. For example, the magnitude of some of the coefficients, such as pro-environment policy (*EnvPolicy*), median household income (*lnMHHI06*), the manufacturing industry (*M_mnf06*), special-purpose local governments (*lnSpeGt02*), concentration in high employment density sub-areas (*ConEmp00*), and weighted average drive-alone commute time (*lnTotCom06*), show an increase in absolute value. The significance of most of the other regression coefficients also changed. The significance of concentration in high employment density sub-areas (*ConEmp00*) changes from $p < 0.0255$ to $p < 0.0039$. The significance of level of specialization in the manufacturing industry (*M_mnf06*) changes from $p < 0.1849$ to $p < 0.070$, or from insignificant to significant. The significance of mixed land use (*LUMix06*) and net population density per square mile (*lnNetP06*) change from $p < 0.082$ to $p < 0.150$ and from $p < 0.046$ to $p < 0.111$, respectively, or from significant to insignificant.

4.2.4 Summary of OLS and Spatial Regression Results for 1990, 2000, and 2006

Through the diagnostics tests for spatial dependence in the ordinary least squares (OLS) estimation with metropolitan spatial structure (MSS) for 1990, 2000, and 2006, as in Tables 4-5, 4-10, and 4-15, respectively, both LM-Lag and LM-Error statistics are highly significant. On the robust forms, both the Robust LM-Lag and the Robust LM-Error statistics are significant. We reject that the null hypothesis that there is no spatial autocorrelation (or dependence) to the OLS regression residuals for 1990, 2000, and 2006, indicating that strong spatial effects are interrelated among neighboring regions. We choose the spatial regression estimation, such as a spatial lag model and a spatial error model, to specify relationships between metropolitan spatial structure and air quality level, while the OLS regression estimation is discarded.

Comparing the results between the spatial lag and error models for 1990, 2000, and 2006, as in the right column in Tables 4-6, 4-11, and 4-16, respectively, we select the spatial error model as a better fit than spatial lag model. The spatial error models 4, 8, and 12 for 1990, 2000, and 2006, respectively, show a higher value of the Log Likelihood and lower values of the Akaike information criterion (AIC) and the Schwarz information criterion (SIC) than the spatial lag models 3, 7, and 11 do.

Through the spatial error models 4, 8, and 12 for 1990, 2000, and 2006, respectively, as shown in Table 4-18, we reject the null hypotheses that there are no relationships between metropolitan spatial structure (MSS) and changes in air quality index values (AQI) across U.S. 610 metropolitan areas while controlling for the major confounding variables (*ceteris paribus*).

The signs of significant predictors in terms of the spatial error models 4, 8, and 12 for 1990, 2000, and 2006, respectively, as shown in Table 4-18, have been consistently identified, either positive or negative. Hypothesis 2 (to identify the effects of developed land on changes in air quality level) is rejected because developed open space (*PctOpen*) is statistically positive and significant as a predictor of changes in air quality level, particularly for 2006. Metropolitan areas with a higher percentage of developed open space produce higher average air quality index values, leading to worsened air quality. Hypothesis 5 (to identify the effects of higher employment concentration on changes in air quality level) is rejected because higher employment concentration (*ConEmp*) is statistically negative and significant as a predictor of changes in air quality level, particularly for 2000 and 2006. Metropolitan areas with a higher percentage of densely employed sub-areas tend to produce lower average air quality index values, resulting in improved air quality. That is, metropolitan areas with polycentric employment centers tend to produce improved air quality level. Hypothesis 7 (to identify the effects of weighted average drive-alone commute times on changes in air quality level) is rejected because weighted average drive-alone commute times (*lnTotCom*) are statistically positive and significant as a predictor of changes in air quality level for 1990, 2000, and

2006. Metropolitan areas with longer weighted average commute times tend to produce higher average air quality index values, leading to worsened air quality.

The following properties of metropolitan spatial structures, including net population density (*lnNetP*, Hypothesis 1), mixed land use (*LUMix*, Hypothesis 3), high population concentration (*ConPop*, Hypothesis 4), average commute time (*AveCom*, Hypothesis 6), centralized population sub-areas (*CenPop*, Hypothesis 8), and centralized employment sub-areas (*CenEmp*, Hypothesis 9), are not statistically significant as predictors of changes in air quality.

In addition to the positive or negative impacts of metropolitan spatial structure on changes in air quality level, as in Table 4-18, we can also reject the null hypotheses that there are no relationships between confounding variables and changes in air quality level for 1990, 2000, and 2006. The statistically positive signs of significant predictors of changes in air quality include more specialized manufacturing industry (*m_mnf*, H10-a) for 2006 and general-purpose local governments per 1,000 persons (*lnGenGt*, H11-a) for 1990. Metropolitan areas with a higher specialized manufacturing industry or more numbers of general-purpose local governments tend to produce higher average air quality index values, leading to worsened air quality. Whereas the statistically negative signs of significant predictors include more specialized research and development (R&D) industry (*m_rd*, H10-c) for 2000, special-purpose local governments per 1,000 persons (*lnSpeGt*, H11-b) for 1990, 2000 and 2006, pro-environment policies (*EnvPolicy*, H12) for 2000 and 2006, median household income level (*lnMHHI*, H15) for 1990, 2000 and 2006, and higher educational attainment (*PctBA*, H16) for 1990. Metropolitan areas with a higher level of specialization in the R&D industry, more numbers of special-purpose governments, environment-centered policies, a higher level of median household income, or a higher percentage of college graduates or higher tend to produce lower average air quality values, leading to improved air quality.

Some of the confounding predictors, such as the service industry (*m_ser*, H10-b), the environmental industry (*m_env*, H10-d), statewide growth management programs (*SGMP*, H13), proportion of Black or Hispanic residents (*PctHisB*, H14), proportion of drive-alone commuters (*PctDriA*, H17), or climate (*lnClimate*, H18), and total population in 1990 (*lnTotPop90*, H19), are not statistically significant as predictors of changes in air quality.

As shown in Table 4-18, we can reject the null hypothesis of Hypothesis 20 that there is no spatial dependence between neighboring regions, because the maximum likelihood spatial error coefficients (λ) are statistically highly significant to predict changes in air quality level for 1990, 2000, and 2006. The magnitude, significance, and sign of the regression coefficients may be affected by the coefficient of spatially weighted average of the regression errors between neighboring regions. This indicates that spatial effects among neighboring regions are statistically significant as predictors of changes in air quality level.

In summary, both the multidimensional properties of metropolitan spatial structures and the major confounding variables are statistically significant as predictors of changes in air quality level. The statistically positive signs tend to produce higher average air quality index values, leading to worsened air quality level. The statistically negative signs tend to produce lower average air quality index values, resulting in improved air quality level.

Table 4-18 Summaries of Spatial Error Models for 1990, 2000, and 2006

Component	Variables	Hypotheses	Signs of Spatial Coefficient Estimates		
			Model 4 - 1990	Model 8 - 2000	Model 12 - 2006
MSS	lnNetP	H1	+	+	+
	PctOpen	H2		+	+*
	LUMix	H3	+	-	-
	ConPop	H4		-	-
	ConEmp	H5		-**	-***
	AveCom	H6	+	+	+
	lnTotCom	H7	+***	+***	+***
	CenPop	H8		-	-
	CenEmp	H9		-	-
Industrial Specialization	m_mnf	H10-a	+	+	+*
	m_ser	H10-b	-	-	-
	m_rd	H10-c	-	-*	-
	m_env	H10-d	+	-	-
Government	lnGenGt	H11-a	+***	+	+
	lnSpeGt	H11-b	-*	-*	-**
	EnvPolicy	H12	+	-*	-***
	SGMP	H13	-	-	+
Socio -demographic	PctHisB	H14	-	+	+
	lnMHHI	H15	-*	-***	-***
	PctBA	H16	-*	+	-
Intermediate	PctDriA	H17	+	+	-
	lnClimate	H18	-	-	-
Initial	lnTotPop90	H19			
Spatial Error	LAMBDA (λ)	H20	+***	+***	+***

Note: + (positive, meaning that AQI increases) ; - (negative, meaning that AQI decreases);

*p-value at a 0.10 level; ** p-value at a 0.05 level; *** p-value at a 0.01 level.

CHAPTER V

DISCUSSIONS AND CONCLUSIONS

5.1 Discussions

The proposed theoretical framework in Figure 3-1 and the OLS and spatial regression models contribute to measure a combination of multidimensional characteristics of metropolitan structure and its confounding factors to predict changes in air quality indices across U.S metropolitan areas for 1990, 2000, and 2006.

Overall, the estimated predictors of air quality improvements are significant, and of the expected sign, in line with empirical evidence in terms of the two major arguments over the relationships between urban structure and air quality in the literature, particularly showing that more compact regions can contribute more to air quality improvements than sprawling regions (Newman & Kenworthy, 1989, 1999; Newton, 1997, 2000; Masnavi, 2000; Williams, 2000; Ewing et al., 2002, 2003; Neuman, 2005; Stone, 2008; Schweitzer & Zhou, 2010). As reviewed by previous empirical works in section 2.3.3.2 (Galster et al., 2001; Cutsinger et al., 2005; Wolman et al., 2005; U.S. Department of Agriculture, 2001; Lang, 2003; Lopez & Hynes, 2003; Tsai, 2005; Torrens, 2008), the effects of compact and sprawling development patterns on changes in air quality co-exist in metropolitan areas for 2000 and 2006.

As displayed by the negative and significant predictors of changes in air quality index values in Table 4-18, metropolitan areas with highly concentrated employment centers show compact development characteristics stated in Table 2-2, leading to improved air quality. The positive and significant predictors of changes in air quality index values, as in Table 4-18, imply that metropolitan areas with more developed open space and longer commute times bring out sprawling development features described in Table 2-2 and Section 2.3.3.3, showing worsened air quality.

The estimated predictors that influence the formation of metropolitan structures and the changes in air quality are also significant and of the expected sign, as seen in the Section 2.3.4. Emerging metropolitan structures can be determined by the spatial distribution of location decisions made by households or firms (specifically in the decisions show to settle outside of central areas). These decisions show in the geographic distributions of employment centers, as either concentrated or dispersed in metropolitan areas. Simultaneously, different forms of governmental structures and local variation in public policies have an important role in the structure of emerging metropolitan areas in terms of the spatial distribution of population or employment. Emerging metropolitan structures determined by the location decisions of households or firms and the impact of political forces may contribute to changes in air quality level in metropolitan areas.

The level of specialization in different industrial sectors shows opposite signs, as seen in Table 4-18. The level of specialization in the manufacturing industry tends to produce higher average air quality index values, while the level of specialization in the R&D industries tends to produce lower average air quality index values. As reported in prior findings (Cooke, 1983; Carlino, 1985; Glaeser & Kahn, 2001; Felsenstein, 2002; Burchfield et al., 2005), more specialized employment in the manufacturing industries tended to be more sprawling, whereas more specialized employment in the services and the idea-intensive industries appeared to be more centralized. In a way, metropolitan areas with more decentralized manufacturing-intensive industries tend to produce worsened air quality, while those with more centralized employment sectors in the R&D-intensive industries tend to produce improved air quality.

Different forms of governmental structure shows opposite signs. Fragmented structure of general-purpose local governments from the polycentric view focusing on voters' preferences and locally service-related problems tends to produce higher average air quality index values. On the other hand, the special-purpose metropolitan governmental structure as viewed from the regionalist's perspective of tackling spillover issues between local communities tends to produce lower average air quality index values. The estimated sign for the fragmented structure of general-purpose local governmental structure, particularly in 1990, is in consistent with prior findings (Carruthers & Ulfarsson, 2002; Carrutheres, 2003), which supports the contention that fragmented governmental structure can contribute to adverse impacts caused by sprawling growth in outlying areas in metropolitan areas, such as environmental pollution and loss of green space. In a way, metropolitan areas with more fragmented general-purpose local governments tend to produce worsened air quality level, while metropolitan areas with more fragmented special-purpose local governments tend to produce improved air quality level.

The role of public policies remains inconclusive, aligned with the debate over which public policies will be beneficial for compact or sprawling development patterns, as highlighted by the debate of Gordon & Richardson (1997) and Ewing (1997). The effect of highly innovative statewide pro-environmental policies on improved air quality level is statistically significant, but the effect of statewide growth management programs (SGMPs) in metropolitan areas on improved air quality level is not statistically significant, as seen in prior findings in the literature (Johnson, 2001; Brueckner, 2001). Metropolitan areas with highly innovative statewide pro-environmental policies tend to produce improved air quality level, but metropolitan areas with statewide growth management programs have little impact on changes in air quality level.

Other confounding forces that shape metropolitan spatial structures, such as income level, racial composition, educational attainment, and regional amenities, have mixed impact on changes in air quality level. Metropolitan areas with more highly educated residents, particularly in 1990, and with a higher level of median household income tend to produce improved air quality level. On the other hand, the effects of a higher percentage of Black or Hispanic residents and regional amenities, such as temperature and geographical location, are not significant in predicting changes in air quality level. This offers contrasts to previous studies in the literature in Section 2.3.4.4 pointing to temperature and geographical location as having an impact on changes in air quality.

Spatial effects among neighboring regions have an impact on changes in air quality level, as seen in Table 4-18 for significant spatial multiplier parameter (λ). The magnitude, significance, and sign of the estimates of the air quality index (AQI) variable and the explanatory variables vary considerably according to the presence of the strength of spatial dependence among neighboring regions to predict changes in air quality level.

5.2 Policy Implications

Based on the empirical results in this dissertation, statewide pro-environmental policy measures to improve environmental performance for a sustainable future, including air quality standards, pollution prevention programs, renewable energy policies, the National Environmental Performance Partnership System (NEPPS) programs, state climate change action plans, state-authored inventories of greenhouse gas emissions, and innovation in comprehensive plan requirements, can contribute to improved air quality level in metropolitan areas. Detailed policy strategies will be required to continue to produce cleaner air quality across metropolitan areas. A regionalist view to tackling negative spillover issues surrounding sprawling development patterns should consider strategies that can account for the presence of the spatial multiplier (or spillover) effects on changes in air quality level among neighboring regions. For example, regional governance, such as collaboration and partnerships between neighboring regions, city-county consolidation, or joint city-suburb strategies, should be emphasized to implement effective outcomes to reduce air quality pollution for the health of people and the environment.

Smart growth strategies, such as mixed land use measures,[46] preservation of open space, transit-oriented development (TOD) including public transit system, and walkable communities, can contribute to improved air quality level in metropolitan areas. Detailed strategies at a regional or state level, such as statewide growth management programs or region-wide growth management programs, should be implemented to tackle negative spillover issues from sprawling development.

Compact development, along with a focus on public transit systems linking clusters of employment, services, research and development (R&D), and environment-friendly industry (Newman & Kenworthy, 1989, 1999; Newton, 2000; Masnavi, 2000), can contribute to improved air quality level.

[46] The effect of mixed land use measures on changes in air quality level in metropolitan areas in terms of a good-fit spatial error model in this dissertation is not statistically significant, but shows a potential effect of mixed land use measures on improved air quality level in metropolitan areas in terms of a spatial lag model, as in Table 4-16.

5.3 Limitations

5.3.1 The Ecological Bias of Spatial Aggregation

When data are aggregated in terms of the mean statistic, it may produce the loss of information leading to lack of identification of parameters at a micro level (or census tracts). We used the mean of a 3-year air quality index at the county level for 1990, 2000, and 2006. The weighted average air quality index of the county-level centroid values may fail to identify the weighted average air quality index of the point-level monitoring sites in the county producing the ecological fallacy problem caused by the difference of spatial units at the census tract level and county levels (Anselin, 2002; Wakefield & Lyons, 2010). In spatial regression models, similarly, the aggregate of the county-level spatial lag terms or the county-level spatial weights will not be consistent with the aggregate of the census tract-level spatial lag terms or the census tract-level spatial weights due to the ecological fallacy problem caused by spatial aggregation (Anselin, 2002; Wakefield & Lyons, 2010). If data are available at a micro level, we can reduce the loss of information caused by the spatial aggregation from a tract level to a county level and provide valid inference for reliable spatial data.

5.3.2. Specification Problems of OLS and Spatial Regression Models

The spatial regression model specifications using an error and lag model are well fitted to the OLS regression estimation in a cross-sectional data for 1990, 2000, and 2006, thus identifying an increase in the log likelihood, a decrease in the AIC, and a decrease in the SIC. These specifications are displayed in Models 3 and 4 in the 1990 maximum likelihood (ML) spatial regression estimation, Models 7 and 8 in the 2000 ML estimation, and Models 11 and 12 in the 2006 ML estimation. However, the alternative model specifications for including new independent variables (or omitted variables) or different spatial weights matrix are needed to create a better fitted model. The reason is that the spatial lag and error models still show specification problems in terms of the high value of non-normality uncovered through Breusch-Pagan test and the strong significance of heteroscedasticity from the Jarque-Bera test (see diagnostics tests in Tables 4-7, 4-12, and 4-17). Additionally, we need to be cautious of limits to interpretation of spatial regression models related with a pseudo-R^2 produced by the presence of spatial dependence in the spatially lagged dependent variables and the regression error terms among neighboring observations.

5.4 Further Studies

5.4.1 Reduction of Methodological Biases of OLS and Spatial Regression Models

Methodological biases in OLS and spatial regression models, such as omitted variable and errors in variables, can bring out misleading results to predict parameter estimates, because the magnitude, significance, and sign of parameter estimates are affected by the presence of the strength of spatial autoregressive coefficients (ρ & λ) for the OLS regression estimation. To reduce these biases of parameter estimates in the OLS and spatial regression estimates, a further work will be needed to specify robust spatial regression models using different spatial weights matrix or different k-nearest functional forms.

5.4.2 Reflection of Air Pollution to Public Health

Future work will be required to reflect impacts of metropolitan spatial structure on population exposures to air quality concentration, particularly relating to minorities or low-income residents in compact and sprawling regions, under an environmental injustice dimension. To perform this analysis, particularly for population exposures to air quality concentration, we can calculate the weighted 3-year average number of days for "unhealthy for sensitive groups" or more categories that correspond to an AQI value above 100, for the county and for 2004-2006 or 1999-2001, using the county-level weighted average AQI grading system developed by the American Lung Association (2012, pp.40-42). The county-level weighting AQI factors will be assigned to each AQI category based on the defined ranges identified by the EPA. The factors can reflect the higher levels of air pollution threatening public health by assigning the following weight factors: a factor of 1 for "Unhealthy for Sensitive Groups"; a factor of 1.5 for "Unhealthy"; a factor of 2 for "Very Unhealthy"; a factor of 2.5 for "Hazardous" or "Very Hazardous." For example, one county had 7 days of "Unhealthy for Sensitive Groups" with a factor of 1 for 2004, 3 days of "Unhealthy" with a factor of 1.5 for 2005, and 2 days of "Hazardous" with a factor of 2.5 for 2006 in the level of AQI. The weighted average an AQI level over 3 years for 2006 for the county would be 5.5, or [(7 days * 1 factor) + (3*1.5) + (2*2.5)]/3. This level reflects that air quality for the county remained unhealthy over the 3 years. The findings can provide policy insights of land development for minorities or low-income population in compact and sprawling regions.

5.4.3 Regional Variations of Effects of MSS on Air Quality

At a regional level, further work will be needed to detect what factors determine air quality in compact and sprawling regions. This work is to compare spatial variation in metropolitan spatial structures and their respective air quality for compact and sprawling regions. This analysis will be required to investigate whether which form is desirable from a regionalist's view of sustainable and environmentally sound design. The practical framework of sustainable urban form matrix proposed by Jabareen (2006) will be employed to test the effects of metropolitan spatial structure on air quality between compact and sprawling regions in the U.S. The findings can provide policy insights of urban structure for a sustainable future.

REFERENCES

Adams, C. T. (1997). *The Philadelphia Experience*. *Annals of the American Academy of Political and Social Science, 551*, 222–234.

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

American Farmland Trust. (2007). *Farming on the Edge: Sprawling Development Threatens America's Best Farmland*. Retrieved from http://www.farmland.org/resources/fote/default.asp.

American Lung Association. *The State of the Air 2012*. Retrieved from http://www.stateoftheair.org/2012/assets/state-of-the-air2012.pdf.

Anas, Alex. (1990). Taste Heterogeneity and Urban Spatial Structure: The Logit Model and Monocentric Theory Reconsidered. *Journal of Urban Economics*, *28*, 318-335.

Anas, Alex, Arnott, Richard, & Small, Kenneth A. (1998). Urban Spatial Structure. In Wassmer, Robert W. (Eds.), *Reading in Urban Economics: Issues and Public Policy* (pp. 65-106). Malden, MA.: Blackwell Publishers Inc.

Anderson, Nathan B., & Bogart, William T. (2001). The Structure of Sprawl: Identifying and Characterizing Employment Centers in Polycentric Metropolitan Areas. *American Journal of Economics and Sociology, 60*(1), 147-169.

Anderson, James R., Hardy, Ernest E., Roach, John T., & Witmer, Richard E. (1976). *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*. U.S. Geological Survey Professional Paper 964. Washington, DC: United States Government Printing Office. Retrieved from http://landcover.usgs.gov/pdf/anderson.pdf.

Anderson, William P., Kanaroglou, Pavlos S., & Miller, Eric J. (1996). Urban Form, Energy and the Environment: A Review of Issues, Evidence and Policy. *Urban Studies, 33*(1), 7-35.

Anselin, Luc. (1988). *Spatial Econometrics: Methods and Models*. Norwell, MA: Kluwer Academic Publishers.

______. (1995). Local Indicators of Spatial Association – LISA. *Geographical Analysis, 27*, 93-115.

______. (2002).Under the Hood issues in the Specification and Interpretation of Spatial Regression Models. *Agricultural Economics, 27*, 247-267.

______. (2003). Spatial externalities: spatial multipliers and spatial econometrics. *International Regional Science Review, 26*, 153-166.

______. (2005). *Exploring Spatial Data with GeoDa: A Workbook*. Urbana, IL.: Center for Spatially Integrated Social Science, University of Illinois, Urbana-Champaign.

______. (2009). Spatial Regression. In A. Stewart Fotheringham and Peter A. Rogerson (Eds.), *The SAGE Handbook of Spatial Analysis* (pp. 255-275). Thousand Oaks, CA.: SAGE Publications Inc.

Anselin, Luc, & Bera, A.K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In A. Ullah and D. Giles (Eds.), *Handbook of Applied Economic Statistics* (pp. 237-289). New York: Marcel Dekker.

Baerwald, T.J. (1982). Land Use Change in Suburban Clusters and Corridors. *Transportation Research Record, 891*, 7-12.

Bank of America. (1996). *Beyond Sprawl: New Patterns of Growth to Fit the New California*. Bank of America Corporation.

Barro, Robert. (1991). Economic Growth in a Cross-Section of Countries. Quarterly *Journnal of Economics, 106*(2), 407-443.

Barro, Robert J., & Sala-i-Martin, Xavier. (1992). Convergence. *Journal of Political Economy, 100*(2), 223-251.

Baumol, W.J. (1986). Productivity Growth, Convergence and Welfare: What does the Long Run Data Show? *American Economic Review, 76*, 1072-1085.

Beatley, Timothy. (2000) *Green Urbanism: Lessons from European Cities.* Washington, DC: Island Press.

Beatley, Timothy, & Newman, Peter. (2009). *Green Urbanism Down Under: Learning from Sustainable Communities in Australia*. Washington, DC: Island Press.

Beckmann, Martin J., & Papageorgiou, Yorgos Y. (1989). Heterogeneous Tastes and Residential Location. *Journal of Regional Science, 29,* 317-323.

Berliant, Marcuse, & Wang, Ping. (May 2004). *Dynamic Urban Models: Agglomeration and Growth*. Unpublished Paper. Retrieved from http://pingwang.wustl.edu/journal_publications/Handbook5-05-04.pdf.

Beeson, Patricia E., DeJong David N., & Troesken, Werner. (2001). Population Growth in U.S. Counties, 1840-1990. *Regional Science and Urban Economics, 31,* 669-699.

Benfield, F. Kaid, Raimi, Matthew D., & Chen, Donald D.T. (1999). *Once There Were Greenfields: How Urban Sprawl is Undermining America's Environment, Economy and Social Fabric*. Washington, DC: Surface Transportation Policy Project, National Resources Defense Council.

Berry, Brian J.L. et al. (1974). *Land Use, Urban Form and Environmental Quality*. Chicago, IL: The University of Chicago.

Bish, Robert L., & Vincent Ostrom (1973). *Understanding Urban Government: Metropolitan Reform Reconsidered.* Washington. DC: American Enterprise Institute for Public Policy Research.

Blair, J. P., & Zhang, Z. (1994). Ties that Bind Revisited. *Economic Development Quarterly, 8,* 73-77.

Bollens, S. A. (1997a). Fragments of Regionalism: The Limits of Southern California Governance. *Journal of Urban Affairs, 19*(1), 105–122.

______. (1997b). Concentrated Poverty and Metropolitan Equity Strategies. *Stanford Law and Policy Review,* 8(2), 11-23.

Bourne, Larry S. (Eds.). (1982). *Internal Structure of the City: Readings on Urban Form, Growth, and Policy*. New York, NY: Oxford University Press.

Boyne, G. A. (1992). Local Government Structure and Performance: Lessons for America? *Public Administration, 70*(3), 333-357.

Bradbury, Katharine L., Downs, Anthony, & Small, Kenneth A. (1982). *Urban Decline and the Future of American Cities*. Washington, DC: The Brookings Institution.

Brasington, David M., & Hite, Diane. (2005). Demand for Environmental Quality: A Spatial Hedonic Analysis. *Regional Science and Urban Economics, 35,* 57-82.

Brenner, Neil. (2002). Decoding the Newest Metropolitan Regionalism in the USA: A Critical Review. *Cities*, *19*(1), 3-21.

Brown, M. (1999). Reconceptualizing Public and Private in Urban Regime Theory: Governance in Aids Politics. *International Journal of Urban and Regional Research*, *23*(1), 70–87.

Brueckner, Jan K. (1983). The Economics of Urban Yard Space: An 'Implicit-Market' Model for Housing Attributes. *Journal of Urban Economics, 13*, 216-234.

______. (1987). The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills Model. In Edwin S. Mills (Eds.), *Handbook of Regional and Urban Economics, Vol. II: Urban Economics* (pp. 821-845). North-Holland, Amsterdam: Elsevier Science Publishers.

______. (1997). Infrastructure Financing and Urban Development: The Economics of Impact Fees. *Journal of Public Economics*, *66*, 383-407.

______. (2000). Urban Sprawl: Diagnosis and Remedies. *International Regional Science Review*, *23*(2), 160-171.

______. (2001). Urban Sprawl: Lessons from Urban Economics. *Brookings-Wharton Papers on Urban Affairs, 2001*, 65-97.

Brueckner, Jan K., & Fansler, David A. (1983). The Economics of Urban Sprawl: Theory and Evidence on the Spatial Sizes of Cities. *The Review of Economics and Statistics, 65*(3), 479-482.

Burchell, Robert W. (2000). Costs and Benefits of Alternative Development Patterns: Sprawl versus Smart Growth. In Lincoln Institute of Land Policy (Eds.), *Metropolitan Development Patterns: 2000 Annual Roundtable* (pp. 40-49). Cambridge, MA.

______. (2002). *Costs of Sprawl-2000*. Transportation Research Board. Report No. 74. Washington, DC: National Academy Press.

Burchell et al. (1998). *Costs of Sprawl-Revisited*. Transportation Research Board. Report No. 39. Washington, DC: National Academy Press. Retrieved from http://ntl.bts.gov/lib/21000/21500/21538/PB99124216.pdf.

Burchfield et al. (2005). *Causes of Sprawl: A Portrait from Space*. Unpublished Paper.

Burgess, Ernest W. (1925). The Growth of the City: An Introduction to a Research Project. In Robert Park, Ernest W. Burgess , & R.D. McKenzie (Eds.), *The City* (pp. 47-62). Chicago: The University of Chicago Press.

Büttler, H. (1981). Equilibrium of a Residential City, Attributes of Housing, and Land-Use Zoning. *Urban Studies, 18*, 23-39.

Byun, Pillsung, & Esparza, Adrian X. (2005). A Revisionist Model of Suburbanization and Sprawl: The Role of Political Fragmentation, Growth Control, and Spillovers. *Journal of Planning Education and Research, 24*, 252-264.

Calthorpe, Peter. (1993). *The Next American Metropolis: Ecology, Community, and the American Dream*. New York: Princeton Architectural Press.

Calthorpe, Peter, & Fulton, William. (2001). *The Regional City: Planning for the End of Sprawl*. Washington, DC: Island Press.

Capello, Robert. (2007). *Regional Economics*. New York, NY: Routledge.

Carlino, Gerald. (1985). Declining City Productivity and the Growth of Rural Regions: A Test of Alternative Explanations. *Journal of Urban Economics, 18*(1), 11-27.

Carlino, G. A., & Mills, E. (1987). The Determinants of County Growth. *Journal of Regional Science*, *27*, 39-54.

Carr, J.B., & Feiock, R.C. (1999). Metropolitan Government and Economic Development. *Urban Affairs Review, 34*(3), 476-488.

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

______. (2003). Growth at the Fringe: The Influence of Political Fragmentation in United States Metropolitan Areas. *Papers in Regional Science, 82*, 475-499.

Carruthers, J. I., & Ulfarsson, G. F. (2002). Fragmentation and Sprawl: Evidence from Interregional Analysis. *Growth and Change, 33*, 312-340.

______. (2008) Does "Smart Growth" Matter to Public Finance? *Urban Studies, 45*, 1791 – 1823.

Cervero, Robert. (1989). Jobs-Housing Balancing and Regional Mobility. *Journal of the American Planning Association, 55*(2), 136-150.

______. (1996). Jobs-Housing Balance Revisited: Trends and Impacts in the San Francisco Bay Area. *Journal of the American Planning Association*, *62*(3), 492-511.

Cervero, R., & Wu, K.L. (1998). Sub-centering and Commuting: Evidence from the San Francisco Bay Area, 1980–90, *Urban Studies, 35*(7), 1059–1076.

Chen, Y., Irwin, E.G., & Jayaprakash, C. (2009). Dynamic Modeling of Environmental Amenity-Driven Migration with Ecological Feedbacks. *Ecological Economics, 68*, 2498–2510.

Chinitz, Benjamin. (1961). Contrasts in Agglomeration: New York and Pittsburgh. *American Economic Review: Papers and Proceedings*, *51*(2), 279-289.

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

Clark, Lara P., Millet, Dylan B., & Marshall, Julian D. (2011). Air Quality and Urban Form in U.S. Urban Areas: Evidence from Regulatory Monitors. *Environmental Science & Technology*, *45*(16), 7028-7035.

Claval, Paul. (2007). The Nature of Cities and the Analysis of Their Cultural Problems. *Tijdschrift voor economische en sociale geografie*, *98*(2), 153-164.

Cliff, A., & Ord, J. K. (1972). Testing for Spatial Autocorrelation among Regression Residuals. *Geographical Analysis, 4*, 267-284.

Congress for the New Urbanism (CNU). (1999). *Charter of the New Urbanism*. New York, NY: McGraw-Hill.

Cooke, Timothy. (1983). Testing a Model of Intraurban Firm Location. *Journal of Urban Economics*, *13*, 257-282.

Cox, Webdell, & Utt, Joshua. (2004). The Costs of Sprawl Reconsidered: What the Data Really Show. *Backgrounder, 1770*.

Craig, S.G., & Ng, P. (2001). Using Quantile Smoothing Splines to Identify Employment Subcenters in a Multicentric Urban Area. *Journal of Urban Economics, 49,* 100-120.

Cutsinger, J., Galster, G., Wolman, H., Hanson, R., & Towns, D. (2005). Verifying the Multi-dimensional Nature of Metropolitan Land Use: Advancing the Understanding and Measurement of Sprawl. *Journal of Urban Affairs*, *27* (3), 235 – 259.

Davis, Jonathan S. (2002). Urban Regime Theory: A Normative-empirical Critique. *Journal of Urban Affairs, 24*(1), 1-17.

Dear, M., & Flusty, S. (1998). Postmodern Urbanism. *Annals of the Association of American Geographers, 88,* 50-72.

Delong, J.B. (1988). Productivity Growth, Convergence and Welfare: Comment. *American Economic Review, 78*, 1138-1154.

Desmet, Klaus, & Fafchamps, Marcel. (2005). Changes in the Spatial Concentration of Employment across US Counties: A Sectoral Analysis 1972-2000. *Journal of Economic Geography, 5*(3), 261-284.

______. (2006). Employment Concentration across U.S. Counties. *Regional Science and Urban Economics, 36*(4), 482-509.

DiGaetano, Alan, & Klemanski**,** John S. (1993). Urban Regime Capacity: A Comparison of Birmingham, England, and Detroit, Michigan. *Urban Affairs Review, 29*(1), 54-83.

Dreier, Peter, Mollenkopf, John, & Swanstrom, Todd (2001). *Place Matters: Metropolitics for the Twenty-First*. Kansas, KS: University Press of Kansas.

Dodge, W. R. (1996). *Regional excellence*. Washington, DC: National League of Cities

Douglass, H. (1925). *The Suburban Trend*. New York: Century Co.

Downs, Anthony. (1994). *New Visions for Metropolitan America*. Washington, DC: The Brookings Institution.

Duany, A., & Plater-Zyberk, E. (Winter 1992). The Second Coming of the American Small Town. *The Wilson Quarterly, 16*(1), pp. 19-50.

Dubin, R. (1988). Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Errors. *Review of Economics and Statistics, 70,* 466-474.

Edmonston, Barry. (1975). *Population Distribution in American Cities*. Lexington, MA: D.C. Health.

Elkin, T., McLaren, D., & Hillman, M. (1991). *Reviving the City: Towards Sustainable Urban Development*. London: Friends of the Earth.

Ellison, G., & Glaeser, E.L. (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy, 105,* 889-927.

Emison, Gerald Andrews. (2001). The Relationships of Sprawl and Ozone Air Quality in United States' Metropolitan Areas. *Regional Environmental Change, 2*, 118-127.

Erickson, Rodney A. (1986). Multinucleation in Metropolitan Economies. *Annals of the Association of American Geographers, 76*, 331-346.

Ewing, Reid H. (1994). Characteristics, Causes, and Effects of Sprawl: A Literature Review. *Environmental and Urban Studies, 21*(2), 1-15.

______. (1997). Is Los Angeles-Style Sprawl Desirable? *Journal of the American Planning Association*, *63*(1), 107-129.

______. (2000). The Future of Land Development. In Lincoln Institute of Land Policy (Eds.), *Metropolitan Development Patterns: 2000 Annual Roundtable* (pp. 66-71).

Ewing, R., Pendall, R., & Chen, D. (2002). *Measuring Sprawl and Its Impact.* Washington, DC: Smart Growth America/U.S. Environmental Protection Agency.

______. (2003). Measuring Sprawl and Its Transportation Impacts. *Journal of the Transportation Research Board, 1831*, 175-183.

Faggian Alessandra., & McCann, Peter. (2009). Universities, Agglomerations and Graduate Human Capital Mobility. *Journal of Economic and Social Geography, 100*(2), 210-223.

Faggian, Alessandra, Olfert, M. Rose, & Partridge, Mark D. (2011). Inferring Regional Well-being from Individual Revealed Preferences: The 'Voting with Your Feet' Approach. Cambridge *Journal of Regions, Economy and Society, 5*(1), 163-180.

Felsenstein, Daniel. (2002). Do High Technology Agglomerations Encourage Urban Sprawl? *The Annals of Regional Science, 36*, 663-682.

Ferguson, M., Ali, K., Olfert, M. R., & Partridge, M. D. (2007). Voting with their Feet: Jobs Versus Amenities. *Growth and Change, 38*, 77-110.

Fingleton, B., & Le Gallo, J. (2010). Endogeneity in a Spatial Context: Properties of Estimators. In Páez, Antonio, Le Gallo, Julie , Buliung, Ron N., & Dall'erba, Sandy (eds.), *Progress in Spatial Analysis. Methods and Application* (pp. 59-73). New York, NY: Springer.

Firey, Walter. (1946). *Social Aspects to Land-Use Planning in the Country-City Fringe: The Case of Flint, Michigan*. East Lansing: Michigan State Agricultural Experiment Station. Special Bulletin No. 339.

______. (1947). *Land Use in Central Boston*. Cambridge, MA: Harvard University Press.

Fishman, Robert. (1987). *Bourgeois Utopias: The Rise and Fall of Suburbia*. New York, NY: Basic Books.

______. (1990). America's New City: Megalopolis Unbound. *Wilson Quarterly, 14*(1), 24-45.

______. (2000). The American Metropolis at Century's End: Past and Future Influences. *Housing Policy Debate, 11*(1), 199-213.

Fitzgerald, Joan, & Leigh, Nancey Green. (2002). *Economic Revitalization: Cases and Strategies for City and Suburbs.* Thousand Oaks, CA: Sage Publications.

Fotheringham, A.S., Brunsdon, C., & Charlton, M. (2000). *Quantitative Geography: Perspectives on Spatial Data Analysis*. London, UK: Sage Publications.

Fotheringham, A. Stewart, & Rogerson, Peter A. (Eds.). (2000). *The SAGE Handbook of Spatial Analysis*. Thousand Oaks, CA: SAGE Publications Inc.

Frank L.D., Sallis J.F., Conway, T.L., Chapman, J.E., Saelens, B.E., & Bachman, W. (2006). Many Pathways from Land Use to Health: Associations between Neighborhood Walkability and Active Transportation, Body Mass Index, and Air Quality. *Journal of the American Planning Association, 72*(1), 75–87.

Frey, William. (1979). Central City White Flight: Racial and Nonracial Causes. *American Sociological Review, 44*, 425-448.

Frisken, F. (1991). The Contributions of Metropolitan Government to the Success of Toronto's Public Transit System: An Empirical Dissent from the Public Choice Paradigm. *Urban Affairs Quarterly, 27*, 268–292.

Frisken, F., & Norris, D. F. (2001). Regionalism Reconsidered. *Journal of Urban Affairs, 23*(5), 467–478.

Fry, J.A., Coan, M.J., Homer, C.G., Meyer, D.K., & Wickham, J.D. (2009). *Completion of the National Land Cover Database (NLCD) 1992–2001 Land Cover Change Retrofit product*. Open-File Report 2008–1379. Reston, VA: U.S. Geological Survey. Retrieved from http://pubs.usgs.gov/of/2008/1379/pdf/ofr2008-1379.pdf.

Fujita, M., & Ogawa, H. (1982). Multiple Equilibria and Structural Transition of Non-Monocentric Urban Configurations. *Regional Science and Urban Economics, 12*, 161-196.

Fujita, M., & Thisse, J.F. (1996). Economics of agglomeration. *Journal of the Japanese and International Economics, 10*, 339-378.

Fujita, M., & Thisse, J.F. (2002). *Economics of Agglomeration*. Cambridge, UK: Cambridge University Press.

Fulton, W., Pendall, R., Nguyen, M., & Harrison, A. (2001). *Who Sprawls Most? How Growth Patterns Differ across the U.S.* Washington, DC: The Brookings Institution.

Galster, G., Hanson, R., Ratcliffe, M., Wolman, H., Coleman, S., & Freihage, J. (2001). Wrestling Sprawl to the Ground: Defining and Measuring an Elusive Concept. *Housing Policy Debate, 12(*4), 681-719. Retrieved from http://www.fanniemaefoundation.org/programs/hpd/pdf/HPD_1204_galster.pdf.

Garreau, Joel. (1991). *Edge City: Life on the New Frontier*. New York: Anchor Books.

George, Darren, & Mallery, Paul. (2011). *SPSS for Windows Step by Step: A Simple Guide and Reference 18.0 Update* (11th eds.). Boston, MA: Pearson Education, Inc.

Gillham, Oliver. (2002). *The Limitless City: A Primer on the Urban Sprawl Debate*. Washington, DC: Island Press.

Giuliano, Genevieve, & Small, Kenneth. A. (1991). Subcenters in the Los Angeles Region. *Regional Science and Urban Economics, 21*(2), 163-182.

Glaeser, Edward L. (1994). Cities, Information and Economic Growth. *Cityscape, 1*, 9-48.

Glaeser, Edward L., Kallal, Hedi D., Scheinkman, José A., & Shleifer, Andrei. (1992). Growth in Cities. *The Journal of Political Economy*, *100*(6), Centennial Issue, 1126-1152.

Glaeser, Edward L., Scheinkman, J., & Shleifer, A. (1995). Economic Growth in a Cross-section of Cities. *Journal of Monetary Economics, 36*(1), 117-143.

Glaeser, Edward L., & Kahn, Matthew E. (2001). Decentralized Employment and the Transformation of the American City. *NBER Working Paper No. 8117*. Cambridge, MA.: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w8117.pdf.

______. (2003). *Sprawl and Urban Growth*. NBER Working Paper No. 9733. Cambridge, MA.: National Bureau of Economic Research. Retrieved from http://www.owlnet.rice.edu/~econ461/papers/glaeser4.pdf.

Glaeser, Edward L., Kolko, Jed, & Saiz, Albert. (2001). Consumer City. *Journal of Economic Geography, 1*(1), 27-50.

Gordon, Peter, Kumar, Ajay, & Richardson, Harry W. (1989). The Influence of Metropolitan Spatial Structure on Commuting Time. *Journal of Urban Economics, 26*(2), 138-151.

Gordon, Peter, & Richardson, Harry W. (1997). Are Compact Cities a Desirable Planning Goal? *Journal of American Planning Association*, *63*(1), 95-106.

Gordon, Peter, Richardson, Harry W., & Jun, Myung Jin. (1991). The Commuting Paradox: Evidence from the Top Twenty. *Journal of the American Planning Association*, *57*, 416-420.

Gore, Charles. (1984). *Regions in Question*. New York: Methuen.

Greene, D.L. (1980). Recent Trends in Urban Spatial Structure. *Growth and Change, 11*, 29-40.

Griffith, D. A. (1981) Modeling Urban Population Density in a Multi-centered City, *Journal of Urban Economics*, *9*, 298–310.

Gujarati, Damodar N. (2003). *Basic Econometrics*. New York, NY: McGraw-Hill Companies, Inc.

Hamilton, David K. (1999). *Governing Metropolitan Areas: Responses to Growth and Change*. New York, NY: Garland Publishing.

______. (2013). Measuring the Effectiveness of Regional Governing Systems: A Comparative Study of City Regions in North America. New York, NY: Springer.

Hamilton, David K., Miller, David Y., & Paytas, Jerry. (2004). Exploring the Horizontal and Vertical Dimensions of the Governing of Metropolitan Regions. *Urban Affairs Review, 40*(2), 147-182.

Hammond, George W., & Tosun, Mehmet S. (2009). The Impact of Local Decentralization on Economic Growth: Evidence from U.S. Counties. *IZA Discussion Paper No. 4574*. Bonn, Germany: Institute for the Study of Labor (IZA). Retrieved from http://ftp.iza.org/dp4574.pdf.

Hanlon, B.F., Vicino, T.J., & Short, J.R. (2006). The New Metropolitan Reality in the US: Rethinking the Traditional Model. *Urban Studies, 43*(12), 2129-2143.

Hanson, Gordon H. (2000). Firms, Workers, and the Geographic Concentration of Economic Activity. In Clark, Gordon L., Feldmann, Maryann P., & Gertler, Meric S. (Eds.), *The Oxford Handbook of Economic Geography* (pp. 477-494). Oxford: Oxford University Press.

Harris, Richard. (1994). Chicago's Other Suburbs. *The Geographical Review, 84*, 394-410.

Harris, C.D., & Ullman, E.S. (1945). The Nature of Cities. *The Annals of the American Academy of Political and Social Science, 242*, 7-17.

Hartwick, J., Schweizer, U., & Varaiya, P. (1976). Comparative Static Analysis of a Residential Economy with Several Classes. *Journal of Economic Theory*, *13*, 396-413.

Hawley, Amos H. (1950). *Human Ecology: A Theory of Community Structure*. New York: Ronald Press.

Hayden, Dolores. (2004). *Building Suburbia: Green Fields and Urban Growth, 1820-2000*. New York: Vintage Books, a division of Random House, Inc.

Heikkila, E., Gordon, P., Kim, J.I., Peiser, R.B., Richardson, H.W., & Dale-Johnson, D. (1989). What Happened to the CBD-Distance Gradient? Land Values in a Polycentric City. *Environment and Planning A*, *21*, 221-232.

Henderson, J. Vernon. (1986). Efficiency of Resource Usage and City Size. *Journal of Urban Economics*, *19*, 47-70.

______. (2003). Marshall's Scale Economies. *Journal of Urban Economics, 53*, 1-28.

Henderson, J.V., Kuncoro, A., & Turner, M. (1995). Industrial Development in Cities. *Journal of Political Economy, 103*(5), 1067-1090.

Hess et al. (Summer 2001). Just What is Sprawl, Anyway? *Carolina Planning, 26*(2), 11-26.

Hill, E., Wolman, H., & Ford III, C. C. (1995). Can Suburbs Survive without their Central Cities? Examining the Suburban Dependence Hypothesis. *Urban Affairs Review, 31*, 147-74.

Howard, Ebenezer. (1898). *To-Morrow: A Peaceful Path to Real Reform*.

Howell-Moroney, Michael. (2008). The Tiebout Hypothesis 50 Years Later: Lessons and Lingering Challenges for Metropolitan Governance in the 21st Century. *Public Administration Review, 68*(1), 97-109.

Hoyt, Homer. (1939). *The Structure and Growth of Residential Neighborhood in American Cities*. Washington, DC: U.S. Government Printing Office.

Hurd, R.M. (1903). *Principles of City Land Value*. New York: The Record and Guide.

IBI Group et al. (1990). *Greater Toronto Area Urban Structure Concept Study*. Prepared for the Government of Ontario, Office for the Greater Toronto Area.

Ihlanfeldt, Keith R. (1995). The Importance of the Central City to the Regional and National Economy: A Review of the Arguments and Empirical Evidence. *Cityscape, 1*(2), 125-150.

Irwin, E., & Bockstael, N. (2002). Interacting Agents, Spatial Externalities, and the Evolution of Residential Land Use Patterns. *Journal of Economic Geography, 2*(1), 31–54.

Jabareen, Yosef Rafeq. (2006). Sustainable Urban Forms: Their Typologies, Models, and Concepts. *Journal of Planning Education and Research, 26*, 38-52.

Jackson, Kenneth. (1985). *Crabgrass Frontier: The Suburbanization of the United States*. New York, NY: Oxford University Press.

Jacobs, Jane. (1961). *The Death and Life of Great American Cities*. New York: Random House and Vintage Books.

______. (1969). *The Economy of Cities*. New York: Random House and Vintage Books.

______. (2001). The Death and Life of Great American Cities. In Scott Campbell and Susan S. Fainstein (Eds), *Readings in Planning Theory* (pp. 103-120). Malden, MA: Blackwell Publishers Inc.

Jaffe, A., Trajtenberg, M., & Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics, 108*, 577-598.

Jarque, C.M., & Bera, A.K. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review, 55*, 163-172.

Jeong, Moon-Gi, & Feiock, Richard C. (2006). Impact Fees, Growth Management, and Development: A Contractual Approach to Local Policy and Governance. *Urban Affairs Review, 41*(6), 749-768.

Johnson, M. (2001). Environmental Impacts of Urban Sprawl: A Survey of the Literature and Proposed Research Agenda. *Environment and Planning, 33*, 717-735.

Kahn, Matthew E. (2001). Does Sprawl Reduce the Black/White Housing Consumption Gap? *Housing Policy Debate, 12*(1), 77-86.

______. (2006). *Green Cities: Urban Growth and the Environment*. Washington, DC: The Brookings Institution.

Kahn, Matthew E., & Schwartz, Joel. (2008). Urban Air Pollution Progress Despite Sprawl: The "Greening" of the Vehicle Fleet. *Journal of Urban Economics, 63,* 775-787.

Kain, John F. (1964). *The Effect of the Ghetto on the Distribution of Nonwhite Employment in Urban Areas*. Washington, DC: National Academy Press.

______. (1968). Housing Segregation, Negro Employment, and Metropolitan Decentralization. *Quarterly Journal of Economics, 82*(2), 175-197.

______. (2004). A Pioneer's Perspective on the Spatial Mismatch Literature. *Urban Studies, 41*(1), 7-32.

Katz, Bruce. (July 2002). *Smart Growth: The Future of the American Metropolis?* (LSE STICERD Research Paper No. CASE058). Retrieved from http://sticerd.lse.ac.uk/dps/case/cp/casepaper58.pdf.

Klosterman, Richard E. (1985). Arguments for and Against Planning. *Town Planning Review, 56*(1), 5-20.

Krugman, Paul. (1991). *Geography and Trade*. Cambridge: MIT Press.

Kruse, Kevin. (2005). *White Flight: Atlanta and the Making of Modern Conservatism*. Princeton, NJ: Princeton University Press.

Kruse, K.M., & Sugrue, T.J. (Eds.). (2006). *The New Suburban History*. Chicago, IL: University of Chicago Press.

Kunstler, James. (1993). *The Geography of Nowhere: The Rise and Decline of America's Man-made Landscape*. New York, NY: Touchstone.

Lang, Robert E. (2003). *Edgeless Cities: Exploring the Elusive Metropolis*. Washington DC: The Brookings Institution.

Lang, Robert E., & Hornburg, Steven. (1997). Planning Portland Style: Pitfalls and Possibilities. *Housing Policy Debate, 8*(1), 1–10.

Le Corbusier. (1935). *La Ville Radieuse (The Radiant City).* Boulogne-Seine.

Lee, Bumsoo, & Gordon, Peter. (February 2007). *Urban Spatial Structure and Economic Growth in US Metropolitan Areas*. Paper presented at the conference on Western Regional Science Association 46th Annual Meeting in the Newport Beach, California.

Lee, Sugie, & Leigh, Nancey Green. (2005). The Role of Inner-ring Suburbs in Metropolitan Smart Growth Strategies. *Journal of Planning Literature, 19*(3), 330-346.

Leinberger, Christpher B. (2007). *The Option of Urbanism: Investing in a New American Dream*. Washington, DC: Island Press.

Lesage, James. (1997). Regression Analysis of Spatial Data. *Journal of Regional Analysis and Policy, 27*(2), 83-94.

______. (1999). Spatial Econometrics: *The Web Book of Regional Science*. Morgantwon, WV: Regional Research Institute, West Virginia University.

Lesage, James, & Pace, R. Kelly. (2009). *Introduction to Spatial Econometrics*. Boca Raton, FL: Chapman & Hall/CRC.

Lincoln Institute of Land Policy. (2000). *Metropolitan Development Patterns: 2000 Annual Roundtable*. Cambridge, MA.

Lockwood, Charles, & Leinberger, Christpher B. (1988). Los Angeles Comes of Age. *Atlantic Monthly*, 43-52.

Lopez, Russ, & Hynes, H. Patricia. (2003). Sprawl in the 1990s: Measurement, Distribution, and Trends. *Urban Affairs Review, 38*(3), 325-355.

Lorenz, Max O. (1905). Methods for Measuring the Concentration of Wealth. *Journal of the American Statistical Association*, *9*, 209-219.

Lucas, R. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics, 22*, 3-42.

Lucy, William H., & Phillips, David L. (1995). Why Some Suburbs Thrive. *Planning, 61*(6), 20-21.

______. (2001). *Suburbs and the Census: Patterns of growth and decline*. Washington, DC: The Brookings Institution.

______. (2003). Suburbs: Patterns of Growth and Decline. In Bruce Katz and Robert E. Lang (Eds.), *Redefining Urban and Suburban America: Evidence from Census 2000* (pp. 117-136). Washington, DC: The Brookings Institution.

Lynch, Kevin. (1981). *A Theory of Good City Form*. Cambridge, MA: MIT Press.

Malpezzi, S. & Guo, W. (2001). *Measuring "Sprawl:" Alternative Measures of Urban Form in US Metropolitan Areas*. Madison, WI: Center for Urban Land Economics Research.

Marlay, Matthew. C., & Gardner, T. K. (August 2010). *Identifying Concentrations of Employment in Metropolitan Areas*. Paper presented at the annual meeting of the American Sociological Association Annual Meeting, Atlanta, GA. Unpublished manuscript.

Marquez, Leorey O., & Smith, Nariida C. (1999). A Framework for Linking Urban Form and Air Quality. *Environmental Modelling and Software with Environmental Data News*, *14*(6), 541-548.

Marshall, Alfred. (1890). *Principles of Economics*. London: Macmillan.

Maryland Office of Planning. (1997). *What You Need to Know About Smart Growth and Neighborhood Conservation*. Baltimore, MD.

Masnavi, Mohammad-Reza. (2000). The New Millennium and the New Urban Paradigm: The Compact City in Practice. In Katie Williams, Elizabeth Burton and Mike Jenks (Eds.), *Achieving Sustainable Urban Form* (pp. 64-73). New York, NY: E & FN Spon.

Massey, D.S., & Denton, N.A. (1988). The Dimensions of Residential Segregation. *Social Forces, 67*, 281-315.

______. (1993). *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard University Press.

McDonald, John F. (1987). The Identification of Urban Employment Subcenters. *Journal of Urban Economics, 21*(2), 242-258.

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

McDonald, John F., & McMillen, Daniel P. (2007). *Urban Economics and Real Estate*. Malden, MA: Blackwell Publishing.

McGinnis, Michael D. (Eds.) (1994). *Polycentricity and Local Public Economies: Readings from the Workshop in Political Theory and Policy Analysis*. Ann Arbor, MI: The University of Michigan Press.

McKenzie, R.D. (1925). The Ecological Approach to the Study of the Human Community. In Robert Park, Ernest W Burgess, & R.D. McKenzie (Eds.), *The City*. Chicago: The University of Chicago Press.

McLoughlin, Brian. (1991). Urban Consolidation and Urban Sprawl: A Question of Density. *Urban Policy and Research, 9*(3), 148-156.

McMillen, Daniel P. (2001). Nonparametric Employment Subcenter Identification. *Journal of Urban Economics, 50*, 448-473.

______. (2006). Testing for monocentric city. In R. Arnott and Daniel P. McMillen (Eds.), *A Companion to Urban Economics* (pp. 128-140). Malden, MA: Blackwell Publishing.

McMillen, Daniel P. & McDonald, John F. (1998). Suburban Subcenters and Employment Density in Metropolitan Chicago. *Journal of Urban Economics, 43*(2), 157-180.

McMillen, Daniel P., & Smith, Stefani C. (2003). The Number of Subcenters in Large Urban Areas. *Journal of Urban Economics, 53*, 321-338.

Mees, Paul. (2010). *Transport for Suburbia: Beyond the Automobile Age*. Sterling, VA: Earthscan.

Mieszkowski, Peter, & Mills, Edwin S. (1993). The Causes of Metropolitan Suburbanization. *Journal of Economic Perspectives, 7*(3), 135-147.

Miller, David Y. (2002). *The Regional Governing of Metropolitan America*. Cambridge, MA: Westview Press.

Mills, Edwin S. (1967). An Aggregative Model of Resource Allocation in a Metropolitan Area. *American Economic Review: Papers and Proceedings, 57*, 197-210.

_____. (1972). *Studies in the Structure of the Urban Economy*. Baltimore, MD: The Johns Hopkins Press.

______. (2000). A thematic history of urban economic analysis. *Brookings-Wharton Papers on Urban Affairs, 1*, 1-52.

Mills, Edwin S., & Tan, Jee Peng. (1980). A Comparison of Urban Population Density Functions in Developed and Developing Countries. *Urban Studies, 17*(3), 313-321.

Moore, Curtis. (2001). *Smart Growth and the Clean Air Act*. Washington, DC: The Northeast-Midwest Institute.

Moran, P. A. P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika 37* (1), 17–23.

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

Nelessen, Anton. (1994). *Visions for a New American Dream: Process, Principles, and an Ordinance to Plan and Design Small Communities*. Chicago, IL: The American Planning Association.

Nelson, Arthur C. (1994). Oregon's Urban Growth Boundary Policy as a Landmark Planning tool. In Carl Abbott, Deborah Howe and Sy Adler (Eds.), *Planning the Oregon Way: A Twenty Year Evaluation* (pp. 25-47). Corvallis, Oregon: Oregon State University Press.

______. (2000). Regulations to Improve Development Patterns. In Lincoln Institute of Land Policy (Eds.), *Metropolitan Development Patterns: 2000 Annual Roundtable* (pp. 72-79).

______. (2001). Exclusionary Housing, Urban Sprawl, and Smart Growth. *Georgia State University Law Review. 17*(4): 1087-1102

Nelson, Arthur C., & Dawkins, Casey J. (2004). *Urban Containment in the United States*. Chicago, IL: The American Planning Association Press.

Nelson, Arthur C., & Duncan, James B. (1995). *Growth Management Principles and Practices*. Chicago, IL: The American Planning Association Press.

Nelson, Arthur C., & Peterman, D. R. (2000). Does Growth Management Matter? The Effect of Growth Management on Economic Performance. *Journal of Planning Education and Research, 19*(3), 277-285.

Nelson, Arthur C., & Poster, Kathryn A. (1999). Metropolitan Governance Structure and Income Growth. *Journal of Urban Affairs, 21*(3), 309-324.

Neuman, Michael. (2005). The Compact City Fallacy. *Journal of Planning Education and Research 25*, 11-26.

Newman, Peter. W. & Kenworthy, Jeff. R. (1989). *Cities and Automobile Dependence: An International Sourcebook.* Aldershot, UK: Gower Publishing.

______. (1992). Is there a Role for Physical Planners? *Journal of the American Planning Assiciation, 58*(3), 353-361.

______. (1999). *Sustainability and Cities: Overcoming Automobile Dependence*. Washington, DC: Island Press.

______. (2000). Sustainable Urban Form: The Big Picture. In Katie Williams, Elizabeth Burton and Mike Jenks (Eds.), *Achieving Sustainable Urban Form* (pp. 109-120). New York, NY: E & FN Spon.

Newton, Peter. W. (1997). *Re-Shaping Cities for a More Sustainable Future: Exploring the Nexus between Urban Form, Air Quality and Greenhouse Gas Emissions*. Melbourne, Australia: Australian Housing and Urban Research Institute.

______. (2000). Urban Form and Environmental Performance. In Katie Williams, Elizabeth Burton and Mike Jenks (Eds.), *Achieving Sustainable Urban Form* (pp. 46-53). New York, NY: E & FN Spon.

Nowak, David J. (2006). Institutionalizing Urban Forestry as a "Biotechnology" to Improve Environmental Quality. *Urban Forestry & Urban Greening*, *5*, 93-100.

Noyelle, Thierry, & Stanback, Thomas Jr. (1983). *The Economic Transformation of American Cities*. Totowa, NJ: Rowan & Allanheld.

Oates, Wallace E. (1972). *Fiscal Federalism*. New York, NY: Harcourt Brace Jovanovich.

Ogburn, W.F. (1937). *Social Characteristics of Cities*. Chicago, IL: International City Managers' Association.

Olewiler, Nancy. (2006). Environmental Sustainability for Urban Areas: The Role of Natural Capital Indicators. *Cities*, *23*(3), 184-195.

Orfield, Myron. (1997). *Metropolitics: A Regional Agenda for Community and Stability*. Washington, DC: Brookings Institution Press and Cambridge, MA: Lincoln Institute of Land Policy.

_____. (2002). *American Metropolitics: The New Suburban Reality*. Washington, DC: Brookings Institution.

Ostrom, Vincent. (1974). The Study of Federalism at Work. *Publius: The Journal of Federalism, 4*(4), 1-18.

Ostrom, Vincent, Tiebout, Charles, & Warren, Robert. (1961). The Organization of Government in Metropolitan Areas: A Theoretical Inquiry. *American Political Science Review, 55*, 831-842.

_____. (2002). *American Metropolitics: Social Segregation and Sprawl*. Washington, D.C.: The Brookings Institution.

Pace, R. Kelley, & Lesage, James. (2010). Spatial Econometrics. In A.E. Gelfand, P. Diggle, M. Fuentes and P. Guttorp (Eds.), *Handbook of Spatial Statistics*. Boca Raton, FL: Chapman & Hall/CRC.

Paehlke, Robert. (2003). Environmental Sustainability and Urban Life in America. In Norman J. Vig and Michael E. Kraft (5th Eds.), *Environmental policy: New Directions for the Twenty-First Century* (pp. 57-77). Washington, DC: CQ Press, a division of Congressional Quarterly Inc.

Papageorgiou, Yorgos Y., & Pines, David. (1989). The Exponential Density Function: First Principles, Comparative Statics, and Empirical Evidences. *Journal of Urban Economics, 26*(2), 264-268.

Park, Robert E. (1915). The City: Suggestions for the Investigation of Human Behavior in the City Environment. *The American Journal of Sociology, 20*(5), 577-612.

Parks, Roger B., & Oakerson, Ronald J. (1993). Comparative Metropolitan Organization: Service Production and Governance Structures in St. Louis and Allegheny Counties. *Publius: The Journal of Federalism 23* (1), 19-39.

Pierce, N. R., Johnson, C., & Hall, J. S. (1993). *Cititstates: How Urban America Can Prosper in a Competitive World*. Washington, DC: Seven Locks.

Porter, Michael. (1990). *The Competitive Advantage of Nations*. New York: Free Press.

Porter, Douglas R. (1997). *Managing Growth in America's Communities*. Washington, DC: Island Press.

Portney, Kent. (2003). *Taking Sustainable Cities Seriously: Economic Development, the Environment, and Quality of Life in American Cities*. Boston, MA: MIT Press.

Queen, S.A., & Thomas, L.F. (1939). *The City: A Study of Urbanism in the United States*. New York, NY: McGraw Hill.

Rajamani, J, Bhat, C. R., Handy, S., Knaap, G., & Song, Y. (2003). *Assessing the Impact of Urban Form Measures in Nonwork Trip Mode Choice After Controlling for Demographic and Level-of Service Effects*. Presented at the 2003 Transportation Research Board Annual Meeting.

Rao, S. T., Zalewsky, E., & Zurbenko, I. G. (1995). Determining Temporal and Spatial Variations in Ozone Air Quality. *Journal of Air and Waste Management Association, 45*, 57-61.

Rappaport, Jordan. (1998). *Local Economic Growth*. Harvard University mimeograph.

______. (2007). Moving to Nice Weather. *Regional Science and Urban Economics, 37,* 375-398.

Rauch, James E. (1993). Does History Matter Only When It Matter Little?: The Case of City-Industry Location. *Quarterly Journal of Economics, 108*(434), 843-867.

Real Estate Research Corporation (1974). *The costs of sprawl: Environmental and economic costs of alternative residential patterns at the Urban Fringe*. Washington, DC: US Government Printing Office.

Ridley, Matt, & Low, Bobbi S. (2001). Can Selfishness Save the Environment? In Scott Campbell and Susan S. Fainstein (Eds.), *Readings in Planning Theory* (pp. 198-212). Malden, MA: Blackwell Publishers Inc.

Robson, Arthur J. (1977). The Effect of Urban Structure on the Concentration of Pollution. *Urban Studies, 14,* 89-93.

Romer, Paul. (1986). Increasing Returns and Long-Run Growth. *Journal of Political Economy, 94*, 1002-1037.

Roseland, Mark. (Eds.). (1997). *Eco-City Dimensions: Healthy Communities Healthy Planet*. Gabriola Island, BC: New Society Publishers.

______. (4th Eds.). (2012). *Toward Sustainable Communities: Solutions for Citizens and their Governments.* Gabriola Island, BC: New Society Publishers.

Rusk, David. (1993). *Cities without Suburbs*. Washington, DC: Woodrow Wilson Center Press.

______. (2003). *Cities without Suburbs: a Census 2000 Update*. Washington, DC: Woodrow Wilson Center Press.

Savageau, David. (7th Eds.). (2007). *PlacesRated Almanac: The Class Guide for Finding Your Best Places to Live in America*. Washington, DC: Places Rated Books LLC.

Savitch, H. V., Collins, D., Sanders, D., & Markham, J. P. (1993). Ties that Bind: Central Cities, Suburbs, and the New Metropolitan Region. *Economic Development Quarterly, 7,* 341–357.

Savitch, H. V., & Vogel, Ronald K. (*2001)*. Paths to New Regionalism. *State and Local Government Review, 32*(3), 158-168.

______. (2004). Suburbs without a City: Power and City-County Consolidation. *Urban Affairs Review, 39*(6), 758-790.

Schnore, L. (1963). The Social and Economic Characteristics of American Suburbs. *The Sociological Quarterly, 4*, 122-134.

Schweitzer, L., & Zhou, J. (2010). Neighborhood Air Quality, Respiratory Health, and Vulnerable Populations in Compact and Sprawled Regions. *Journal of the American Planning Association, 76*(3), 363-371.

Shannon, Claude E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*, 379-423, 623-656.

Sharpe, William, & Wallach, Leonoard. (1994). Bold New City or Built-Up 'Burb'? *American Quarterly, 46*(1), 1-30.

Schuler, R.E. (1974). The Interaction between Local Government and Urban Residential Location. *American Economic Review*, *64*, 682-696.

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

Simon, C., & Nardinelli, C. (1996). The Talk of the Town: Human Capital, Information and the Growth of English Cities, 1861-1961. *Explorations in Economic History, 33*(3), 384-413.

Siy, Eric, Koziol, Leo, & Rollins, Darcy. (August 2001). *The State of the States: Assessing the Capacity of States to Achieve Sustainable Development through Green Planning*. San Francisco, CA: Resource Renewal Institute. Retrieved from www.rri.org.

Small, K. A., & Song, S.F. (1992). 'Wasteful' Commuting: A Resolution, *Journal of Political Economy, 100*, 888–898.

Song, Yan, & Knaap, Gerrit-Jan. (2004). Measuring urban form: Is Portland Winning the War on Sprawl? *Journal of the American Planning Association, 70*(2), 210-225.

Squires, Gregory D. (2002). Urban Sprawl and Uneven Development of Metropolitan America. In G. Squires (Eds.), *Urban Sprawl: Causes, Consequences and Policy Responses*. Washington, DC: The Urban Institute Press.

Stanback, Thomas. (2002). *The Transforming Metropolitan Economy*. New Brunswick, NJ: Center for Urban Policy Research, Rutgers University.

Stansel, Dean. (2005). Local Decentralization and Local Economic Growth: A Cross-sectional Examination of US Metropolitan Areas. *Journal of Urban Economics, 57*, 55–72.

Stead, Dominic, Williams, Jo, & Titheridge, Helena. (2000). Land Use, Transport and People: Identifying the Connections. In Katie Williams, Elizabeth Burton and Mike Jenks (Eds.), *Achieving Sustainable Urban Form* (pp. 174-186). New York, NY: E & FN Spon.

Stephens, G. Ross, & Wikstrom, Nelson. (2000). *Metropolitan Government and Governance: Theoretical Perspectives, Empirical Analysis, and the Future*. New York, NY: Oxford University Press.

Stone, Clarence N., & Sanders, Heywood. T. (1987). *The Politics of Urban Development*. Lawrence, KS: University of Kansas Press.

Stone, Brian Jr. (2005). Urban Heat and Air Pollution: An Emerging Role for Planners in the Climate Change Debate. *Journal of the American Planning Association 71*(1), 13–25.

______. (2008). Urban Sprawl and Air Quality in Large US Cities. *Journal of Environmental Management, 86*, 688-698.

Spivey, Christy. (2008). The Mills-Muth Model of Urban Spatial Structure: Surviving the Test of Time? *Urban Studies, 45*(2), 295-312.

Thurmaier, K., & Wood, C. (2002). Interlocal Agreements as Overlapping Social Networks: Picket-fence Regionalism in Metropolitan Kansas City. *Public Administration Review, 62*(5), 585-596.

Thurston, L., & Yezer, A. (1994). Causality in the Suburbanization of Population and Employment. *Journal of Urban Economics, 35*, 105-118.

Tiebout, Charles M. (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy, 64*(5), 416-424.

Torrens, Paul M. (2008). A Toolkit for Measuring Sprawl. *Applied Spatial Statistics, 1*, 5-36.

Tsai, Yu-Hsin. (2005). Quantifying Urban Form: Compactness versus ‘Sprawl’. *Urban Studies, 42*(1), 141-161.

U.S. Bureau of Census. *Census 1990 (SF1 and SF3), Census 2000 (SF1 and SF3), and 2005-2009 American Community Survey.* Retrieved from http://factfinder2.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t.

______. (2003). *Metropolitan and Micropolitan Statistical Areas Main*. Washington, D.C. Retrieved from http://www.census.gov/population/metro/.

______. (2003). *Statistical Abstract of the United States*. Washington, DC: U.S. Government Printing Office.

U.S. Bureau of Labor Statistics. *Green Job Definition*. Retrieved from http://www.bls.gov/green/green_definition.pdf.

U.S. Bureau of Transportation Statistics. *The 2000 Census Transportation Planning Package (CTPP)*. Retrieved from http://www.census.gov/geo/www/cob/tz_metadata.html#gad.

U.S. Department of Agriculture. (2001). *National Resources Inventory: Highlights*. Washington, D.C.

U.S. Department of Housing and Urban Development. (1999). *State of the Cities Report*. Retrieved from www.hud.gov/soc/soc.html.

U.S. Environmental Protection Agency. *Air Quality Index (AQI) data from Air Quality System (AQS)*. Retrieved from http://www.epa.gov/ttn/airs/airsaqs/.

U.S. Geological Survey. *National Land Cover Databases (NLCD1992, NLCD2001, NLCD2006) from The Multi-Resolution Land Characterization (MRLC) consortium*. Washington, DC. Retrieved from http://www.mrlc.gov/index.php.

Vicino, Thomas J. (2008). The Political History of a Postwar Suburban Society Revisited. *History Compass, 6*(1), 364-388.

von Thünen, Johann Heinrich. (1826). *Der Isolierte Staat in Beziehung auf Landwritschaft und Nationalokonomie*. Hamburg: F. Perthens.

Wachter, Susan M. (2000). Cities and Regions: Findings from the 1999 State of the Cities Report. In Lincoln Institute of Land Policy (Eds.), *Metropolitan Development Patterns: 2000 Annual Roundtable* (pp. 20-25). Cambridge, MA.

Wakefield, Jonathan, & Lynos, Hillary. (2010). Spatial Aggregation and the Ecological Fallacy. In A. Gelfand, P.J. Diggle, P. Guttorp, & M. Fuentes (Eds.), *Handbook of Spatial Statistics*. Chapman & Hall/CRC Press.

Wallis, A. D. (1995). *Inventing Regionalism*. Denver, CO: National Civic League.

Ward, R.C. (1987). The Impact of Metropolitan Fragmentation on Economic Development. *Urban Resources, 4*, SL1-SL4.

Ward, Michael D., & Gleditsch, Kristian Skrede. (2008). *Spatial Regression Models*. Thousand Oaks, CA: Sage Publications, Inc.

WCED. (1987). *Our Common Future*. The World Commission on Environment and Development. Oxford and New York: Oxford University Press.

Wheaton, William C. (1974). A Comparative Static Analysis of Urban Spatial Structure. *Journal of Economic Theory, 9*(2), 223-237.

Wheeler, Stephen M. (2000). Planning for Metropolitan Sustainability. *Journal of Planning Education and Research, 20*(2), 133-145.

______. (2002). The New Regionalism: Key Characteristics of an Emerging Movement. *Journal of the American Planning Association*, *68*(3), 267-278.

White, M.J. (1976). Firm Suburbanization and Urban Subcenters. *Journal of Urban Economics, 3,* 323-343.

Wiersinga, W. (1997). *Compensation as a Strategy for Improving Environmental Quality in Compact Cities*. Amsterdam: Bureau SME.

Williams, Katie. (2000). Does Intensifying Cities Make them More Sustainable? In Katie Williams, Elizabeth Burton and Mike Jenks (Eds.), *Achieving Sustainable Urban Form* (pp. 30-45). New York, NY: E & FN Spon.

Wolman, G., Galster, G., Hanson, R., Ratcliffe, M., Furdell, K., & Sarzynski, A. (2005). The Fundamental Challenge in Measuring Sprawl: Which Land Should Be Considered? *The Professional Geographer*, *57* (1), 94 – 105.

Wood, C. (2006). Scope and Patterns of Metropolitan Governance in Urban America: Probing the Complexities in the Kansas City Region. *American Review of Public Administration, 36,* 337-353.

Wright, Frank Lloyd. (1932). *The Disappearing City*. New York, NY: W.F. Payson.

Yang, C., & Fujita, M. (1983). Urban Spatial Structure with Open Space. *Environment and Planning A, 15*, 67-84.

Yin, M., & Sun, J. (2007). The Impacts of State Growth Management Programs on Urban Sprawl in the 1990s. *Journal of Urban Affairs*, *29*(2), 149-179.

1000 Friends of Oregon. (1997). *Making the Connections: Summary of LUTRAQ*. Portland, Oregon. Retrieved from http://www.friends.org/sites/friends.org/files/reports/LUTRAQ%20Volume%207%20Making%20the%20Connections%20-%20A%20summary%20of%20the%20LUTRAQ%20project%2C%201997.pdf.