

**URBAN FORM AND TRAVEL PATTERNS AT THE REGIONAL SCALE  
CONSIDERING POLYCENTRIC URBAN STRUCTURE**

A Dissertation

by

YOUNG-JAE YI

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

DOCTOR OF PHILOSOPHY

August 2012

Major Subject: Urban and Regional Sciences

Urban Form and Travel Patterns at the Regional Scale

Considering Polycentric Urban Structure

Copyright 2012 Young-Jae Yi

**URBAN FORM AND TRAVEL PATTERNS AT THE REGIONAL SCALE  
CONSIDERING POLYCENTRIC URBAN STRUCTURE**

A Dissertation

by

YOUNG-JAE YI

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

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee, Ming-Han Li

Chanam Lee

Committee Members, Josias Zietsman

Raghavan Srinivasan

Head of Department, Forster Ndubisi

August 2012

Major Subject: Urban and Regional Sciences

## ABSTRACT

Urban Form and Travel Patterns at the Regional Scale

Considering Polycentric Urban Structure. (August 2012)

Young-Jae Yi, B.E., Urban Engineering, University of Seoul;

M.U.P., University of Michigan at Ann Arbor

Co-Chairs of Advisory Committee: Dr. Ming-Han Li  
Dr. Chanam Lee

Increasing concerns about climate change have attracted global interests in reducing auto travel. Regional average vehicle miles traveled (VMT) vary across the urbanized areas in the U.S., suggesting a potential influence of development patterns on greenhouse gas emission.

To explore the contribution of development control to driving reduction at the regional scale, this dissertation estimated impacts of urban form on two travel outcomes at the metropolitan scale: daily vehicle miles traveled (DVMT) per capita and daily transit passenger miles (DPMT) per capita. To overcome major problems of previous studies, i.e., lack of generalizability and multicollinearity, a cross-sectional analysis of 203 U.S. urbanized areas was conducted, using directed acyclic graph and structural equation modeling.

A literature review revealed gaps in the previous research: while individual-level behavioral studies have identified distance from the center as the most influential factor on VMT, regional-level studies have not reflected this relationship and failed to deliver effective implications for land use policies. A method to identify regional centers was evaluated to appropriately measure polycentric urban structure of contemporary metropolitan areas. The evaluation found that lower density cutoff, wider reference area, and equal treatment between central business district (CBD) and subcenters yielded

better performance in McMillen's two-stage nonparametric method. Results also showed that for polycentric areas, the use of a polycentric model produced a better model fit than the monocentric model.

Major findings of this dissertation include 1) higher regional concentration, greater local density and less road supply per capita lowered VMT, and 2) higher local density and more transit supply per capita increased PMT. These results imply that different approaches to development control are needed for different sustainable transportation goals – intensifying regional centers such as infill developments for VMT reduction, and compact neighborhood development approaches, such as transit oriented development for transit promotion.

However, CBD has a limited capacity and indiscreet compact developments at the urban fringe can lead to decentralization from the regional perspective, and consequently result in increased VMT. This study suggests polycentricism as a potential solution for the contradictory development principle. By allowing dispersion and concentration at the same time, urban form control at the regional level will be more beneficial than conventional local-level control.

**DEDICATION**

To Seoki and my loves: Ji-young, Jung-min, and Jung-Hwan.

To father and mother-in-law, wishing your health.

To father-in-law and my mom..., you are always in my mind.

## ACKNOWLEDGEMENTS

My sincere appreciation to my committee co-chairs, Dr. Ming-Han Li and Dr. Chanam Lee. Dr. Li, you are my true mentor as an academic advisor as well as a life advisor. I have always respected your advice. I still feel regretful that I could not embrace your expert area for my dissertation research but will search for ways to comprise what I learned from you in my future research. Dr. Lee, thanks for your dedication to my study and kind considerations throughout the course of this research. I could finish this long way, thanks to you.

I appreciate my committee members, Dr. Josias Zietsman and Dr. Raghavan Srinivasan for their constructive guidance and support for this research. I thank my colleagues at Texas Transportation Institute, Mr. Jett McFalls, Mr. Derrold Foster, Ms. Beverly Storey, Mr. Rodney Jackson, and Mr. Alex Ferrazas for their support and consideration of my works and family. Thanks to my friends, Dr. Chan-Yong Sung, Dr. Sangyun Lee, and Dr. JaeSu Lee for their encouragement and joyful discussion.

Finally, my deep appreciation to my parents for their support and to my wife, Seoki for her patience and love ever.

**NOMENCLATURE**

CBD	Central Business District
CTPP	Census Transportation Planning Package
DAG	Directed Acyclic Graph
DCBD	Distance to CBD
DPMT	Daily Passenger Miles of Travel
DVMT	Daily Vehicle Miles of Travel
EPA	Environmental Protection Agency
FHWA	Federal Highway Administration
FTA	Federal Transit Administration
NTD	National Transit Database
PMT	Passenger Miles Traveled
SEM	Structural Equation Modeling
TOD	Transit Oriented Development
VMT	Vehicle Miles Traveled

## TABLE OF CONTENTS

	Page
ABSTRACT .....	iii
DEDICATION .....	v
ACKNOWLEDGEMENTS .....	vi
NOMENCLATURE .....	vii
TABLE OF CONTENTS .....	viii
LIST OF FIGURES .....	xi
LIST OF TABLES .....	xiii
1. INTRODUCTION .....	1
1.1 Background .....	1
1.2 Objectives .....	3
1.3 Significance .....	4
1.4 Dissertation Structure .....	5
2. IMPACTS OF URBAN FORM ON VEHICLE MILES TRAVELED – A LITERATURE REVIEW .....	6
2.1 Introduction .....	6
2.2 Micro Approach .....	6
2.2.1. Unit of analysis .....	6
2.2.2. Data source .....	7
2.2.3. Urban form measures .....	7
2.2.4. Control variables .....	8
2.2.5. Findings and implications .....	9
2.3 Macro Approach .....	11
2.3.1. Unit of analysis .....	11
2.3.2. Data source .....	11
2.3.3. Urban form measures .....	12
2.3.4. Findings and implications .....	13
2.4 Discussions: Gaps and Links between Micro and Macro Approaches	15
2.5 Conclusion .....	18

	Page
3. EVALUATING THE PERFORMANCE OF CENTER IDENTIFICATION METHOD .....	19
3.1 Outline .....	19
3.2 Introduction .....	19
3.3 Previous Research on Center Identification Methods .....	20
3.4 Methodology .....	23
3.4.1. Data and study areas .....	23
3.4.2. Subcenter identification .....	25
3.4.3. Density function .....	27
3.4.4. Performance evaluations .....	28
3.5 Results and Discussions .....	29
3.5.1. Performance changes made by using polycentric model .....	30
3.5.2. Attributes affecting center identification performance .....	31
3.5.3. Empirical application test .....	33
3.6 Conclusions .....	36
4. INFLUENCES OF URBAN FORM ON VEHICLE MILES TRAVELED AT THE METROPOLITAN SCALE .....	38
4.1 Outline .....	38
4.2 Introduction .....	38
4.3 Previous Research .....	40
4.4 Research Approach .....	43
4.5 Methodology .....	44
4.5.1. Data .....	45
4.5.2. Study areas .....	45
4.5.3. Variables and measurements .....	46
4.5.4. Analytical methods .....	50
4.5.5. Descriptive statistics .....	52
4.6. Results and Implications .....	55
4.6.1. Directed acyclic graph (DAG) .....	55
4.6.2. Structural equation modeling (SEM) .....	57
4.6.3. Structural equation modeling (SEM) without control variables .....	60
4.7 Discussion .....	62
4.8 Summary and Conclusion .....	64

	Page
5. INFLUENCES OF URBAN FORM ON TRANSIT PASSENGER MILES TRAVELED AT THE METROPOLITAN SCALE .....	67
5.1 Outline .....	67
5.2 Introduction .....	67
5.3 Research Approach .....	71
5.4 Methodology .....	72
5.4.1. Data .....	72
5.4.2. Study areas .....	73
5.4.3. Variables and measurements .....	74
5.4.4. Analytical methods .....	78
5.4.5. Descriptive statistics .....	80
5.5 Results and Implications .....	82
5.5.1. Directed acyclic graph (DAG) .....	82
5.5.2. Structural equation modeling (SEM) .....	85
5.6 Summary and Conclusion .....	89
6. SUMMARY AND CONCLUSIONS .....	92
6.1 Summary .....	92
6.2 Limitations .....	93
6.3 Conclusions .....	94
REFERENCES .....	98
VITA .....	108

## LIST OF FIGURES

	Page
Figure 3.1 Study areas for the center identification study.....	24
Figure 3.2 Actual job densities and smoothed density estimates .....	26
Figure 3.3 A simple operation to translate polycentricity using monocentric model (center identification study) .....	28
Figure 3.4 Correlation between population gradients of the best-performing polycentric model (25%/0.10/Distance) and monocentric gradients, 41 polycentric urbanized areas in the U.S. in 2000.....	35
Figure 3.5 Change of monocentric gradient by the strength of subcenter .....	36
Figure 4.1 Conceptual framework for the DVMT study.....	44
Figure 4.2 Study areas for the DVMT study.....	46
Figure 4.3 A simple operation to translate polycentricity using monocentric model (DVMT study) .....	49
Figure 4.4 Directed acyclic graphs of variables influencing vehicle miles traveled per capita among 203 U.S. urbanized areas, 2002 .....	57
Figure 4.5 Structural equation modeling of simplified urban form-VMT models....	61
Figure 4.6 Summary of findings from the DVMT study .....	66
Figure 5.1 Conceptual framework for the DPMT study .....	71
Figure 5.2 Study areas for the DPMT study.....	73
Figure 5.3 A simple operation to translate polycentricity using monocentric model (DPMT study).....	77
Figure 5.4 Directed acyclic graphs of factors influencing passenger miles traveled per capita among 203 U.S. urbanized areas, 2002 .....	83
Figure 5.5 Structural equation model of DPMT per capita (log transformed data) ..	86

	Page
Figure 5.6 Correlation between local density and population density gradient by population size.....	89
Figure 5.7 Summary of findings from the DPMT study .....	90

## LIST OF TABLES

	Page
Table 2.1 Partial elasticity of VMT with respect to built environment variables .....	10
Table 2.2 Comparison of micro and macro approaches .....	16
Table 3.1 Variables used for empirical tests and data sources .....	24
Table 3.2 Center identification results by each monocentric and polycentric model	30
Table 3.3 Performances of eight polycentric urban structure models for 41 polycentric urbanized areas in the U.S. in 2000 .....	31
Table 3.4 Paired-samples mean differences and t-values in three center identification performance benchmarks by matching parameter values .....	32
Table 3.5 Results of empirical application test using a stepwise regression model of vehicle miles of travel (VMT) .....	33
Table 4.1 Variable descriptions, measures, and data sources and geographic unit of data .....	47
Table 4.2 Descriptive statistics of variables .....	54
Table 4.3 Standardized correlations (Pearson's R) among variables .....	55
Table 4.4 Effects on DVMT/c and model fit statistics, estimated using SEM .....	58
Table 4.5 Effects on DVMT/c in simplified urban form-VMT models, estimated using SEM .....	61
Table 5.1 Variable descriptions, measures, and data sources and geographic unit of data .....	75
Table 5.2 Descriptive statistics of variables .....	81
Table 5.3 Standardized correlations (Pearson's R) among variables .....	82
Table 5.4 Diagnosis of spurious relationships using DAG search algorithms .....	84
Table 5.5 Effect size (elasticity) on DPMT/c, estimated by SEM .....	87

## 1. INTRODUCTION

### 1.1. Background

Increasing concerns about climate change have attracted global interests in reducing auto travel as a way to reduce greenhouse gas. The transportation sector producing 30% of CO<sub>2</sub> is the second largest cause of greenhouse gas after electricity generation in the U.S. (2008 data reported by U.S. EPA 2011). Moreover, vehicular greenhouse gas emissions are growing at a faster rate than overall greenhouse gas emissions. Regional average vehicle miles traveled (VMT) varies significantly across the urbanized areas in the U.S., suggesting a potential influence of development patterns on greenhouse gas emission. According to Highway Statistics 2000, daily VMT per capita of 401 U.S. urbanized areas greatly varies from 6 to 56 miles with an average of 23 miles.

It has long been recognized that the amount of aggregated regional-level travels is determined by the interaction between urban form and transportation infrastructure (Handy 1996). In many urbanized areas developed after World War II in the U.S., strict zoning regulations (e.g., exclusive land uses, minimum lot size requirement, etc.) have caused dispersed urban form and increased road constructions. Automobile-oriented policies and enormous investments on highway construction have promoted migrations toward suburbs and consequently led to low density, dispersed development patterns. Dispersed developments in turn have intensified dependency on private vehicles (Mieszkowski and Mills 1993). This cyclical urban phenomenon, called sprawl, has been blamed for many urban and environmental problems, including increased fuel use and emission, congestion, deforestation, and degraded ecological quality (Burchfield et al. 2006). Remedies have been sought for managing and guiding the urban development toward more compact urban form. Various sustainable development concepts (e.g., compact city, new urbanism, transit oriented development, traditional neighborhood development, growth management, and smart growth) have emerged. Many U.S.

---

This dissertation follows the style of *Growth and Change*.

governments have adopted such policy tools to fight against sprawl (CPDR 2008; U.S. SCCST 2009).

Despite the supposedly strong relationship between urban form and travel, over 60 years of studies have not provided an agreement in the effects of urban form control on driving reduction. Many studies found compact neighborhoods are conducive to sustainable travel behaviors (Ewing and Cervero 2001; 2010; Holtzclaw 1994; Holtzclaw 1990; Frank and Engelke 2005; Kockelman 1997; Cervero and Kockelman 1997), while others argued that polycentric regional urban structure is a more sustainable development pattern (Jenks et al. 1996). Some found that dense and/or mixed use developments reduced driving distance or increased transit use (Frank and Engelke 2005; Holtzclaw 1994; Kockelman 1997). Others denied these effects and further insisted that paying more attention to gas price control or the technological improvement in fuel efficiency would be more effective for emission control (Pisarski 2009). This lack of agreements in theories and empirical findings weakened the justification of plausibly strong urban form-travel relationship, and therefore led to discourage development control efforts in many growing metropolitan areas.

The inconsistent findings are attributed to a lack of standardized methodology (Susan Handy 1996; Crane 2000; Gomez-Ibanez 1996). The methodological limitations include the lack of generalizability, collinearity problems among explanatory factors, and inconsistent urban form measures. Due to the scarcity of data, generalizability is the most frequent problem in many micro approach studies that are carried out in one or several metropolitan areas. The high quality travel survey data for the individual-level studies are available only in a limited number of large areas (e.g., New York, NY; Los Angeles, CA; San Francisco, CA; Seattle, WA; Oregon, PO. etc.). However, each of these areas has a unique regional setting and is different in its size, population and job distribution patterns, road and transit supply, and relative travel costs among travel alternatives, compared to other areas. The magnitude of urban form impacts estimated in a unique regional condition is therefore not applicable to other regions with different conditions.

Collinearity among explanatory variables is an inherent problem in most travel studies. Places with a low population density tend to have dispersed land uses, more roads per capita, and consequently greater VMT. Transit use tends to be higher in denser areas, and such areas tend to be populated by more low-income people with a lower level of automobile access (Taylor and Fink 2002). At the regional scale, regions with greater population tend to have a greater density, more compact neighborhoods, and more supply of transportation infrastructure. Unraveling the relationships among multiple variables correlated with each other is a significant methodological challenge applicable to most urban form-travel studies (Crane 2000; Gomez-Ibanez 1996).

Urban form measures have been inconsistent in previous studies. Four Ds (i.e., density, diversity, design and destination accessibility) is a popular classification system of urban form attributes in micro-level studies (Ewing and Cervero 2001). While extensive literature review studies found destination accessibility is the most influential attribute determining VMT (Ewing and Cervero 2001; 2010), many studies excluded this variable. The destination accessibility at the individual scale is comparable to the centeredness at the regional collective scale because trip destinations tend to be concentrated in regional centers. Few macro studies have used an effective measure of centeredness. Also, density is the most frequent measure of urban form in macro studies; however, this heavily aggregated-level density is conceptually different from the local-level density used in micro studies. The local-scale density relevant variable has been rarely treated in previous macro studies.

## **1.2. Objectives**

To explore the contribution of development control to auto driving reduction at the regional scale, the primary objective of this dissertation is to estimate impacts of urban form on two aggregate travel outcomes: daily vehicle miles traveled (DVMT) per capita and daily transit passenger miles (DPMT) per capita, in U.S. metropolitan areas. Two secondary objectives are established to facilitate the methodological specification necessary to carry out the primary objective. One is to identify methodological limitations of previous macro-scale studies. This task focuses on identifying gaps and

correspondences between micro and macro studies and assisting the development of the methodological framework for this study. The other secondary objective is to identify appropriate measures for polycentric urban structure which is an important urban form attribute influencing the regional aggregate level of travel but has been neglected in many previous studies. This dissertation accomplishes these objectives by searching for a valid method to identify regional centers including both the central business district and other subcenters.

### **1.3. Significance**

This study can contribute to urban form-travel knowledge at the conceptual, methodological, and policy implication perspectives. At the conceptual perspective, this study provides a systematic review of micro-level behavioral studies and macro-level phenomenal studies, highlighting the differences and agreements between them. Despite the distinctive tradition in interests, assumptions, methodology and implications, the two types of studies have frequently been interpreted within the same conceptual and theoretical context. The mixed interpretations of study findings can confuse the understanding of urban form-travel relationships. This study distinguishes both approaches from the methodological perspectives, while maintaining conceptual connections between them.

This study addresses the methodological limitations that are described in the background section as possible reasons for the disagreement among previous studies. Few studies have been able to address all the limitations. A total of 203 urbanized areas were used in this dissertation to gain more generalizable results applicable to mid to small size metropolitan areas in the U.S. Directed acyclic graph (DAG) and structural equation modeling (SEM) are employed to disentangle the complex relationships among correlates of VMT. The urban form measures used in this study help fill the gap between micro and macro studies. Particularly, the gradient-based measures effectively represent the degree of concentration of population and jobs toward urban employment centers and give a straightforward implication to regional urban form controls. Gradient-based

measures reflect the finding from micro studies that the accessibility to regional centers has a strong effect on individual VMT. This study improves the gradient measures to further account for the polycentric urban structure.

Two contributions are made from the policy implication perspective. While previous studies tend to deliver implications of urban form controls at the neighborhood scale, this study supports that regional scale controls of development patterns (e.g., density gradient and poverty gradient from regional employment centers) can contribute to auto travel reduction. Also, this study evaluates urban form impacts in consideration of the polycentric urban structure. Many contemporary metropolitan areas have multiple regional centers significantly restructuring trip origins and destinations and affecting the collective travel distance within a region. Nevertheless, most of the previous studies were based on the monocentric assumption.

#### **1.4. Dissertation Structure**

This dissertation is composed of six sections. Section 1 describes the background, objectives, significances and the organization of the dissertation. Section 2 provides the literature review in regard to the impacts of urban form on VMT and transit use. The review focuses on identifying the gaps between individual-based micro and region-based macro scale studies. Searching for an appropriate method to quantify regional urban form, Section 3 tests a regional center identification method, named McMillen's two-stage non-parametric method, with multiple combinations of associated parameters. The best performing parameter combination identified from the test is used to identify the regional centers and measure the regional urban form, which are used in the subsequent sections. Section 4 examines the influences of urban form on regional average VMT. Section 5 explores the impacts of urban form on the regional average of transit passenger miles traveled (PMT). Sections 3, 4, and 5 are structured as an independent journal manuscript, including introduction, research framework, methodology, results, and conclusions. Section 6 summarizes findings from each section and delivers overall conclusions.

## 2. IMPACTS OF URBAN FORM ON VEHICLE MILES TRAVELED – A LITERATURE REVIEW

### 2.1. Introduction

Studies examining urban form impact on VMT and transit use can be classified into two discrete approaches – micro and macro. The micro approach is people oriented and focuses on urban form as locational attributes that influence individual behavior. The macro approach is place oriented and recognizes urban form as characters of the place, which represent the patterns of how people and resources are distributed across the region. This review (1) explained micro and macro approaches in terms of the unit of analysis, variables used, data sources, major findings and policy implications, and (2) discussed gaps and links between them toward a comprehensive understanding of the relationship between urban form and VMT.

### 2.2. Micro Approach

#### 2.2.1. Unit of analysis

The purpose of micro approach is to estimate built environmental impacts on individual travel behavior. Travel behavior studies typically have tested not only VMT but also other travel outcomes such as car ownership, trip frequency, trip length and mode choice. The ideal unit of analysis for micro approach is an individual or a household that enables a control of individual travel decision factors. Many earlier behavioral studies have been conducted at the neighborhood unit (e.g., zip code in Holtzclaw 1990; 1994; traffic analysis zone in Friedman, Gordon, and Peers 1994; census tract in Frank and Pivo 1994) because individual datasets were hardly available at that time. Such aggregated methods are less preferred for travel behavior studies because they bear an unrealistic assumption that people in the same neighborhood have the same sociodemographic characteristics and travel behavior. For that reason, most of recent studies have been conducted at the individual or household unit.

### 2.2.2. Data source

Individual travel data are obtained from travel surveys and include information about individual sociodemographic characteristics, household size and structure, vehicle ownership, and a diary of trips on one or two given days including origin and destination, start and end time, mode of travel, accompaniment, and the purpose of travel. Travel surveys have provided researchers with rich datasets, but they have limitations as well; “The most pressing problem is that these region-wide surveys do not include many respondents in any one neighborhood” (Handy 1996). For example, national household travel survey (NHTS) in 2001 had the sample size of about 60,000 from 300,000,000 households. Regional level travel surveys probably provide a better response rate, but those surveys are available in a limited number of large metropolitan areas due to high cost. Given the limited sources of high quality surveys, many travel behavior studies were conducted at one or several regional scales. This brings up the generalizability issue particularly when regional factors (e.g., size, number of job centers, regional infrastructure, sprawl level, etc.) are considered significant for travel patterns.

### 2.2.3. Urban form measures

Micro approach typically measures the built environment at the local scale such as census tract, block group, traffic analysis zone and a certain distance of radius from an individual object. Cervero and Kockelman (1997) classified built environment attributes into three Ds – density, diversity and design. Density is measured in terms of persons, households, jobs, or floor area per unit area. The area can be the whole unit area or a developed part of the unit area. Diversity refers to land use mix and balance. Land use mix is typically characterized by the dissimilarity index which measures the degree of land use difference among adjacent parcels or grid cells. Land use balance is measured by the degree of land use share, called entropy index, wherein low values indicate single-use environments and higher values more varied land uses. Design refers to street

network characteristics often measured with average block size, the number of intersections, or the proportion of four-way intersections.

The neighborhood-based three Ds have been criticized for the scale mismatch with typical activity space (Badoe and Miller 2000; Ewing et al. 2008). For example, the average length of U.S. trips is 6.8 miles which is far beyond the limits of a neighborhood (Ewing et al. 2008). Studies considering three Ds of trip origin only (e.g., Dunphy and Fisher 1996; L. D. Frank and Engelke 2005; J Holtzclaw et al. 2002; Schimek 1996) might deliver limited implications as they fail to control other potential built environment factors beyond the residential environment. Distance to destination clearly has a significant impact on VMT. Destination accessibility, the fourth D, was considered to address this problem. The measure of accessibility should be cautiously selected depending on the purpose of trip (Handy 1993). Local accessibility is the accessibility for local activity (e.g., grocery shopping, school commuting, etc.), and is typically measured by distance from home to closest store or other services of interest. Regional accessibility is the accessibility to regional activity destination such as job and regional recreation place. Its measures are extensive including distance to central business district, the number of jobs accessible within a given travel time, or more complicated measures based on gravity model or utility theory.

#### 2.2.4. Control variables

Two types of control variable can be defined for travel behavior studies – socioeconomic and self-selection. Socioeconomic factors (e.g., age, sex, income, etc.) have been considered in most of recent studies since the information is readily available. They have shown an impact on trip frequency but provided an insignificant or marginal impact on VMT (Ewing and Cervero 2001). Self-selection issue has emerged with a question that the relationship between built environment and travel behavior might be spuriously created by individual preference of certain travel patterns (e.g., people who prefer less travel may choose to live in such built environment, and thus travel less). Cao et al. (2009) classified 38 studies by nine self-selection control methodologies (i.e.,

direct questioning, statistical control, sample selection model, instrumental variables model, propensity score, joint discrete choice models, nested logit model, structural equation model, and longitudinal model), and reviewed them focusing on the impact of built environment and self-selection on individual travel behavior (not limited to VMT). This extensive literature review concludes that built environment has a direct and strong impact even after controlling self-selection regardless of the methodology used. The degree of self-selection impact somewhat varied by study but was found to be weaker than built environmental impact.

#### 2.2.5. Findings and implications

There are at least eight review studies of the literature on the built environment and motorized travel (Badoe and Miller 2000; Cao et al. 2009; Cervero 2003; Crane 2000; Ewing and Cervero 2001, 2010; Handy 1996; Stead and Marshall 2001). A common finding from the literature is that regional accessibility is the most influential factor on VMT or travel distance. In contrast, the impacts of other local-based three Ds are inconsistent among studies. These review studies commonly point out that this inconsistency is due to the large variance among studies in their study site, data, measures and methodology used. As an attempt to generalize the results of different studies, a series of meta-analyses were conducted by Ewing and Cervero (2001; 2010). They selected 3 to 10 articles from over 200 studies based on the methodological rigor, and calculated the average effect size of travel behavior outcomes with respect to four Ds using selected 10 studies. The effect size is measured by partial elasticity, which estimate the effect size of any given built environment variable after controlling for other variables. Table 2.1 presents average elasticities by built environment factor in both studies. They agree that regional accessibility is most strongly associated with VMT, showing about -0.2 of elasticity with respect to “job accessibility by auto,” and “the inverse of the distance to downtown.” The other consistent finding is the small effect of household/population density with about -0.05 of elasticity in controlling for regional accessibility. The variable most dramatically changed between 2001 and 2010 studies is

“design,” which is quantified by street/intersection density or the fraction of 4-way intersections. The elasticity was -0.03 in the earlier study but -0.12 in the later study. The effect size of diversity was somewhat different between the two meta-analyses, presenting -0.05 and -0.09 each.

**Table 2.1**  
**Partial elasticity of VMT with respect to built environment variables**

Ewing and Cervero (2001)			Ewing and Cervero (2010)		
Variable	Number of studies	Partial elasticity	Variable	Number of studies	Weighted avg. elasticity
Local density	10	-0.05	Household/population density	9	-0.04
			Job density	6	0.00
Local diversity	6	-0.05	Land use mix (entropy index)	10	-0.09
			Jobs-housing balance	4	-0.02
Local design	4	-0.03	Intersection/street density	6	-0.12
			% 4-way intersections	3	-0.12
Regional Accessibility	5	-0.20	Job accessibility by auto	5	-0.20
			Distance to downtown	3	-0.22

As pointed out in the studies, the elasticities presented are still not generalizable due to small sample size. However, the meta-analyses reveal an important finding from micro approach: regional accessibility has a stronger impact on VMT than local built environment. In other words, people who live closer to regional center produce lower VMT. This is probably because job, shopping and recreational opportunities at the local level cannot fully satisfy the needs of residents. Agglomeration economies are still working even in this sprawling era. The strong attractive power of regional centers implies that a key urban form attribute influencing regional VMT is the spatial pattern of population distribution around regional centers. Without considering this regional configuration patterns, denser development at local level may have a limited influence on regional VMT reduction.

## **2.3. Macro Approach**

### 2.3.1. Unit of analysis

The purpose of macro approach is to estimate how regional development patterns influence collective VMT levels (e.g., VMT per capita). Even if significant factors influencing individual travel behavior were identified, it would be somewhat difficult to interpret the individual-level implications as collective-level policy goals. Built environment cannot be ideally controlled to be consistent across every individual in a region. Regional average VMT depends on distribution patterns of residents and jobs as well as people and jobs with certain characteristics (e.g., low income people, quality jobs, etc.). The ideal unit of analysis for macro approach studies is a travel-shed that includes all origins and destinations of daily-basis trips (e.g., commuting and shopping). Since the travel-shed is a conceptual definition, macro studies have been typically conducted at the scale of metropolitan area, urbanized area, or city, depending on data availability of outcome variables.

### 2.3.2. Data source

A major data source of collective VMT in the U.S. is daily vehicle miles traveled (DVMT) from *Highway Performance Maintenance System (HPMS)*. DVMT refers to daily traffic volume calculated by traffic counts multiplied by lane length of highways and arterials. The count-based traffic information does not include details of travel patterns (e.g., origin/destination, vehicle type, trip purpose, etc.); thus, DVMT data is reliable only at a highly aggregate scale that includes origin/destination of most trips in the region. Inability to separate vehicle type and trip purpose is an inherent limitation of the data. Nevertheless, DVMT data, covering major streets of 401 urbanized areas, are so far the most complete nationwide dataset of regional collective VMT. Some studies have used survey sample data to calculate regional averages (e.g., average of sample individual VMT or gasoline use), but the reliability of the estimated mean is questionable due to small sample sizes.

### 2.3.3. Urban form measures

A number of measures have been developed to quantify regional urban form (Clifton et al. 2008; Galster et al. 2001; Ewing et al. 2008; Knaap, Song, and Nedovic-Budic 2007; Longley and Mesev 2000; Srinivasan 2002; Torrens 2008; Tsai 2005). Existing measures can be broadly classified into three attributes -- density, diversity, and urban structure. Density is most frequently used among the three. Density is measured in terms of population, workers, and jobs at the regional scale. Some studies (e.g., Ewing et al. 2003) made a composite index using multiple density associated variables (e.g., regional density, local density, regional density gradient, etc.). Diversity at the regional scale has been measured by the average of locally measured accessibility, for example, percentage of residents with satisfactory neighborhood shopping within one mile (Ewing et al. 2003), or per grid cell accessibility based on cumulative count of jobs within a given travel time (Cervero and Murakami 2010). Land use balance indices based on Shannon's entropy theory have often been used for quantifying diversity (Ewing et al. 2003; Torrens 2008), though their implication at the regional scale is not as clear as the local scale. Urban structure refers to distribution and texture of the development. Measures of urban structure are extensive. For example, Galster et al. (2001) offered five measures of urban structure (i.e., continuity, concentration, clustering, centrality, and nuclearity) among their eight measures of sprawl. Spatial autocorrelation measures such as Moran's I, Geary coefficient, and Getis-Ord G statistic were used to quantify the degree of clustering (Torrens 2008; Tsai 2005). Torrens (2008) also offered fractal dimension, contagion, and interspersed and juxtaposition index to measure how fragmented or scattered a development is. To characterize centrality as well as overall degree of distribution of development, Gini coefficient of local density (Eidlin 2005; Tsai 2005), and density gradient (Clifton et al. 2008; Torrens 2008; Lee 2007; Kim 2007) have been widely used. These urban structure measures differ in implications, thus should be carefully selected depending on the interest of study.

#### 2.3.4. Findings and implications

Relatively few studies explored the relationship between urban development patterns and regional VMT/fuel consumption/greenhouse gas emission. A pioneering study by Newman and Kenworthy (1989) presents an inverse relationship between regional density and gasoline use in 32 major cities in North America, Europe, Asia, and Australia. This study simply explores correlations of city average gasoline use with respect to densities at three different scales (i.e., citywide, inner city and outer area) and a centrality measured by the proportion of population and jobs in the inner city. The original study has been improved by controlling for a number of variables on transport service level and socioeconomic characteristics (Newman and Kenworthy 1999). Van de Coevering and Schwanen (2006) updated this study by employing a multiple regression model including additional housing variables (e.g., housing tenure, age and size) and socio-economic variables (e.g., gender, age, education, and income level). They found that VMT was significantly associated with population density, job centrality and employment rate. They also found that housing variables had stronger relationships with mode choice than regional urban form. Overall, this series of studies reconfirmed population density as the key factor on gasoline use and modal split.

Ewing et al. (2002; 2003) developed a series of composite sprawl indices and estimated the impacts of those indices on selected travel outcomes. To create the sprawl indices, they conducted a principal components analysis to extract four factors out of 22 land use and street network variables – density, land use mix, centeredness, and street accessibility. Individual factor scores were standardized to have a mean value of 100 and standard deviation of 25. A multiple regression estimated the impacts of these sprawl factors on DVMT per capita after controlling for population, per capita income, household size and the percentage of working age. Only density factor showed a significant impact. Specifically, a 50 unit (or two standard deviation) increase in the density factor was associated with a decrease of 10.75 DVMT per capita. This amount of decrease indicates 40% reduction in DVMT per capita in areas with an average-level density (Ewing et al. 2008). For example, San Francisco (155) and Washington, D.C

(106) showed the 50 unit difference in the density factor. Density factor also showed strong and significant relationships with vehicle ownership and transit share, but was not significantly associated with mean commuting time. This implies that density factor contributes to VMT reduction through influencing car ownership and mode choice but does not affect travel distance.

Cervero and Murakami (2010) employed structural equation modeling (SEM) to explore the complex causal relationships between regional urban form, DVMT per capita and other variables in 370 U.S. urban areas. They measured regional development patterns with gross population density, job densities, and accessibilities for each retail and basic job. Accessibility was measured based on the cumulative count of jobs that can be reached within 30 minutes over a transportation network. Other variables included transportation supply, mode choice and sociodemographic factors. A primary finding was that population density had a strong direct relationship with DVMT per capita (direct elasticity -0.604), but the effect was offset by the traffic-encouraging effects of higher density such as denser road networks and a higher access to retail shopping (indirect elasticity 0.223, yielding a net elasticity -0.381). Accessibility to basic jobs had relatively modest effects.

All these macro approach studies concluded that density is the most influential factor on regional VMT, though their estimations of effect size do not agree with each other. Density dominates over other land-use configuration attributes. This is probably because density at the regional scale is a composite measure as a product of other sophisticated urban form measures, and therefore, absorbs the effects of those measures in a predictive model (Ewing, Pendall, and Chen 2002; Cervero and Murakami 2010). However, regional density provides limited implications for local land use policy. An increase of density at the regional scale is like regional population growth, which is not an attribute that local land use policy can control. Hence, it might be appropriate for macro approach studies to exclude regional density variables and estimate real effects of other policy-relevant urban form measures.

#### **2.4. Discussions: Gaps and Links between Micro and Macro Approaches**

Micro and macro approaches are distinctive in the focus of study, methodology, findings and policy implications (Table 2.2). The only agreement is “denser development” as a policy implication. However, this policy implication is marginally supported in both approaches. In micro approach studies, the influence of local density on VMT is weaker than accessibility. Macro approach has been dealing with density at the regional scale, and therefore cannot clearly explain how locally denser development influences regional density. Ironically, micro approach studies imply that denser development may increase regional VMT if it occurs at a location remote from regional centers. This contradiction is a major gap between both approaches. Micro approach recognizes “distance from job centers” as the key variable influencing VMT, whereas, macro approach has identified “centeredness” of population (a macro approach version of accessibility) as an insignificant variable. Two reasons for this difference can be considered.

The first reason is that the two approaches use different density (i.e., local density for micro approach vs. regional density for macro approach). The implication between local and regional density is entirely different. Regional density is a proxy variable of multiple regional characteristics that micro approach inherently cannot deal with. Thus, the differences in findings between micro and macro approach might reveal the limitation of micro approach, that is, it cannot consider regional variances. For example, regional density may influence the level of transit development; or the elasticity between accessibility and individual VMT might vary by region depending on other regional factors such as regional distribution patterns of population and jobs, and regional supply of roads and associated congestion level. This is why each metropolitan planning organization has its own regional accessibility coefficient for operating traffic simulation model.

The other reason for the difference between the micro and macro approach is found from the inappropriate definition of urban form made by the macro approach. First, it might be improper to regard regional density as an urban form measure.

Regional density is a composite characteristic that cannot be controlled by locally denser development but by regional policies including urban containment and population control. In a VMT predictive model, regional density may be useful as a variable representing transportation infrastructure, but not as an urban form variable. Other urban form measures also should be chosen or developed more carefully. Linking with findings from micro approach, it would be ideal that the measures used for macro approach can represent the average accessibility to regional activity and local activity separately.

**Table 2.2**  
**Comparison of micro and macro approaches**

	Micro approach	Macro approach
Purpose	Built environment impacts on individual VMT	Regional development pattern impacts on collective VMT
Unit of analysis	Individual/household	Urban area (travel shed)
Dependent variable	VMT/VHT/ Gasoline use (collected from survey)	DVMT per capita (collected from traffic count)
Major independent variables	<ul style="list-style-type: none"> <li>• Local density</li> <li>• Local diversity (entropy)</li> <li>• Local design</li> <li>• Regional accessibility</li> <li>• Individual socioeconomic variables</li> <li>• Individual self-selection variables</li> </ul>	<ul style="list-style-type: none"> <li>• Regional density</li> <li>• Regional diversity</li> <li>• Centeredness</li> <li>• Other regional development texture</li> <li>• Aggregate socioeconomic variables</li> </ul>
Major findings	<ul style="list-style-type: none"> <li>• Regional accessibility is the most influential factor on individual VMT</li> <li>• Local built environment influences mode choice</li> </ul>	<ul style="list-style-type: none"> <li>• Regional density is the most influential factor on regional collective VMT</li> </ul>
Policy implications	<ul style="list-style-type: none"> <li>• Infill development</li> <li>• Denser development</li> </ul>	<ul style="list-style-type: none"> <li>• Urban containment</li> <li>• Denser development</li> </ul>
Strength	<ul style="list-style-type: none"> <li>• Detailed control for sociodemographic factors</li> </ul>	<ul style="list-style-type: none"> <li>• Ability to control for diverse regional factors</li> </ul>
Limitations	<ul style="list-style-type: none"> <li>• Generalizability</li> </ul>	<ul style="list-style-type: none"> <li>• Limited control for sociodemographic factor</li> </ul>

The insufficient consideration in selecting urban form measures can be found frequently. For example, all the macro approach studies reviewed here have defined

regional density as a main urban form attribute despite its inappropriateness as an urban form measure. Newman and colleagues' studies (Newman and Kenworthy 1989; 1999; Van de Coevering and Schwanen 2006) used only one simple centeredness measure (i.e., proportion jobs in CBD in the total number of jobs) and did not consider the accessibility to local activity. Ewing et al. (2002; 2003; 2008) could not explain influences of specific variables straightforwardly, because each factor was quantified with a composite measure. For example, density factor was composed of seven density indicators including gross population density, urban population density, lot size, density of population centers, percentage of population living in a lower density area, percentage of population living in a higher density area, and CBD density estimated by exponential density function. Cervero and Murakami (2010) measured the accessibility to basic jobs and local retails separately, but those measures' effectiveness was doubtful. They calculated the accessibility at every grid cell by multiplying the cell population by the amount of jobs that can be reached from the cell in 30 minutes. These values were supposed to have a strong correlation with local densities. Thus their regional average value likely had a strong correlation with regional density rather than representing the average accessibility to regional centers. Actually in this study, the elasticity of regional density was 0.98 with respect to the accessibility to basic jobs, and 0.81 with respect to the accessibility to local jobs. Note that other studies showed a weak correlation between density and centeredness.

To measure relevant regional-level urban form attributes influential to VMT, the macro approach can derive ideas from the micro approach. Micro studies have hypothesized that VMT increases as people live farther from jobs and everyday services, and employed "regional accessibility" and "local-level three Ds" as built environment measures. These individual-level measures can be converted to the aggregated level with "density gradient" and "population-weighted average of local three Ds." Population density gradient toward job centers is one of the effective measures to present the accessibility to jobs. As the bid-rent theory (Alonso 1964) implied, developments tend to be concentrated in specific locations, like central business districts and subcenters, rather

than dispersed constantly across a region. In addition, a gradient is easy to interpret and delivers straightforward policy implications. The population density gradient shows the changing trend of density from center to periphery. If this measure has a significant influence on VMT, then it can suggest location specific density control guidelines based on distance from the center, such as rural-to-urban transect. Population-weighted average of tract density directly follows the concept of local density measure of the micro approach – measuring built environment in individual perspective. The regional average of individual-level local density increases as more people live in denser neighborhoods.

## **2.5. Conclusion**

This review identified that the inconsistency between micro and macro approach is originated from seemingly similar but conceptually different urban form measures. By using inappropriate urban form measures such as regional density, macro studies have failed to support that urban form control can contribute to VMT reduction. Micro approach studies, despite minor inconsistency in methodology, have collected solid evidence that “distance from job centers” and “local built environment” have a significant influence on VMT. This study suggested urban form measures for the macro approach, based on findings from the micro approach studies. The use of appropriate urban form measures in macro studies will help fill the gap with micro studies and enhance the generalizability of the implications.

### **3. EVALUATING THE PERFORMANCE OF CENTER IDENTIFICATION METHOD**

#### **3.1. Outline**

Center identification is the first step to understanding the influence of a polycentric urban structure. Although several center identification methods have been developed, few studies have evaluated the performance of these methods. McMillen's two-stage nonparametric method is fully flexible in controlling three major parameters to define urban centers – density threshold of a center, the extent of the nearby area affected by a center, and the relative power between a central business district and other subcenters. This study evaluates which parameter values perform better in identifying the centers using the data of 348 U.S. urbanized areas. The results showed that the uses of polycentric models for the polycentric areas yielded a better model fit than the use of a monocentric model. The use of monocentric model for polycentric areas could underestimate the job concentration level. The model with the lower density cutoff, the wider reference area, and the equal treatment between CBD and subcenters overall yielded the better performance in center identification.

#### **3.2. Introduction**

Most of the contemporary large metropolitan areas are recognized as polycentric, having multiple employment centers (Cervero and Wu 1997; Giuliano and Small 1991; McDonald and Prather 1994; McMillen and McDonald 1998). Research found that a polycentric urban structure has prominent effects on commuting (Cervero and Wu 1997), population density (Giuliano and Small 1991), and property value (Heikkila et al. 1989; McDonald and McMillen 1990). Building a polycentric urban structure model is necessary to explore the influences of polycentricism. Center identification is the first step to constructing a polycentric urban structure model. While several center identification methods were developed by researchers, including Giuliano and Small (1991), McDonald (1989), McMillen (2001), Craig and Ng (2001) and Redfean (2007),

few studies evaluated the performance of the method they used. Most of the existing evaluations were conducted based on local knowledge for the limited number of metropolitan areas. This approach would be more accurate for these specific areas but may not validate the general uses for cross-sectional studies of multiple metropolitan areas.

A major reason for the lack of evaluation studies is the difficulty in making objective definitions of urban centers to be compared with estimated results. Researchers agreed with two general characteristics of centers – 1) a significantly larger employment density than nearby locations, and 2) a significant effect on the density of nearby areas. However, the detailed definitions are not clear, as McMillen (2001) pointed out: “how large is large? What is the appropriate definition of nearby? Should we condition on distance from the central business district (CBD) or consider each site only within its local context? (p.449)” McMillen provided a fully-flexible method to allow users to control these parameters without suggesting which parameters are more appropriate. Studies employing this method arbitrarily selected the parameter values, thus the identified centers for the same areas varied by study.

Using McMillen’s method, this study evaluated which parameter values performed better in terms of the gradient and goodness-of-fit of density structure model. A simple monocentric density function was employed for the ease of interpretation. In the monocentric function, the polycentric structure was considered by assuming that all centers are located at an imaginary central location. The better model fit and the steeper gradient in this model implies the better positioning of centers. These polycentric models based on various combinations of associated parameter values were compared with each other and with the traditional monocentric model.

### **3.3. Previous Research on Center Identification Methods**

Center identification methods fall into two general categories: methods based on absolute and relative density guidelines (Lee 2007). Absolute methods identify candidate centers by an absolute minimum density cutoff. Areas (typically, census tracts or traffic

analysis zones) with higher density than the density cutoff are identified as center candidates. Giuliano and Small (1991) proposed this method, applying 10 jobs per acre of employment density as well as 10,000 jobs of total employment as the thresholds for Los Angeles, CA. This approach was employed by many studies with different density criteria in different cities (Anderson and Bogart 2001; Bogart and Ferry 1999; Peter Gordon and Richardson 1996; Pfister et al. 2000). Despite the popularity, the arbitrary nature of the density and total employment thresholds has been a major criticism against the absolute methods. Setting an absolute density threshold was not only sensitive to the unit of analysis but also required the detailed knowledge about study areas (McMillen 2001).

Methods based on relative density guidelines focused on capturing areas that have relatively higher employment density than nearby locations. This approach typically relied on statistical tests to identify significant residuals from a density function. Earlier methods employed parametric-regression-based density functions to model the exponential decay of densities by distance from CBD. Subcenters were identified as positive residuals from the monocentric models (McDonald 1989; McDonald and Prather 1994). This approach was criticized for its assumption that the topography of employment densities is symmetric about the CBD. Two different tracts in a same distance from the CBD might have different nearby conditions. In polycentric regions with strong subcenters, the local rises in density by the subcenters would make the monocentric density gradient flatter. This could reduce the probability of identifying some mid-to-small subcenters. To solve these problems, more recent methods employed nonparametric regressions to estimate polycentric density models. These include quantile regression (Craig and Ng 2001) and locally weighted regression (McMillen 2001). While these nonparametric methods considered both the distance and direction from the CBD, the latter provides a more flexible procedure that can be easily applicable to many different regions. Locally weighted regression estimates a smoothed density surface, considering only nearby areas for any data point (e.g., tract) with more weight

given to closer locations. It can control the threshold of relative density as well as the extent of a nearby area.

In recent center identification methods, the absolute and relative methods explained above serve as the initial stage to identify candidate centers because they consider only the first characteristic of centers – a significantly larger employment density than nearby locations. The next stage typically involves examining the second characteristic of centers – the significant influence on the densities of nearby areas. A two-stage procedure was proposed by Gordon et al. (1986). They identified candidate sites for Los Angeles via visual inspection of density maps. They used distance from 57 candidate sites as explanatory variables for density functions and concluded that only six subcenters have statistically significant effects on densities. McMillen and McDonald (1998) used Giuliano and Small's (1991) procedure to identify candidate centers for Chicago and selected 17 centers from 20 candidates in the second stage. The first stages of these earlier methods were somewhat arbitrary. More recent methods such as Craig and Ng (2001), McMillen (2001; 2003) and Redfearn (2007) improved this issue by employing nonparametric regression procedures for the first stage.

The studies cited above attempted to evaluate the performance of their center identification methods. While most of these existing evaluations included visual inspections based on local knowledge, McMillen (2001) used a gravity variable proposed by Shukla and Waddle (1991) to estimate the aggregate effects of proximity to identified subcenters on densities of individual traffic analysis zones (TAZs). The gravity was designed to increase as a site is located closer to subcenters. The gravity variable, however, was not used for comparing between different center identification methods, but used for comparing his polycentric model to the traditional monocentric model. For six cities, the regression model with both gravity and distance from the CBD (DCBD) yielded a better goodness-of-fit than the model with only DCBD. The study found that the addition of the gravity variable caused the coefficients of DCBD to turn positive in some cities including Dallas, Houston, and San Francisco, though density and DCBD are expected to have a negative relationship. The study concluded that it is

because the traditional CBD is no longer the significant determinant of the broad spatial trend in densities. He suggested that a more realistic specification would be to consider the CBD as simply another of the multiple centers.

### **3.4. Methodology**

#### ***3.4.1. Data and study areas***

The data came from two sources – Census Transportation Planning Package (CTPP) for employment information and Census TIGER for geographic boundaries. While U.S. Census dataset organizes demographic data by place of residence, CTPP dataset reorganizes the raw census data based on place of work and provides the number of employments at the various geographic scales nationwide. Since this study's performance evaluations included an application to vehicle miles traveled (VMT) estimation models, travel-associated information was obtained from Highway Statistics. In the VMT model, independent variables included two density gradient variables (i.e., job and population) estimated from the identified centers and five control variables (i.e., lane miles per capita, population, population density, job density, and median household income).

Given the time and geographical inconsistencies across these databases, a major consideration was given to join the datasets and to decide time and geographic definitions for the study. This study employed CTPP 2000 and Highway Statistics 2002. Year 2000 data was the most recent version for CTPP available as of 2012. Meanwhile, year 2002 dataset is the first Highway Statistics data using year 2000 urbanized area definition. This study assumed that the regional demographic structure did not change significantly during the two-year period. The urbanized area population estimations by Highway Statistics 2002 and the estimations by CTPP 2000 showed 0.998 of correlation coefficient. Table 3.1 presents variables used for empirical performance tests and their data sources.

**Table 3.1**  
**Variables used for empirical tests and data sources**

	Variable	Data source; geographic unit of data
Dependent	Daily vehicle miles of travel per capita	Highway Statistics 2002, FHWA; urbanized area
Independent	Population density gradient Job density gradient	CTPP 2000, Part 1; census tract CTPP 2000, Part 2; census tract
Control	Lane mile per capita Population Population density Job density Median household income	Highway Statistics 2002, FHWA; urbanized area U.S. Census 2000; urbanized area U.S. Census 2000; urbanized area U.S. Census 2000; urbanized area U.S. Census 2000; urbanized area

Note: CTPP – Census Transportation Planning Package; FHWA – Federal Highway Administration

Among 384 U.S. urbanized areas in Highway Statistics 2002 (except Alaska, Hawaii, and Puerto Rico), 36 cases were dropped due to missing data and geographic discrepancies. Figure 3.1 shows the selected 348 study areas.



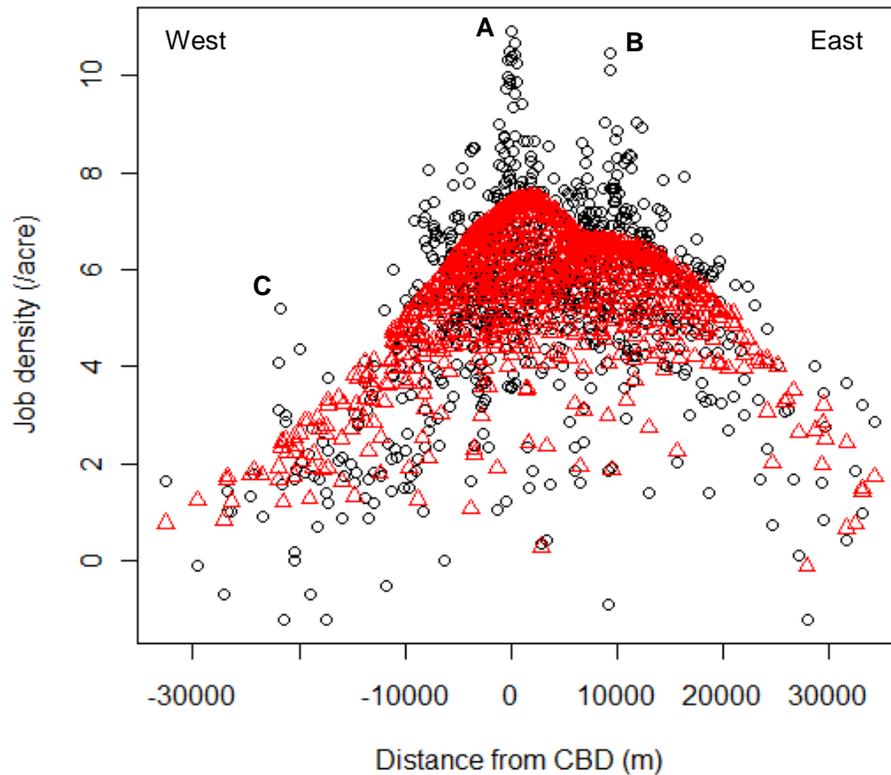
**Figure 3.1 Study areas for the center identification study (Source: U.S. Census 2000)**

### 3.4.2. Subcenter identification

This study tested eight modified versions of McMillen's (2001) two-stage nonparametric method. As mentioned above, the first stage is to identify candidate centers that show a significant residual from a job density estimation model. The second stage is to select the final list of centers from the candidates that have a significant influence on the density of nearby areas. This procedure is composed of three steps and each step had modifiable parameters that affect the final list of centers.

The first step is to estimate a density surface using locally weighted regression (LWR). This non-parametric regression conducts a series of local regressions at each observation point using a cluster of nearby observations, which are often called a "moving window." Within the window, closer observations receive more weight. The modifiable parameter in this step is the span of the window. Window size can be determined by the number of nearby observations or the percent of nearby subset of all observations. The larger window size leads to a smoother surface as more observations are used in local regressions. While the original study by McMillen (2001) used 50% subset, this study tested 25% and 50% windows to estimate the impact of window size on center identification performance. Figure 3.2 presents the actual and estimated job densities of census tracts in San Antonio urbanized area in Texas.

The second step selects candidate centers by identifying significantly greater residuals than 0 at the designated statistical significance level. McMillen suggested a normalized residual for the statistical test, and the significance level is modifiable in this step. A stricter threshold would yield the fewer but stronger candidates. The original study used the 0.05 significance level only, while this study added a more generous threshold of 0.1 to examine the influence of different cutoffs on the performance. The normalized residual and the thresholds can be described as follows:  $(y_i - \hat{y}_i) / \hat{\sigma}_i > 1.96$  or  $1.64$ , where  $\hat{y}_i$  is the LWR estimator of  $y$  at tract  $i$ , and  $\hat{\sigma}_i$  is the estimated standard error for the prediction. In addition to this relative threshold, tracts with the density less than 10 jobs/acre were dropped off.



**Figure 3.2 Actual job densities and smoothed density estimates.**

A horizontal view from the south side in San Antonio, TX. Circles represent actual job densities at the census tract level and triangles represent their locally weighted regression estimates. Two obvious peaks (A and B) over the smoothed density estimation surface indicate potential centers. The peak C would be excluded by an absolute minimum density threshold although it has a higher density than nearby area.

The last step selects centers from the candidates that show a significant explanatory power on job densities of nearby areas. This final list of centers is obtained by the stepwise (backward) regression of job density on distances to candidate centers, where the candidates with an insignificant effect will be omitted from the final list. The regression model can be described as follows:  $y_i = g(DCBD) + \sum_j (\delta_j d_{ij}) + u_i$ , where  $y_i$  is logarithm of job density at site  $i$ ,  $d_{ij}$  is distance between site  $i$  and subcenter  $j$ ,  $g(DCBD)$  is a function to control for the influence of CBD on job density,  $\delta_j$  is regression weight, and  $u_i$  is error term. This model tests if the distances from each candidate center have a significant influence on the job densities of nearby tracts. Stepwise regression excludes

candidates that have an insignificant influence. In addition, candidates showing a positive coefficient need to be excluded because the expected relationship is negative (i.e., the density of a site increases as getting closer to a center). It is noteworthy that the original method attempted to consider CBD and other subcenters differently. The variable representing the influence of CBD is the density estimator based on the distance from CBD, while the influence of a subcenter is represented by the distance from the subcenter itself. The logic behind this is that CBD has a predominant influence over other centers. This approach is more advantageous for identifying subcenters in a more sophisticated manner particularly in regions with one very strong CBD. However, it may not be appropriate for the regions with multiple stronger centers (as the author already recognized in conclusion). So this study also tested identification models with “distance from CBD” as the CBD control variable. The models consider CBD in the same manner with other subcenters.

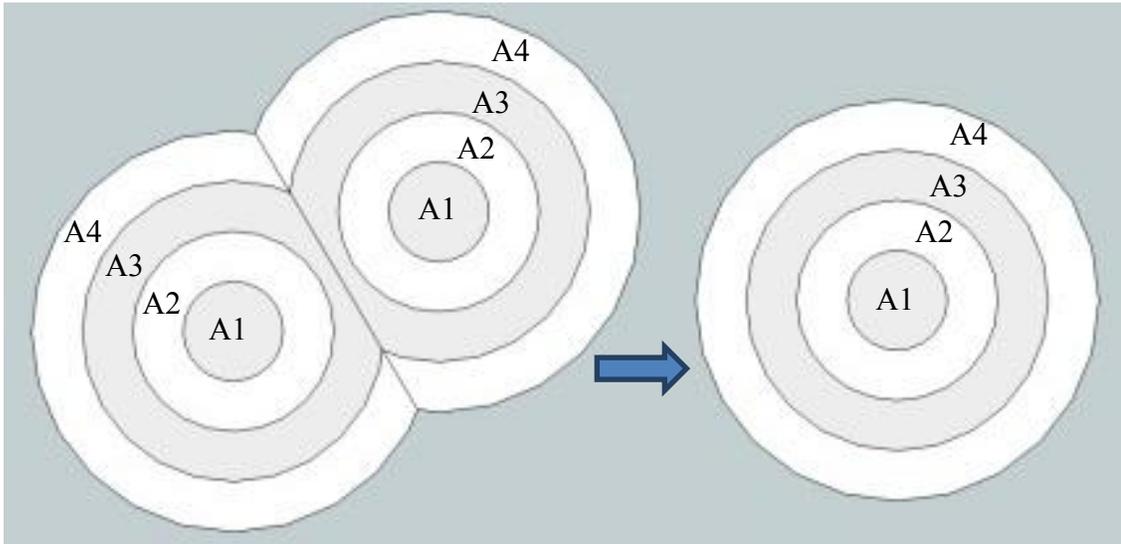
Hence, with two alternatives for each of three steps (i.e., 25% and 50% window for the LWR estimation; 0.05 and 0.1 significance levels for the candidate center threshold; and density estimator by the distance from CBD and the distance itself for the CBD influence control), a total of eight center identification models were tested and compared.

### 3.4.3. Density function

The performances of the eight center identification models were examined based on how the identified centers improve job density estimation models. To represent multiple centers to density estimation, an appropriate polycentric model is necessary. There are several polycentric models, which are based on spline density curve (Craig and Ng 2001), gravity (McMillen 2001), and locally weighted regression (McMillen 2004). However, these non-parametric models are technically complicated to calculate and difficult to translate the estimated parameter values into policy guidelines.

This study used a monocentric model to take advantage of its simplicity and added an operational assumption to apply the monocentric model to polycentric urban

structures. A polycentric structure with multiple concentric circles is translated into a monocentric structure by assuming that all employment centers (including CBD) are located at the same location (Figure 3.3). A simple exponential decay function was used for the monocentric model as follows:  $D = \text{EXP} [\alpha \cdot d + \beta]$ , where  $D$  is job density at distance  $d$  from the imaginary center,  $\alpha$  is the gradient, and the exponential of  $\beta$  is the estimated density of the center. The gradient  $\alpha$  can be calculated by a linear regression with the transformation of the original function:  $\text{LN}(D) = \alpha \cdot d + \beta$ , where the gradient  $\alpha$  is the slope of regression model.



**Figure 3.3 A simple operation to translate polycentricity using monocentric model (center identification study)**

#### 3.4.4. Performance evaluations

This study used two criteria to estimate the center identification performance – the goodness-of-fit and the gradient estimated from the job density model based on each center identification model. By removing outliers located beyond CBD, the model based on a better center identification is expected to have a better model fit and steeper gradient. The goodness-of-fit is represented by the r-square of the regression type of the

density model. Center identification models showing a statistically insignificant gradient at the 0.05 level were excluded from evaluations.

Evaluations were conducted in four ways, 1) comparison of the performance criteria, 2) examination of performance criteria changes made by using a polycentric model instead of a monocentric one, 3) a series of paired t-tests to identify which treatments in window size, candidate threshold and CBD control have significant influences on the performance criteria, and 4) an empirical application of the estimated polycentric urban structure models to a transportation model. As many travel behavior studies found that the distance from CBD is the most influential factor of vehicle miles traveled (Ewing and Cervero 2001; 2010), a regression model of daily vehicle miles of travel (DVMT) per capita was used for the empirical test. Independent variables included two density gradient variables (i.e., job and population) estimated from each center identification model and five common variables, including lane miles per capita, population, population density, job density, and median household income. Stepwise procedure was employed to avoid the multicollinearity problem. As a result, population density gradient survived from the stepwise procedure in eight of the nine models (i.e., one monocentric and eight polycentric). The coefficients and statistical significances of the population density gradients were compared to examine the empirical performance of each center identification scheme.

### **3.5. Results and Discussions**

Results showed that 331 of 348 (95%) urbanized areas had a significant job density gradient in both monocentric and polycentric models. While the number of identified polycentric areas varied by the model, about 73 to 82% of all urbanized areas were classified as monocentric form (Table 3.2). The stricter models using the smaller window size (25%), the stricter candidate cut threshold (0.05) and the CBD control with density estimation tended to recognize the fewer numbers of polycentric areas. The average number of identified centers and the average r-square of job density gradients were similar among the models because the majority of observations were monocentric.

**Table 3.2**  
**Center identification results by each monocentric and polycentric model**

Model			Number of UAs with a significant gradient	Number of polycentric UAs identified	Number of centers by UA		Average R <sup>2</sup> of job density gradient	Average job density gradient
					Avg.	Max.		
Monocentric			331	NA	1	1	.44	.44
Polycentric								
Window Size	Cut Threshold	CBD Control						
25%	0.05	Density	331	48	1.5	22	.44	.38
50%	0.05	Density	331	52	1.5	18	.45	.38
25%	0.10	Density	331	56	1.6	33	.45	.39
50%	0.10	Density	331	68	1.7	30	.45	.39
25%	0.05	Distance	331	57	1.5	19	.45	.39
50%	0.05	Distance	331	70	1.5	19	.45	.39
25%	0.10	Distance	331	66	1.7	30	.45	.39
50%	0.10	Distance	331	91	1.8	29	.45	.39

Note: Calculations were performed using 331 U.S. urbanized areas that show a significant job density gradient at the 0.05 level.

### 3.5.1. Performance changes made by using polycentric model

To highlight the performances of polycentric versus monocentric models, further examinations were performed only with the areas that every center identification model recognized as polycentric. Table 3.3 presents the performances of polycentric models for the 41 polycentric areas. Polycentric areas tend to show a worse model fit than monocentric areas (i.e., 0.45 of r-square with all areas versus about 0.3 with polycentric areas).

Meanwhile, the use of polycentric models for the polycentric areas yielded a better model fit than the use of monocentric model (i.e. 43 to 60% increase in r-square), indicating the superiority of the polycentric model. Polycentric models also resulted in a much steeper gradient (i.e., 85 to 112% increases). This implies that the use of monocentric model to polycentric areas may underestimate job concentration level. The least strict model (50%/0.10/Distance) yielded the largest improvements in both r-square

and gradient, and 25%/0.10/Distance model performed the second best. The least performing model was the strictest model (25%/0.05/Density). The results indicate that less strict models have the better chance to identify appropriate centers.

**Table 3.3**  
**Performance changes made by using polycentric models instead of monocentric model for 41 polycentric urbanized areas in the U.S. in 2000**

Model			Average Number of Centers by UA	Average R <sup>2</sup>	Performance Change	
Window Size	Cut Threshold	CBD Control			Average Change in R <sup>2</sup>	Average Change in Gradient
25%	0.05	Density	4.6	0.26	45%	85%
50%	0.05	Density	4.7	0.26	45%	86%
25%	0.10	Density	5.9	0.29	57%	103%
50%	0.10	Density	5.8	0.28	56%	102%
25%	0.05	Distance	4.5	0.26	43%	90%
50%	0.05	Distance	4.7	0.27	47%	95%
25%	0.10	Distance	5.6	0.28	57%	109%
50%	0.10	Distance	5.7	0.28	60%	112%

### 3.5.2. Attributes affecting center identification performance

The next evaluation examined whether the change in each of the three parameters of center identification made a meaningful difference. This evaluation was conducted by a paired t-test between two models that have one different parameter value and two common parameter values. For example, the test between 50%/0.05/Density and 25%/0.05/Density could show if the difference in window size made a statistically significant difference in the performance criteria.

Table 3.4 presents mean differences and their significances in three criteria (i.e., the number of centers, the fit of density model, and the estimated job gradient) by matching pair. The results explain why 50%/0.10/Distance model was the best performing model in the previous test. The change in window size has negligible impacts on the benchmarks, while the candidate cut threshold was revealed as the most influential parameter. The less strict (0.10) cut threshold identified more centers in more

appropriate locations which makes the density gradient steeper. It also yielded a better model fit. The stricter threshold might exclude some good center candidates in the initial stage.

CBD control was also recognized as a significant parameter determining the location of centers. Although CBD control did not make any significant distinction in the number of centers and the goodness-of-fit of the density model, the distance-based CBD control yielded a steeper gradient. This result differs from the traditional way of controlling the CBD impact. McMillen's original method and its successors controlled the CBD impact using density estimation by the distance from CBD rather than the distance itself. The superiority of distance-based CBD control over the density-based control implies that some subcenters have a regional influence equivalent to CBD in many U.S. urbanized areas.

**Table 3.4**  
**Paired-samples mean differences and t-values in three center identification performance criteria by matching parameter values**

		Paired-samples Difference in:					
		Number of Centers		R-square		Density Gradient	
		Mean	t	Mean	t	Mean	t
Difference in Window Size (50% vs. 25%)	50%-25%, 0.05/Den.	0.098	0.585	0.000	0.000	0.001	0.244
	50%-25%, 0.10/Den.	-0.073	-0.453	-0.005	-0.884	-0.001	-0.467
	50%-25%, 0.05/Dist.	0.171	0.943	0.009	1.567	0.006	1.301
	50%-25%, 0.10/Dist.	0.073	0.393	0.000	0.000	0.001	0.360
Difference in Cut Threshold (0.10 vs. 0.05)	0.10-0.05, 50%/Den.	1.146**	3.290	0.020*	2.479	0.012*	2.620
	0.10-0.05, 25%/Den.	1.317**	3.265	0.024**	3.274	0.015**	3.149
	0.10-0.05, 50%/Dist.	1.000**	2.938	0.016*	2.515	0.009	1.950
	0.10-0.05, 25%/Dist.	1.098**	3.100	0.025**	3.592	0.014**	2.925
Difference in CBD Control (Distance vs. Density)	Dist.-Den., 50%/0.05	0.000	0.000	0.003	0.486	0.013*	2.477
	Dist.-Den., 25%/0.05	-0.073	-0.368	-0.006	-0.821	0.008	1.423
	Dist.-Den., 50%/0.10	-0.146	-0.771	0.000	0.000	0.009*	2.064
	Dist.-Den., 25%/0.10	-0.293	-1.550	-0.005	-0.806	0.007	1.302

\*\* significant at the 0.01 level

\* significant at the 0.05 level

### 3.5.3. Empirical application test

Results from the empirical application test partially agreed with the two performance tests above. For the empirical test, stepwise regression estimations of DVMT per capita were conducted for nine center identification models, including one monocentric and eight polycentric models. Among seven independent variables (i.e., both job and population density gradient measures estimated from each center identification model and five common variables), the stepwise procedure identified only two or three variables as significant at the 0.05 level (Table 3.5). “Lane miles per capita” and “median household income” were recognized as most influential and significant in every model. Population density gradient was significant in eight of nine models, reconfirming the insight from travel behavior studies.

**Table 3.5**  
**Results of empirical application test using a stepwise regression model of vehicle miles of travel (VMT)**

- Dependent variable: DVMT per capita

- Independent variables: job density gradient, population density gradient, lane miles per capita, population, population density, job density, and median household income

- Samples: 41 polycentric urbanized areas in the U.S.

	Mono-centric	25%/ 0.10/ Dist.	50%/ 0.10/ Dist.	25%/ 0.05/ Dist.	50%/ 0.05/ Dist.	25%/ 0.10/ Den.	50%/ 0.10/ Den.	25%/ 0.05/ Den.	50%/ 0.05/ Den.
R-square	.613	.645	.632	.628	.615	.606	.606	.550	.622
Std. coefficient									
<i>PD.grd</i>	-.254*	-.310**	-.293*	-.283**	-.259*	-.240*	-.241*	NA	-.270**
<i>LM/c</i>	.820**	.816**	.814**	.833**	.821**	.825**	.826**	.787**	.817**
<i>Income</i>	.316**	.330**	.303**	.338**	.314**	.323**	.312**	.320**	.316**

Note: PD.grd – Population density gradient; LM/c – Lane miles per capita;

Income – median household income

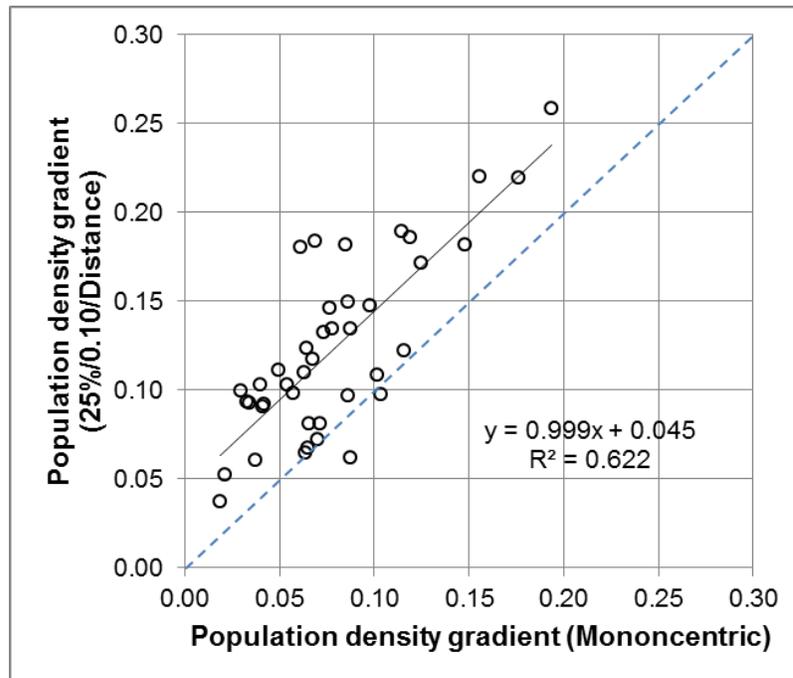
\*\* significant at the 0.01 level

\* significant at the 0.05 level

Basically, seven of the nine center identification models yielded similar results in terms of model fit, variables identified as most influential, and their coefficients. None

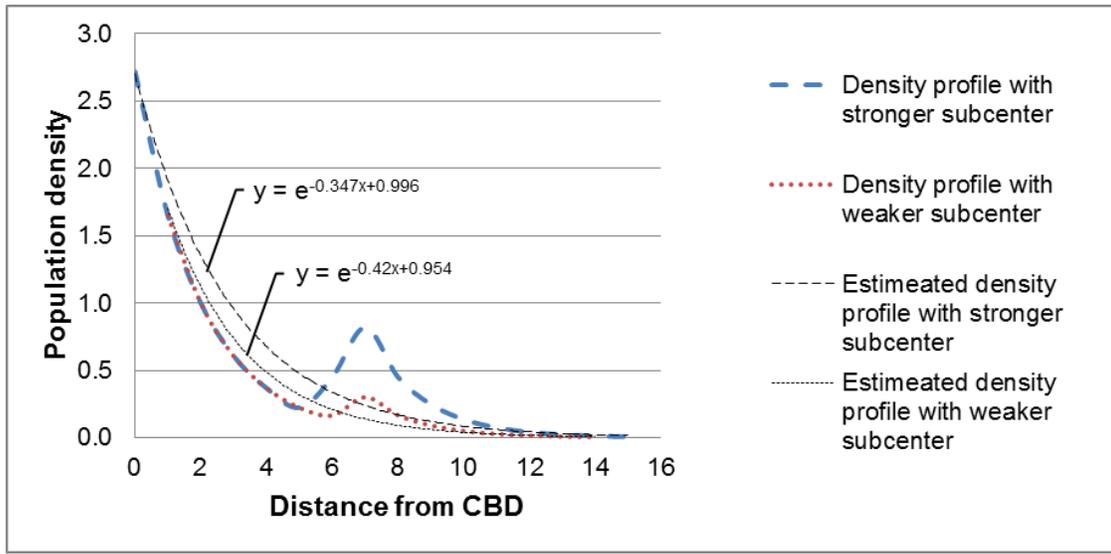
of these models produced a distinctively better performance in the empirical travel model, because the explanatory power of population density gradient was not strong enough to make a significant change in overall estimation results. Despite the minimal influence in the empirical test, different center identification models caused subtle changes in the values of model fit and coefficients. Also, their trend tended to be consistent with previous tests. Models with a weaker candidate cutoff and distance-based CBD control showed larger r-square values and stronger coefficients of population density gradient. Smaller window size also slightly improved these performance criteria. The best performing model was revealed as 25%/0.10/Distance model in which the significance of population density gradient was accepted at the 0.01 level. The only case inconsistent with this tendency is 50%/0.05/Density model, which was expected to perform least in the tendency. This result is unexpected but no specific reasons could be identified.

Another unexpected finding in the empirical test was that polycentric models did not yield significantly different results from the monocentric model despite the clear distinction between the two. Polycentric models have 43 to 60 percent larger goodness-of-fit values and 85 to 112 percent larger gradient values than monocentric model (Table 3.3). This lack of distinction in the empirical test is attributed to the strong correlation between them. The correlation between population gradients of the best-performing polycentric model (25%/0.10/Distance) and monocentric gradients has an r-square of 0.622 and a slope of 0.999 (Figure 3.4).



**Figure 3.4 Correlation between population gradients of the best-performing polycentric model (25%/0.10/Distance) and monocentric gradients, 41 polycentric urbanized areas in the U.S. in 2000**

This strong correlation between monocentric and polycentric gradients is contradictory to the expectation, considering that those gradients are changed to opposite directions by the strength of subcenters. That is, stronger subcenters tend to result in steeper polycentric gradients but gentler monocentric gradients (Figure 3.5). Hence the relative strength of subcenters is the key variable to differentiate between monocentric and polycentric gradients – the stronger the subcenters, the larger the gradient gap. According to this tendency, the strong correlation between the two gradients would be explained by the consistency of relative subcenter strength among polycentric areas (i.e., areas with stronger CBD tend to have stronger subcenters). The empirical test result could not provide a clear answer to which policy option between strengthening CBD versus promoting subcenters is more effective for the reduction of regional average VMT.



Note: The power of the exponential function indicates density gradient.

**Figure 3.5 Change of monocentric gradient by the strength of subcenter.**

two areas with the same CBD strength: stronger subcenters result in gentler monocentric gradients (while steeper polycentric gradients).

### 3.6. Conclusions

Polycentricity or multi-centered urban structure became a common characteristic of contemporary larger metropolitan areas. There have been many attempts to accurately capture the complexity of polycentric urban structure. Center identification is the first step to study polycentricity. While several center identification methods were developed, few studies evaluated the performance of the methods. Most of the existing evaluations were conducted based on local knowledge for the limited number of metropolitan areas.

A major reason for the shortage of evaluations is the difficulty in making objective definitions of urban centers. Studies agreed with two general characteristics of centers – significantly larger employment density than nearby locations and a significant effect on the density shapes of nearby areas. However, the details are not clear in definitions of “larger” and “nearby,” and the way to treat central business district (CBD). McMillen’s two-stage nonparametric method is fully flexible in controlling these three parameters. Its first stage selects candidate centers based on the size of residual from a smoothed density estimate surface. The larger residual has the better chance to be selected as a center. The cut-off of residual size determines the degree of “larger.” The

extent of “nearby” can be controlled by the size of reference area for the smoothed density estimation function. The second stage determines the final list of centers based on their influences on densities of nearby areas. A candidate that has a larger impact on densities of closer areas is considered as a center. In this stage, CBD can be treated as another center equally with other subcenters or can be specially treated as a control for estimating the influence of other subcenters.

Using the McMillen’s method, this study evaluated which parameter values help identify more appropriate centers. Eight combinations of parameter values (i.e., two alternatives by each of three parameters) were tested. The performances were evaluated with two criteria – the gradient and the model-fit statistic estimated from the polycentric density model based on each center identification model.

For polycentric areas, polycentric models yielded a better model fit than monocentric models. Polycentric models also resulted in a steeper gradient, which implies that the use of monocentric model to polycentric areas may underestimate the job concentration level. The least strict model (i.e., the wider reference area, the lower cutoff for candidate centers, and the equal treatment of CBD) overall yielded the best performance in center identification. According to the paired t-tests between different values in each parameter, the less strict candidate cut threshold identified more centers in more appropriate locations. The extent of nearby reference area did not make a significant difference in the identification performance. While CBD control did not make any significant distinction in the number of centers and the model fit, the distance-based CBD control yielded the steeper gradient, indicating the more appropriate positioning of centers.

Empirical application tests of the eight center identification scenarios on a travel model did not find any significant difference among the scenarios. Even the monocentric scenario model yielded similar parameter estimates and model fit with the best performing polycentric model. This was unexpected since the two models showed a clear distinction in density gradient values. The study found this is due to the strong correlation in population density gradient between the two models. That is, the areas having the greater monocentric gradient tend to have the greater polycentric gradient.

## **4. INFLUENCES OF URBAN FORM ON VEHICLE MILES TRAVELED AT THE METROPOLITAN SCALE**

### **4.1. Outline**

This study evaluated the influence of regional development patterns on collective vehicle miles traveled. A cross-sectional analysis of 203 U.S. urbanized areas was conducted. Directed acyclic graph (DAG) and structural equation modeling (SEM) were used to analyze complex relationships among urban form, transportation infrastructure, and income segregation. Considering the polycentric urban structure of many metropolitan areas, regional urban forms were measured with changing trends of population, job and poverty rate by the distance from multiple centers. The study found: 1) Vehicle miles traveled (VMT) is lower in areas with more concentrated urban structure and more compact neighborhoods, and these two effects were independent; 2) VMT was strongly affected by road supply; while, transit supply had no significant influence; and 3) poverty gradient had a minimal effect on VMT. Overall, the study indicates the “rural-to-urban transect” approach is superior to traditional “compact neighborhood” approach for reducing VMT, because it accounts for both regional and local urban form effects.

### **4.2. Introduction**

Growing concerns over climate change have attracted keen interests in reducing the amount of vehicle use. Producing 30% of CO<sub>2</sub>, the transportation sector was the second largest source of greenhouse gas after electricity generation in the U.S. in 2008 (U.S. EPA 2011). Many governmental bodies established policy initiatives targeting auto travel reduction, and their efforts typically included redesigning urban form to make residents travel shorter and introducing alternative travel options to reduce driving (U.S. SCCST 2009; CPDR 2008).

Planners are beginning to recognize that local urban form should be organized in the regional context to reduce vehicle travels more effectively. Common planning

guidelines (e.g., transit oriented development or traditional neighborhood development) tend to pursue the “compact neighborhood” approach, focusing on denser and mixed-use developments at the neighborhood scale. This local approach does not consider regional development patterns and often can lead to intense developments in suburban areas. Some argue that there is no clear evidence for the relationship between compact city and sustainable travel behavior (Neuman 2005). On the contrary, recent land use initiatives emphasize regional frameworks to classify local urban form by regional location, such as rural-urban transect, form-based zoning, etc. (Parolek et al. 2008).

Studies of regional development patterns in association with collective vehicle miles traveled (VMT) are limited in terms of amount, methodology and implications. These studies provided limited insight into the urban form-VMT relationship. Most of the collective VMT studies concluded that density is the most influential factor (Cervero and Murakami 2010; Ewing et al. 2003; Newman and Kenworthy 1989; Newman and Kenworthy 1999; Van de Coevering and Schwanen 2006). However, an increase of density at this large scale simply implies regional population growth, which is not an attribute land use plans can control. Meanwhile, the majority of urban form-travel studies was conducted at the individual level in only one or several metropolitan areas, and thus, could not control for a wide variety of regional settings. Their main interest was to evaluate the “compact neighborhood” approach, while their finding was that location (i.e., distance to central business district or regional accessibility to jobs) is a better predictor of individual VMT than the surrounding built environment (Ewing and Cervero 2001; 2010). This implies that regional development patterns (i.e., how many people live closer to job centers) have a potentially strong influence on collective VMT rather than local urban form (i.e., how many people live in denser neighborhoods).

The bid-rent theory is useful to synthesize the findings from the disaggregate studies. The theory explains settlement decisions made by the tradeoff between land price and transportation cost (Alonso 1964). In this concept, the notion that VMT increases with the distance from centers infers two ideas: 1) VMT is lower in denser neighborhoods because denser developments occur near the job center and 2) VMT is

greater among higher income people because they are less sensitive to transportation costs and tend to prefer suburban living. These insights can be applied to aggregate studies at the regional scale too and can connect the disaggregate and aggregate studies.

### **4.3. Previous Research**

The studies examining urban form impact on VMT can be classified into two discrete approaches – micro and macro. Both approaches are distinctive in the subject of analysis and the associated perspective on urban form. The micro approach is intended to analyze people and understand urban form as a locational attribute that influences individual behavior. The macro approach focuses on place and recognizes urban form as characters of the place presenting how people and resources are distributed across the region. Urban form was frequently called built environment in micro studies, while it was often called regional development pattern in macro studies. Because of this difference, they differ as well in detailed methodology; thus, the findings need to be interpreted differently.

There are several literature review studies on built environment and travel behavior (Badoe and Miller 2000; Cao et al. 2009; Cervero 2003; Crane 2000; Ewing and Cervero 2001, 2010; Handy 1996; Stead and Marshall 2001). A common finding from the literature reviews was that regional accessibility was the most influential factor on VMT. In contrast, the impacts of local-based three Ds (i.e., density, diversity, and design) were inconsistent among studies. These review studies commonly pointed out that this inconsistency was due to the broad variance among studies in their study sites, data, measures and methodologies used. Ewing and Cervero (2001; 2010) attempted to generalize the results of different studies using a series of meta-analyses. They selected literature from over 200 studies based on their own methodological rigor and calculated the average effect size of travel behavior outcomes with respect to four Ds (i.e., three Ds + destination accessibility). These meta-analyses confirmed that destination accessibility (or distance to central business district) is the strongest factor on VMT. They also concluded that density has a marginal impact but the combined impact of the three Ds

was larger. A review by Cao et al. (2009) particularly focused on studies that consider the self-selection issue. The study reconfirmed that the impacts of the built environment measures remain valid even after controlling for self-selection.

Relatively few studies explored the relationship between development patterns and travel outcome at the regionally aggregated scale. A pioneering study by Newman and Kenworthy (1989) presented the inverse relationship between regional density and gasoline use in 32 major cities in the world. Their study simply explored the impacts of densities at three different scales (i.e., citywide, inner city and outer area) and centrality. The centrality was measured by the proportion of population and jobs in the inner city. The original study has been improved by considering a number of additional variables, including transport service level and socioeconomic characteristics (Newman and Kenworthy 1999; Van de Coevering and Schwanen 2006). This series of studies reconfirmed that population density is the key factor to determine average gasoline use.

Ewing et al. (2002; 2003) developed a series of composite sprawl indices and estimated the impacts of those indices on selected travel outcomes. To create the sprawl indices, they conducted a principal components analysis to condense 22 land use and street network variables to four factors – residential density, land use mix, centeredness, and street accessibility. Individual factor scores were standardized to have a mean value of 100 and a standard deviation of 25. A multiple regression estimated the impacts of these sprawl factors on daily vehicle miles of travel (DVMT) per capita in controlling for population, per capita income, household size and the percentage of working age population. The studies found that the density factor and the percentage of working age population have a significant impact on VMT. However, these studies provide limited policy implications about what specific urban form attributes influence VMT because the urban form measures were composite. For example, the density factor is composed of four attributes as follows: 1) gross population density in persons per square mile, 2) the percentage of population living at densities less than 1,500 persons per square mile, 3) the percentage of population living at densities greater than 12,500 persons per square

mile, and 4) the estimated density at the center of the metropolitan area derived from a negative exponential density function.

Cervero and Murakami (2010) employed structural equation modeling (SEM) to explore causal relationships between urban form, daily vehicle miles traveled (DVMT) per capita and other control variables of 370 US urban areas. They measured regional development patterns with gross population density, gross job density, and accessibilities for each retail and basic job. Accessibility was measured based on the cumulative count of jobs that can be reached within 30 minutes over a transportation network. A primary finding is that population density has a strong direct relationship with DVMT per capita (direct elasticity -0.604), but the effect is offset by the traffic-encouraging effects of higher density such as denser road networks and a higher access to retail shopping (indirect elasticity 0.223, yielding a net elasticity -0.381). Accessibility to basic jobs showed relatively modest effects. However, these accessibility measures' effectiveness is doubtful because these values are expected to have a strong correlation with local densities. Thus, it is likely that their regional average value has a strong correlation with regional density and does not represent the average accessibility to regional centers. Actually, in their study the elasticity of regional density is high (0.98) with respect to the accessibility to basic jobs. Note that other studies show a weak correlation between density and centeredness, another development pattern measure.

The only agreement between the micro and macro approaches is “denser development” as a policy implication. However, this policy implication was marginally supported in both approaches. In micro approach studies, the influence of local density on VMT is only marginal in comparison to accessibility. The macro approach has been dealing with density only at the regional scale, and yet it cannot clearly explain how local denser developments are related to regional density. Ironically, the micro approach studies imply that denser development may increase regional VMT if it occurs at a location remote from regional centers. This contradiction is a major gap between the both approaches. The micro approach recognizes “accessibility” (or distance from job centers) as the key variable influencing VMT, whereas the macro approach has

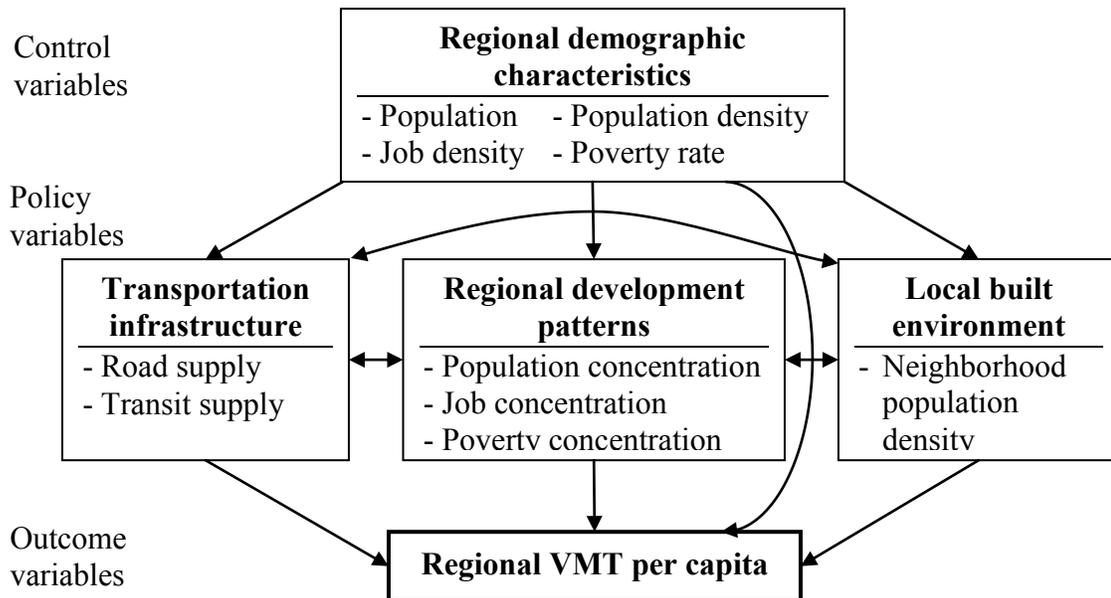
identified “centeredness” of population (a macro approach version of accessibility) as a less significant variable than density.

#### **4.4. Research Approach**

To explore the contribution of development control to auto driving reduction at the regional scale, this study aims to estimate impacts of urban form on DVMT per capita in U.S. metropolitan areas. Figure 4.1 is the conceptual framework that shows how this study comprehends relationships of associated variables. This study classifies influential factors of average VMT into three constructs—regional development patterns, local built environment, and transportation infrastructure. The framework has a hierarchical structure that includes urban area characteristics (i.e., population, population density, job density, and poverty rate<sup>1</sup>) as control variables at the higher level over the three policy factor constructs of VMT. This structure is devised to serve two purposes: 1) to control various sizes and economic conditions of urbanized areas, and 2) to set the directions of relationships between these control variables and the key constructs. The constructs at the lower level include attributes that can be controlled by transportation and land use policies. On the other hand, regional demographic characteristics are not a product of urban form and transportation supply, but they may influence planning policies and shape distributions of population, job, and infrastructure. Except this structural condition, all possibilities of relationships are allowed between variables in the model, reflecting the relationships in the real world. Regions with stronger centers are expected to show a higher concentration of jobs, population and poverty. They also have more dense neighborhoods, and require fewer roads. The key questions of this study are to identify how these variables are interrelated with each other and with VMT, and to identify what factors are more influential than others in predicting the average VMT.

---

<sup>1</sup> We also tested with median household income, but its higher correlations with all other explanatory variables in the model distorted the overall relationship structure and produced an unacceptable level of model fit. Poverty rate had insignificant correlations with other control variables and yet showed a high correlation with the income variable. Poverty rate also showed a significant relationship with the VMT variable, while, the income variable did not.



**Figure 4.1 Conceptual framework for the DVMT study**

Recognizing “distance to regional centers” as the key correlate of average VMT, this study hypothesized that average VMT is lower in these cases: 1) as residents are located more closely to regional centers, 2) as regional employment centers are stronger, and 3) as income segregation is less severe. The study also tested a hypothesis that previous studies identified: 4) VMT is lower in areas with more dense neighborhoods. The first three hypotheses represent development control at the regional scale, such as “rural-to-urban transect” approach; while, the fourth hypothesis supports a local-scale control, such as “compact neighborhood” approach. This study evaluated both policy options by testing the hypotheses.

#### 4.5. Methodology

A cross-sectional analysis of 203 U.S. urbanized areas was conducted with two methodological emphases. First, gradient-based measures were used to quantify the changes of interested urban form attributes (i.e., population, job, and poverty rate) by distance from centers. While traditional gradient measures were based on a monocentric

urban model, this study further considered a polycentric urban structure commonly found in many metropolitan areas. Also, directed acyclic graph (DAG) and structural equation modeling (SEM) were used to control intervening factors (e.g., transportation infrastructure, and income segregation) between urban form and VMT.

#### 4.5.1. Data

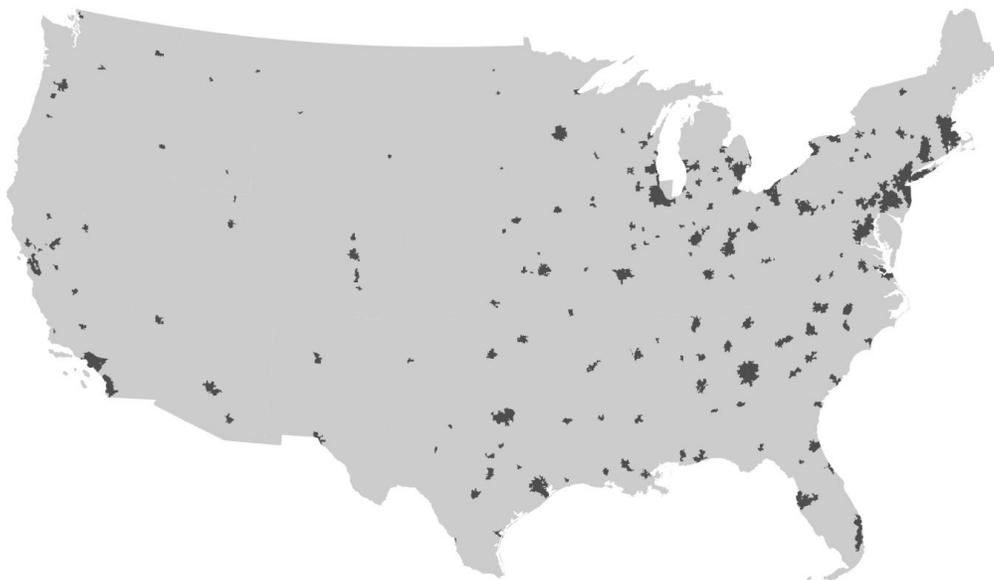
Data for this study came from four different sources: Highway Statistics for VMT and road information by urbanized area, National Transit Database (NTD) for transit information by transit provider, Census Transportation Planning Package (CTPP) for jobs and population information by census tract, and U.S. Census for poverty by tract and urbanized area boundaries. Given the time and geographical inconsistencies across these databases, a major consideration was taken to join variables from different sources and to determine time and geographic definitions for the study. This study employed CTPP 2000 and Highway Statistics 2002. The year 2000 data is the most recent version for CTPP available for this study. Meanwhile, the year 2002 dataset is the first Highway Statistics data using the year 2000 urbanized area definition. Here, the study assumed that the regional demographic structure did not change much during the two years. NTD 2002 was selected for the consistency of transportation variables. The urbanized area population estimates in NTD 2002 showed a correlation coefficient of 0.998 with this study's population estimates based on CTPP 2000.

#### 4.5.2. Study areas

Among 384 U.S. urbanized areas (except Alaska, Hawaii, and Puerto Rico) in Highway Statistics 2002, 26 cases were dropped due to missing data and geographic discrepancies among data sources. Then, 203 areas that showed consistent distribution patterns of jobs, population, and poverty—these were assumed to be concentrated more toward regional centers—were selected. Population and jobs showed this distribution pattern in 331 areas, but only 203 areas showed this pattern in poverty. This selection based on regional development patterns was to consider assumptions of the standard

urban model. The model predicted constantly decreasing job and population density with distance from the centers. The model also predicted lower income households would locate closer to the dense centers while higher income households would locate in the low density outskirts (Giuliano et al. 2008). These hypotheses enable establishing predictive relationships between regional development patterns and regional average VMT. For the other areas with unpredictable regional development patterns, a different research framework is necessary.

Out of the selected areas, 122 areas were identified as monocentric and 81 areas as polycentric based on the center identification results of this study. Figure 4.2 presents the selected 203 urbanized areas.



**Figure 4.2 Study areas for the DVMT study** (Source: U.S. Census 2000)

#### 4.5.3. Variables and measurements

Table 4.1 presents variable names, data sources, and the geographical units of original data. Regional average VMT was measured with daily vehicle miles traveled (DVMT) per capita. DVMT refers to daily traffic volume calculated by traffic counts

multiplied by lane length of major roads, including highways, arterials, collectors, and local roads. The data was provided by the Federal Highway Administration (FHWA) at the urbanized area scale. Since the VMT were estimated from local traffic counts, the FHWA dataset is not perfect. However, it would be the most reliable regional-scale VMT dataset among the available in the U.S. The FHWA has made great efforts to maintain consistent data quality among urbanized areas using a standardized monitoring system, called the Highway Performance Monitoring System.

**Table 4.1**  
**Variable descriptions, measures, and data sources and geographic unit of data**

Variable description	Variable measure	Variable name	Data source; geographic unit of data
<u>Travel outcome</u> Regional VMT	Daily vehicle miles of travel per capita	DVMT/c	Highway Statistics 2002, FHWA; urbanized area
<u>Transportation infrastructure</u> Road supply	Street lane feet per capita	Road/c	Highway Statistics 2002, FHWA; urbanized area
Transit service supply	Daily actual transit revenue feet per capita	Transit/c	NTD 2002, FTA; transit service provider
<u>Regional development patterns</u> Population concentration	Population density gradient	PD.grd	CTPP 2000, Part 1: population at place of residence; census tract CTPP 2000, Part 2: population at place of work; census tract CTPP 2000, Part 1: population with poverty status determined; census tract
Job concentration	Job density gradient	JD.grd	
Poverty concentration	Gradient of poverty rate	Pvt.grd	
<u>Local built environment</u> Neighborhood density	Population-weighted average of tract population density	Local.den	CTPP 2000, Part 1: population at place of residence; census tract
<u>Controls for size and demographic condition</u> Urban area size	Population	Pop	Census 2000; urbanized area
Population profile	Population density	Pop.den	Census 2000; urbanized area
Employment profile	Job density	Job.den	Census 2000; urbanized area
Economic profile	Poverty rate	Pvt.rate	Census 2000; urbanized area

Note: FHWA-Federal Highway Administration; FTA-Federal Transit Administration;  
NTD - National Transit Database; CTPP-Census Transportation Planning Package

Policy-relevant variables include three regional development pattern variables (representing the concentrations of residents, jobs and poverty), one local built environment variable (representing neighborhood compactness), and two transportation infrastructure variables (representing road and transit supply). Other variables in the model—population, population density, job density and poverty rate—served as statistical controls for regional demographic aspects.

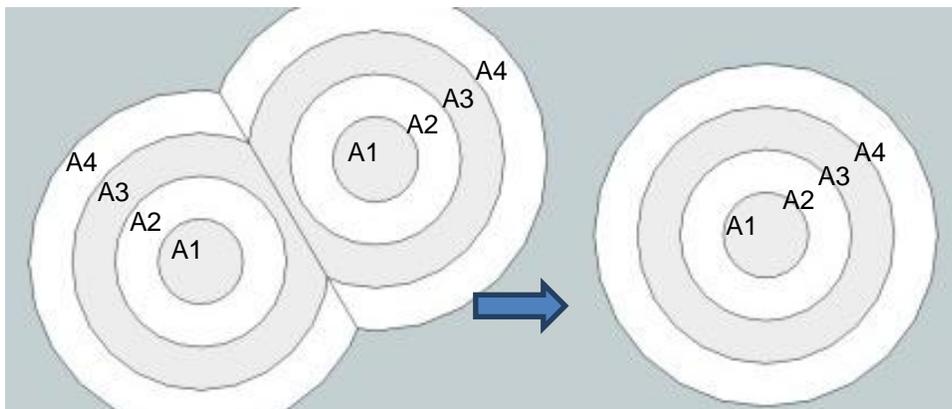
Road supply was measured with “lane feet per capita.” While this study interpreted this measure as a transportation infrastructure variable, some micro studies used this as an urban form design factor. Transit supply is measured with “actual vehicle revenue feet,” which refers to the distance that buses or railway vehicles travel while in their revenue service. NTD provided the data by service provider. All provider information in a same urbanized area was summed up.

Local built environment was measured with population density only, while micro approach studies typically used three Ds (i.e., density, diversity, and design). The other two Ds could not be measured in this study due to data limitations and the aggregate nature of this study. However, density is considered an effective proxy for the remainders at the large scale (Ewing 1994; Cervero 1993). To aggregate the densities of multiple neighborhoods at the regional scale, the population-weighted average of tract population density was used. This regional index is designed to increase as more people live in denser neighborhoods. It is calculated by  $\sum_i (P_i \cdot D_i) / \sum_i (P_i)$ , where  $P_i$  is population at tract  $i$ , and  $D_i$  is population density at tract  $i$ .

Regional development patterns were measured in terms of spatial distributions of residents, jobs, and poverty, each of which characterizes trip origins, destinations, and travel behavior, respectively. This study is particularly interested in how these features are concentrated near regional employment centers because micro studies found the accessibility to regional centers is a key VMT variable. Gradient measures were used to quantify regional development patterns due to two advantages they provide. A gradient is effective to present the concentration toward a specific place like an employment center. Also, gradients are easy to interpret and convey straightforward policy implications.

For example, population density gradient shows the changing trend of density from center to periphery. If this measure has a significant influence on VMT, then it can suggest location specific density control guidelines based on distance from the center, such as rural-to-urban transect. This study used the simple exponential decay function to calculate gradients,  $y = \text{EXP}[\alpha x + \beta]$ , where  $y$  is the variable of interest (i.e., job density, population density, or poverty rate) at distance  $x$  from center,  $\alpha$  is gradient, and the exponential of  $\beta$  is the estimated density of center. The gradient  $\alpha$  can be calculated by a linear regression with the transformation of the original function,  $\text{LN}(y) = \alpha x + \beta$ , where the gradient  $\alpha$  is the slope of regression analysis. Urbanized areas with statistically insignificant  $\alpha$  at the 0.05 level were excluded from this study.

A major limitation of traditional gradient measurement is that it can only be used in monocentric urban models, while many contemporary metropolitan areas have polycentric urban structure. There are several gradient estimation methods for polycentric models (e.g., Craig and Ng 2001 and McMillen 2001; 2004), but these methods are technically complicated to calculate and translate into policy guidelines. To take advantage of the simplicity of the monocentric model, this study translated a polycentric structure into a monocentric structure by assuming that all employment centers (including CBD) are placed at the same location (Figure 4.3). A gradient value measured in this assumption is identical to the average of gradients from each center.



**Figure 4.3 An operation to translate polycentricity using monocentric model (DVMT study)**

Center identification is the first step to estimate gradients of polycentric regions. Among a number of center identification methods (e.g., McDonald 1989; Gordon et al. 1989; Giuliano and Small 1991; McMillen 2001; 2003; and Redfearn 2007), this study employed the two-stage non-parametric method of McMillen (2001). McMillen's method is one of the most commonly used center identification methods, and allows for flexible applications to any areas without requiring local knowledge. Therefore, it was selected to be appropriate for this study involving a large number of urbanized areas across the U.S. Following McMillen's process, center candidates were first identified as tracts having significantly higher densities than a smoothed density surface. The second stage determined centers among the candidates based on the significance of influence on job densities of nearby locations. Specifically, the method includes three steps: (1) estimate a job density surface using locally weighted regression (LWR); (2) select candidate centers as significantly greater residuals from the LWR estimates; and (3) exclude insignificant candidates based on the influence of proximity to each candidate on density. In this procedure, center identification results depend on three modifiable parameters: (1) window size for the LWR, (2) p-value to determine significant residuals, and (3) way to control CBD influence compared to other subcenters. Section 2 compared eight combinations of different parameter values for their performance in identifying employment centers, based on the model fit and the gradient of density model estimated from each parameter combination and the application to a VMT estimation model. This study used the best performing parameter combination out of the eight combinations—the window size of 25%, the p-value of 0.1, and the same control between CBD and subcenters.

#### 4.5.4. Analytical methods

A directed acyclic graph (DAG) was used to identify the structure of relationships. Directed graph stems from the field of artificial intelligence and computer science to present causal relationships among a set of variables using an arrow graph (Spirtes et al. 2001). DAG is a directed graph with no cyclic path, where there is at least

one ultimate outcome variable, such as VMT used in this study. Several search algorithms are available to calculate paths of the structural relationships. These algorithms statistically determine the path between any two variables in a model among four possible paths: one-direction, correlation, no relation, or unable to be determined. The algorithms can be broadly classified into the PC algorithm family (e.g., conservative PC, PCD, and FCI) and the greedy equivalent search (GES) algorithm.

The PC algorithm starts the search with a complete, undirected graph which has a line (called “edge”) with no arrowhead connecting each variable (called “node”) with every other variable (Spirtes et al. 2000). Edges between nodes are removed sequentially based on conditional independence. Edges that survive this removal process are then directed using simple rules. For example, imagine a triangle of variables X, Y, and Z, and it is known that Z is correlated with both X and Y. If the edge between X and Y are conditionally independent given Z, then we can direct X—Z—Y as  $X \rightarrow Z \leftarrow Y$ . Any full path structure can be read as a set of multiple triangles. Since this algorithm determines path by individual edge, each variable should have a normal distribution in order to determine the independence between any two variables in the triangular relationship.

The GES algorithm is a stepwise search over alternative DAGs using Bayesian posterior scores (Chickering 2002). The algorithm consists of two stages. It begins with a DAG in the condition that all variables are independent with each other (i.e., no edges between nodes). Edges are added and/or edge directions reversed in a systematic search across classes of equivalent DAGs if the Bayesian posterior score is improved. The first stage ends when a local maximum of the Bayesian score is found such that no further edge additions or reversals improve the score. After this first stage, the second stage commences to delete edges and reverse directions, if such actions result in improvement of the Bayesian posterior score. The algorithm terminates if no further deletions or reversals improve the score.

These algorithms determine relationships based purely on statistical procedures and may produce unreasonable results. To avoid this problem, users can set up specific conditions based on rational hypotheses or knowledge. As described in the conceptual

framework (Figure 4.1), this study classified variables into three hierarchical tiers (i.e., control, policy, and outcome variables) and established a rule that variables in the lower tier do not influence variables in the higher tier (the lowest tier is the outcome variable, DVMT per capita).

Using the Tetrad IV software, this study attempted multiple algorithms, searching for models with better model fit. Models with logarithm-transformed data were also tested, particularly because PC-based algorithms may require stricter normality for each individual variable. To determine reasonable DAGs, every tested model was evaluated with the relative chi-square, for which value closer to 1 indicates a better-fitted model (Carmines and McIver 1981). Then final DAG models were selected after excluding worse-fitted models and models showing a duplicative relationship structure.

Direct and indirect influences of explanatory variables on regional VMT were estimated by structural equation modeling (SEM) based on the DAGs selected. SEM is useful for tracking relative effects of variables on each other through hypothesized relationship paths. SEM estimates parameter coefficients by solving simultaneous equations of a series of hypothesized relationships in a model. The technique uses iterative methods (e.g., maximum likelihood method, generalized least square, etc.) that involve a series of attempts to obtain estimates of unknown parameters until it finds the model covariance matrix best fitted with the actual covariance matrix representing the hypothesized relationships (Hoyle 1995). Therefore, SEM can be regarded as a general method for constructing a predictive model with relatively few restrictions (e.g., regression analysis, factor analysis, and ANOVA are special cases of SEM). The AMOS 20 software package was used to conduct SEM analyses in this study.

#### 4.5.5. Descriptive statistics

Table 4.2 shows descriptive statistics of variables used, including normality indices. Since both DAG and SEM calculate model-fit statistics using normal theory-based estimation methods, one of the main concerns about the data in this study is

whether the sample has a multivariate normal distribution. In a strict perspective, multivariate normality is not necessarily achieved by the normality of every single variable (Mardia 1974). However, researchers often found that a looser standard can be allowed for structural equation models (Gao et al. 2008; Kline 1998; Voortman and Druzdzel 2008). For example, Klein (1998) suggested that a distribution with a skewness less than 3 and kurtosis less than 10 is allowable. Most variables met with Klein's normality criteria except population (Pop) and local density (Local.den). Both variables were extremely skewed because the majority of study areas had a population of less than 200,000. Three extreme outliers (i.e., New York, NY; Los Angeles, CA; and Chicago, IL) also might influence their normality. Among the two extreme non-normal variables, population does not have to satisfy the normality assumption since it is an exogenous variable in the model (i.e., population is not influenced by any other variable). According to regression theory, the normality assumption is applied to residuals, and therefore, endogenous variables. The local built environment variable (Local.den) still has the non-normality problem. Common suggestions to solve this problem include variable transformation to improve the univariate normality of each individual variable (Andreassen et al. 2006; Bollen 1989; Yuan et al. 2000) and the deletion of outliers to improve multivariate normality (Bagley and Mokhtarian 2002). However, as Gao et al. (2008) pointed out, "the pursuit of a multivariate normal distribution by the deletion of observations should be consciously weighed against the loss of model power and generalizability in the interpretation of the results (p. 116)." This study weighed model interpretation more heavily and did not delete data. Instead, logarithm-transformed models were tested. Natural logarithm transformation made all variables pass the normality test based on the Kolmogorov-Smirnov index (Lilliefors 1967) and Klein's (1998) criteria, though it still did not meet the multivariate normality condition of less than 3 of Mardia's statistic (Mardia 1974).

**Table 4.2**  
**Descriptive statistics of variables**  
(values in parenthesis are calculated from the natural log-transformed data)

Construct	Variable	Unit	Min.	Max.	Mean	Std. Dev.	Skewness	Kurtosis
Outcome	DVMT/c	mi/person	12.7 (2.54)	43.4 (3.77)	22.8 (3.1)	5.3 (0.23)	0.6 (-0.1)	0.6 (-0.2)
Regional demographic characteristics	Pop	thousand	51 (3.93)	17485 (9.77)	782 (5.72)	1777 (1.22)	6.1 (0.8)	46.8 (0.2)
	Job.den	/sq. mi	634 (6.32)	3615 (8.06)	1293 (7.11)	441 (0.30)	2.0 (0.3)	6.7 (0.3)
	Pop.den	/sq. mi	1337 (7.14)	7514 (8.86)	2475 (7.75)	926 (0.30)	2.4 (0.8)	7.9 (0.9)
	Pov.rate	percent	4.4 (1.48)	34.5 (3.54)	13.4 (2.53)	4.8 (0.33)	1.5 (0.2)	3.5 (0.9)
Transportation infrastructure	Road/c	ft/person	11 (2.40)	68.8 (4.23)	26 (3.22)	7.7 (0.30)	1.0 (-0.4)	4.2 (0.6)
	Transit/c	ft/person	31.3 (3.44)	671.2 (6.51)	165.5 (4.95)	99.7 (0.56)	1.9 (0.0)	5.7 (0.0)
Local built environment	Local.den	/sq mi	974 (6.88)	33540 (10.42)	3628 (8.06)	2799 (0.48)	6.7 (0.9)	65.8 (2.7)
Regional development patterns	JD.grd		0.062 (1.82)	1.205 (4.79)	0.373 (3.43)	0.229 (0.62)	1.0 (-0.1)	0.6 (-0.7)
	PD.grd		0.037 (1.31)	0.936 (4.54)	0.265 (3.06)	0.178 (0.67)	1.2 (0.0)	1.2 (-0.7)
	Pvt.grd		0.012 (0.18)	0.559 (4.02)	0.192 (2.75)	0.116 (0.70)	0.8 (-0.7)	0.0 (0.4)

Note 1: Refer to Table 4.1 for detailed descriptions of variables.

Note 2: The log-transformations of gradient measures were done after multiplying the original value with 100 to avoid a negative value.

DVMT/c is correlated with almost all variables (except population and transit supply) in the model. Many factor variables also are significantly correlated with each other as shown in Table 4.3. This raises the multicollinearity problem that can distort estimations of explanatory powers of factor variables in a multiple regression model. Structural equation modeling (SEM) is not free of multicollinearity, as the model is a simultaneous equation form of several regression models. However, SEM often can relieve the multicollinearity problem by decomposing a single regression model with many independent variables into many simpler regression models.

**Table 4.3**  
**Standardized correlations (Pearson's R) among variables**  
 (values in parenthesis are calculated from the natural log-transformed data)

Variables	DVMT/c	Pop	Job.den	Pop.den	Pvt.rate	Road/c	Transit/c	Local.den
DVMT/c	1							
Pop	-.047 (.312**)	1						
Job.den	-.163* (-.247**)	.429** (.309**)	1					
Pop.den	-.214** (-.329**)	.525** (.431**)	.876** (.846**)	1				
Pvt.rate	-.168* (-.153*)	-.096 (-.242**)	.131 (.032)	.151* (.066)	1			
Road/c	.445** (.439**)	-.347** (-.450**)	-.489** (-.538**)	-.568** (-.671**)	.043 (.064)	1		
Transit/c	-.117 (-.114)	.511** (.379**)	.490** (.486**)	.448** (.465**)	-.027 (-.005)	-.343** (-.358**)	1	
Local.den	-.230** (-.320**)	.819** (.521**)	.618** (.735**)	.691** (.869**)	.048 (.044)	-.495** (-.696**)	.621** (.563**)	1
JD.grd	-.403** (-.403**)	-.341** (-.740**)	.085 (.092)	-.074 (-.059)	.387** (.302**)	.110 (.155*)	-.070 (-.103)	-.098 (-.111)
PD.grd	-.416** (-.427**)	-.318** (-.750**)	.016 (.031)	-.113 (-.101)	.327** (.249**)	.085 (.171*)	-.056 (-.100)	-.065 (-.070)
Pvt.grd	-.337** (-.314**)	-.333** (-.622**)	.055 (.103)	-.071 (-.024)	.333** (.243**)	.180 (.185**)	-.152* (-.116)	-.096 (-.038)

\*\* Correlation is significant at the 0.01 level

\* Correlation is significant at the 0.05 level

Note: Refer to Table 4.1 for detailed descriptions of variables

## 4.6. Results and Implications

### 4.6.1. Directed acyclic graph (DAG)

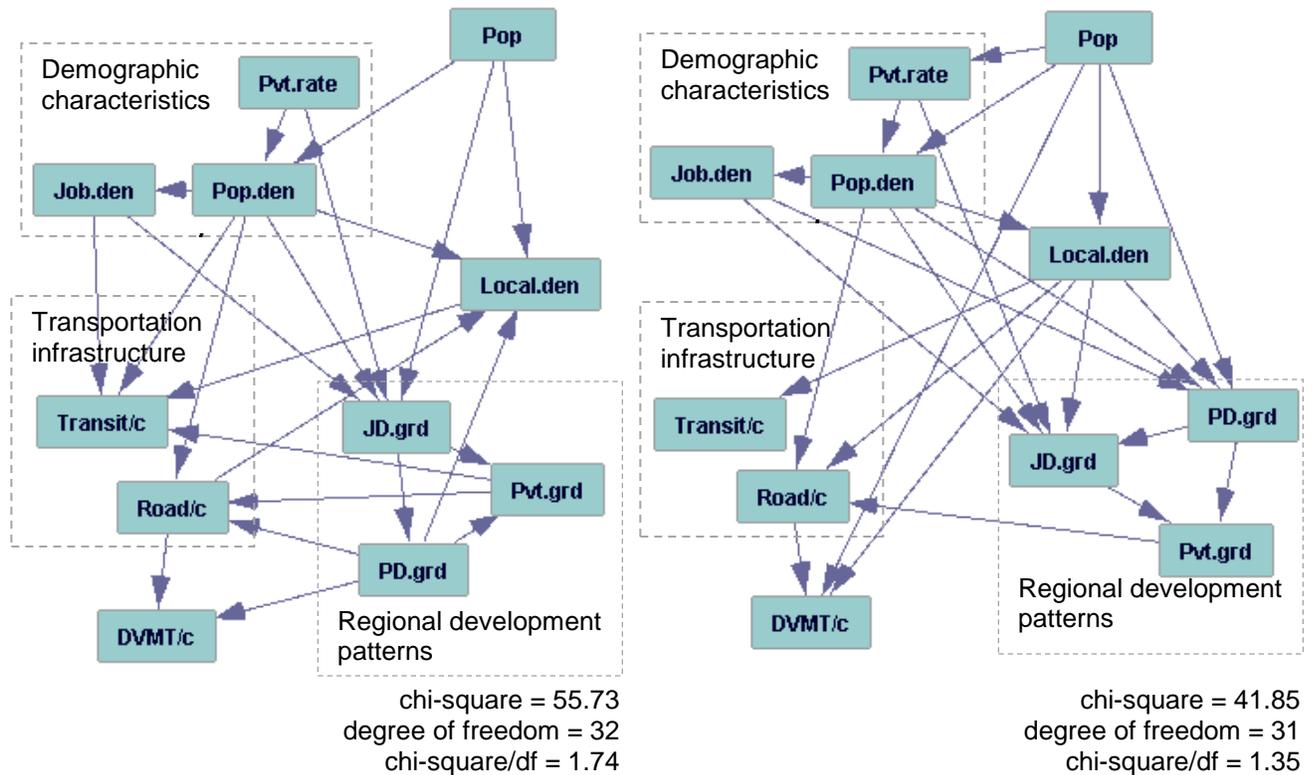
To find more reasonable structural relationships, multiple algorithms (i.e., GES and PC and its modifications) were tested to draw DAGs. Both non-transformed and transformed data were tested in the DAG procedure. Model fit was estimated using the relative chi-square (chi-square divided by the degree of freedom), in which a value less than 2 is regarded as reasonably fitted (Carmines and McIver 1981).

The two different types of reasonably fitted DAGs were found: the GES with non-transformed data (GES-non) and the GES with log-transformed data (GES-log)

(Figure 4.4). Both the GES-non and the GES-log agreed that road supply (Road/c) had a direct effect on VMT. A major distinction between the two was the paths between urban form variables and VMT. The GES-non model identified that VMT was directly influenced by population density gradient (PD.grd) while local density (Local.den) had no effect. The model also implies more compact neighborhoods were achieved by the regional development patterns, showing local density level was a product determined by regional development patterns and road supply. On the contrary, the GES-log model supports a counter theory that local density has a direct impact on VMT while regional development patterns only have an indirect impact. This is an interesting distinction in planning perspective particularly because the GES-log model reinforces the traditional “compact neighborhood” principle whereas the GES-non model supports the “rural-to-urban transect” concept. The clear distinction between the two models was presented later through parameter estimates by SEM.

Another noteworthy implication found in both models is that transit supply (Transit/c) has no significant effect on VMT per capita. This would not indicate the uselessness of transit development, but rather indicates that most U.S. urbanized areas are primarily dependent on private vehicles and transit supply is not sufficient enough to affect vehicle usages. Since this study focuses on influences on VMT, the transit supply variable was excluded in SEM analyses.

There also exist some differences in relationships between control and regional development pattern variables. For example, control variables have direct relationships with both job and population density gradients in the GES-log model, while the GES-non model gives direct paths between control variables and job density gradient only. However, this difference does not seem to suggest meaningful policy implications but to simply indicate the overall association between the two constructs. Areas with a larger population tend to have a higher density and a steeper density gradient, but planning policies cannot control population and regional density to manage density gradient.



(a) DAG searched by GES algorithm with non-transformed data

(b) DAG searched by GES algorithm with log-transformed data

Note: Refer to Table 4.1 for detailed descriptions of variables

**Figure 4.4 Directed acyclic graphs of variables influencing vehicle miles traveled per capita among 203 U.S. urbanized areas, 2002**

#### 4.6.2. Structural equation modeling (SEM)

The path models derived from the DAG searching process were estimated using SEM. Since the two models used different types of data, their results also need to be interpreted differently. The non-transformed model has a wide variety of variable units; thus, standardized effect size is useful to compare the relative explanatory power among factor variables. In the log-transformed model, non-standardized effect size represents elasticity that reflects the relative sensitivity (percentage change) of DVMT/c to a 1% increase in each factor variable, holding other factors constant. Table 4.4 summarizes SEM results by model, including direct and indirect effects of factor variables on

DVMT/c and model-fit statistics. Goodness-of-fit was measured with five indices as follows. The values in parentheses indicate critical thresholds to determine a good model fit (Hu and Bentler 1999).

- Relative chi-square ( $< 2.0$ )
- Comparative fit index: CFI ( $> 0.90$ ),
- Normed fit index: NFI ( $> 0.95$ ),
- Tucker-Lewis Index: TLI (or non-normed fit index: NNFI) ( $> 0.90$ ), and
- Root mean square error of approximation: RMSEA ( $\approx 0.05$ ).

**Table 4.4**  
**Effects on DVMT/c and model fit statistics, estimated using SEM**

Variables		GES-non (standardized effects)			GES-log (elasticity)		
		Total Effect	Direct Effect	Indirect Effect	Total Effect	Direct Effect	Indirect Effect
PD.grd	Pop. density gradient	-0.500	-0.448	-0.053	0.024	-	0.024
JD.grd	Job density gradient	-0.389	-	-0.389	0.015	-	0.015
Pvt.grd	Poverty rate gradient	0.160	-	0.160	0.030	-	0.030
Local.den	Local density	-	-	-	-0.257	-0.155	-0.102
Road/c	Lane feet per capita	0.484	0.484	-	0.424	0.424	-
Pop	Population	-0.036	-	-0.036	0.060	0.138	-0.079
Pop.den	Pop.density	-0.292	-	-0.292	-0.449	-	-0.449
Job.den	Job density	-0.240	-	-0.240	0.016	-	0.016
Pvt.rate	Poverty rate	-0.198	-	-0.198	-0.070	-	0.070
<i>Model statistics</i>							
Chi-square		52.67			22.02		
Degrees of freedom		26			22		
Relative chi-square		2.026			1		
CFI ( $> 0.90$ )		0.983			1.000		
NFI ( $> 0.95$ )		0.967			0.988		
TLI ( $> 0.90$ )		0.971			1.000		
RMSEA ( $\approx 0.05$ )		0.071			0.002		

Note 1: GES-non is a DAG model searched by GES algorithm with non-transformed data, and GES-log is a model searched by GES algorithm with log-transformed data.

Note 2: Refer to Figure 4.4 for detailed relationship paths in each model. All paths in the two models are statistically significant at the 0.01 level.

The GES-log model, passing all these criteria, showed a better model fit. The GES-non model also showed a good fit overall but only marginally satisfied the relative chi-square and RMSEA criteria. All path coefficients in the two models were statistically significant at the 0.01 level. The models agreed that the strongest positive predictor was road supply (Road/c) – the more road supply, the more VMT. They also agreed that population has a minimal effect, indicating the size of region is less important in determining the level of VMT.

Regional population density had a relatively strong influence on VMT in both models (i.e., -0.292 of standardized effect in the GES-non and -0.449 of elasticity in the GES-log). This result is consistent with many previous studies (Cervero and Murakami 2010; Ewing et al. 2008; Newman and Kenworthy 1989; 1999). However, it should be noted that its effect is indirect in both models. For example in the GES-non, population density influenced VMT through two indirect paths, “Pop.den→Road/c →DVMT/c” and “Pop.den→JD.grd→PD.grd→ DVMT/c” (Figure 4.4-a). An indirect path can be interpreted in two different ways – an actual causality path or a spurious relationship. A relationship between X and Y can be regarded as spurious when a third intervening factor W affects both variables. This can be described as “X←W→Y” in a DAG. Determining the type of relationship is a major function of the DAG searching process. In both the GES-non and the GES-log models, DAG searching algorithms agreed with the pattern of “Pop.den → Road/c → DVMT/c,” which indicates the actual indirect relationship between regional population density and VMT.

Meanwhile, the two models were inconsistent results regarding VMT reduction factors: population density gradient (PD.grd) and job density gradient (JD.grd) in the GES-non and regional population density (Pop.den) and local density (Local.den) in the GES-log. This indicates that both models agreed on the significant influence of urban form on VMT but disagreed about the rationale of the influence. Considering that population density gradient and local density has a direct effect in GES-non and GES-log, respectively, this disagreement indicates the theoretical competition between the “compact neighborhood” principle versus the “rural-to-urban transect” concept.

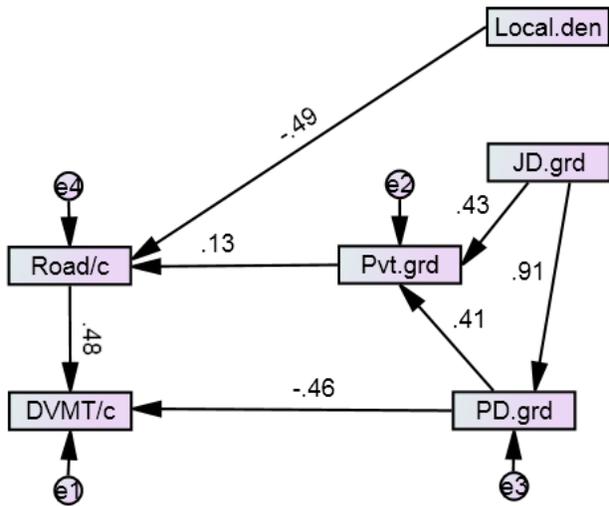
#### 4.6.3. Structural equation modeling (SEM) without control variables

The full model results could not provide a clear decision about which theory is more plausible. To evaluate the relative applicability of the competing theories (i.e., compact neighborhood development versus regional development pattern control), another series of SEM analyses were conducted after excluding indirect control variables, such as population, population and job density, and poverty rate. The removal of these variables could be justified because they have only indirect or minimal impacts. These simple models also would help avoid possible model estimation errors by making the non-normal variable exogenous. Considering a possible distortion by the control variables onto the whole relationship structure, another set of path searching was conducted for the simplified models. All algorithms identified three discrete patterns of relationships in common – "Local.den - Road/c - DVMT/c," "PD.grd - DVMT/c," and "Road/c - Pvt.grd" – in both non-transformed and log-transformed models. No link was identified between PD.grd and Local.den. A path diagram was developed based on the identified relationships, and SEM analyses were conducted with both non- and log-transformed data using the same path diagram.

Figure 4.5 presents the results. The models showed a good fit, and every path was significant at the 0.01 level. The non- and log-transformed models yielded similar results in terms of model fit and coefficient estimates. Different from the full-models, local density and regional development pattern variables showed similar effects on VMT in both simple models

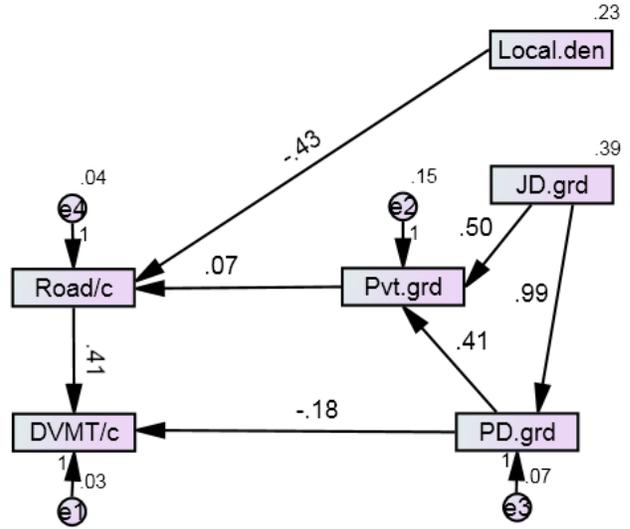
The concentration of population toward centers was the primary player to lower average VMT. Population density gradient (PD.grd) had a direct effect, and its net elasticity was relatively higher (-0.169) among urban form variables (Table 4.5). This conceptually agreed with previous studies which found a significant relationship either between regional average gasoline use and job centrality (Van de Coevering and Schwanen 2006) or between individual VMT and distance to CBD (Ewing and Cervero 2001, 2010).

(a) Non-transformed model  
(number indicates standardized effect)



Chi-square = 10.84  
 Degrees of freedom = 8  
 Relative chi-square = 1.355  
 CFI (> 0.90) = 0.996  
 NFI (> 0.95) = 0.986  
 TLI (> 0.90) = 0.993  
 RMSEA ( $\approx$  0.05) = 0.042

(b) Log-transformed model  
(number indicates elasticity)



Chi-square = 12.20  
 Degrees of freedom = 8  
 Relative chi-square = 1.53  
 CFI (> 0.90) = 0.995  
 NFI (> 0.95) = 0.986  
 TLI (> 0.90) = 0.991  
 RMSEA ( $\approx$  0.05) = 0.051

Note 1: All path coefficients are significant at the 0.01 level.

Note 2: DVMT/c – daily VMT per capita; Road/c – lane feet per capita; Local.den – population-weighted average of tract density; PD.grd – population density gradient; JD.grd – job density gradient; Pvt.grd – poverty rate gradient

**Figure 4.5 Structural equation modeling of simplified urban form-VMT models**

**Table 4.5 Effects on DVMT/c in simplified urban form-VMT models, estimated using SEM**

		Non-transformed model (standardized effects)			Log-transformed model (elasticity)		
		Total Effect	Direct Effect	Indirect Effect	Total Effect	Direct Effect	Indirect Effect
PD.grd	Pop. density gradient	-0.433	-0.459	0.026	-0.169	-0.181	0.012
JD.grd	Job density gradient	-0.368	-	-0.368	-0.153	-	0.015
Pvt.grd	Poverty rate gradient	0.065	-	0.065	0.028	-	0.030
Local.den	Local density	-0.235	-	-0.235	-0.178	-	-0.178
Road/c	Land feet per capita	0.484	0.484	-	0.412	0.424	-

Note: Refer to Figure 4.5 for detailed relationship paths in each model. All paths in two models are significant at the 0.01 level.

In addition to the direct effect, population density gradient also exerted its influence under close relationships with other regional development pattern variables. The path of “JD.grd→PD.grd→ DVMT/c” can be translated as follows: the concentration of jobs toward centers tends to attract more population near to the centers reducing the average travel distance. The elasticity between JD.grd and PD.grd is very high, 0.99. Also, the path of “JD.grd→Pvt.grd→Road/c →DVMT/c” can be interpreted that the stronger centers tend to attract low income people more strongly and push out higher income population to the outskirts. This process would result in an increase in the demand for roads and travels by the higher income group who tend to drive more. An interesting point here is that the dispersion of poverty showed an effect to reduce road construction and VMT (0.028 of elasticity). This is somewhat surprising because all gradients of poverty, population density, and job density had a positive relationship with each other, and the dispersion of population and jobs significantly increased VMT. The impact of poverty gradient suggests that the efforts to relieve income segregation will be in sum beneficial to a region’s sustainability goals. It reduces VMT, reduces road construction, and makes many equity goals achievable, such as lessening racial segregation and the inequality in labor market, educational, and health outcomes (Ananat 2007; Cutler and Glaeser 1995; Ellen et al. 2000; Mayer 2002)

The path of “Local.den → Road/c → DVMT/c” suggests that areas having more high-density neighborhoods tend to require fewer road supply and yield less VMT. This finding agrees with the claim of “compact city” advocates. The net elasticity of local density (-0.178) is slightly higher than the elasticity of population density gradient (-0.169), while the non-transformed model indicates that population density gradient has a higher effect than local density (-0.433 versus -0.235 of standardized effect).

#### **4.7. Discussion**

At first glance, the test results of the simplified models appear to present a balanced view between the “rural-to-urban transect” and the “compact neighborhood” principles, because the influences of local density and regional development patterns are

independent of each other. Initially, this study tested models conditioning that local density and job density gradient have a covariate relationship, but the relationship was confirmed insignificant. However, this independence is somewhat odd in policy perspective. If we consider urban form change in a single region, it is reasonable that an increase of population density gradient means more people living in denser neighborhoods thus causing an increased local density measure. The opposite relationship can be considered too. Density gradient will become gentler as more compact neighborhoods are developed further from a center. Hence, population density gradient and local density are expected to have a significant relationship in any direction. The insignificant relationship between these two actually confuses policy decision making between “transect” (or infill development) and “compact neighborhood.”

This seemingly contradictory result has to be interpreted differently considering that this study is cross-sectional rather than longitudinal. A more valid implication of the independence between regional development patterns and local density in this cross-sectional study is that many urbanized areas have compact neighborhoods not only near centers but also in the outskirts. Consequently, the variances in development patterns in suburbs weakened the relationship between the regional and local variables. Compact development in suburbs must be an extra benefit in comparison to dispersed development there. This would be the reason why local density has a significant effect after controlling for regional variables. Here the effects of regional variables are regarded as additional, and thus an area with compact neighborhoods near centers has much lower average VMT by the summed effects of local density and regional density gradient.

Hence, in terms of policy implication, the test results indeed indicate “rural-to-urban transect” is superior to “compact neighborhood” for reducing VMT. The “transect” becomes an important element of New Urbanism, which provides a regional framework for organizing human habitats in a range of intensity from the most rural environment to the most urban. Although the Charter for the New Urbanism includes regional principles from the early stage, the early New Urbanism concept was generally

organized in the emphasis of compact developments at the neighborhood scale, such as traditional neighborhood developments (TNDs) and transit-oriented developments (TODs) (Bohl and Plater-Zyberk 2006). Data shows an interesting tendency that implies the consequence of this fashion – larger urbanized areas tend to have more high-density neighborhoods but lower density gradients. That is, many compact neighborhoods have been developed in suburbs. Based on the results of this study, the compact neighborhood approach appears to achieve a partial success and rather might lessen the chance to reduce VMT.

Two reasons for the limited success of TOD can be explained by this study. One is the limited uses of transit. The provision of transit services was enlarged as dense neighborhoods increased (the path of “Local.den→ Transit/c” in Figure 4.4-a), but the effect of transit supply on VMT was not significant. The other is the urban nature of unequal spatial distribution of jobs and population. Quality jobs and markets tend to be clustered at limited areas, such as CBD or other subcenters, as supported by urban economic theories such as bid-rent theory and agglomeration of economy. Among 348 urbanized areas identified in this study, 333 areas have a significant job density gradient, and 324 areas showed a significant population density gradient. The strong effect of density gradient on DVMT/c indicates that denser development at a location remote from regional centers may just make residents increase their VMTs, particularly in the private vehicle oriented world.

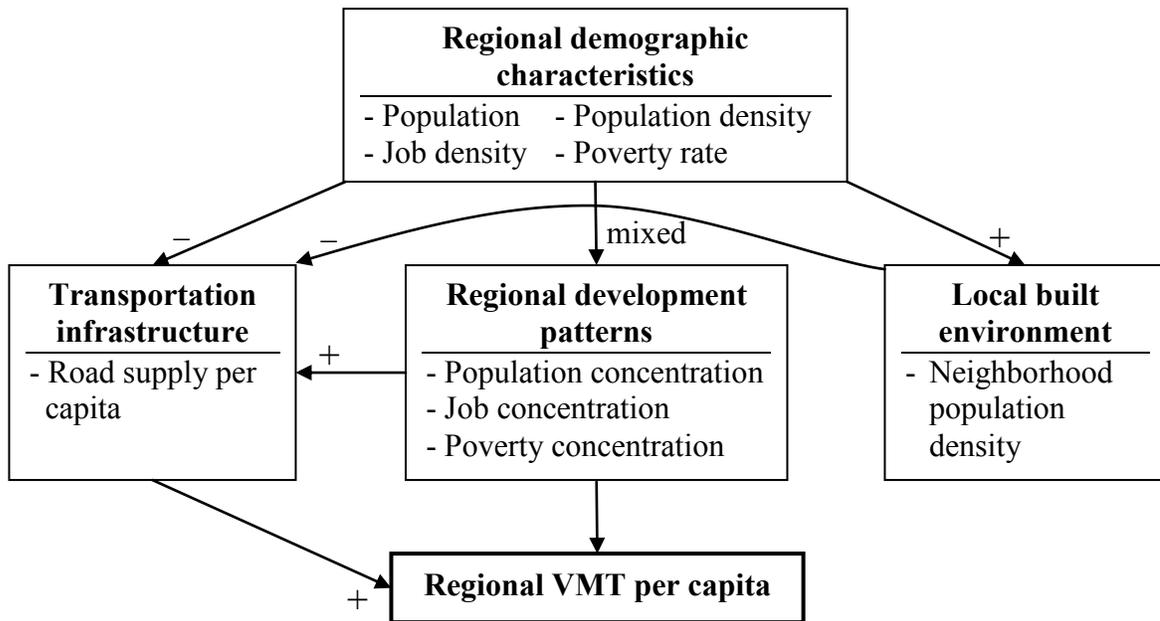
#### **4.8. Summary and Conclusion**

This study examined the influence of local and regional development patterns on collective VMT, considering polycentric urban structure and complicatedly intervening factors (e.g., transportation infrastructure, and income segregation). A series of cross-sectional analyses of 203 U.S. urbanized areas was conducted, using directed acyclic graph (DAG) and structural equation modeling (SEM), which are particularly useful to address complex relationships. Recognizing “distance to regional centers” as the key dimension to quantify regional development patterns conducive to reducing regional

VMT, this study measured regional development patterns with changing trends of population, jobs and poverty rate by the distance from centers. In the calculation process, multiple centers were identified to consider the polycentric urban structure of modern metropolitan areas, and gradients were calculated after operationally transferring the polycentric structure into the monocentric one in the assumption that all employment centers are placed at a same location.

As illustrated in Figure 4.6, major findings of this study are as follows: 1) VMT is lower as more people and jobs are located more closely to centers (or steeper population and job density gradients); 2) the higher concentration also increases VMT by requiring more road supply; 3) VMT is lower as more people live in denser neighborhoods (or higher population-weighted average of tract density); 4) road supply has a strong positive impact on VMT; 4) transit supply is not related to VMT; 5) poverty gradient has a minimal effect on VMT; and 6) the effects of local versus the regional urban form on VMT were independent of each other. These findings collectively suggest that the “rural-to-urban transect” approach that considers both regional and local urban form effects is superior to traditional “compact neighborhood” approach that only considers local urban form effects.

This study encompasses a comprehensive picture of regional sustainable development principles such as smart growth, New Urbanism, and compact city. The analyses were intended to explore if these principles and methods are promising in terms of a concrete sustainability performance indicator, VMT. Specific attention was paid to clarifying a few potential factors that may distort the urban form-VMT relationship, such as polycentricism, regional vs. local urban form, and spatial income segregation. The use of aggregate data limits detailed controls, but this study helps fill the gap in the disaggregate travel behavior studies that rely on only one or several metropolitan areas and provides additional insights on how to balance travel-associated factors in the regional planning context.



**Figure 4.6 Summary of findings from the DVMT study**

## **5. INFLUENCES OF URBAN FORM ON TRANSIT PASSENGER MILES TRAVELED AT THE METROPOLITAN SCALE**

### **5.1. Outline**

This paper examines the influence of urban form on transit passenger miles traveled (PMT) at the regional scale, particularly focusing on policy-interpretable urban form measures and structural relationships among predictor variables. A series of cross-sectional analyses of 203 U.S. urbanized areas was conducted using directed acyclic graph (DAG) and structural equation modeling (SEM). The study found that: 1) PMT per capita is greater in urbanized areas that supply more transit service hours and have a greater amount of dense neighborhoods; 2) more concentrated regional development patterns were associated with moderate increase of PMT per capita, but this effect may not be generalized throughout the U.S.; 3) while regional population density has been known as the strongest transit increasing factor, its effect might be spurious or only indirect; and 4) road supply did not cause a difference in PMT per capita among U.S. urbanized areas because transit demand and supply is too low to be a viable substitute travel option for driving.

### **5.2. Introduction**

Growing concerns over climate change have attracted keen interests in reducing the amount of vehicle use. Producing 30% of CO<sub>2</sub>, the transportation sector was the second largest source of greenhouse gas after electricity generation in the U.S. in 2008 (U.S. EPA 2011). Promoting transit, with restructuring urban form toward more compact form, has been a major policy means to achieve the auto travel reduction goal. Transit oriented development (TOD) is a representative example. The main idea of TOD is a combination of transit infrastructure constructions and compact neighborhood developments near transit stations to promote transit use and reduce vehicle travel (Cervero et al. 2002). This idea has been accepted by many State/local governments in the U.S. as a sustainable planning strategy (U.S. SCCST 2009; CPDR 2008).

Many studies have examined factors affecting transit uses at both the micro and the macro scales. Micro-scale studies examined individuals' transit choices using disaggregate variables such as individual socioeconomic characteristics, car ownership, and local built environment (Cervero and Radisch 1996 Bento et al. 2005 Cervero and Duncan 2006 Ewing et al. 2009 Frank and Pivo 1994 Kitamura et al. 1997 Rodriguez and Joo 2004; Zhang 2004). Macro scale studies were interested in understanding how the aggregate scale of transit use is affected by macro scale influences such as population and employment, regional density, transit fares and service levels (Chen et al. 2011; Taylor et al. 2009; Litman 2004; Gomez-Ibanez 1996; Hendrickson 1986). These two kinds of studies were based on different assumptions, and thus provided different implications. Martel (1996) pointed out that choice theory-based studies at the individual scale have few implications for the behavior of large-scale aggregates.

Micro and macro studies have treated urban form factors differently. Focusing on built environment impact on transit use, micro studies measured various types of urban form attributes, including residential density, employment density, retail density, transit stop density, land use mix, distance to transit stop, distance to central business district, job accessibility and so on. The geographic unit of measurement is also diverse (e.g., census tract, block group, and two-mile distance near travel origins and/or destinations). While many micro studies overall agreed that dense, compact development is more conducive to promoting transit ridership, there exists some inconsistencies about what and how specific urban form attributes affect transit uses (Taylor and Fink 2002). For example, some found that residential density has a significant effect (Dunphy and Fisher 1996; Baker 1994; Ross and Dunning 1997; Cervero 2002; Cervero and Duncan 2006; Zhang 2004), while others negated the effect (Frank and Pivo 1994; Ewing et al. 2009; Pushkar et al. 2000). This inconsistency supposedly came from variances in study area, variables and analysis framework used. Ewing and Cervero (2010) attempted to generalize the results of micro studies with meta-analysis. They found that the distance to nearest transit stop, percent of 4-way intersection and road density have the strongest effect on transit mode choice (0.29, 0.29 and 0.23 of elasticity, respectively). Effects of

neighborhood-level diversity and density were found moderate (0.12 and 0.07 of elasticity, respectively).

Macro studies were designed to evaluate the influences of regionally aggregated attributes on the collective transit ridership. However, focusing on evaluating the effects of fare and service level (e.g., vehicle revenue miles, etc.) change on collective transit ridership, these aggregated level studies tended to inadequately consider urban form effects. Many studies ignored urban form factors in their estimation models (Kain 1996; Lane 2010; Agthe and Billings 1978; Gomez-Ibanez 1996; Chen et al. 2011). Although some macro studies used regional (e.g., city or metropolitan) scale densities of population and employment as proxy measures of urban form (Spillar and Rutherford 1998; Taylor et al. 2009; Bento et al. 2005), these large scale density measures do not deliver land use policy implications that micro studies' urban form measures can provide. Increasing regional density depends on the growth of population and employment at the regional scale, and this is not a land use policy issue. A compact neighborhood development may not contribute to increasing regional density but only lead migrations from one neighborhood to another within a same region. While many macro studies found significant impacts of fare and service level (Gomez-Ibanez 1996; Kain 1996; Agthe and Billings 1978; Lane 2010; Chen et al. 2011; Taylor et al. 2009), some studies found that employment and population change significantly affect transit ridership (Hendrickson 1986; McLeod Jr et al. 1991; Gomez-Ibanez 1996; Litman 2004; Taylor et al. 2009; Chen et al. 2011). This implies a potentially strong effect of urban form on transit use, considering that population and employment size tend to be strongly correlated with population and employment densities, road density as well as local-level density.

Generalizability is a problem in both micro and macro studies. Due to the scarcity of quality disaggregate data, micro studies have been conducted with samples in one or several metropolitan areas and thus could not sufficiently consider regional attributes that might significantly influence transit uses (e.g., the size of region, the size of transit infrastructure, service level, and fare). A majority of macro studies also were

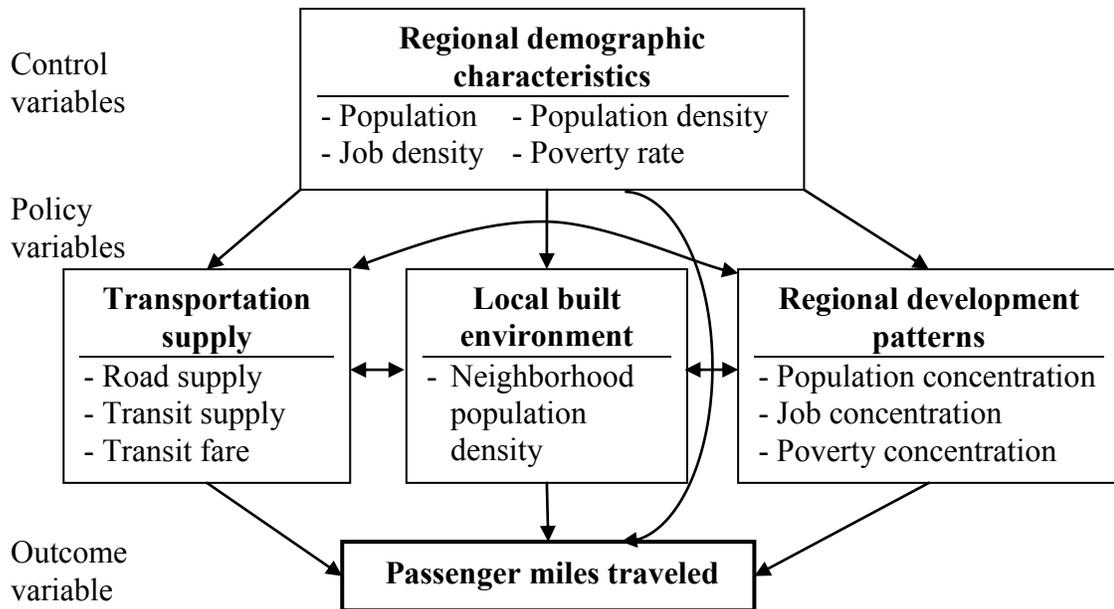
conducted for one or a few regions because they prefer longitudinal analysis to examine the change in transit ridership in a given area. One of few cross-sectional studies at the macro level was done by Taylor et al. (2009). They tested 23 variables measuring geographic, economic, population, road system, and transit system characteristics for 265 U.S. urbanized areas. After eliminating independent variables that were insignificant or highly collinear with other variables, they established a simultaneous equation with two regression models representing transit supply and demand each. The study found that per capita transit supply (vehicle revenue hours) was lower in areas with lower density and lower percent carless households, and located in the South of the U.S. Transit ridership per capita was influenced by transit supply, land area, median household income, non-transit trips (walking, biking, etc.), transit fares, and service frequency.

Another challenge in urban form-transit studies is the generally high level of collinearity among factor variables. Many macro studies identified population, employment, fare, and service level (or transit supply) as significant factors on transit uses. All these variables tend to be highly correlated with each other because transit use level is a balanced outcome between demand and supply. For example, areas with a greater population tend to show greater employment and density, and greater population and density are typically accompanied by more transit supply. In this situation, spurious relationships might exist between some variables and transit use; thus, it is difficult to identify what factors have actual influences. Resolving the structural influence of these various factors on one another and on transit use is a significant methodological challenge (Crane 2000; Gomez-Ibanez 1996).

This study attempted to examine the influence of urban form on transit use with three emphases: 1) focusing on transit passenger miles traveled (PMT) rather than the number of passenger trips because PMT is a more relevant indicator of reduction in fuel usage and emission, 2) using urban form measures that can be translated to land use policy implications, and 3) searching for a structural relationship among variables.

### 5.3. Research Approach

Figure 5.1 is the conceptual framework that shows how this study hypothesized the relationships among study variables. This study classifies influential factors of PMT per capita into three policy-relevant study constructs – regional development patterns, local built environment, and transportation supply factor. The framework has a hierarchical structure that includes urban area characteristics (i.e., population, population density, job density and income) as control variables at the higher level over the three influential constructs of VMT. This structure is devised to serve two purposes: 1) to control various sizes and economic conditions of urban areas, and 2) to set the directions of relationships between these control variables and the key constructs. Regional demographic characteristics are not a product of urban form and transportation supply, but they may influence planning policies and shape distributions of population, job, and infrastructure. Except this structural condition, all possibilities of relationships are allowed between variables in the model, reflecting the relationships in the real world.



**Figure 5.1 Conceptual framework for the DPMT study**

Regions with stronger centers are expected to show the higher concentration of jobs, population and poverty. They also have more dense neighborhoods, and require fewer roads. Key questions of this study are to identify how these variables are interrelated with each other and with PMT, and what variables are more influential than others in predicting PMT per capita at the urbanized area scale.

#### **5.4. Methodology**

A cross-sectional analysis of 203 U.S. urbanized areas was conducted with two methodological emphases. First, gradient-based measures were used to quantify the changes of interested urban form attributes (i.e., population, job, and poverty rate) by distance from centers. While traditional gradient measures were based on a monocentric urban model, this study further considered a polycentric urban structure commonly found in many metropolitan areas. Also, directed acyclic graph (DAG) and structural equation modeling (SEM) were used to control intervening factors (e.g., transportation infrastructure, and income segregation) between urban form and VMT.

##### 5.4.1. Data

Data for this study came from four different sources: Highway Statistics for VMT and road information by urbanized area, National Transit Database (NTD) for transit information by transit provider, Census Transportation Planning Package (CTPP) for jobs and population information by census tract, and U.S. Census for poverty by tract and urbanized area boundaries. Given the time and geographical inconsistencies across these databases, a major consideration was taken to join variables from different sources and to determine time and geographic definitions for the study. This study employed CTPP 2000 and Highway Statistics 2002. The year 2000 data is the most recent version for CTPP available for this study. Meanwhile, the year 2002 dataset is the first Highway Statistics data using the year 2000 urbanized area definition. Here, the study assumed that the regional demographic structure did not change much during the two years. NTD 2002 was selected for the consistency of transportation variables. The urbanized area

population estimates in NTD 2002 showed a correlation coefficient of 0.998 with this study's population estimates based on CTPP 2000.

#### 5.4.2. Study areas

Among 374 U.S. urbanized areas (except Alaska, Hawaii, and Puerto Rico) in Highway Statistics 2002, 26 cases were dropped due to missing data and geographic discrepancies among data sources. Then, 203 areas that showed consistent distribution patterns of jobs, population, and poverty—these were assumed to be concentrated more toward regional centers—were selected. Population and jobs showed this distribution pattern in 331 areas, but only 203 areas showed this pattern in poverty. Out of the selected areas, 122 areas were identified as monocentric and 81 areas as polycentric based on the center identification results of this study. Figure 5.2 presents the selected 203 urbanized areas.



**Figure 5.2 Study areas for the DPMT study (Source: U.S. Census 2000)**

This selection based on regional development patterns was to consider assumptions of the standard urban model. The model predicted constantly decreasing job and population density with distance from the centers. The model also predicted lower income households would locate closer to the dense centers while higher income households would locate in the low density outskirts (Giuliano et al. 2008). These hypotheses enable establishing predictive relationships between regional development patterns and regional average PMT. For the other areas with unpredictable regional development patterns, a different research framework is necessary.

#### 5.4.3. Variables and measurements

Table 5.1 presents variable names, data sources, and the geographical units of original data. Regional average PMT was measured with daily transit passenger miles traveled (DVMT) per capita. Policy-relevant variables include three regional development pattern variables (representing the concentrations of residents, jobs and poverty), one local built environment variable (representing neighborhood compactness), and three transportation policy variables (representing road supply, transit supply, and transit fare). Other variables in the model—population, population density, job density and poverty rate—served as statistical controls for regional demographic aspects.

Road supply was measured with “lane feet per capita.” While this study interpreted this measure as a transportation infrastructure variable, some micro studies used this as an urban form design factor. Transit supply is measured with “actual vehicle revenue feet,” which refers to the distance that buses or railway vehicles travel while in their revenue service. Transit fare is estimated with the passenger revenues divided by unlinked passenger trips. NTD provided the transit data by service provider. All provider information in a same urbanized area was summed up.

Local built environment was measured with population density only, while micro approach studies typically used three Ds (i.e., density, diversity, and design). The other two Ds could not be measured in this study due to data limitations and the aggregate nature of this study. However, density is considered an effective proxy for the

remainders at the large scale (Ewing 1994; Cervero 1993). To aggregate the densities of multiple neighborhoods at the regional scale, the population-weighted average of tract population density was used. This regional index is designed to increase as more people live in denser neighborhoods. It is calculated by  $\sum_i (P_i \cdot D_i) / \sum_i (P_i)$ , where  $P_i$  is population at tract  $i$ , and  $D_i$  is population density at tract  $i$ .

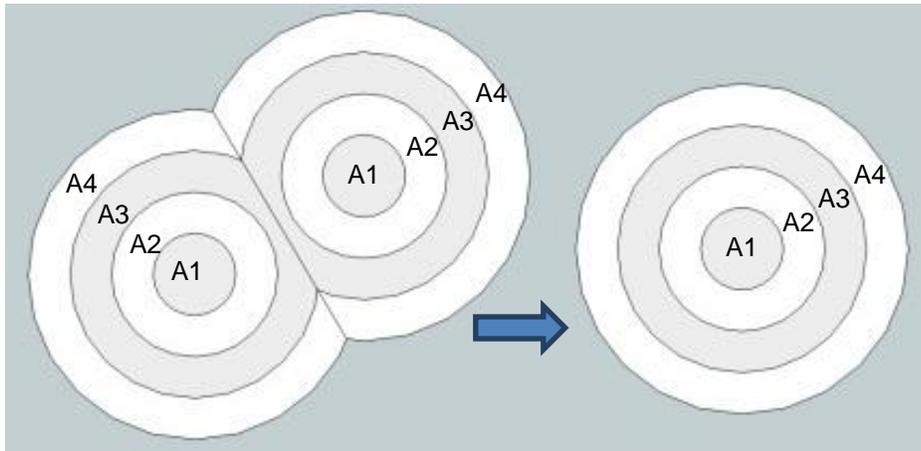
**Table 5.1**  
**Variable descriptions, measures, and data sources and geographic unit of data**

Variable description	Variable measure	Variable name	Data source; unit of measurement
<u>Travel outcome</u> Regional PMT	Daily transit passenger miles traveled	DPMT/c	NTD 2002, FTA; transit service provider
<u>Transportation infrastructure</u>			
Road supply	Street lane mile per capita	Road/c	Highway Statistics 2002, FHWA; urbanized area
Transit service supply	Daily actual transit revenue miles per capita	Transit/c	NTD 2002, FTA; transit service provider
Transit fare	Passenger revenue divided by unlinked passenger trip	Fare	NTD 2002, FTA; transit service provider
<u>Regional urban form</u>			
Population concentration	Population density gradient	PD.grd	CTPP 2000, Part 1: population at place of residence; census tract CTPP 2000, Part 2: population at place of work; census tract CTPP 2000, Part 1: population with poverty status determined; census tract
Job concentration	Job density gradient	JD.grd	
Low income concentration	Gradient of poverty rate	Pvt.grd	
<u>Local built environment</u>			
Neighborhood density	Population-weighted population density	Local.den	CTPP 2000, Part 1: population at place of residence; census tract
<u>Controls for size and demographic condition</u>			
Urban area size	Population	Pop	Census 2000; urbanized area
Population profile	Population density	Pop.den	Census 2000; urbanized area
Employment profile	Job density	Job.den	Census 2000; urbanized area
Economic profile	Poverty rate	Pov.rate	Census 2000; urbanized area

Note: FHWA-Federal Highway Administration; FTA-Federal Transit Administration;  
NTD - National Transit Database; CTPP-Census Transportation Planning Package

Regional development patterns were measured in terms of spatial distributions of residents, jobs, and poverty, each of which characterizes trip origins, destinations, and travel behavior, respectively. This study is particularly interested in how these features are concentrated near regional employment centers because micro studies found the accessibility to regional centers is a key VMT variable. Gradient measures were used to quantify regional development patterns due to two advantages they provide. A gradient is effective to present the concentration toward a specific place like an employment center. Also, gradients are easy to interpret and convey straightforward policy implications. For example, population density gradient shows the changing trend of density from center to periphery. If this measure has a significant influence on VMT, then it can suggest location specific density control guidelines based on distance from the center, such as rural-to-urban transect. This study used the simple exponential decay function to calculate gradients,  $y = \text{EXP} [\alpha x + \beta]$ , where  $y$  is the variable of interest (i.e., job density, population density, or poverty rate) at distance  $x$  from center,  $\alpha$  is gradient, and the exponential of  $\beta$  is the estimated density of center. The gradient  $\alpha$  can be calculated by a linear regression with the transformation of the original function,  $\text{LN}(y) = \alpha x + \beta$ , where the gradient  $\alpha$  is the slope of regression analysis. Urbanized areas with statistically insignificant  $\alpha$  at the 0.05 level were excluded from this study.

A major limitation of traditional gradient measurement is that it can only be used in monocentric urban models, while many contemporary metropolitan areas have polycentric urban structure. There are several gradient estimation methods for polycentric models (e.g., Craig and Ng 2001 and McMillen 2001; 2004), but these methods are technically complicated to calculate and translate into policy guidelines. To take advantage of the simplicity of the monocentric model, this study translated a polycentric structure into a monocentric structure by assuming that all employment centers (including CBD) are placed at the same location (Figure 5.3). A gradient value measured in this assumption is identical to the average of gradients from each center.



**Figure 5.3 A simple operation to translate polycentricity using monocentric model (DPMT study)**

Center identification is the first step to estimate gradients of polycentric regions. Among a number of center identification methods (e.g., McDonald 1989; Gordon et al. 1989; Giuliano and Small 1991; McMillen 2001; 2003; and Redfean 2007), this study employed the two-stage non-parametric method of McMillen (2001) because it can be applied to any area without local knowledge. Its first stage identified center candidates as tracts having significantly higher densities than a smoothed density surface. The second stage determined centers among the candidates based on the significance of influence on job densities of nearby locations. Specifically, the method includes three steps: (1) estimate a job density surface using locally weighted regression (LWR); (2) select candidate centers as significantly greater residuals from the LWR estimates; and (3) exclude insignificant candidates based on the influence of proximity to each candidate on density. In this procedure, center identification results depend on three modifiable parameters: (1) window size for the LWR, (2) p-value to determine significant residuals, and (3) way to control CBD influence compared to other subcenters. Section 2 compared eight combinations of different parameter values for their performance in identifying employment centers, based on the model fit and the gradient of density model estimated from each parameter combination and the application to a VMT estimation model. This study used the best performing parameter combination out

of the eight combinations—the window size of 25%, the p-value of 0.1, and the same control between CBD and subcenters.

#### 5.4.4. Analytical methods

Directed acyclic graph (DAG) was used to identify the structure of relationships among study variables. Directed graph stems from the field of artificial intelligence and computer science to present causal relationships among a set of variables using an arrow graph (Spirtes et al. 2001). DAG is a directed graph with no cyclic path, where there is at least one ultimate outcome variables, such as VMT used in this study. Several search algorithms are available to calculate paths of the structural relationships. These algorithms statistically determine the path between any two variables in a model among four possible paths: one-direction, correlation, no relation, or inability to determine. The algorithms can be broadly classified into PC algorithm family (e.g., conservative PC, PCD, and FCI) and greedy equivalent search (GES) algorithm.

The PC algorithm starts the search with a complete, undirected graph which has a line (called “edge”) with no arrowhead connecting each variable (called “node”) with every other variable (Spirtes et al. 2000). Edges between nodes are removed sequentially based on conditional independence decisions. Edges that survive this removal process are then directed using simple rules. For example, imagine a triangle of variables X, Y, and Z, and it is known that Z is correlated with both X and Y. If the edge between X and Y are conditionally independent given Z, then we can direct  $X—Z—Y$  as  $X→Z←Y$ . Any full path structure can be read as a set of multiple triangles. Since this algorithm determines path by individual edge, each variable should have a normal distribution in order to determine the independence between any two variables in the triangular relationship.

The GES algorithm is a stepwise search over alternative DAGs using Bayesian posterior scores (Chickering 2002). The algorithm consists of two stages. It begins with a DAG in the condition that all variables are independent with each other (i.e., no edges between nodes). Edges are added and/or edge directions reversed in a systematic search

across classes of equivalent DAGs if the Bayesian posterior score is improved. The first stage ends when a local maximum of the Bayesian score is found such that no further edge additions or reversals improve the score. After this first stage, the second stage commences to delete edges and reverse directions, if such actions result in improvement of the Bayesian posterior score. The algorithm terminates if no further deletions or reversals improve the score.

These algorithms determine relationships based purely on statistical procedures and may produce unreasonable results. To avoid this problem, users can set up specific conditions based on rational hypothesis or knowledge. As described in the conceptual framework (Figure 5.1), this study classified variables into three hierarchical tiers (i.e., control, policy, and outcome variables), and established a rule that variables in the lower tier do not influence variables in the higher tier (the lowest tier is the outcome variable, DVMT per capita).

Using the Tetrad IV software, this study attempted multiple algorithms, searching for models with better model fit. Models with logarithm-transformed data were also tested, particularly because PC-based algorithms may require stricter normality for each individual variable. To determine reasonable DAGs, every tested model was evaluated with the relative chi-square, for which value closer to 1 indicates better-fitted model (Carmines and McIver 1981). Then final DAG models were selected after excluding worse-fitted models and models showing a duplicative relationship structure.

Direct and indirect influences of explanatory variables on regional VMT were estimated by structural equation modeling (SEM) based on the DAGs selected. SEM is useful for tracking relative effects of variables on each other through hypothesized relationship paths. SEM estimates parameters by solving simultaneous equations of a series of hypothesized relationships in a model. The technique uses iterative methods (e.g., maximum likelihood method, generalized least square, etc.) that involve a series of attempts to obtain estimates of unknown parameters until it finds the model covariance matrix best fitted with the actual covariance matrix representing the hypothesized

relationships (Hoyle 1995). Therefore, SEM can be regarded as a general method for constructing a predictive model with relatively few restrictions (e.g., regression analysis, factor analysis, and ANOVA are special cases of SEM). The AMOS 20 software package was used to conduct SEM analyses in this study.

#### 5.4.5. Descriptive statistics

Table 5.2 shows descriptive statistics of variables used. The non-normality of several variables would be a potential problem because parameter estimations for DAG and SEM are mostly based on the assumption of normality. A major effect of normality violation is an overestimation of chi-square value. The higher the chi-square means the worse the model fit. The overestimated chi-square could lead researchers to think that their models were worse fitted than the actual fit. The lack of multivariate normality also tends to yield the underestimation of standard errors and can result in the misreading of regression paths to be statistically significant more than the actual significance (Kline 1998).

Meanwhile, several simulation studies found that SEM parameter estimates are still fairly accurate under conditions of severe non-normality of data (Gao et al. 2008; Kline 1998; Voortman and Druzdzel 2008). Klein (1998) suggested that a distribution with skewness less than 3 and kurtosis less than 10 is allowable. Most variables in this study meet this normality criteria except daily passenger miles traveled per capita (DPMT/c), population (Pop) and population-weighted average of tract densities (Local.den). Among these three variables, population does not have to be normally distributed since it is an exogenous variable by definition in this study (i.e., population is not influenced by any other variable). Transit use variable (DPMT/c) and local built environment variable (Local.den) still have the non-normality problem though. Common suggestions to solve this problem include data transformation (Andreassen et al. 2006; Bollen 1989; Yuan et al. 2000) and outlier removal (Bagley and Mokhtarian 2002). This study avoided data removal approach, and instead transformed the data.

**Table 5.2**  
**Descriptive statistics of variables**  
(values in parenthesis are calculated from the natural log-transformed data)

Construct	Variable	Unit	Min.	Max.	Mean	Std. Dev.	Skewness	Kurtosis
Outcome	DPMT/c	mi/person	0.012 (2.48)	2.884 (7.97)	0.201 (4.78)	0.305 (0.96)	5.2 (0.4)	3.0 (0.4)
Regional demographic characteristics	Pop	thousand	51 (3.93)	17485 (9.77)	782 (5.72)	1777 (1.22)	6.1 (0.8)	46.8 (0.2)
	Job.den	/sq. mi	634 (6.32)	3615 (8.06)	1293 (7.11)	441 (0.30)	2.0 (0.3)	6.7 (0.3)
	Pop.den	/sq. mi	1337 (7.14)	7514 (8.86)	2475 (7.75)	926 (0.30)	2.4 (0.8)	7.9 (0.9)
	Pov.rate	percent	4.4 (1.48)	34.5 (3.54)	13.4 (2.53)	4.8 (0.33)	1.5 (0.2)	3.5 (0.9)
Transportation infrastructure	Road/c	ft/person	11 (2.40)	68.8 (4.23)	26 (3.22)	7.7 (0.30)	1.0 (-0.4)	4.2 (0.6)
	Transit/c	ft/person	31.3 (3.44)	671.2 (6.51)	165.5 (4.95)	99.7 (0.56)	1.9 (0.0)	5.7 (0.0)
	Fare	US\$/trip	0.05 (1.53)	2.35 (5.46)	0.64 (4.06)	0.31 (0.489)	1.7 (-1.09)	6.2 (4.7)
Local built environment	Local.den	/sq mi	974 (6.88)	33540 (10.42)	3628 (8.06)	2799 (0.48)	6.7 (0.9)	65.8 (2.7)
Regional development patterns	JD.grd	ratio	0.062 (1.82)	1.205 (4.79)	0.373 (3.43)	0.229 (0.62)	1.0 (-0.1)	0.6 (-0.7)
	PD.grd	ratio	0.037 (1.31)	0.936 (4.54)	0.265 (3.06)	0.178 (0.67)	1.2 (0.0)	1.2 (-0.7)
	Pvt.grd	ratio	0.012 (0.18)	0.559 (4.02)	0.192 (2.75)	0.116 (0.70)	0.8 (-0.7)	0.0 (0.4)

Note 1: Refer to Table 5.1 for detailed descriptions of variables.

Note 2: The log-transformations of fare and gradient measures were done after multiplying the original values with 100 to avoid a negative value.

DPMT/c is correlated with all variables in the model. Many predictor variables also are significantly correlated with each other as shown in Table 5.3. It raises multicollinearity problem that can distort estimations of explanatory powers of factor variables in a multiple regression model. SEM is not free of multicollinearity because it is a regression-based method. However, SEM often can relieve multicollinearity problem by decomposing a single regression model with many independent variables into many simpler regression models.

**Table 5.3**  
**Standardized correlations (Pearson's R) among variables**  
(values in parenthesis are calculated from the natural log-transformed data)

Variables	DPMT/c	Pop	Job.den	Pop.den	Pvt.rate	Road/c	Transit/c	Fare	Local.den
DPMT/c	1								
Pop	.789** (.632**)	1							
Job.den	.508** (.519**)	.429** (.309**)	1						
Pop.den	.528** (.567**)	.525** (.431**)	.876** (.846**)	1					
Pvt.rate	-.027 (.027)	-.096 (-.242**)	.131 (.032)	.151* (.066)	1				
Road/c	-.403** (-.514**)	-.347** (-.450**)	-.489** (-.538**)	-.568** (-.671**)	.043 (.064)	1			
Transit/c	.791** (.835**)	.511** (.379**)	.490** (.486**)	.448** (.465**)	-.027 (-.005)	-.343** (-.358**)	1		
Fare	.159* (.035)	.167* (.239**)	-.028 (-.058)	-.004 (-.036)	-.254** (-.344**)	-.043 (-.063)	.205** (.089)	1	
Local.den	.853** (.680**)	.819** (.521**)	.618** (.735**)	.691** (.869**)	.048 (.044)	-.495** (-.696**)	.621** (.563**)	.130 (.073)	1
JD.grd	-.214** (-.308**)	-.341** (-.740**)	.085 (.092)	-.074 (-.059)	.387** (.302**)	.110 (.155*)	-.070 (-.103)	-.217** (-.292**)	-.098 (-.111)
PD.grd	-.205** (-.305**)	-.318** (-.750**)	.016 (.031)	-.113 (-.101)	.327** (.249**)	.085 (.171*)	-.056 (-.100)	-.160* (-.222**)	-.065 (-.070)
Pvt.grd	-.246** (-.284**)	-.333** (-.622**)	.055 (.103)	-.071 (-.024)	.333** (.243**)	.180 (.185**)	-.152* (-.116)	-.158* (-.199**)	-.096 (-.038)

\*\* Correlation is significant at the 0.01 level

\* Correlation is significant at the 0.05 level

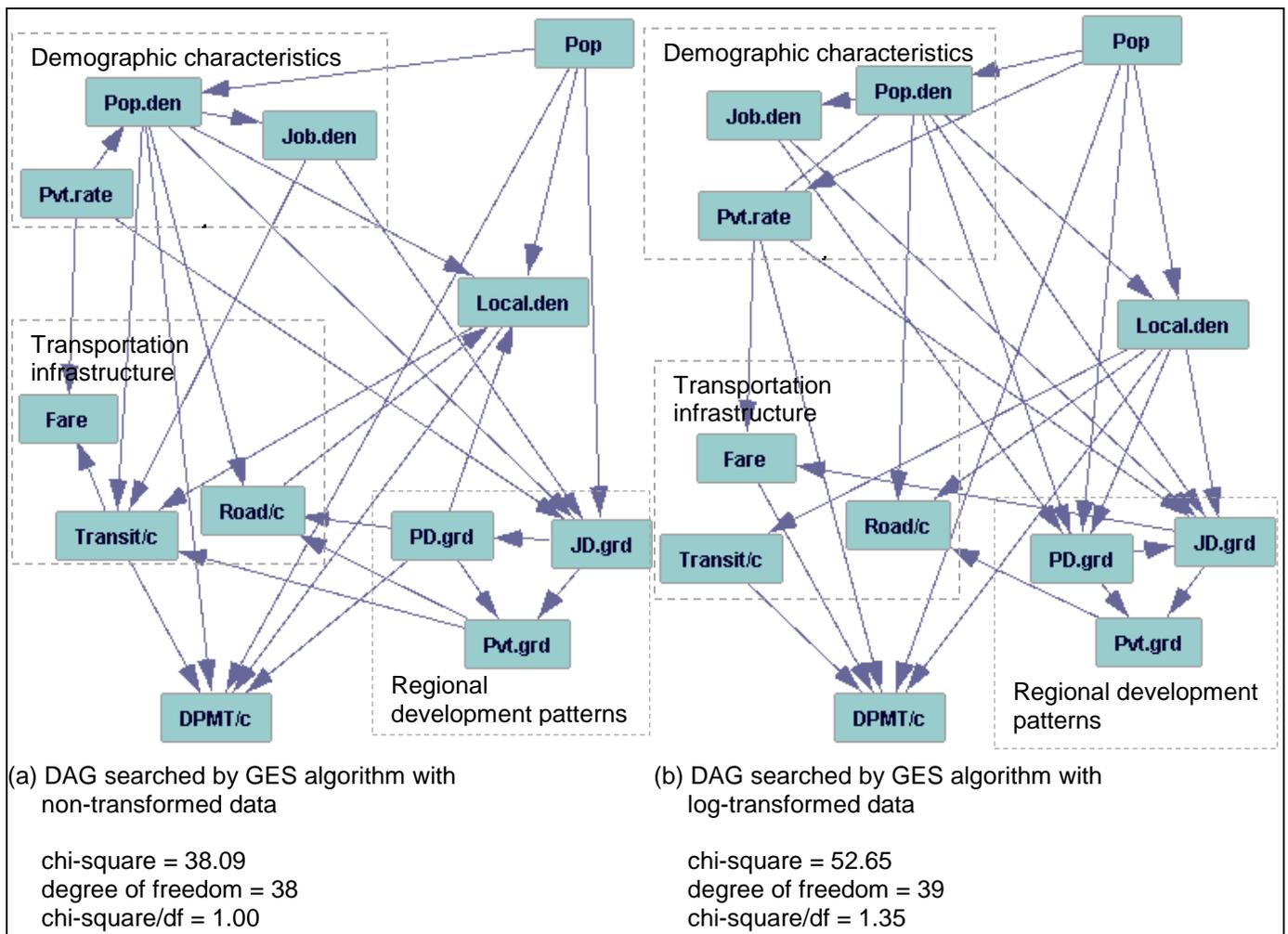
Note: Refer to Table 5.1 for detailed descriptions of variables

## 5.5. Results and Implications

### 5.5.1. Directed acyclic graph (DAG)

To find more reasonable structural relationships, multiple algorithms (i.e., GES and PC and its modifications) were tested to draw DAGs. Both non-transformed and transformed data were tested in the DAG procedure. Model fit was estimated using the relative chi-square (chi-square divided by the degree of freedom), in which a value less than 2 is regarded reasonably fitted (Carmines and McIver 1981).

Two different types of reasonably fitted DAGs were found: the GES with non-transformed data (GES-non), and the GES with log-transformed data (GES-ln). Figure 5.4 shows their path diagrams and model fit statistics. Both GES-non and GES-ln are overall similar to each other. Local density, transit supply and population have direct effects on PMT per capita. Fare shows a direct relationship in the GES-ln, but not in the GES-non.



Note: Refer to Table 5.1 for detailed descriptions of variables

**Figure 5.4 Directed acyclic graphs of factors influencing passenger miles traveled per capita among 203 U.S. urbanized areas, 2002**

A major distinction between the two models is that GES-non identified regional development patterns as direct and indirect factors, while GES-In negated these influences. There also exist some differences between the two models in the relationship between demographic characteristics and regional development patterns. These gaps might be caused by the presence of collinearity, particularly between direct and indirect factors. An indirect path can be interpreted in two different ways – an actual causality path or a spurious relationship. A relationship between X and Y can be regarded as spurious when their relationship becomes insignificant by adding a third intervening factor W. This can be described as “ $X \leftarrow W \rightarrow Y$ ” in a DAG. Diagnosing the type of relationship is a major function of the DAG searching process. Table 5.4 presents the diagnosis result for indirect paths to PMT per capita (DPMT/c) in each model.

**Table 5.4**  
**Diagnosis of spurious relationships using DAG search algorithms**

Model	Path	Relationship	Search algorithm
GES-non	Job.den → Transit/c → DPMT/c	Spurious	GES
	Pvt.rate → PD.grd → DPMT/c	Actual	PC, GES
	Road/c → Local.den → DPMT/c	Spurious	PC, GES
	Pop.den → Local.den → DPMT/c	Actual	PC, GES
GES-In	Pop.den → Local.den → DPMT/c	Spurious	PC
		Actual	GES

Note 1: GES-non is a DAG model searched by GES algorithm with non-transformed data, and GES-In is a model searched by GES algorithm with log-transformed data.

Note 2: Refer to Table 5.1 for detailed descriptions of variables

Removing the identified spurious variables made the two path diagrams more similar. In the GES-non model, job density (Job.den) and road supply (Road/c) had a spurious relationship with PMT per capita. These variables showed no relationship with PMT per capita in the GES-In model. Poverty rate (Pov.rate) was found to have an actual

indirect influence in the non-transformed model, and the log-transformed model identified the variable as a significant direct factor. The only disagreement in the decision on spurious factor between the non- and log-transformed models was about regional population density (Pop.den). The GES-non recognized that the regional density has an actual influence, while the variable was determined as being spuriously related through local density factor (Local.den) in the GES-ln.

The two models have minor differences in the relationship between demographic characteristics and regional development patterns. These differences do not seem to suggest meaningful policy implications but to simply indicate an association tendency between the two constructs. In other words, areas with a larger population tend to have a higher density and a steeper density gradient; however, regional population and density control policies do not aim to manage the spatial configuration of any specific population, job or poverty.

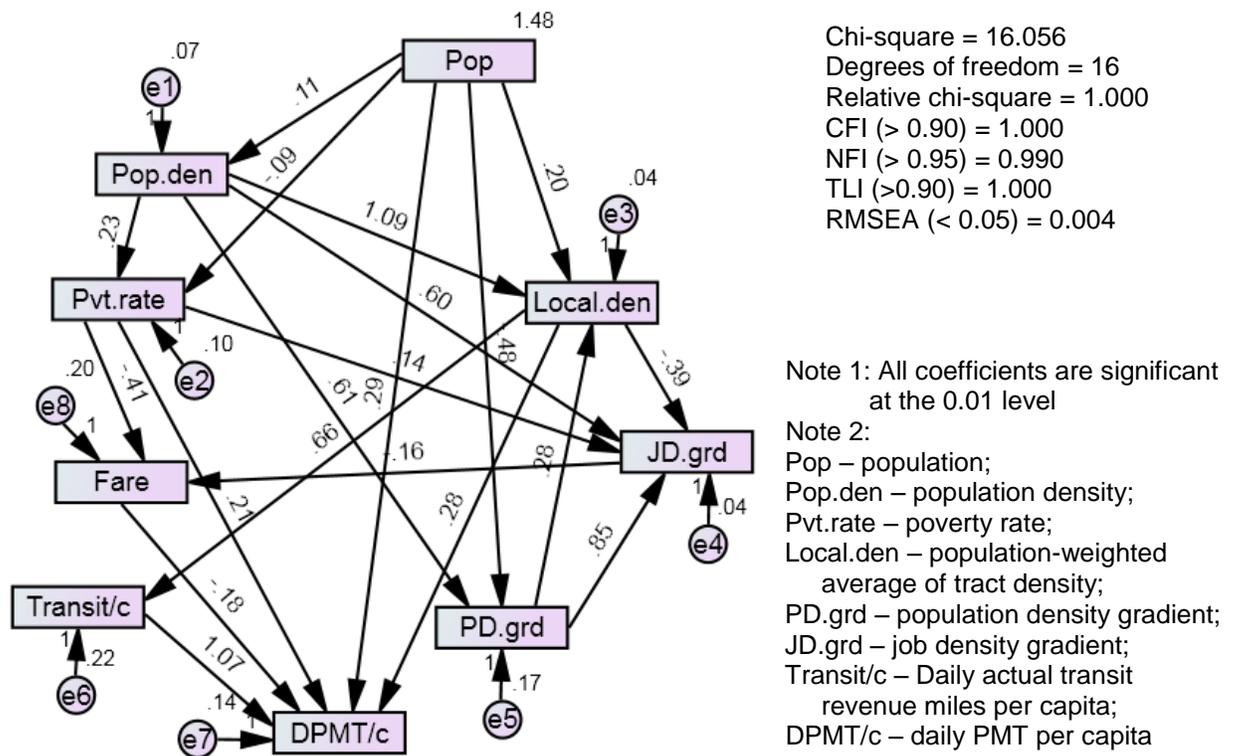
Based on the DAG searching and the diagnosis of spurious relationships, variables of job density, road supply, fare and poverty gradient were excluded in further structural equation modeling analyses. Considering a possible distortion by these variables onto the whole relationship structure, another set of DAG searching was conducted without these variables.

#### 5.5.2. Structural equation modeling (SEM)

The path models derived from the DAG searching process were estimated using structural equation modeling. Since the two models produced a similar structural relationship and model fit level, this study selected to present the log-transformed model (GES-ln) results for two advantages: 1) this log-log model satisfies the normality assumption, and 2) the model's parameter estimates represent elasticities, reflecting the relative sensitivity (percentage change) of DPMT/c to a 1% increase in each factor variable holding other factors constant.

Figure 5.5 shows the SEM results. Every path in the model showed a statistical significance at the 0.01 level. It also had an almost perfect model fit based on five

goodness-of-fit indices. The indices and their critical thresholds to determine a good fit are as follows (Hu and Bentler 1999): relative chi-square (< 2.0), comparative fit index: CFI (> 0.90), normed fit index: NFI (> 0.95), Tucker-Lewis index: TLI (or non-normed fit index: NNFI) (> 0.90), and root mean square error of approximation: RMSEA ( $\approx$  0.05).



**Figure 5.5 Structural equation model of DPMT per capita (log transformed data)**

Table 5.5 summarizes direct and indirect effects of explanatory variables on PMT per capita. In terms of total effect, regional population density (Pop.den) showed the highest elasticity. However, it should be noted that all of its effects are indirect. Referring to Figure 5.5, population density (Pop.den) influences PMT per capita (DPMT/c) through local density (Local.den). In the diagnosis of indirect paths, PC-based algorithms agreed with the pattern of “Pop.den ← Local.den → DPMT/c” in this log-

transformed model, indicating the spurious relationship between regional population density and PMT per capita. This reveals that the areas with a greater amount of dense neighborhoods tend to have not only a higher level of transit use but also a higher regional density. This also implies that regional density could serve as a good proxy measure of local density level in transit use models in studies on U.S. urbanized areas. The relationship between regional population density and local density was almost unit-elastic (i.e., the elasticity between them is 1.26). Results also indicate that a higher transit fare moderately reduce PMT per capita (-0.183 of elasticity) under the control for other variables (particularly, transit supply and population).

**Table 5.5**  
**Effect size (elasticity) on DPMT/c, estimated by SEM**

Variable	Description	Total effect	Direct effect	Indirect effect
Pop	Population	.455	.285	.170
Pop.den	Population density	1.317	.000	1.317
Pvt.rate	Poverty rate	.285	.206	.079
Transit/c	Daily actual transit revenue miles per capita	1.066	1.066	.000
Fare	Passenger revenue divided by unlinked passenger trip	-.183	-.183	.000
Local.den	Population-weighted average of tract density	.966	.276	.690
PD.grd	Population density gradient	.293	.000	.293
JD.grd	Job density gradient	.029	.000	.029

Note: Refer to Figure 5.5 for detailed relationship paths in the model. All paths in the model are statistically significant at the 0.01 level.

Excluding population density, the strongest factor is transit supply, indicating that a 1% larger actual transit revenue miles per capita is contributed to an 1.066% more transit passenger miles traveled per capita. The second strongest factor is local density which has 0.966 of elasticity. The strong effect of these two factors reflects the success of continuous efforts for transit oriented development (TOD) in the U.S. While our results indicate that increasing local density alone has a moderate direct effect on

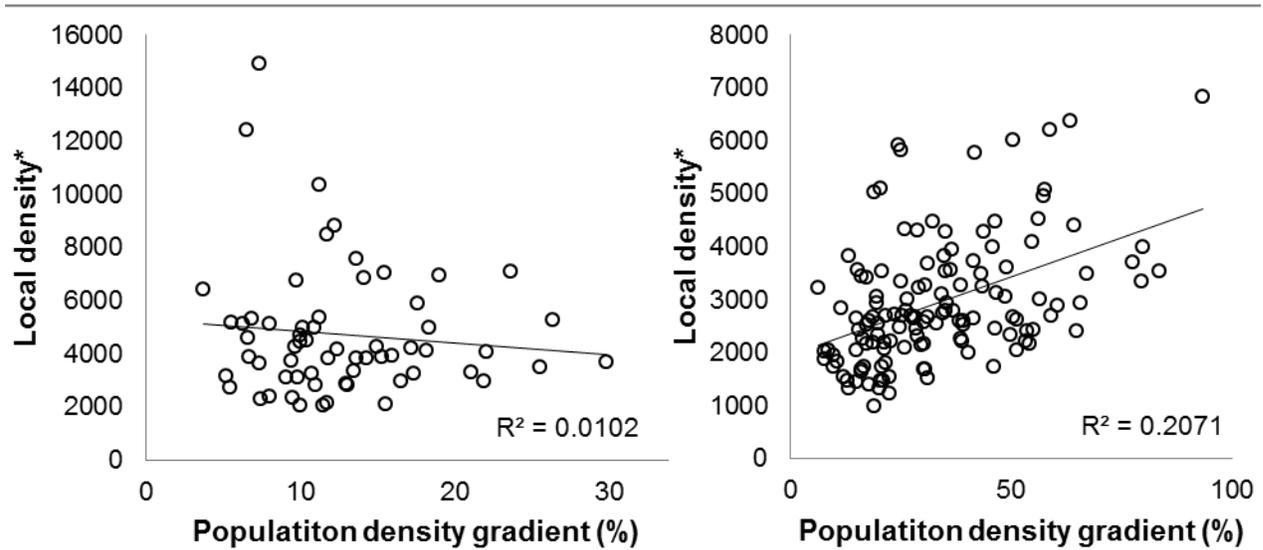
increasing PMT per capita (i.e., 0.276 of elasticity), greater effect was achieved when more transit service was provided. The elasticity of this indirect effect through the transit supply is 0.690.

Population and poverty rate showed a moderate effect (i.e., 0.455 and 0.285 of elasticity, respectively), reflecting that transit use is greater in urbanized areas with greater population and more low income people. Their effects are fairly independent of each other showing near zero elasticity between them.

Under the control of population and poverty rate, regional development patterns (particularly the population density gradient in this model) also showed a moderate effect, 0.293 of elasticity. This influence was not direct but was achieved indirectly through the local density factor. To check the spuriousness of this path, a DAG searching was conducted among population density gradient, local density and passenger miles traveled. All search algorithms yielded the same result of “PD.grd → DPMT/c ← Local.den” which indicates an actual influence. This inference appears to make sense because the increased population density gradient means more compact local development near regional centers.

However, it should be noted that the bivariate correlation between regional development patterns and PMT per capita were negative (Table 5.3). It indicates that the higher concentration of population, job and poverty have a PMT reduction effect too. This effect was found significant in the non-transformed model (not presented in this study), but became insignificant after the log-transformation. The inconsistency of the relationship between regional development patterns and transit use can be observed in previous individual-level studies. Cervero and Duncan (2006) found that the distance to CBD has a significant positive influence on weekday boardings per station, while others (e.g., Bento et al. 2005, Baker 1994, and Pushkar et al. 2000) failed to find this influence. This inconsistency might be due to positive and negative relationships of regional development patterns with local urban form depending on population size. The gentler population gradient may imply either fewer compact neighborhoods near centers

in most smaller urbanized areas or more compact developments at the outskirts in other larger areas (Figure 5.6).



(a) UAs with population over 500,000  
(except New York-Newark urbanized area, NY)

(b) UAs with population less than 400,000

\* Local density: population-weighted average of tract densities

**Figure 5.6 Correlation between local density and population density gradient by population size**

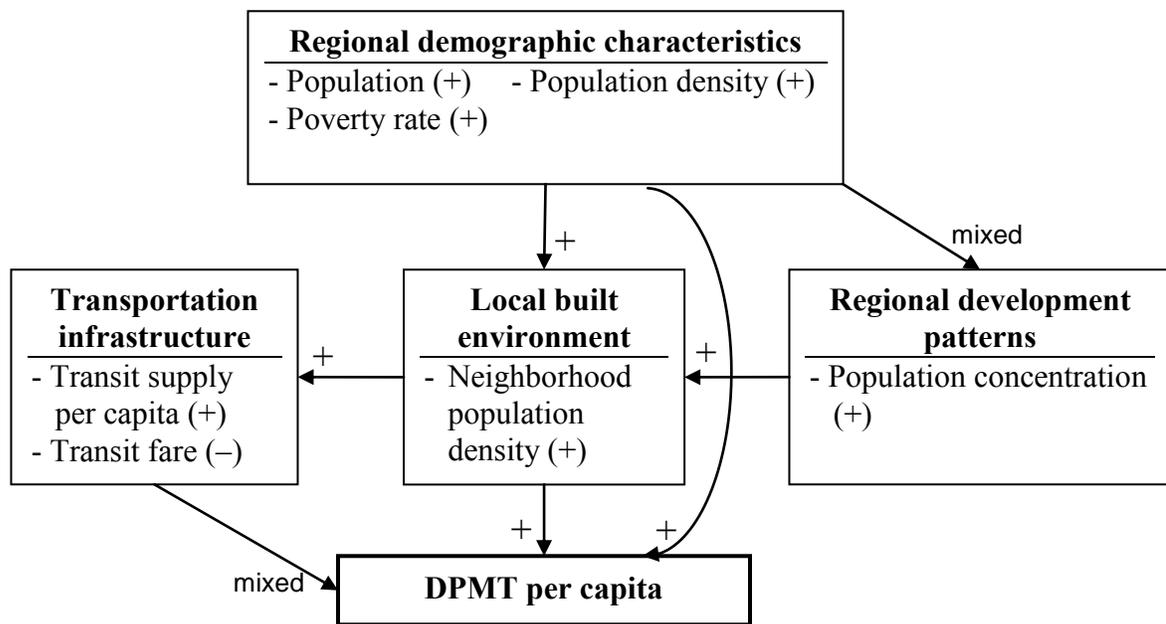
## 5.6. Summary and Conclusion

This paper examined the influence of urban form on transit passenger miles traveled (PMT) at the urbanized area scale, emphasizing policy-interpretable urban form measures and structural relationships among predictor variables. A series of cross-sectional analyses on 203 U.S. urbanized areas was conducted, using directed acyclic graph (DAG) and structural equation modeling (SEM).

As illustrated in Figure 5.7, this study found that PMT per capita is greater in urbanized areas that supply more transit service hour and include a greater amount of dense neighborhoods. These two variables are closely associated with key characteristics

of transit oriented development (TOD), and their strong effects reflect the success of TOD in the U.S.

The results for this study present that a more concentrated regional development pattern (or steeper density gradient) led to a moderate increase of PMT per capita. However, this study also found that the regional development pattern effect may not be generalized throughout U.S. urbanized areas because the lower level of regional concentration implies two opposite effects on transit use by population size – 1) the fewer compact neighborhoods near centers and thus the fewer transit use in most smaller urbanized areas, or 2) the more compact neighborhoods at the outskirts and thus the greater transit use in some larger urbanized areas.



**Figure 5.7 Summary of findings from the DPMT study**

Not surprisingly, population and poverty rate were identified as positive factors of PMT per capita. Population density showed the strongest positive effect as found from many previous studies, while this study found that this effect is spurious or only

indirect. Hence, transit promoting efforts, including urban form restructuring and transit supply expansion, can be justified for most urbanized areas regardless of the gross density level.

Road supply was found to have no effect on PMT per capita. Considering that road supply is a strongest predictor of vehicle miles traveled (Section 4), the lack of road supply effect on PMT implies that transit is not a substitute travel option to private vehicle. In other words, an increase of transit use does not necessarily mean a decrease of driving. Transit fare showed a significant relationship with PMT per capita as found in many previous studies.

During the past several years, we have observed several signs of increasing demand for transit, including a rise in fuel price, growing transit ridership, growing immigrant population, increasing traffic congestion and limited parking supply (APTA 2011; Blumenberg and Evans 2007 therein Chen et al. 2011). If these transit demand forces are expected to continuously grow in the future, the remaining challenge will be to provide transit service sufficiently and appropriately to people in need of affordable travel mean. This study reconfirmed that ongoing efforts of urban form control and transit service expansion are effective measures to promote regional aggregate transit use regardless of given population size and socioeconomic condition. The use of aggregate data limits detailed controls, but this study helps fill the gap in the disaggregate travel behavior studies of only one or several metropolitan areas.

## 6. SUMMARY AND CONCLUSIONS

### 6.1. Summary

Section 2 reviewed urban form-travel studies, focusing on the distinction between micro and macro approaches in terms of the unit of analysis, variables used, data sources, major findings and policy implications. This review showed a gap in the previous research that micro studies have identified distance from the center as the most influential factor on VMT, while macro studies have not reflected this relationship.

Section 3 evaluated parameters influencing the performance of regional center identification methods. This study found that lower density cutoff, and equal treatment between CBD and subcenters yielded better results. Results also showed that for polycentric areas, the use of a polycentric model produced a better model fit than the monocentric model. This section provided a methodological support for the measure of polycentric regional urban form.

Section 4 examined the influence of local and regional development patterns on collective VMT, considering the polycentric urban structure and multiple intervening factors (e.g., transportation infrastructure, and income segregation). Major findings of this section is as follows: 1) VMT is lower as more people and jobs are located more closely to centers; 2) the higher concentration toward centers also increases VMT by requiring more road supply; 3) VMT is lower as more people live in denser neighborhoods; 4) road supply has a strong positive impact on VMT; 4) transit supply is not related to VMT; 5) poverty gradient has a minimal effect on VMT; and 6) the effects of local versus regional urban form on VMT are independent of each other.

Section 5 examined the influence of urban form on transit passenger miles traveled (PMT) at the urbanized area scale, particularly focusing on applying policy-relevant urban form measures and searching for a structural relationship among predictor variables. This section found that: 1) PMT per capita is greater in urbanized areas that supply more transit service hours and have denser neighborhoods; 2) more concentrated patterns of regional development result in a moderate increase of PMT per capita; 3)

regional population density only has an indirect effect on transit use ; and 4) road supply do not cause differences in PMT per capita among U.S. urbanized areas because transit demand and supply are too low to be an alternative to driving.

## **6.2. Limitations**

This dissertation has several limitations. First, this study was based on only 203 out of 347 urbanized areas that show statistically significant gradients in job density, population density and poverty rate. Further studies are needed for remaining 144 urbanized areas with a different methodological framework.

The second limitation lies in urban form measures. Gradient measures used in this study aggregated multiple centers into one imaginary center, and thus cannot distinguish areas with a dominant CBD and smaller subcenters from areas with multiple strong centers. Although the latter would be a more desirable form in terms of polycentricism concept, this study could not evaluate the difference. Local built environment measures were captured by census tract-level density only. Inclusion of other sophisticated measures that can represent land use diversity and design at the local level may provide more findings with additional policy implications.

Third, although the 203 samples are not small, urbanized areas in the U.S. tend to have insufficient variability in some potentially significant characteristics affecting VMT, such as transit infrastructure, fuel price and polycentricity, which limits the statistical power to detect their potential association with VMT. Transit is underdeveloped in most U.S. urbanized areas. Fuel price in the U.S. is fairly constant across the states and lower than many European and Asian countries. Only 81 among 203 areas have a polycentric urban structure. An addition of more study areas with diverse regional settings will improve generalizability of this study and offer more solid policy implications.

Fourth, data limitations need be noted. Since a DVMT is estimated by summing up traffic counts multiplied by lane lengths of road sections, the data has three limitations inherited from traffic survey procedure. Traffic surveys usually do not cover

local and minor roads, and often estimated based on a statistical model. Therefore, minor errors are inevitable for those model-estimated counts. Also, traffic surveys cannot differentiate visitors' travels. Urbanized areas that attract many outside visitors would have overestimated DVMT estimations. Finally, traffic survey methods are inconsistent among urbanized areas. While every state follows a general federal-level standard for the traffic survey (FHWA 2010), data quality might vary by urbanized area.

### **6.3. Conclusions**

Major findings of this dissertation research can be summarized as 1) higher regional concentration, greater local density and less road supply lowered VMT; and 2) higher local density and more transit increased PMT. These results imply that different approaches to development control are needed for different sustainable transportation goals—1) intensifying regional centers such as infill developments for VMT reduction, and 2) compact neighborhood development approaches such as transit oriented development for transit promotion.

This study could not clearly present which development control approach is most effective in achieving the sustainable transportation goal because compact neighborhood approach has both positive and negative potentials for VMT reduction. Transit is underutilized in most U.S. metropolitan areas and shows no significant influence on VMT per capita in this study (Section 4). Considering the strong impact of local density on PMT per capita, continuing compact neighborhood development efforts may promote transit use and consequently reduce VMT in the future. On the other hand, indiscreet compact developments at the urban fringe may lead to decentralization (or dispersed development) from the regional perspective and consequently result in increased VMT per capita. Infill development approach satisfies both the compact local urban form and the centralized regional development; however, containing growth in a very restricted boundary will encounter a physical limitation in capacity.

Polycentricism can explain the seemingly contradictory development principle between compact neighborhood and regional centralization by allowing dispersion and

concentration at the same time (Jenks et al. 1996; Ewing 1997). This study suggests a potentially positive impact of polycentric development on sustainable travel outcomes. A regional center can emerge naturally through any historic or economic processes, such as edge cities. Meanwhile, its development can be controlled by appropriate policies and planning methods such as the form-based code and the urban-rural transect approach. This implies that land use control policies can make a significant contribution to VMT reduction.

These general implications are not new. Over 40 years of smart growth efforts have promoted sustainable communities and regions, fighting against auto-oriented, dispersed development patterns. The Smart Growth Network (2002a; 2002b) announced 10 principles of smart growth and 200 policies for implementation. This dissertation provided additional empirical evidence that confirmed and expanded the potential for these smart growth solutions to help reduce automobile dependency in many urbanized regions in the U.S. Specific implications from this study include the following.

First, findings suggest that more effective infill development strategies should focus on areas near existing employment centers, and consider the overall regional development configurations. Some of current infill development strategies, such as growth boundary or green belt, tend to overlook the development patterns within the boundary. This type of conventional approach is effective in containing urban expansion but does not include any mechanism to regulate dense and high-impact developments along the edge of the boundary. Further, the green belt approach has shown to bring leapfrog developments outside of the outer edge of the belt (Amati and Taylor, 2010).

Second, the jobs-housing balance needs to be stressed in the regional context. Many sustainable development principles emphasize mixed land uses, but most of them are discussed and applied to the neighborhood-level planning and site-level projects. The jobs included in such small scale projects may be limited to service and commercial jobs to support the immediately surrounding neighborhoods. The regional distribution of quality jobs and housing developments close to the jobs is also very important. In the regional jobs-housing balance perspective, infill developments in CBDs might need to

focus more on residential uses, while the developments in subcenters need to focus on achieving the balance between jobs and housing. Proper strategies to locate housing developments to improve the jobs-housing balance are recommended.

Third, plans to manage and regulate regional urban form need to precede plans for local urban form. Growing number of governments have adopted form-based codes based on the rural-to-urban transect (Parolek et al. 2008). The New Urbanism offered the transect concept suggesting the need for different urban design guidelines for diverse local environmental settings within the regional context. However, it does not provide specific guidelines for transect planning. This dissertation study suggests a simple guideline for the transect (i.e., decreasing development intensity as the distance from employment centers is increased); however, more research (e.g., how to determine actual development intensity thresholds based on regional settings, how to consider different local settings in the same distance from centers, etc.) is required to develop a complete set of guidelines of the transect planning.

Finally, TOD has shown to be a useful strategy to promote transit uses, but its plans still need to consider regional urban form for the VMT reduction purpose simultaneously. This study indicates that compact neighborhood developments accompanied by transit developments increase the PMT per capita. Meanwhile, this study also implies that the increased transit developments do not necessarily have a VMT reduction impact in many small and medium size urbanized areas. The TOD emphasizes compact developments near transit stops and may lead to a dense, monotonic urban form that is isolated and does not respond to the regional configuration. This pattern of development does not help reduce the regional average VMT. Therefore, establishing a hierarchy of TOD nodes within the context of regional development patterns will be an efficient transitional step until cities grow enough that the impact of transit developments on VMT reduction will become more significant in the future.

Despite the ongoing efforts of smart growth, barriers to attracting developments to inner cities and subcenters are persistent in four dimensions: marketability, land use policies, residents, and financing. The first two dimensions are planning issues; therefore

the solutions for them depend on the will and capacity of governments and developers. Governments can improve the marketability of urban centers by establishing infill developments as goals, conducting associated plans and actions, and promoting the plans. Governments favorable to redevelopments of urban centers would be more willing to accept creative ideas to overcome existing barriers and introduce policy tools to facilitate development process, such as relieving restrictive regulations (e.g., zoning, parcel size, parking, road width, etc.), streamlining permit process (e.g., limiting review period and public hearings), tax incentives for developers and future residents, facilitating land assembly processes, establishing focused public investment area, and lessening development fees (MRSC 1997). Developers are market-oriented thinkers and strive to improve the marketability through more attractive design. They negotiate with governments for more favorable policies and government incentives. These planning issues would be relatively easier to resolve than the tangible problems of financing and resistance from residents and stakeholders. Redevelopment plans typically require land assembly, gentrification, infrastructure improvement, high-density development, or longer construction periods. Solutions for the stakeholder opposition are agreed among studies – educating them and involving them in the plan (TMRPA 2005; MRSC 1997). Various financing solutions have been offered, such as supporting with governmental grants (e.g., urban development action grants, community development block grants, tax exempt industrial development bonds, city infrastructure grants, etc.), tax-increment financing, bank land, issuing project bonds, revolving loan funds, and providing loan guarantees for private developers (MRSC 1997). However, the ultimate success of every financing effort would depend on the marketability and feasibility of plans. The market potential of urban centers has been stated (Levine 2006; Porter 1997) and this is partly supported by the fact that the majority of urbanized areas show a higher density near centers. A successful urban center development depends on the collaboration among governments, stakeholders, developers and citizens.

## REFERENCES

- Agthe, D.E., and R.B. Billings. 1978. The Impact of Gasoline Prices on Urban Bus Ridership. *The Annals of Regional Science* 12(1): 90–96.
- Alonso, W. 1964. *Location and Land Use: Toward a General Theory of Land Rent*. Cambridge, MA: Harvard University Press.
- Amati, M., and L. Taylor. 2010. From Green Belts to Green Infrastructure. *Planning Practice and Research*. 25(2): 143-155
- Ananat, E.O. 2007. The Wrong Side(s) of the Tracks Estimating the Causal Effects of Racial Segregation on City Outcomes. *National Bureau of Economic Research Working Paper Series* No. 13343. Cambridge, MA: National Bureau of Economic Research.
- Anderson, N.B., and William T.B. 2001. The Structure of Sprawl: Identifying and Characterizing Employment Centers in Polycentric Metropolitan Areas. *American Journal of Economics and Sociology* 60(1): 147–169.
- Andreassen, T., B. Lorentzen, and U. Olsson. 2006. The Impact of Non-Normality and Estimation Methods in SEM on Satisfaction Research in Marketing. *Quality & Quantity* 40: 39–58.
- APTA. 2011. *2011 Public Transportation Fact Book*. Washington, DC: American Public Transportation Association.
- Badoe, D.A., and E.J. Miller. 2000. Transportation-land-use Interaction: Empirical Findings in North America, and Their Implications for Modeling. *Transportation Research Part D: Transport and Environment* 5(4): 235–263.
- Bagley, M.N., and P.L. Mokhtarian. 2002. The Impact of Residential Neighborhood Type on Travel Behavior: A Structural Equations Modeling Approach. *The Annals of Regional Science* 36: 279–297.
- Baker, B. 1994. *Building Orientation - A Supplement to the Pedestrian Environment*. Vol. 4B. Portland, OR: 1000 Friends of Oregon.
- Bento, A.M., M.L. Cropper, A.M. Mobarak, and K. Vinha. 2005. The Effects of Urban Spatial Structure on Travel Demand in the United States. *Review of Economics and Statistics* 87(3): 466–478.

- Blumenberg, E., and A.E. Evans. 2007. *Growing the Immigrant Transit Market: Public Transit Use and California Immigrants*. Paper presented at the 86th Annual Meeting of Transportation Research Board, Washington, DC.
- Bogart, W.T., and W.C. Ferry. 1999. Employment Centres in Greater Cleveland: Evidence of Evolution in a Formerly Monocentric City. *Urban Studies* 36(12): 2099–2110.
- Bohl, C.C., and E. Plater-Zyberk. 2006. Building Community across the Rural-to-urban Transect. *Places* 18(1): 4–17.
- Bollen, K. 1989. *Structural Equations with Latent Variables*. Hoboken, NJ: Wiley.
- Burchfield, M., H.G. Overman, D. Puga, and M.A. Turner. 2006. Causes of Sprawl: A Portrait from Space. *Quarterly Journal of Economics* 121(2): 587–633.
- Cao, X., P.L. Mokhtarian, and S.L. Handy. 2009. Examining the Impacts of Residential Self-Selection on Travel Behaviour: A Focus on Empirical Findings. *Transport Reviews: A Transnational Transdisciplinary Journal* 29(3): 359–395.
- Carmines, E., and J. McIver. 1981. Analyzing Models with Unobserved Variables: Analysis of Covariance Structures. In *Social Measurement: Current Issues*, 65–115. Beverly Hills, CA: Sage Publication, Inc.
- Cervero, R. 1993. *Transit-supportive Development in the United States: Experiences and Prospects*. Washington, DC: U.S. Department of Transportation, Federal Transit Administration.
- Cervero, R. 2002. Built Environments and Mode Choice: Toward a Normative Framework. *Transportation Research Part D: Transport and Environment* 7(4): 265–284.
- Cervero, R. 2003. The Built Environment and Travel: Evidence from the United States. *European Journal of Transport and Infrastructure Research* 3(2): 119–137.
- Cervero, R., and M. Duncan. 2006. Which Reduces Vehicle Travel More: Jobs-Housing Balance or Retail-Housing Mixing? *Journal of the American Planning Association* 72(4): 475–490.
- Cervero, R., C. Ferrell, and S. Murphy. 2002. Transit-oriented Development and Joint Development in the United States: A Literature Review. *TCRP Research Results Digest* 52: 1–144

- Cervero, R., and K. Kockelman. 1997. Transport and Environment : Travel Demand and the 3Ds: Density, Diversity, and Design. *Transportation Research Part D* 2(3): 199–219.
- Cervero, R., and J. Murakami. 2010. Effects of Built Environments on Vehicle Miles Traveled: Evidence from 370 US Urbanized Areas. *Environment and Planning A* 42(2): 400–418.
- Cervero, R., and C. Radisch. 1996. Travel Choices in Pedestrian versus Automobile Oriented Neighborhoods. *Transport Policy* 3(3): 127–141.
- Cervero, R., and K. Wu. 1997. Polycentrism, Commuting, and Residential Location in the San Francisco Bay Area. *Environment and Planning A* 29(5): 865–886.
- Chen, C., D. Varley, and J. Chen. 2011. What Affects Transit Ridership? A Dynamic Analysis Involving Multiple Factors, Lags and Asymmetric Behaviour. *Urban Studies* 48(9): 1893–1908.
- Chickering, D.M. 2002. Optimal Structure Identification with Greedy Search. *Journal of Machine Learning Research* 3: 507–554.
- Clifton, K, R. Ewing, G-J. Knaap, and Y. Song. 2008. Quantitative Analysis of Urban Form: a Multidisciplinary Review. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability* 1(1): 17–45.
- CPDR. 2008. *CP&DR's SB 375 Resources Page*. Ventura, CA: California Planning & Development Report. <http://www.cp-dr.com/node/2185>.
- Craig, S.G., and P.T. Ng. 2001. Using Quantile Smoothing Splines to Identify Employment Subcenters in a Multicentric Urban Area. *Journal of Urban Economics* 49(1): 100–120.
- Crane, R. 2000. The Influence of Urban Form on Travel: An Interpretive Review. *Journal of Planning Literature* 15(1): 3–23.
- Cutler, D.M., and E.L. Glaeser. 1995. Are Ghettos Good or Bad? *National Bureau of Economic Research Working Paper Series* No. 5163. Cambridge, MA: National Bureau of Economic Research.
- Dunphy, R, and K. Fisher. 1996. Transportation, Congestion, and Density: New Insights. *Transportation Research Record* 1552: 89–96.
- Eidlin, E. 2005. The Worst of All World: Los Angeles, California, and the Emerging Reality of Dense Sprawl. *Transportation Research Record* 1902: 1–9.

- Ellen, I.G., D.M. Cutler, and W. Dickens. 2000. Is Segregation Bad for Your Health? The Case of Low Birth Weight [with Comments]. *Brookings-Wharton Papers on Urban Affairs*: 203–238.
- Ewing, R. 1994. Characteristics, Causes, and Effects of Sprawl: A Literature Review. *Environmental and Urban Issues* 11(2): 1–15.
- Ewing, R. 1997. Is Los Angeles-Style Sprawl Desirable? *Journal of the American Planning Association* 63(1): 107-126.
- Ewing, R., K. Bartholomew, S. Winkelman, J. Walters, and D. Chen. 2008. *Growing Cooler: The Evidence on Urban Development and Climate Change*. Washington, DC: Urban Land Institute.
- Ewing, R., and R. Cervero. 2001. Travel and the Built Environment: A Synthesis. *Transportation Research Record* 1780: 87–114.
- Ewing, R., and R. Cervero. 2010. Travel and the Built Environment -- A Meta-Analysis. *Journal of the American Planning Association* 76(3): 265–294.
- Ewing, R., A.C. Nelson, and K. Bartholomew. 2009. *Response to Special Report 298 Driving on the Built Environment: The Effects of Compact Development on Motorized Travel, Energy Use, and CO2 Emissions*. Salt Lake City, UT: Metropolitan Research Center.
- Ewing, R., R. Pendall, and D. Chen. 2002. *Measuring Sprawl and Its Impact*. Technical report. Washington, DC: Smart Growth America.
- Ewing, R., R. Pendall, and D. Chen. 2003. Measuring Sprawl and Its Transportation Impacts. *Transportation Research Record* 1831: 175–183.
- FHWA. 2010. *Highway Performance Monitoring System: Field Manual*. Washington, DC: Federal Highway Administration.  
<http://www.fhwa.dot.gov/policyinformation/hpms/fieldmanual/>
- Frank, L.D., and P. Engelke. 2005. Multiple Impacts of the Built Environment on Public Health: Walkable Places and the Exposure to Air Pollution. *International Regional Science Review* 28(2): 193–216.
- Frank, L.D, and G. Pivo. 1994. Impacts of Mixed Use and Density on Utilization of Three Modes of Travel: Single-Occupant Vehicle, Transit, and Walking. *Transportation Research Record* 1466: 44–52.

- Galster, G., R. Hanson, M.R. Ratcliffe, and H. Wolman. 2001. Wrestling Sprawl to the Ground: Defining and Measuring an Elusive Concept. *Housing Policy Debate* 12(4): 681–717.
- Gao, S., P.L. Mokhtarian, and R.A. Johnston. 2008. Nonnormality of Data in Structural Equation Models. *Transportation Research Record* 2082: 116–124.
- Giuliano, G., A. Agarwal, and C. Redfearn. 2008. Metropolitan Spatial Trends in Employment and Housing. *Special Report 298: Driving and the Built Environment: The effects of Compact Development on Motorized Travel, Energy Use, and CO<sup>2</sup> Emissions*. Washington, DC: Transportation Research Board.
- Giuliano, G., and K.A. Small. 1991. Subcenters in the Los Angeles Region. *Regional Science and Urban Economics* 21(2): 163–182.
- Gomez-Ibanez, J.A. 1996. Big-city Transit Ridership, Deficits, and Politics: Avoiding Reality in Boston. *Journal of the American Planning Association* 62(1): 30-50.
- Gordon, P., A. Kumar, and H.W. Richardson. 1989. Congestion, Changing Metropolitan Structure, and City Size in the United States. *International Regional Science Review* 12(1): 45 –56.
- Gordon, P., and H.W. Richardson. 1996. Beyond Polycentricity: The Dispersed Metropolis, Los Angeles, 1970-1990. *Journal of the American Planning Association* 62(3): 289–295.
- Gordon, P., H.W. Richardson, and H.L. Wong. 1986. The Distribution of Population and Employment in a Polycentric City: The Case of Los Angeles. *Environment and Planning A* 18(2): 161 – 173.
- Handy, S. 1993. Regional Versus Local Accessibility: Implications for Nonwork Travel. *Transportation Research Record* 1400: 58–66.
- Handy, S. 1996. Methodologies for Exploring the Link Between Urban Form and Travel Behavior. *Transportation Research Part D: Transport and Environment* 1(2): 151–165.
- Heikkila, E., P. Gordon, J.I. Kim, R.B. Peiser, H.W. Richardson, and D. Dale-Johnson. 1989. What Happened to the CBD-distance Gradient?: Land Values in a Policentric City. *Environment and Planning A* 21(2): 221–232.
- Hendrickson, C. 1986. A Note on Trends in Transit Commuting in the United States Relating to Employment in the Central Business District. *Transportation Research Part A* 20(1): 33–37.

- Holtzclaw, J. 1990. *Explaining Urban Density and Transit Impacts on Auto Use*. San Francisco, CA: Sierra Club.
- Holtzclaw, J. 1994. *Using Residential Patterns and Transit to Decrease Auto Dependence and Costs*. San Francisco, CA: Natural Resources Defense Council.
- Holtzclaw, J., R. Clear, H. Dittmar, D. Goldstein, and P. Haas. 2002. Location Efficiency: Neighborhood and Socio-Economic Characteristics Determine Auto Ownership and Use - Studies in Chicago, Los Angeles and San Francisco. *Transportation Planning & Technology* 25(1): 1-27.
- Hoyle, R.H. 1995. *Structural Equation Modeling: Concepts, Issues, and Applications*. Beverly Hills, CA: Sage Publication, Inc.
- Hu, L., and P.M. Bentler. 1999. Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* 6: 1–55.
- Jenks, M., E. Burton, and K. Williams. 1996. *Compact City: A Sustainable Urban Form?* 1st ed. London, GB: E & FN Spon.
- Kain, J.F. 1996. *Cost-Effective Alternatives to Atlanta's Costly Rail Rapid Transit System*. Harvard Institute of Economic Research Working Paper. Cambridge, MA: Harvard Institute of Economic Research.  
<http://econpapers.repec.org/paper/fthharver/1762.htm>.
- Kim, S.. 2007. Changes in the Nature of Urban Spatial Structure in the United States, 1890-2000. *Journal of Regional Science* 47(2): 273–287.
- Kitamura, R., P.L. Mokhtarian, and L. Daidet. 1997. A Micro-analysis of Land Use and Travel in Five Neighborhoods in the San Francisco Bay Area. *Transportation* 24(2): 125–158.
- Kline, R.B. 1998. *Principles and Practice of Structural Equation Modeling*. New York, NY: The Guilford Press.
- Knaap, G-J., Y. Song, and Z. Nedovic-Budic. 2007. Measuring Patterns of Urban Development: New Intelligence for the War on Sprawl. *Local Environment* 12: 239–257.
- Kockelman, K. 1997. Travel Behavior as Function of Accessibility, Land Use Mixing, and Land Use Balance: Evidence from San Francisco Bay Area. *Transportation Research Record* 1607: 116–125.

- MRSC. 1997. *Infill Development: Strategies for Shaping Livable Neighborhoods*. Washington, DC: Municipal Research & Services Center of Washington. <http://www.mrsc.org/publications/textfill.aspx>
- Lane, B.W. 2010. The Relationship Between Recent Gasoline Price Fluctuations and Transit Ridership in Major US Cities. *Journal of Transport Geography* 18(2): 214-225.
- Lee, B. 2007. 'Edge' Or 'Edgeless' Cities? Urban Spatial Structure In U.S. Metropolitan Areas, 1980 To 2000. *Journal of Regional Science* 47(3): 479–515.
- Levine, J. 2006. *Zoned Out: Regulation, Markets, and Choices in Transportation and Metropolitan Land-Use*. Washington, DC: Resources for the Future.
- Lilliefors, H.W. 1967. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. *Journal of the American Statistical Association* 62(318): 399-402.
- Litman, T. 2004. Transit Price Elasticities and Cross-elasticities. *Journal of Public Transportation* 7(2): 37–58.
- Longley, P.A, and V. Mesev. 2000. On the Measurement and Generalisation of Urban Form. *Environment and Planning A* 32(3): 473 – 488.
- Mardia, K.V. 1974. Applications of Some Measures of Multivariate Skewness and Kurtosis in Testing Normality and Robustness Studies. *Sankhya: Indian Journal of Statistics, Series B* 36: 115–128.
- Martel, R. 1996. Heterogeneity, Aggregation, and a Meaningful Macroeconomics. In *Beyond Microfoundations: Post Walrasian Macroeconomics*. New York, NY: Cambridge University Press.
- Mayer, S.E. 2002. How Economic Segregation Affects Children's Educational Attainment. *Social Forces* 81(1): 153–176.
- McDonald, J.F. 1989. Econometric Studies of Urban Population Density: A Survey. *Journal of Urban Economics* 26: 361–385.
- McDonald, J.F, and D.P. McMillen. 1990. Employment Subcenters and Land Values in a Polycentric Urban Area: The Case of Chicago. *Environment and Planning A* 22(12): 1561 – 1574.
- McDonald, J.F., and P.J. Prather. 1994. Suburban Employment Centres: The Case of Chicago. *Urban Studies* 31(2): 201–218.

- McLeod Jr, M.S., K.J. Flannelly, and R.W.Behnke. 1991. Multivariate Time-series Model of Transit Ridership Based on Historical, Aggregate Data: The Past, Present and Future of Honolulu. *Transportation Research Record* 1297: 76–84.
- McMillen, D.P. 2001. Nonparametric Employment Subcenter Identification. *Journal of Urban Economics* 50(3): 448–473.
- McMillen, D.P. 2003. Identifying Sub-centres Using Contiguity Matrices. *Urban Studies* 40(1): 57-69.
- McMillen, D.P. 2004. Employment Densities, Spatial Autocorrelation, and Subcenters in Large Metropolitan Areas. *Journal of Regional Science* 44(2): 225–244.
- McMillen, D.P., and J.F. McDonald. 1998. Suburban Subcenters and Employment Density in Metropolitan Chicago. *Journal of Urban Economics* 43(2): 157–180.
- Mieszkowski, P., and E.S. Mills. 1993. The Causes of Metropolitan Suburbanization. *Journal of Economic Perspectives* 7(3): 135–47.
- Neuman, M. 2005. The Compact City Fallacy. *Journal of Planning Education and Research* 25(1): 11–26.
- Newman, P., and J. Kenworthy. 1989. Gasoline Consumption and Cities -- A Comparison of U.S. Cities with a Global Survey. *Journal of the American Planning Association* 55(1): 24-37.
- Newman, P., and J. Kenworthy. 1999. *Sustainability and Cities: Overcoming Automobile Dependence*. 1st ed. Washington, DC: Island Press.
- Parolek, D.G., K.Parolek, and P.C. Crawford. 2008. *Form Based Codes: A Guide for Planners, Urban Designers, Municipalities, and Developers*. 1st ed. Hoboken, NJ: Wiley.
- Pfister, N., R. Freestone, and P. Murphy. 2000. Polycentric or Dispersion?: Changes in Center Employment in Metropolitan Sydney, 1981 to 1996. *Urban Geography* 21(5): 428–442.
- Pisarski, A. 2009. The Nexus of Energy, Environment and the Economy: A Win, Win, Win Opportunity. *ITE Journal* January: 30–41.
- Porter, E.M. 1997. New Strategies for Inner-City Economic Development. *Economic Development Quarterly* 11(1): 11-27.

- Pushkar, A., B. Hollingworth, and E. Miller. 2000. *A Multivariate Regression Model for Estimating Greenhouse Gas Emissions from Alternative Neighborhood Designs*. Paper Presented at the 79th Annual Meeting of the Transportation Research Board, Washington, DC.
- Redfearn, C.L. 2007. The Topography of Metropolitan Employment: Identifying Centers of Employment in a Polycentric Urban Area. *Journal of Urban Economics* 61(3): 519–541.
- Rodr uez, D.A., and J. Joo. 2004. The Relationship Between Non-motorized Mode Choice and the Local Physical Environment. *Transportation Research Part D: Transport and Environment* 9(2): 151–173.
- Ross, C.L., and A.E. Dunning. 1997. *Land Use Transportation Interaction: An Examination of the 1995 NPTS Data*. Atlanta, GA: Georgia Institute of Technology.
- Schimek, P. 1996. Household Motor Vehicle Ownership and Use: How Much Does Residential Density Matter? *Transportation Research Record* 1552: 120–125.
- Shukla, V., and P. Waddell. 1991. Firm Location and Land Use in Discrete Urban space☆ A Study of the Spatial Structure of Dallas-Fort Worth. *Regional Science and Urban Economics* 21: 225–253.
- Smart Growth Network. 2002a. *Getting to Smart Growth I: 100 policies for implementation*. Washington, DC: International City/County Management Association.
- Smart Growth Network. 2002b. *Getting to Smart Growth II: 100 more policies for implementation*. Washington, DC: International City/County Management Association.
- Spillar, R.J., and G.S. Rutherford. 1998. The Effects of Population Density and Income on Per Capita Transit Ridership in Western America Cities. *ITE 1990 Compendium of Technical Papers*. Washington, DC: Institute of Transportation Engineers. 327–331
- Spirtes, P., C. Glymour, and R. Scheines. 2001. *Causation, Prediction, and Search, 2nd ed.* Cambridge, MA: The MIT Press.
- Srinivasan, S. 2002. Quantifying Spatial Characteristics of Cities. *Urban Studies* 39(11): 2005–2028.

- Stead, D., and S. Marshall. 2001. The Relationships between Urban Form and Travel Patterns: An International Review and Evaluation. *European Journal of Transport and Infrastructure Research* 1(2): 113-141
- Taylor, B., and C. Fink. 2002. *The Factors Influencing Transit Ridership: A Review and Analysis of the Ridership Literature*. Los Angeles, CA: UCLA Institute of Transportation Studies.
- Taylor, B., D. Miller, H. Iseki, and C. Fink. 2009. Nature and/or Nurture? Analyzing the Determinants of Transit Ridership across US Urbanized Areas. *Transportation Research Part A: Policy and Practice* 43(1): 60–77.
- TMRPA. 2005. *Infill Development: Barriers and Incentives*. Reno, NV: Truckee Meadows Regional Planning Agency
- Torrens, P.M. 2008. A Toolkit for Measuring Sprawl. *Applied Spatial Analysis* 1: 5–36.
- Tsai, Y-H. 2005. Quantifying Urban Form: Compactness versus ‘Sprawl’. *Urban Studies* 42(1): 141–161.
- U.S. EPA. 2011. *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2009*. Washington, DC: U.S. Environmental Protection Agency. <http://epa.gov/climatechange/emissions/usinventoryreport.html>.
- U.S. SCCST. 2009. *Federal Surface Transportation Policy and Planning Act of 2009*. Washington, DC: U.S. Senate Committee on Commerce, Science & Transportation. <http://commerce.senate.gov/public/index.cfm>.
- Van de Coevering, P., and T. Schwanen. 2006. Re-evaluating the Impact of Urban Form on Travel Patterns in Europe and North-America. *Transport Policy* 13(3): 229–239.
- Voortman, M., and M.J. Druzdzal. 2008. Insensitivity of Constraint-based Causal Discovery Algorithms to Violations of the Assumption of Multivariate Normality. In *Proceedings of the Twenty-First International Florida Artificial Intelligence Research Society Conference, 2008*. 690–695. Palo Alto, CA: Association for the Advancement of Artificial Intelligence.
- Yuan, K-H, W. Chan, and P.M. Bentler. 2000. Robust Transformation with Applications to Structural Equation Modelling. *British Journal of Mathematical and Statistical Psychology* 53(1): 31–50.
- Zhang, M. 2004. The Role of Land Use in Travel Mode Choice - Evidence from Boston and Hong Kong. *Journal of the American Planning Association* 70(3): 344–360.

**VITA**

Name: Young-Jae Yi

Address: Texas Transportation Institute, Environmental Management Program  
2929 Research Parkway, TAMU Research Park  
College Station, TX 77843

Email Address: [y-yi@tamu.edu](mailto:y-yi@tamu.edu)

Education: B.E., Urban Engineering, University of Seoul, 1996  
M.U.P., Urban Planning, University of Michigan at Ann Arbor, 2004  
Ph.D., Urban and Regional Sciences, Texas A&M University at  
College Station, 2012