CONNECTING LAND USE AND TRANSPORTATION

TOWARD SUSTAINABLE DEVELOPMENT:

A CASE STUDY OF THE HOUSTON-GALVESTON METROPOLITAN AREA

A Dissertation

by

JAE SU LEE

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

December 2009

Major Subject: Urban and Regional Planning

CONNECTING LAND USE AND TRANSPORTATION TOWARD SUSTAINABLE DEVELOPMENT: A CASE STUDY OF THE HOUSTON-GALVESTON METROPOLITAN AREA

A Dissertation

by

JAE SU LEE

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
	Josias Zietsman
Committee Members,	David A. Bessler
	Christopher D. Ellis
Head of Department,	Forster Ndubisi

December 2009

Major Subject: Urban and Regional Planning

ABSTRACT

Connecting Land Use and Transportation toward Sustainable Development:

A Case Study of the Houston-Galveston Metropolitan Area. (December 2009)

Jae Su Lee,

B.S., University of Seoul, Korea;

M.S., Seoul National University, Korea

Co-Chairs of Advisory Committee: Dr. Ming-Han Li
Dr. Josias Zietsman

How do land use characteristics affect individual and household travel behavior in a regional context? Can the investigation justify the land use policies to reduce automobile dependence and achieve the goals of sustainable development in the metropolitan areas? Previous research enhanced our understanding of the connections between land use and travel behavior. It also provided implications for managing automobile-dependent travel behavior. However, there are questions still left unanswered about the causal connections between them, and the effectiveness of the land use policies to manage travel demand.

To address the issues, attention is focused on the effects of land use measures on travel behavior outcomes from different modeling perspectives. The travel demand modeling explores the associations between land use and travel behavior. In addition, the causal modeling helps clarify the causal connections between them. It includes the structural equation models (SEMs) and the directed acyclic graphs (DAGs). The study

focuses on six counties of the Houston-Galveston Area Council (HGAC) area. Travel behavior outcomes contain individual mode choice, household automobile trip generation and household total vehicle miles traveled (VMT). Three dimensions (i.e., density, diversity and design) of six land use measures are considered, which are computed using quarter-mile buffers for both trip origins and destinations. Different travel outcomes and modeling strategies are examined for different travel purposes.

The significance of land use measures in affecting travel behavior is found to be evident, while varying to a certain degree according to trip purposes, travel outcomes and methodologies. For individual model choice, multinomial logit (MNL) models, the SEMs and the DAGs for different trip purposes support the hypothesis that land use measures directly affect individual mode choice behavior when other factors are kept constant. There is also evidence from causal models that land use factors indirectly influence it through travel time. For household automobile trip generation, there is no evidence to assert that land use measures at origin significantly affect household automobile trip rates when travel cost and socioeconomic variables are controlled. However, it is confirmed that land use measures have indirect causal connections with automobile trips through travel costs for all trip purposes. For household total VMT, it is found that land use patterns around residential locations are not only significantly associated, but also causally connected with household VMT. To summarize, compact development with high density and improved network design generally contribute to the reduction in automobile dependent travel patterns in the HGAC region.

DEDICATION

To my lovely family, Seung Mi, Jae Ha, and Seong-Uk,

and

my mother, Mrs. Ye Soon Kim, and late father, Mr. Seung Gon Lee

ACKNOWLEDGEMENTS

I have been deeply indebted to my co-chairs, committee members, colleagues and family for my dissertation. First of all, I would like to thank Dr. Ming-Han Li, my wonderful advisor and co-chair of my dissertation committee. He provided financial and mental support, and consistently encouraged me to complete my study. I am also deeply grateful to Dr. Joe Zietsman, my co-chair, who has been giving me financial and academic support while I have been working at the Texas Transportation Institute. Without their help, my dissertation would have never come into the world. I am fortunate in having great committee members. I am grateful to Dr. David Bessler for his expertise in analytical methods and interpretations, and to Dr. Christopher Ellis for his valuable insights into measurement and methodology. Many thanks go to a number of people who helped me obtain the data for my study: Mr. Charlie Hall from the Texas Department of Transportation, Dr. David Pearson and Mr. Edwin Hard from TTI, and Mr. Chris Van Slyke, Ms. Heng Wang, and Ms. Sharon Ju from the Houston-Galveston Area Council. I wish to acknowledge many excellent colleagues in the Urban and Regional Science program at Texas A&M University: Dr. Jae Bum Jun, Dr. Jung Eun Kang, Chan Yong Sung, Young-Jae Yi, Joong-Hyuk Choi, and Jung-Jae Yoon. I am indebted to my family for their love and support: Seung Mi, Jae Ha and Seong-Uk. I also want to express my heartfelt gratitude to my parents (Seung Gon Lee in heaven and Ye Soon Kim), brothers (Dong Su in heaven and Hak Su Lee), sisters (Gyung Ja, Hyun Sook, Keum Ei, and Yoon Ja Lee), my parents-in-law and other family members.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

FIGURE Page

LIST OF TABLES

TABLE Page

CHAPTER I

INTRODUCTION

1.1 Background

Automobile dependence has been intensifying over past decades in the United States. Between 1960 and 2006, total number of registered vehicles has grown by 120%. Vehicle miles traveled (VMT) and passenger miles traveled have increased by 187% and 132%, respectively. Total number of residents and households during the same period, however, has only augmented by 66% and 116% each. Although automobile dependence has improved the economic efficiency and competitiveness greatly, it has had harmful impacts on the economic, societal and environmental system including traffic congestion, traffic accidents, air and water pollution, energy and land consumption, ecological disruption and public health problems.

The U.S. has experienced rapid urban growth and suburbanization as well during this period. As a consequence, land use and development patterns are characterized as detached low-density residential communities, segregated commercial and industrial sites, and automobile-oriented urban and transportation planning, which is termed urban sprawl. A self-reinforcing pattern of growing automobile dependence, automobile-oriented planning and development and segregated and sprawling land use have brought detrimental effects on our economy, society, and environment (VTPI 2008a).

This dissertation follows the style of *Journal of the American Planning Association.*

These concerns, combined with growing awareness of the consequences of automobile dependence, have led the public to pay attention to a comprehensive framework called sustainable development and transportation. Sustainable transportation is an applied concept of sustainable development to the transportation field. Sustainable transportation has been prevalent as the aforementioned issues in transportation and land use should be addressed in comprehensive and integrated manners. Consequently, policies and strategies for increasing transportation system efficiency as well as decreasing negative impacts are the most effective ways for achieving the goals and objectives of sustainable transportation. One of the main academic efforts is to investigate the relationship between land use and travel behavior patterns (Zietsman and Rilett, 2002; Litman and Burwell, 2006).

Land use and transportation are closely connected with each other. There have been a number of studies on the impact of land use measures on individual and household travel behavior. The studies are significant in that they suggest policy implications for reducing automobile dependence and achieving the goals of sustainability. Significant improvements have been made in land use measurement, model estimation methods and methodological framework.

However, the adequacy of land use policies still remains questionable for reducing automobile dependence and accomplishing the goals of sustainability. This is mainly due to lack of consistent results and an integrated approach toward sustainability of previous studies. This study can make some contributions as follows. First, sustainability measures related to land use attributes are developed. Land use attributes are measured in detailed spatial level with the Geographic Information System (GIS) techniques. Second, causal relationships between land use and travel behavior are examined beyond conventional travel demand models. Lastly, Houston-Galveston metropolitan area is one of the biggest regions in the U.S. Little research, however, has been conducted to understand the connections between land use and travel behavior.

1.2 Objectives of the Study

The objectives of the study are fourfold.

First, land use measures in terms of sustainable transportation will be examined and developed. Three dimensions of land use characteristics, density, diversity and design will be formulated in order to be applied to the metropolitan area.

Second, the associative connections between land use measures and travel behavior outcomes will be investigated using conventional analytical methods based on economic behavior theory for utility maximization. The impacts of land use measures are also estimated and compared with different travel purposes.

Third, the causal relationships between land use and travel behavior will be further investigated to understand the causal connections among land use measures, travel time and cost variables, socioeconomic characteristics, and travel behavior outcomes. They will be estimated for different travel purposes.

Last, policy implications for integrating land use and transportation and thus reducing the negative effects of automobile dependence will be suggested. Implications for improving current regional travel demand models will also be addressed.

1.3 Scope of the Study

Land use and transportation are closely intertwined. Transportation investments and policies influence land development patterns. Land use attributes also affect individual and household travel behaviors (Handy 2002). This study primarily focuses on the effects of land use characteristics on travel behavior in terms of both associative and causal relationship. Figure 1.1 describes the relationship between land use patterns and travel behavior patterns.

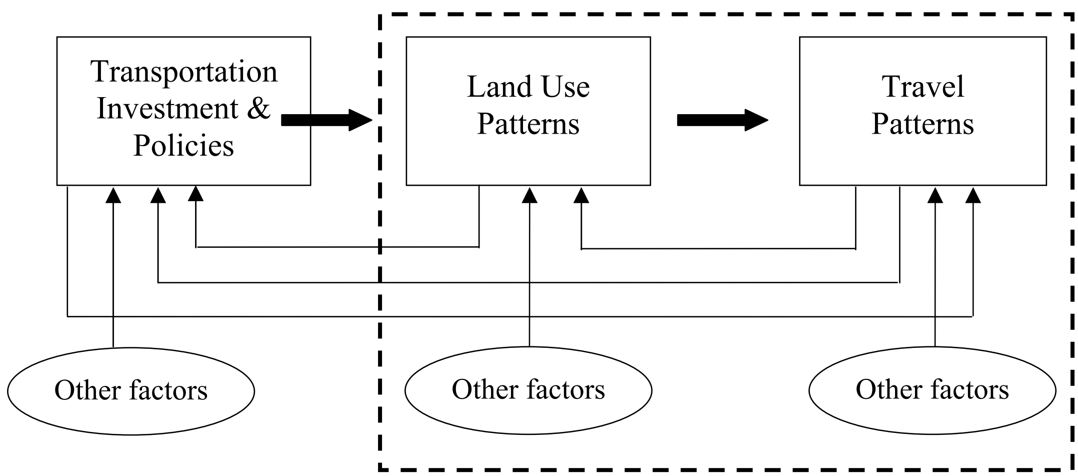


Figure 1.1 Connections between Land Use and Transportation.

1.4 Organization of the Study

The study is organized into six chapters. Chapter I addresses research background, objectives, and scope. Chapter II reviews literature focusing on the issues of automobile dependence and sustainable transportation, and the effects of land use and travel behavior pattern. Chapter III presents the framework and research design for analyzing travel demand and causal relationship. Research hypotheses are also addressed according to both travel demand models and causal models. Chapter IV introduces the study area and data sources, and discusses how the variables of interest are measured and applied for the study. Chapter V examines overall household travel pattern and land use characteristics. Model estimation results are also presented and interpreted for individual mode choice, household auto trip generation, and household total VMT. The last chapter makes conclusions based on major findings of the investigation. Then, policy implications are explored to deal with automobile dependence as well as to achieve the objectives of sustainable development. Limitations and possible improvements of this research are also discussed.

CHAPTER II

LITERATURE REVIEW

This chapter reviews the literature related to automobile dependence and sustainable transportation, and land use on transportation behavior pattern. The first section, automobile dependence and sustainable transportation, discusses the definitions, general trends and causes and consequences of automobile dependence. In addition, concepts and objectives, issues and challenges, performance measurement and the role of land use related to sustainable transportation are examined. The second section, land use impact on transportation behavior pattern, provides a synopsis of related research and examines relevant issues and efforts in detail.

2.1 Automobile Dependence and Sustainable Transportation

2.1.1 Automobile Dependence: A Problem

2.1.1.1 Definitions

Automobile dependence is a social trend indicating that an automobile has been indispensable with sustaining a wide variety of human activities including commute, business, and social gathering. It can also be defined and measured as higher proportion of automobile use and ownership, fewer numbers of available alternative modes, and automobile-oriented land use or urban form (Newman and Kenworthy 1999; Litman and Burwell 2006). There are a number of studies on automobile dependence and its impacts on our economy, environment and society. Nonetheless, its definition and measurement have been varied according to the purposes and approaches of related researches.

Newman and Kenworthy (1989a, 1989b) specified automobile dependence as the interrelation of land use and transportation. The intensity of automobile dependence was measured using the correlation between the density of an urban area and gasoline consumption per person. It was found that there was a negative relationship between them (Mindali et al. 2004). This is thought to be the most important finding for a series of following studies (Lee 2006). They made an important contribution to the understanding of the nature of automobile dependence and how it can be structured into the urban dimension. They argued in their later work that transportation priorities, explained as high propensity for automobiles and the supply of relevant infrastructure, together with economic and cultural priorities are primary factors creating automobile-dependent cities. They are characterized as low-density and detached land use, and a high proportion of automobile use and ownership (Newman and Kenworthy 1999). Some studies have expressed sharp criticism of the research. They pointed out that it would not be appropriate to analyze the relationship between aggregate urban density and average per capita gasoline consumption, and apply a simple method of clustering and correlation between them to explain complex system of the urban structure (Gordon and Richardson 1989; Gomez-Ibanez 1991; Goodwin 1997; Mindali et al. 2004).

Automobile dependence has also been explained with a high percentage of auto driving and less available travel modes which are caused by the interaction between automobile transport and land use patterns (Litman 2002; Litman and Laube 2002). In particular, Litman (2002) interpreted the social phenomenon from an economic perspective. Household internal and external economic costs due to increased automobile dependence were compared with those due to balanced transportation in our communities.

Goodwin (1997) introduced a different approach to automobile dependence, which is described as a dynamic and developmental process of personal and social behavior by times. When it comes to travel modal split in an urban area, automobile dependence can be explained by personal mode choice based on individual preference for an automobile mainly due to better convenience and mobility. It can also be resulted from the unavailability of alternative modes related to personal attitude, land use patterns, and other conditions. In a similar vein, Stradling (2001) defined it as a degree for satisfying individual travel needs. Both absolute and relative measures of automobile dependence were suggested. The former included vehicle trip frequency, travel time and distance, while the latter focusing on the personal attitudes toward an automobile including vehicle use rate in mixed mode choices and activities.

2.1.1.2 General Trends

People in the U.S. have been more and more depending on automobiles over past decades as shown in Table 2.1 as they have been keeping up their growing demands on various activities including commuting, recreation and shopping. Between 1960 and 2006, total population, households and housing units have grown by about 66%, 116% and 116%, respectively. During the same time period, the numbers of vehicle registration and licenses have increased by 120% and 132% each, which indicates that automobile ownership and related demand have become greater than the net increases of socio-demographic figures. In addition, total vehicle miles traveled (VMT) and total passenger miles traveled (PMT) have become longer by 187% and 132%, respectively. They imply that automobile use has expanded more than socio-demographic growth over the decades.

Furthermore, net increases of yearly total VMT per household and total PMT per person are 33% and 40%, respectively. Total VMT per vehicle has grown by 31% per year. An economic indicator, total expense related to personal automobiles has also increased by more than eleven times during the decades. It suggests that automobile related expenditures including purchase and maintenance costs have rapidly increased in the U.S. although the growth rates of the population and vehicle registration are considered. The intensity of these indicators has decreased compared with the time period between 1970 and 2006; the trends of growing automobile dependence, however, are still significant in the U.S.

Table 2.1 Summary Statistics of the Trends of Automobile Dependence in the U.S.

	1960	1970	1980	1990	2000	2006	1960-2006 [5)]
Total population [1)]	179.3	203.2	226.5	248.7	281.4	298.4	66
Total households [1)]	53.0	63.4	80.4	91.9	105.5	114.4	116
Total housing units [1)]	58.3	68.7	88.4	102.3	115.9	126.2	116
Registered vehicles [1) 3)]	61.7	89.2	121.6	133.7	133.6	135.4	120
Vehicle license [1)]	87.3	111.5	145.3	167.0	190.6	202.8	132
Total VMT [2) 3)]	587.0	919.7	1121.8	1417.8	1600.3	1682.7	187
Total PMT [2) 3)]	1145.0	1754.2	2024.2	2140.9	2544.5	2658.6	132
VMT / household	11,071	14,495	13,955	15,420	15,171	14,711	33
PMT / person	6,385	8,632	8,935	8,608	9,041	8,911	40
VMT / vehicle [3)]	9,518	9,989	8,813	10,277	11,976	12,427	31
Personal auto expense [4)]	222	361	925	1,518	2,235	2,778	1,149

Note: 1) millions; 2) billions; 3) only for passenger cars; 4) million dollars; 5) net increase (%) compared with base year.
Sources: 1) U.S. Census Bureau (2009); 2) U.S. Census Bureau (2008); 3) U.S. Census Bureau (2002); 4) U.S. Census Bureau (2007); 5) BTS (2008).

These trends of growing automobile dependence in the U.S. have also been observed in other ways. Two economic indicators are measured on a yearly basis: average total automobile cost per mile and total transportation expenditures by governments. First indicator shows that every American has been spending more and more upon owning and operating automobiles for several decades (see Figure 2.1).[1] Another measure reveals how much money the federal, state and local governments spend in the transportation field. Figure 2.2 suggests that both total and highway expenditures have consistently augmented over 20 years. In particular, the increasing

[1] BTS notes that it is not sound to make direct comparison before and after 1985 and 2004 due to major changes in calculation method in these years.

expense for highway mode supports the argument of growing automobile dependence in the U.S.

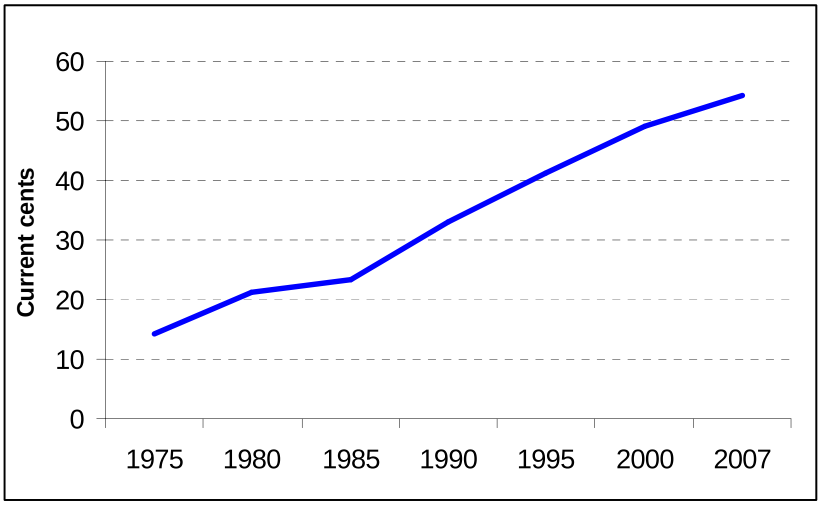


Figure 2.1 Average Total Automobile Cost per Mile. Source: BTS (2008).

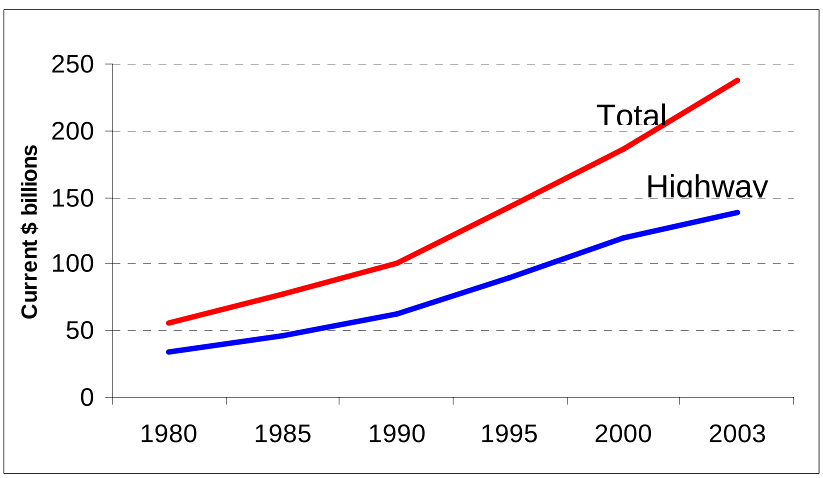


Figure 2.2 Total Transportation Expenditures by Governments. Source: BTS (2008).

In addition, the survey result of principal commuting modes reinforces the evidence. As presented in Figure 2.3, the share of driving mode is dominant; on the other hand, the number of workers using non-automobile modes is very small. In short, automobile dependence in the U.S. has been growing for many decades when various indicators of automobile ownership and use, economic spending and modal splits are taken into consideration. It is a result of a self-reinforcing cycle of increased automobile ownership and use, decreased alternative modes and automobile-oriented transportation and land use policies (VTPI 2008a).

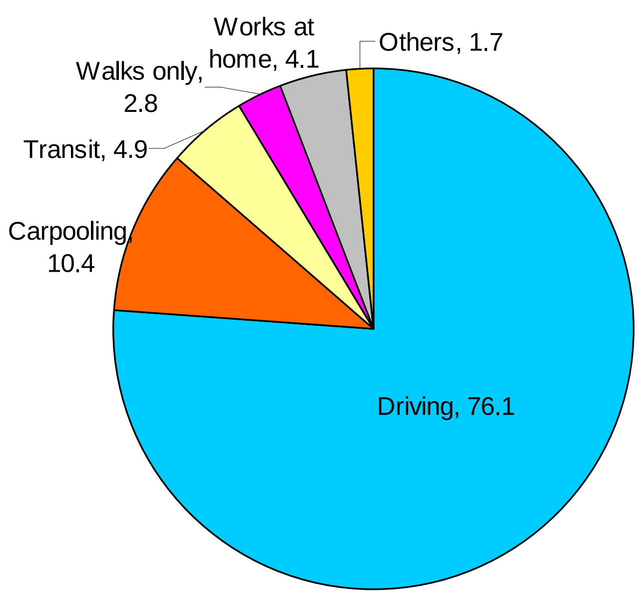


Figure 2.3 Principal Modes of Commuting Trips. Source: BTS (2008).

2.1.1.3 Causes and Consequences

There are a number of causes of growing automobile dependence. Lee (2006) identified some factors by which automobile dependence in the U.S. has been aggravated: progress in transportation technology, improvement of transportation infrastructure, land use patterns, reduced availability of alternative modes, socioeconomic characteristics, and personal attitudes.

VTPI (2008a) also examined some factors in terms of transportation practices: conventional transportation planning, evaluation, and current investment. Conventional transportation planning practices forecasted vehicle traffic demand in the future, and execute projects for constructing and improving roadway and parking capacity (Litman and Burwell 2006). It made transportation system and land use more automobile-dependent. Transportation evaluation practices mainly focused on automobile traffic, while little consideration is given to other modes. Also, dominant portion of current investment and funding to road and parking construction and improvement accelerated automobile dependence in the U.S.

Its positive influences on our economy and society have been also documented. It has increased automobile mobility and convenience, affordability of vehicle travel for both low-income households and disadvantaged people. Increased mobility has positive impact on economic productivity and efficiency. Economic development is relevant to fuel and vehicle production and services, and some places accessible to automobiles (VTPI 2008a). Dupuy (1999) argued that higher level of automobile dependence is a natural consequence of more positive effects than negative effects. He contended that policies focusing on demand and supply of vehicles, and changes in network system have a positive influence on decreasing automobile dependence. Land use factor, however, was not considered in his research (Lee 2006).

On the other hand, it has had negative effects on our economic, societal and environmental systems. Its diseconomies include infrastructure construction and maintenance cost, traffic congestion, traffic accident damages, automobile ownership and maintenance cost, fewer travel mode choice options and less accessible land use patterns (Litman 2008a; Lee 2006). Negative social effects encompass public health, equity and segregation. Negative effects of automobile dependence on environmental system have also been extensively reported. They incorporate water and air pollution, energy depletion, loss of lands for agricultural and ecological production, vehicle disposal, and habitat disruptions (Raad 1998; WHO 2000; Black 2005; Lee 2006; BTS 2008; Litman 2008a). Litman and Burwell (2006) classified the impacts into three dimensions of sustainability as summarized in Table 2.2.

Table 2.2 Transportation Impact on Sustainable Development.

Economic	Social	Environmental
Traffic congestion Mobility barriers Accident damages Facility costs Consumer costs Depletion of Non-Renewable Resources	Inequity of impacts Mobility disadvantaged Human health impacts Community interaction Community livability Aesthetics	Air and water pollution Habitat degradation Hydrologic impacts Depletion of Non-Renewable Resources

Source: Litman and Burwell (2006).

An annual cost of congestion, for instance, was estimated at $67.5 billion for 75 U.S. metropolitan areas (Schrank and Lomax 2002). Figure 2.4 illustrates how the traffic congestion in urbanized areas has been growing as a consequence of increased automobile dependence. The annual road congestion index (RCI) measures vehicle travel density on major roads in different types of urban areas.[2] As shown in the figure, traffic congestion has been continuously increased in all types of urban areas over 20 years in the U.S.

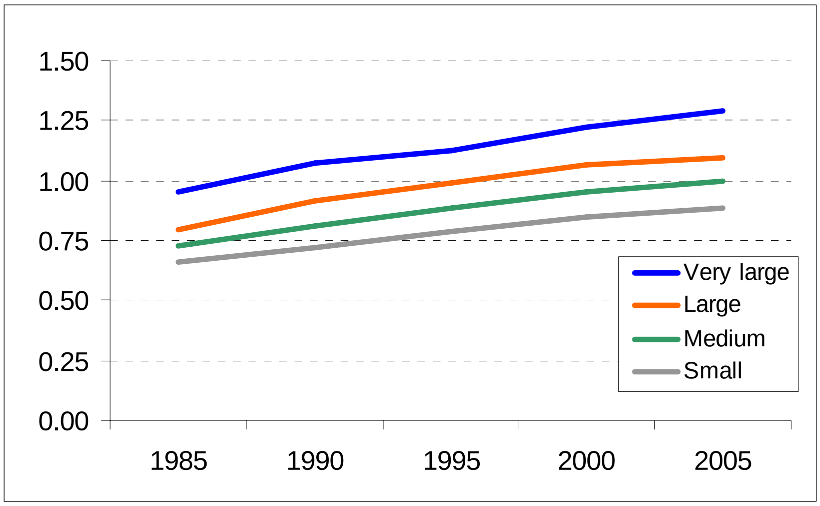


Figure 2.4 Annual Road Congestion Index. Source: Schrank and Lomax (2007).

2.1.2 Sustainable Transportation: A Solution

2.1.2.1 Concepts and Objectives

Consequences caused by growing automobile dependence are linked to the tripod of sustainability: economic, environmental and social dimensions. In addition, growing concern about the negative impacts of automobile dependence and policy changes in the U.S. have required comprehensive framework and actions in transportation (Newman and Kenworthy 1999; Litman and Burwell 2006; Litman 2008a). These challenges and issues have led to the introduction of sustainability into the transportation sector.

[2] An RCI over 1.0 implies an urban area is undesirable on an average in terms of congestion level on major roadways during the peak period. Study areas are those with more than 500,000 population and some smaller areas (Schrank and Lomax 2007).

There is no general agreement on the definition of sustainable development; rather it has been defined and applied according to the goals and objectives of each agent or organization (Beatley 1995; Litman and Burwell 2006). There is, however, a widely used concept of sustainable development defined by the World Commission on Environment and Development (WCED): sustainable development "meets the needs of the present without compromising the ability of future generations to meet their own needs" (WCED 1987). It is a changing and progressive concept considering people's growing demands for various dimensions in our society (Zietsman and Rilett 2002). It has been embodied into the transportation field, called sustainable transportation or transportation sustainability. There is no standard definition for transportation sustainability as well. OECD (1999), for example, defined sustainable transport as "transportation that does not endanger public health or ecosystems and meets needs for access consistent with 1) use of renewable resources below their rates of regeneration, and 2) use of non-renewable resources below the rates of development of renewable substitutes."

European Council of Ministers of Transport (ECMT 2004) proposed a definition of sustainable transportation. In addition to the idea of OECD (1999), ECMT (2004) specified sustainable transport system "allows the basic access and development needs of individuals, companies and society to be met safely and in a manner consistent with human and ecosystem health, and promotes equity within and between successive generations," and "is affordable, operates fairly and efficiently, offers a choice of transport mode and supports a competitive economy, as well as balanced regional development."

Jeon et al. (2006) also pointed out sustainable transportation did not have a unanimous definition for its own. Sustainable transportation system, as they defined, "should be effective and efficient in providing its users with equitable and safe access to basic social and economic services, should promote economic development, and not be harmful to the environment." Based on the concepts, it is confirmed that transportation sustainability is connected with the three dimensions of sustainability to accomplish its goals and objectives (Zietsman and Rilett 2002; Jeon and Amekudzi 2005).

2.1.2.2 Issues and Challenges

As sustainability generally incorporates economic growth, environmental conservation and social welfare, transportation sustainability also reflects a lot of related issues. It should be understood that sustainability in transportation per se can be achieved only when the three elements are fully addressed altogether (Zietsman and Rilett 2002). The issues of sustainable transportation can be categorized into three dimensions. Economic growth includes issues on productivity, business activity, employment, tax burden and trade; environmental preservation comprises issues on pollution prevention, climate protection, biodiversity and habitat preservation; social welfare encompasses issues on equity, public health, community livability, cultural and historical values, and public involvement (Litman and Burwell 2006; Litman 2008a).

Litman and Burwell (2006) proposed some policy approaches from a comprehensive point of view. They include: technological innovation such as alternative fuel and fuel-efficient vehicles, and Intelligent Transportation System; transportation demand management for improving traffic flow and increasing travel choices; economic reform including full-cost pricing and congesting pricing; alternative modes such as transit, ridesharing, and non-motorized modes; and land use and community design changes to decrease trip distance and increase mode choice. Wachs (2005) presented seven issues and questions with regard to sustainability in the future transportation: sustainable transportation indicators, changes in technology, the effect of government regulation, direct control of individual travel behavior, the effect of pricing policy, public education, and regional planning.

Schipper (2002) placed emphasis on governance sustainability in addition to other three elements in sustainable transportation. The key issues of the governance sustainability is to make an agreement and balance among stakeholders, and to develop effective policy measures for addressing transportation problems. In the same way, Zietsman and Rilett (2002) reviewed institutional and policy frameworks in the U.S. Detailed policies were examined to achieve the goals of sustainable transportation. Policy measures were presented including pricing, technology, regulation, traffic management, non-motorized transportation, behavior and education, and land use and transportation.

Table 2.3 summarizes the issues and challenges of transportation sustainability (STI 2008).

Table 2.3 Issues of Sustainable Transportation.

Economic	Social	Environmental
Accessibility quality Traffic congestion Infrastructure costs Consumer costs Mobility barriers Accident damages Depletion of Non-Renewable Resources	Equity and fairness Mobility disadvantaged Affordability Human health impacts Community cohesion Community livability Aesthetics	Air pollution Climate change Noise pollution Water pollution Hydrologic impacts Habitat/ecological degradation Depletion of Non-Renewable Resources

Source: STI (2008).

2.1.2.3 Performance Measurement

Sustainability in transportation can be assessed using a combination of indicators which is useful for setting up baselines, tracking changing patterns, evaluating alternatives, assessing and comparing particular regions or organizations, and establishing future performance objectives (CST 2000; Litman and Burwell 2006; Litman 2008a). Litman and Burwell (2006) argued that conventional and simple performance measures were not helpful for achieving sustainable transportation goals because they did not take into consideration the variety of related issues and concerns. Litman (2008a) defined sustainable transportation indicators with three broad categories, and proposed a group of indicators by each dimension.

Zietsman and Rilett (2002) claimed that little research on sustainable transportation has been done due to lack of understanding transportation sustainability and quantifying performance measures. They introduced advanced technologies for data collection and measurement at a disaggregate level, and the decision-making process. It was found that the final decision on project selection could be varied with the introduction of sustainable transportation concept and measures instead of economic feasibility analysis. Zietsman et al. (2003) have applied similar methodology to the previous research into two corridors: one in South Africa, a developing country, and another in the U.S. a developed country. They maintained that the implementation of the goals of sustainable transportation is important; therefore, they should be appropriately defined, measured, and employed into the decision-making process. It was argued that the same method could be implemented to decide transportation project priorities, and to compare different corridors regardless of their classification, goals, mode, time and spatial boundary.

Jeon and Amekudzi (2005) examined the characteristics of definitions, measurements and indicators of sustainable transportation system. They determined three frameworks for measuring transportation sustainability using indicator systems. It is found that sustainable transportation has been primarily assessed by effectiveness and efficiency of transportation system and the environmental impacts. Jeon et al. (2006) criticized that sustainability concepts have not been fully incorporated into the regional planning process including long-range regional plans and transportation improvement projects. The multi-criteria decision making approach was employed to evaluate a current and future transportation and land use plans in Atlanta Metropolitan Region. Indicators were classified into four groups: system effectiveness, economic, environmental, and social welfare indicators. They concluded that the method was useful for integrating sustainable transportation measures into transportation planning and decision-making process, and assessing plans with regard to sustainability goals and objectives.

2.1.2.4 The Role of Land Use in Transportation Sustainability

Studies of sustainable transportation have focused on performance measures and decision-making process. But, little research on the role of land use in transportation sustainability has been conducted. Land use effects on travel behavior patterns have been mainly studied. Therefore, it is reasonable to assert that sustainable transportation issues have been connected with land use in most American regions (Litman and Burwell 2006).

Litman and Burwell (2006) summarized transportation objectives and solutions that are consistent with the goals and objectives of sustainability. In particular, many solutions related to land use and development were proposed. They included efficient land use for freight mobility; neotraditional street planning and mixed land use for mobility of non-drivers; multi-modal community and land use; and pedestrian planning and livable community design.

STI (2008) listed potential indicators for achieving sustainable transportation goals within a number of categories and subcategories of sustainability concerns. Two main categories linked to the role of land use are overall accessibility and land use impacts. The former includes land use accessibility; the latter consists of three subcategories: sprawl, transport land consumption, and ecological and cultural degradation. Litman (2008a) also identified sustainable transportation indicators. Land use and development plays an important role in a set of economic, social and environmental indicators. They cover employment accessibility, land use mix, land use planning, non-motorized transport, and land use impact indicators.

The role of land use in transportation sustainability cannot be overstated. It is also important to coordinate land use and transportation planning in order to make them compatible (Litman 2008b). The impacts of land use on travel behavior will be reviewed in more detail in the following section.

2.2 Land Use Impact on Travel Behavior Pattern

2.2.1 A Synopsis of Related Research

Land use or urban form[3] and transportation are closely connected with each other in two major and more minor ways (Handy 2002). Transportation investments and policies influence land use and development patterns; land use and development also affect transportation and travel behavior patterns. A number of studies examining the effect of land use and development on travel behavior outcomes have been mainly conducted with regard to theoretical framework and methods, practical analyses and applications (Badoe and Miller 2000; Crane 2000; Cervero 2002). The research started from the late 1980s in response to the public interest in how and to what extent land use measures can reduce automobile dependence. Badoe and Miller (2000), Crane (2000), and Ewing and Cervero (2001) provide great reviews from various perspectives.

Academic investigations of this discipline germinated from a pivotal research conducted by Newman and Kenworthy in 1989 (Newman and Kenworthy 1989b). They analyzed the simple relationship between transportation and land use in 32 major

---

[3] Urban form is often recognized as more comprehensive than land use pattern in a spatial boundary. In this point of view, land use pattern is an aspect of urban form involving a variety of spatial characteristics. However, this study considers land use to be the same concept as urban form as already did in many studies. Built environment introduced in some studies is taken into account in the same way.

international cities. It was claimed that urban density had negative impact on average annual gasoline use. An important contribution has been made to enhancing our understanding of how land use could systematize automobile dependence. The research has opened a ground for policy debates among the experts of planning and development fields. During the early 1990s, an interest has been increased in land use policies to manage transportation demand, which resulted in policy debates on the effectiveness of land use policies (Zhang 2004; Lee 2006). The arguments were originally developed from two different viewpoints: "get the price right" based on price-based mechanism in the transportation markets (Gomez-Ibanez 1991; Giuliano and Small 1993; Giuliano 1995), and "get the land use right" mainly depending on physical planning and design (Cervero 1991; Jacobs 1992; Cervero and Landis 1995; Newman et al. 1995).

A group of professionals supporting the former point of view argued that the connection between land use and transportation has consistently diminished in the U.S. and other developed countries. It was, they maintained, due to decreasing travel costs, well-developed transportation systems, and structural shifts to an information-based economy (Giuliano 1995). In response, others claimed that the transportation and land use connection should be still considered an important matter (Cervero and Landis 1995). There has been strong evidence that land use patterns significantly affected travel demand; land use and development, therefore, remained an important measure and policy to manage travel demand. Litman (2000) also stated that transportation market has been distorted with violated free market principles. Limited choices and increased automobile dependence due to the market distortions resulted in economic inefficiency, social inequity and environmental disruption. To address them, feasible and cost-effective market reforms should be established.

Great advances have been made in land use measurement and methodology until late 1990s. Land use measures such as density, diversity or land use mix, and accessibility were significantly increased. They enlarged the capacity to evaluate the built environment efficiently and effectively in both quantitative and qualitative ways. In addition, studies examined the relationship between transportation and land use using the regression analysis methods by employing various land use variables, while controlling for other economic and individual factors (Cervero and Gorham 1995; Cervero 1996; Handy 1996a; Cervero and Kockelman 1997; Kockelman 1997; Levinson and Kumar 1997; Boarnet and Sarmiento 1998; Handy et al. 1998; Crane 2000).

Furthermore, academic efforts have been made to establish an analytical framework based on consumer behavior theory for utility maximization of microeconomics that originated from the work of Domencich and McFadden (1975). The travel demand models have been elaborated to incorporate the full set of explanatory variables such as travel time and cost variables, individual and household socioeconomic factors, and land use measures (Crane and Crepeau 1998; Boarnet and Greenwald 2000; Boarnet and Crane 2001a; Cervero 2002; Zhang 2004; Lee 2006).

Recently, some issues are still being discussed and investigated. They include theory and modeling framework, land use measurement, causal relationship and self-selection, substitution effect, automobile captivity, and application of empirical results into the real travel model. These six issues will be discussed in greater details.

2.2.2. Relevant Issues and Efforts

2.2.2.1 Theory and Modeling Framework

There are three broad groups of researches from the standpoint of analytical and modeling framework: simulation, description and multivariate statistical studies (Crane 2000; Boarnet and Crane 2001a). Multivariate statistical methods applied for the majority of recent studies are specified and estimated with enough consideration of other factors, external validity and policy implications. However, they often suffer from lack of a conceptual framework and theory to explain the linkage of land use and travel behavior (Crane 2000; Cervero 2002).

Most of the estimated models in the previous studies have originated from the theory of economic behavior for utility maximization (McFadden 1974; Domencich and McFadden 1975; Ben-Akiva and Lerman 1985; Ben-Akiva and Bierlaire 1999). In this sense, the models should reflect individual behavior and motivations (McFadden 1974). But many studies have failed to consider transportation cost and system factors mainly due to the lack of behavioral framework, which led to biased estimates (Boarnet and Sarmiento 1998; Crane 2000; Boarnet and Crane 2001a; Cervero 2002).

Cervero and Kockelman(1997) presented conventional travel demand models with the utility based theory in their study of the San Francisco Bay Area. Three dimensions of built environment, density, diversity and design features were introduced with socio-demographics and transportation system in the model. The analytical efforts gave a significant impact on subsequent researches. They found that built environment significantly reduced the number of trips and the probability of auto choice. The effects of built environment in combination were substantial. However, a weakness from the theoretical perspectives still existed.

Boarnet and Sarmiento (1998) examined land use influence on non-work trip rates with consumer demand framework based on neo-classical economic theory. The demand model of non-work automobile trips was specified as a function of travel time, income and socio-demographic variables. They concluded the results did not clarify the connections between them. However, it is questionable if the estimation methods were appropriate (Lee 2006).

Crane (2000) raised several questions about why the results and the arguments of related literature have been debatable, and how they could be enhanced in terms of modeling framework. After reviewing numerous studies, he argued that studies containing demand variables based on economic theory were more appealing than others. In addition, he maintained the linkage of design factors to price variables (Boarnet and Crane 2001b), application of appropriate scale of geography, and incorporation of residential decision into the model (Boarnet and Crane 2001a). Boarnet and Crane (2001a) asserted with a critical eye that many past studies have poorly applied the behavioral theory and estimation of travel demand. They employed the demand theory into different model frameworks and specifications. It is concluded that land use measures influenced non-work automobile trip rates through prices using speed and distance; if there would be no significant relationship between land use and trip prices, on the other hand, the connection should be no more significant. This assumption would serve misleading results in estimation of price elements using land use variables (Lee

2006).

Cervero (2002) criticized most previous researches into land use impact on mode choice for lack of sound theories and methodologies partly due to modeling conventions and data constraints. Attention has been paid to establish a normative framework based on discrete choice theory. He found that inclusion of land use factors into the models significantly improve the mode choice models. Elasticity estimates for built environment suggested that density and diversity variables had stronger effect than design factors. Another inquiry into the influence of land use on mode choice has been conducted for both work and non-work trips in Boston and Hong Kong (Zhang 2004). Models were specified on the basis of discrete choice theory with modal attributes, socioeconomic characteristics and land use measures. The study concluded that travel demand models gave considerable benefit for model estimation process; land use variables were important while other variables were controlled; land use effect on travel mode choice was as strong as driving cost when their elasticity estimates were combined.

After exploring extensive relevant literature, Badoe and Miller (2000) identified disagreement in our current knowledge of the connections between land use and transportation. It was mainly caused, they claimed, by data and methodological limitations including aggregation bias and exclusion of transportation system variables in the model estimation, which lead, in turn, to erroneous results of the model estimation.

More recently, Lee (2006) investigated both correlation and causal relationships between land use and travel behavior. For the purposes, conventional travel demand models were specified for individual mode choice, household trip rates and household vehicle miles of travel. As travel demand models, three important categories of independent variables were included: travel price, socioeconomic characteristics and land use attributes. From the case study of Dallas-Fort Worth metropolitan area, he concluded that land use attributes were statistically significant.

2.2.2.2 Measurement and Unit of Analysis

Empirical models estimated with disaggregate rather than aggregate travel data were well consistent with the theory of economic behavior. They have also improved explanatory and forecasting power, and avoided the aggregation problem in the models (Ben-Akiva and Lerman 1985; Boarnet and Crane 2001a). Modal attributes such as travel time and cost were calibrated for traffic analysis zone (TAZ) usually with the support of regional travel demand model (Cervero 2002; Zhang 2004; Lee 2006). Land use variables were also measured mainly for either TAZ (Cervero 2002; Zhang 2004; Lee 2006) or other geographically predetermined zones such as census tract and zip-code area (Cervero and Kockelman 1997; Kockelman 1997; Boarnet and Sarmiento 1998; Boarnet and Crane 2001a; Litman 2008b).

Kockelman(1997) and Cervero and Kockelman(1997) introduced an innovative way of land use measurement. The 3Ds, density, diversity and design were categorized with many specific land use measures. Density included population and employment density and accessibility; diversity consisted of dissimilarity index, entropy, vertical mixture and so on; design encompassed street measures, pedestrian and cycling provisions, and site design measures. Boarnet and Crane (2001a) gauged land use characteristics with different measurement units for Orange County and Los Angeles travel data: neighborhood level such as one quarter-mile circular area, census block group and tract, and postal code area. It was inferred that the effect of land use variables would depend partly on different geographical scale of measurement.

More recently, a different study design was introduced with more detailed measurement of urban from and travel outcomes (Krizek 2003). For travel behavior outcomes, not only were conventional travel behavior measures, but tour-based variables were also computed including number of tours and number of trips per tour. Urban form measures were computed based on each 150-meter grid cell which formed the whole area of interest. They were then averaged over one quarter mile of walking distance to calculate neighborhood accessibility.

Land use measures computed in these spatial extents inevitably cause spatial aggregation bias or ecological fallacy (Boarnet and Crane 2001a; Krizek 2003). It would not be best if they were quantified on a very detailed level of geographical area such as a residential lot. Not only does this approach need a lot of time and cost, but it also brings about loss of important spatial information. This scale of measurement could not appropriately reflect surrounding context influencing travelers' decision making. Therefore, certain level of geographical unit of analysis should be at least maintained such as census block or one quarter-mile boundary of both trip ends. Location factors in the regional context could also be considered as they had a significant effect on personal travel decision (Handy 1996b; Krizek 2003). However, it is difficult, often impossible, to generate land use variables in a detailed spatial level due to lack of data and measurement tools. They can be accomplished thanks to the availability of parcel-based of land use information and advanced Geographic Information System (GIS) techniques.

2.2.2.3 Causal Relationship and Self-selection

The majority of previous studies have been limited to describe the correlation between land use characteristics and travel outcomes. They were not able to explain causal connections between them (Crane 2000; Boarnet and Crane 2001a; Lee 2006). The issue of causality between land use and travel behavior has recently drawn public attention together with the improvement of modeling travel demand. Although some studies address the issue based on the analytical framework such as travel demand modeling, academic interests have increased in the causal relationship between them and causal notion for the explanatory variables (Lee 2006).

Badoe and Miller (2000) presented a simplified figure illustrating the interactions between land use and travel behavior. Although the seemingly causal relationship was developed not from the empirical analysis, but from an overview of related literature, it suggested that there were a number of connections within the whole structure. In addition, it was recommended that modeling interactions among them should adhere to a comprehensive and integrated perspective.

Bagley and Mokhtarian (2002) have taken advantage of the structural equation modeling (SEM) approach to examine the causal connections between neighborhood type and travel behavior, while including residential and lifestyle attitudes, and socioeconomic characteristics. The results showed that residential attitudes and lifestyle factors had the greatest effect. Residential location did not play a significant role in explaining travel behavior when attitudes and socioeconomic factors were controlled. It was inferred that the relationship between land use and travel behavior was not an outcome of direct causality, but a simple reflection of complicated associations among them. This study is noteworthy in that it examined causal framework and implied multiple causal directions although they were predetermined and assumed (Lee 2006).

Krizek (2003) investigated how urban form changes causally influence travel behavior changes, while considering other variables. An innovative research was designed employing longitudinal data of Puget Sound Transportation Panel survey between 1989 and 1998. It was assumed that movers were in the state of total equilibrium in regard to neighborhood type over a short period; thus preferences could be controlled. It was found that urban form factors significantly influenced the decreases in vehicle miles traveled, person miles traveled, and trips per tour; only neighborhood accessibility significantly increased the number of tours. The study supported the causality between urban form and travel behavior despite its modest impact.

Handy et al. (2005) criticized that previous researches neither analyzed statistical association nor controlled for the effect of self-selection or travel attitudes; therefore, they failed to understand the effect of self-selection and causal relationship between neighborhood land use and travel behavior. The analysis of cross-sectional data showed that both objective and perceived neighborhood factors were not significant when travel attitudes were introduced. The quasi-longitudinal study presented that there was strong evidence of a negative causal relationship between changes in accessibility and changes in driving distance. Therefore, land use policies for higher density, more mixed uses and better accessibility were expected to decrease automobile dependence. The causal connection between built environment and walking behavior was investigated with same quasi-longitudinal design and similar sets of explanatory variables (Handy et al. 2006). They concluded that there was clear evidence, though incomplete, of the causal relationship between them.

Schwanen and Mokhtarian (2005) raised a question about the exogeneity of residential location choice and thus the direct causal relationship between land use and travel behavior. They examined whether the mismatch between a commuter's preferences and living neighborhood conditions encourages residents to travel more than the match between them does. They found that built environment together with travel attitudes has significant impact on the probability of commuting mode choice. However, this investigation considered neither travel price variables nor detailed land use measures in the model estimation.

Lee (2006) claimed that many studies based on assumed causal structure have frequently failed to clearly explain the connections within theoretical framework; therefore, causal relationship between land use and travel behavior need to be carried by observed data. A new method called the directed acyclic graphs (DAGs) was applied to investigate their causation for individual mode choice, number of household trips and household VMT by different trip purposes. Base on the case study of Dallas-Fort Worth Metropolitan Area, population and employment density, and regional accessibility were found to be causally connected to reducing automobile choice for non-work trips. Regional accessibility caused the decrease in automobile trips and VMT. It was suggested that land use densification accompanied by mixed development would be effective to reduce automobile dependence in the region.

People who like to walk or take transit may choose to live in a neighborhood where well designed sidewalks and good transit services are available, while making their attitudes or preferences satisfied. It implies that residents having specific travel preferences are self-selective in the neighborhood in which they live. In terms of the travel behavior model, it is significant because the inclusion of travel attitudes or preferences into the model could change the observed relationship between residential location and travel behavior (Boarnet and Sarmiento 1998; Bagley and Mokhtarian 2002; Handy et al. 2005). It is also important for land use planning and policy in that observed differences of travel outcomes could not be due to land use patterns only, but due to both land use and other factors (Krizek 2003). When it comes to modeling travel demand, residential self-selection process is contrary to the plausible assumption that land use variables causally influence personal travel behavior. Therefore, this potential bias should be properly addressed in the empirical model estimation.

There have been some remarkable endeavors to tackle the self-selection bias. Kitamura et al. (1997) were concerned about whether land use really affects travel behavior, while controlling for other factors. They concluded that total number of trips had stronger and more direct association with attitudes than land use variables. Nonetheless, it has potential problems in model choice and specification without travel price variables. Another effort has been made by using instrumental variables (Boarnet and Sarmiento 1998). They raised a question on the possibility that individuals choose to live in a neighborhood partly because of their travel preferences. Instrumental variables were introduced for replacing land use variables with residential location as a function of individual socio-demographics and location attributes. They found that land use impact would depend partly on different geographical scale of measurement.

Khattak and Rodriguez (2005) studied the role of residential self-selection in increasing alternative mode choice in neo-traditional and conventional neighborhoods in Chapel Hill, NC. As a result, the neo-traditional neighborhoods showed lower external and higher internal trips, higher share of non-driving modes, and fewer VMT after household characteristics and self-selection factors were controlled. However, it did not consider travel price, household income and objective land use measures, which might lead to biased estimates in the empirical models. Handy et al. (2005) claimed that significance of objective and perceived neighborhood measures disappeared after travel attitudes were incorporated in modeling household VMT. However, it is concluded in their later work that built environment was still meaningful for estimating walk and bike trip rates together with travel attitudes and residential preferences (Handy et al. 2006).

Cao et al. (2006) addressed the issues related to the linkage between built environment and walking behavior. Analyzing strolling and shopping trips surveyed at six communities located in Austin, TX, they found that residential self-selection factor significantly affected both types of travels. Neighborhood characteristics were also significant for both types of travels even when residential preference was kept constant.

This study, however, did not consider important variables such as travel cost and objective land use measures.

2.2.2.4 Substitution Effect

Substitution effect primarily concerns about whether automobile mode can be replaced by other modes including transit, walk and bike in the neo-traditional and transit-oriented neighborhoods (Cervero and Radisch 1996; Ewing and Cervero 2001; Krizek 2003). In more detail, studies investigate whether and how people living in much dense, mixed-use and pedestrian-friendly communities are inclined to substitute public and alternative mode trips for driving trips.

Cervero and Radisch (1996) found that the residents living in these areas showed higher number of non-work trips on foot and thus lower number of non-walk trips by automobile. It was argued from the finding that internal walking trips substituted for external automobile trips. A theoretical approach based on the behavioral framework was introduced to examine the effects of different design elements such as grid, traffic calming and mixed and intensive land use. Traffic calming clearly reduced automobile trips, driving choice and VMT; the effects of other elements, however, were not clear (Crane 1996). It was suggested that the elasticity of trip demand by travel mode and purpose, and cross-elasticity among modes would be useful to figure them out.

Handy (1996b) argued that urban form was an important factor in decision to walk. There was clear evidence of replacing driving with walking to the store in the neighborhoods. Even though residents tended to substitute walking trips for driving trips, total amount of savings in travel length would not be large enough to assert automobile dependence to be reduced. In a similar vein, a study found that neighborhood accessibility increased household daily VMT with others being equal. However, there was not conclusive evidence indicating that non-automobile trips were substituted for automobile trips in highly accessible neighborhoods (Krizek 2003).

Recently, Khattak and Rodriguez (2005) examined the substitution of walking trips for driving trips in neo-traditional and conventional neighborhoods. They found that the external trips decreased and internal trips increased in the neo-traditional neighborhood after other factors were controlled. The proportion of non-driving modes became higher and total VMT decreased. As total trips were not significantly different, they concluded substitution effect existed between driving and non-driving modes.

There are several limitations in the researches for examining substitution effect of land use measures. First, most of them have only focused on a certain number of small neighborhoods with different land use patterns. It is also ambiguous to classify the communities into pedestrian- and auto-oriented areas, which makes their external validity questionable. For the regional level, the simulation method based on the estimated models, while considering travel mode and trip rates, can be helpful.

2.2.2.5 Automobile Captivity

Automobile captivity is an outcome caused by excessive automobile dependence. In specific, a person choosing a mode among available choice options does not make use of others except automobile mode due to some reasons. They include transportation system factors such as travel time and cost and transit availability, socioeconomic characteristics such as income and vehicle availability, and land use attributes such as availability of pedestrian and bike road and single-family residence segregated from employment or shopping centers.

It is important to establish reasonable analytical framework and method. In terms of choice set formation, conventional discrete choice models such as multinomial logit and probit models assume that choice options are equally distributed to every choice maker. It does not often make sense; rather, an individual is more likely to make a choice based on different choice set determined by restrictions such as income, attitude and surrounding land use pattern. Shocker et al. (1991) provided precise definitions of latent constructs including universal set, consideration set and choice set. They maintained that individual choice set generation should be specified in the modeling process.

Manski (1977) discussed the decision-making rule and lack of information about choice formation process. A significant contribution was made in the probabilistic choice theory. Two-stage choice process was suggested: choice set generation and choice making based on given set. Gaudry and Dagenais (1979) developed a classical captivity model, called the dogit model that incorporated both captive choice and free choice components.

Two different approaches are generally available to capture choice captivity factors related to land use characteristics (Lee 2006). First method divides individual travel data into captive trips and free choice trips according to some land use conditions (Beimborn et al. 2003). Alternative way is to parameterize captive factors in the choice model. The logit captivity model was first proposed by McFadden (1976), and applied to different travel data collected in many areas (Swait and Ben-Akiva 1986a, 1986b, 1987a, 1987b); then it has been developed to either the probabilistic choice set model (Swait and Ben-Akiva 1986b Ben-Akiva and Boccara 1995; Zhang 2005, 2006; Lee 2006) or the choice set generation model (Swait and Ben-Akiva 1987a, 1987b; Swait 2001; Basar and Bhat 2004).

Zhang (2005) found that land use density and accessibility were significant for increasing travel choices and substituting alternative modes for automobile mode. Another study (Zhang 2006) also confirmed that density, transit access and network connectivity helped reduce the probability of being captive to automobile. Lee (2006) also specified multinomial logit captivity models and found that dominance of residential use at trip origin for driving mode was significant for home-based work trips. Sometimes, evidence of choice captivity is unintentionally observed in the studies of land use impact on mode choice (for example, see the results and conclusion of Schwanen and Mokhtarian (2005)).

2.2.2.6 Application of Empirical Models

Last issue of importance is about how the empirical results can be embodied into the real situation of transportation market. It is interested in the application of the estimated model results of land use impact on travel behavior into the practical transportation planning and forecasting.

It has been argued that the estimators of the conventional travel demand models such as trip generation and mode choice models tended to be underestimated and biased. It was because full array of land use measures has been ignored although disaggregated travel survey data was used in the modeling process (Cervero 2002).

Previous research suggested that elasticity estimates of land use variables represented the degree of connections between land use and transportation in the travel demand models. As they are transferable and applicable from a region to others, land use effects can be addressed with elasticity estimates in metropolitan areas where conventional travel models were used. EPA's Smart Growth Index (SGI) model, for instance, incorporated elasticity values of density, diversity, design and regional accessibility measures. It was also recognized that elasticity of each measure was not substantial; however, their total value was quite substantial (Ewing and Cervero 2001).

Cervero (2006) investigated alternative modeling methods for applying land use effects on travel demand. Two approaches, post-processing and direct modeling were examined. The former incorporated elasticity estimates into the existing travel demand model; the latter, on the other hand, has directly specified travel model for neighborhoods, most of which has been estimated for ridership of transit-oriented development projects. He argued that the efforts were effective and efficient because of inclusion of significant land use effects and reduction of time and cost for model estimation. Those alternative approaches do not substitute for the labored four-step demand models; it supplements the traditional forecasting models.

CHAPTER III

ANALYTICAL FRAMEWORK AND HYPOTHESES

This chapter presents analytical framework, research design and hypotheses. The analytical framework covers travel demand analysis and causal relationship analysis in which the correlations and causalities between land use and travel behavior are of main concerns. It serves as a foundation of setting up research design and hypotheses.

3.1 Analytical Framework

Urban transportation system shows some distinct characteristics as a consequence of its use by individual travelers including residents and visitors. Meyer and Miller (2001) presented six attributes of which four are noteworthy for the study: trip purpose, temporal distribution, spatial placement and modal split distribution.

Trip purposes of passenger transportation have been classified commonly into work, shopping, social or recreation, school and business. Considering a home as a trip end, trips are used to be categorized into five or sometimes fewer groups: home-based work, home-based shop, home-based school, home-based other and nonhome-based trips. Despite challenging alternative approaches, the trip-based model is still more applicable than others. Urban travels have shown temporal distribution over the day. It commonly shows "double peaking" indicating that most work trips occur in the early morning and evening. This trip-making feature affects roadway congestion and transit operation. In addition, every trip has both an origin and a destination that are spatially located; therefore, land use pattern and transportation network layout are associated. Finally, travel modes include driving-alone, shared-ride, transit, bike and walk modes. Conventional transportation system has often ignored alternative modes due to their small share and auto-oriented system (Meyer and Miller 2001).

In general, individual trip-making process involve trip purpose, time of day, origin and destination, travel mode, route from origin to destination, and frequency (Meyer and Miller 2001). When a person starts to make trips, some essential decisions should be made on them. Ideally, each trip-maker takes some important factors into consideration to make effective, efficient and comfortable trips. They include trip-maker's needs, transportation system, socioeconomic characteristics, and land use attributes of trip ends.

Individual travel behavior based on the travel decisions needs to be modeled to explain and forecast those decisions and travel outcomes. In terms of land use impact on travel behavior, modeling travel behavior has additional purposes. They include testing hypotheses set up from the theory, and understanding the causal relationship between them (Lee 2006). To address the issues, two approaches are introduced. One is the framework of travel demand analysis; and another is the framework of causal relationship analysis. They are necessarily connected and complementary with each other.

3.1.1 Travel Demand Analysis

Estimating and forecasting travel demand is one essential step of urban transportation planning process. The urban transportation modeling system (UTMS) is composed of four main steps: trip generation, trip distribution, modal split and trip assignment. However, UTMS has been criticized because it is not based on the theory of travel behavior (Meyer and Miller 2001). [4]

Transportation demand models are employed to examine current system and to forecast the future according to some changes. Three basic assumptions are addressed. First, the important characteristics are specified as observed variables. Second, a presumed functional relationship between the observed variables and the travel outcomes exists. In other words, there is assumed causal connections between them explained by the theory. Last, the functional relationship is essentially consistent for all individuals over time (Meyer and Miller 2001).

Conventional travel demand models are specified, estimated and evaluated based on the theory of consumer behavior. The basic concept is that an individual chooses a combination of goods and services over others for maximizing his or her utility, subject to a budget constraint as follows (Ben-Akiva and Lerman 1985; McCarthy 2001; Meyer and Miller 2001). [5]

---

[4] Many efforts to improve UTMS have been made since the 1970s. As a result, two major developments have been achieved: individual choice or random utility models and activity-based models (Meyer and Miller 2001). The former will be discussed later; however, the latter is beyond the scope of this study.

[5] Discrete choice theory is similar with the economic theory in that the consumer choice for utility maximization is still effective; however, they employ different functional specifications due to discrete dependent variables of the discrete choice models (Ben-Akiva and Lerman 1985; Meyer and Miller 2001).

$$\max U = U(q_1, q_2, \ldots, q_{n-1}, q_n)$$
$$\textit{subject to } I = p_1 q_1 + p_2 q_2 + \ldots + p_{n-1} q_{n-1} + p_n q_n$$

where $q_1, \ldots, q_n$ are the quantities of goods and services; and $p_1, \ldots, p_n$ are the prices of goods and services for a constraint income I.

Even though this illustration implies that the utility is a function of the quantity of goods, transportation services, for example, are a matter of their attributes rather than their quantities. In other words, the consumers of transportation services are more concerned about the properties of the services by which they create their own utilities (Lancaster 1966). The theory suggests that the demand for goods and services is conditional on a range of trip characteristics, attributes of comparably available modes, and the consumer's socioeconomic characteristics.

In a similar vein, most of the estimated models of land use effects on travel behavior outcomes mainly stem from the theory of economic consumer behavior (Domencich and McFadden 1975; Ben-Akiva and Lerman 1985). The travel demand models should thus take into account travel prices, comparative characteristics and socioeconomic attributes. In addition, land use attributes should be taken into the demand models because spatial features of trip ends affect travel decision-making theoretically and empirically.

Consider, for example, the number of household trips under the theory of consumer behavior. Travel is a derived demand, which means that people travel to satisfy the demands for various activities at different destinations. It can be assumed that a household makes a choice of a combination of trips by different modes to maximize a utility function, subject to total budget (Crane 1996). When focusing only on automobile trips, we can define total number of driving trips in a household as a function of trip cost, income, household socioeconomic characteristics and land use measures as follows (Boarnet and Crane 2001a). [6]

$$Auto\ trips = f(p,\ y,\ y^2,\ L;\ S)$$

where p is a vector of relative travel prices such as the generalized cost; y and y^2 are total household income and income squared, respectively; L is a vector of land use characteristics; and S is a vector of socioeconomic attributes.

In recent years, the travel demand models have been improved greatly by incorporating full set of explanatory variables based on the travel demand theory. Individual travel is influenced by travel prices, personal and household socioeconomic characteristics and land use attributes. Traditional demand variables help estimate the short-term impact on travel outcomes; on the other hand, land use measures enable to gauge the long-term effect on travel behavior (Boarnet and Greenwald 2000; Boarnet and Crane 2001a; Cervero 2002; Zhang 2004; Lee 2006).

3.1.2 Causal Relationship Analysis

The travel demand model assumes that there are direct causal relationships between explanatory variables and travel behavior outcomes. These connections depend on the theory and empirical research. This assumption stimulates modeling professionals

[6] There are three different model specification strategies proposed in Boarnet and Crane (2001a) with regard to land use measures. One strategy with minor modifications is employed which is similar to the second proposed model.

to specify and estimate models representing their causal linkage. However, the majority of them only reflect the associative relationships between them (Meyer and Miller 2001).

Studies on the causal relationships between land use and travel behavior addressed the issue of cause-and-effect interactions in terms of design of data collection, variable selection such as attitudes, causal notion and test, and causal structure between explanatory variables (Lee 2006). Most of them have failed to discover the causal linkages beyond the correlations between them. Some shortcomings are inherent in explaining causal connections between them mainly due to assumed causal linkages established by the theory (Boarnet and Crane 2001a; Crane 2000; Lee 2006).

Nevertheless, many significant implications are provided for exploring causal relationships between land use and travel behavior. First, the structural equation modeling approach makes it easier to investigate complex interactions simultaneously between endogenous variables as well as between endogenous and exogenous variables (Bagley and Mokhtarian 2002). Second, the studies introducing longitudinal or quasi-longitudinal data help examine causal connections between them (Krizek 2003; Handy et al. 2005, 2006). These research designs make it possible to address time order criterion for establishing causality (Handy et al. 2005). These investigations can control for travel attitudes and residential preferences as well as other factors. Last, an advanced analytical framework using cross-sectional data can be applied to deal with the issue of causality (Lee 2006). The approach supports the causal structure carried by observed data rather than by the theoretical foundations common in most studies.

A study illustrates complex causal connections among land use measures and travel outcomes as presented in Figure 3.1. Although it is not structured based on empirical studies, the causal relationships seem to be clear among the variables and helpful for further research.

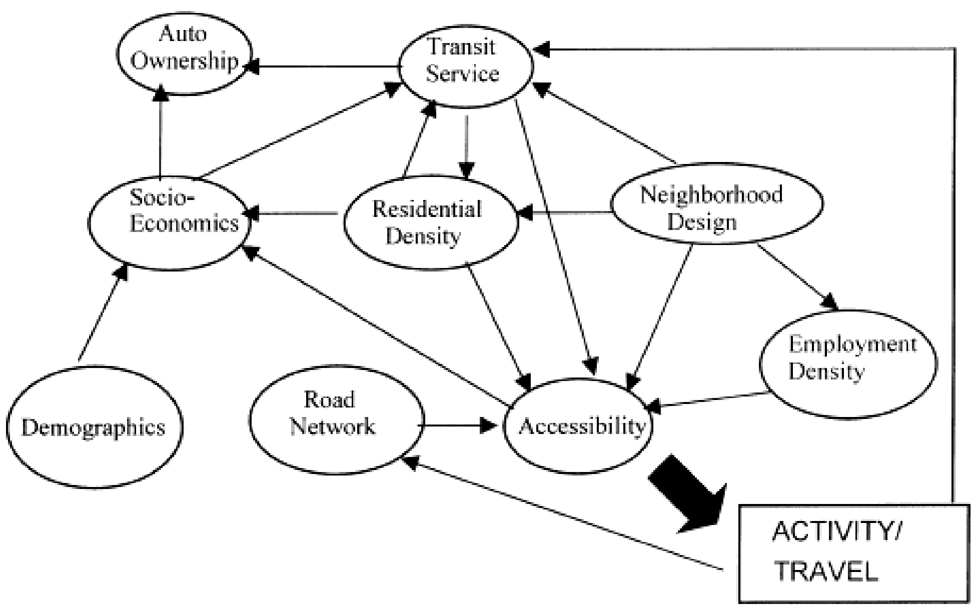


Figure 3.1 The Impact of Land Use on Travel Behavior.
Source: Badoe and Miller (2000).

There are several reasons why it is difficult for the studies to employ either longitudinal or quasi-longitudinal designs. Above all, they require at least two travel surveys in an area for relatively short terms (Krizek 2003), or the information on whether a resident has been moved recently or not (Handy et al. 2005). These efforts are hard to be achieved for the conventional household travel survey for metropolitan regions. Another reason is related with measuring attitudes and residential preferences.

The metropolitan household travel survey does not include related questions because they are not helpful for forecasting regional travel demand. Rather, an additional household travel survey for selected neighborhoods should be conducted to address the issue of self-selection (Handy et al. 2005, 2006). It is criticized that each trip-maker becomes a 'black box' because these unobserved preferences play a role in the middle of the built environment and travel behavior (Lee 2006). The study does not consider the self-selection issue.

To address the issue of causality, two different approaches are introduced: the structural equation modeling (SEM), and the directed acyclic graphs (DAGs). They have their own characteristics for the analysis of causal relationships among the variables which will be explained later. They have in common in many points; thus, they are complementary with each other, not opposites.

The SEM does not indicate single analytical approach; rather, it incorporates a number of modeling frameworks, which allows evaluating the entire models. It is useful for experimental and observational data, and cross-sectional and longitudinal data. The SEM is basically depending on assumed causal structure among variables. Many concerns and questions have been raised on whether it could be used for assessing causal connections among variables. Arguments were made that causal inferences based on the SEM results would be controversial (Thompson 2000; Lee 2006). Despite the concerns, it can be introduced to justify causal inferences as long as certain assumptions on causality are rendered in advance (Pearl 2000; Kaplan 2009). The counterfactual theory with the manipulative perspectives delivers theoretical foundations and methods for examining causal inferences. Consequently, the SEM requires that the variables be scrutinized in terms of manipulation and control (Kaplan 2009).

Studies of land use impacts on travel behavior relying upon the assumptions of causal relationships are not successful in addressing neighborhood self-selection and interdependence of explanatory variables, which leads to biased parameters. They are frequently suffering from lack of valid interpretation and theory of land use and travel behavior interaction. Therefore, their causal relationships should be inferred by observed data unless the theories are useful for explaining those relations (Lee 2006). The DAG approach is not dependent on assumed causality. Rather, it is employed to make clear the causal connections primarily based on observed data. It is intended to handle the independent relations among variables that are established with statistical association. Correlation does not imply causation; however, the statistical associations of independence and dependence based on observed data frequently suggest causal relations among those variables (Cooper 1999).

3.2 Research Design

This research introduces regression methods for modeling travel demand, and SEM and DAG methods for analyzing causality. The regression methods consist of the multinomial logit (MNL) model for individual mode choice, the negative binomial model for household automobile trips, and ordinary least squares (OLS) regression for household total VMT.

Travel behavior can be defined and measured in many ways. Three common travel outcomes are selected: individual mode choice (i.e., driving-alone, shared-ride, transit, and walk and bike), the number of household automobile trips and household total VMT. They are explained and compared with different travel purposes including work trip and non-work trips.

Travel demand is affected by travel price variables such as travel time and cost; individual and household socioeconomic characteristics, including age, sex, auto ownership and availability, household size and income; and land use attributes covering density (i.e., population and employment density), land use diversity (i.e., entropy index and dissimilarity measure), and design factors (i.e., connectivity measure and roadway length). Land use variables are measured in a quarter-mile boundary for each trip end. It intends to reduce the bias caused by spatial aggregation as well as to maintain spatial information affecting trip-maker's decisions.

The entire structure of the causal relationship models include same groups of variables: travel prices, socioeconomic characteristics and land use attributes. As maintained earlier, travel attitudes and residential preferences are not taken into consideration in the model estimation process. In addition, same measures of travel behavior are taken into the analysis: individual mode choice, household automobile trip frequency, and household VMT. In order to compare the results of the SEM with those of the DAG, two mode choice options (automobile vs. non-automobile) are only considered for mode choice analysis. The results are compared with different travel purposes including total trips, home-based work trips and home-based other trips.

In specific, assumed causal inferences are investigated on 1) whether travel prices (exogenous variables) significantly cause each travel outcome (endogenous variable); 2) whether three dimensions of land use (exogenous variables) are causally connected with each travel behavior outcome (endogenous variable); 3) whether socioeconomic characteristics (exogenous variables) causally influence each travel outcome (endogenous variable); and 4) whether either a set of land use measures or socioeconomic variables have indirect causal relations with each travel outcome (endogenous variable) through travel prices (endogenous variables). Overall, model specification strategies vary with different travel outcomes due to different theoretical foundations; however, same specification strategy is applied to different trip purposes.

The DAG approach involves same categories of travel prices, socioeconomic attributes, land use measures and travel behavior outcomes. Some constraints need to be imposed for making the final outputs more reasonable even though the entire causal structure is obtained from observational data. One is that land use measures at trip origin cannot cause those at destination, and vice versa. Another condition is that socioeconomic characteristics cannot be caused by other groups of variables.

3.3 Research Hypotheses

How do land use characteristics affect individual or household travel behavior in a regional context? It is the main research question raised in the study. Based on the research question, a number of hypotheses are identified with regard to land use measures. They are classified into two broad topics: conventional travel demand and causal models. They are also divided into different travel outcomes and travel purposes.

3.3.1 Hypotheses for Travel Demand Models

Hypotheses for travel demand models are set forth for three different travel outcomes. Main hypothesis is that various land use variables have significant effects on travel outcomes, but in different ways. For individual mode choice, it is assumed each trip-maker has four choice options: driving-alone, shared-ride, transit and walk and bike.

3.3.1.1 Individual Mode Choice Models

1) *Population density at both origin and destination is significantly associated with the probability of travel mode choice*. High population density increases the probability of choosing driving-alone mode at origin, and decreases the likelihood of choosing automobile modes (driving-alone and shared-ride) at destination. High population density at origin indicates a single-family residential neighborhood, which promotes driving-alone mode choice. The increase in population density at destination implies mixed land uses, which discourages people to choose automobile modes.

2) *Employment density at both trip ends is significantly correlated with the likelihood of travel mode choice especially for work trips.* Employment density has a positive impact on the probability of non-automobile mode choice (transit and walk and bike) at origin; however, it has a negative effect on automobile choice probability at destination. This density measure is particularly important for home-based work trips.

3) *Dissimilarity measure at trip ends has significant relationship with the probability of travel mode choice for nonwork trips.* Dissimilarity index at origin is positively associated with the chance of choosing alternatives to automobile modes.

4) *Connectivity measure at trip ends is positively associated with an increase in the probability of both non-automobile choice at origin and automobile choice at destination.* Connectivity measure augments the likelihood of choosing transit and walk and bike modes at origin especially for nonwork trips. It also raises the probability of taking driving-alone and shared-ride modes at destination for work trips.

5) *Roadway length variables at both trip ends significantly associated with the likelihood of travel mode choice for nonwork trips.* It is likely to increase not only the probability of choosing alternatives to automobile modes at origin, but it also heightens the chance of automobile mode choice at destination for home-based other trips.

6) *The extended models significantly improve the base models without land use variables for both home-based work and home-based other trips*.

3.3.1.2 Household Automobile Trip Models

1) *Density measures (population and employment density) at origin are significantly associated with household total automobile trips*. It is generally assumed that these measures have negative impact on automobile trip rates as they increase in a residential area.

2) *Entropy measure at origin is significantly associated with household automobile trip rates*. It is argued that residents living in an area with balanced land uses are less likely to make automobile trips.

3) *Design measures (connectivity and roadway length measures) at origin have significant association with household automobile trip rates*. It is thought that

neighborhoods with well-organized networks encourage residents to reduce automobile trips.

4) *The extended models show significant improvement compared with the base models without land use variables for different travel purposes*.

3.3.1.3 Household VMT Models

1) *Density measures (population and employment density) at origin are significantly associated with household total VMT.* Both population and employment density variables have negative effects on household total VMT for total trips. On the other hand, population density is not significant for work trips, and employment density is not significant for nonwork trips.

2) *Entropy index at origin is significantly correlated with household total VMT.* It is assumed that travel distance is shorter in a neighborhood with balanced land uses than in a neighborhood with single residential use.

3) *Design measures (connectivity and roadway length measures) at origin significantly affect a decrease in household total VMT*. Obviously, people living in a neighborhood with well-organized and designed road network reduce their automobile travel distance.

4) *Land use measures in the extended models significantly contribute to model improvement for all travel purposes.*

3.3.2 Hypotheses for Causal Models

Hypotheses for causal models are presented for three travel behavior outcomes. They hold the same hypotheses for the SEM and DAG approaches. Each trip-maker is assumed to have only two choice options: automobile or non-automobile. Main hypothesis is that various land use variables not only have direct causal influences, but also show indirect causal effects on travel outcomes through travel price.

3.3.2.1 Individual Mode Choice Models

1) *Employment density is a direct cause of automobile choice for home-based work trips*. Employment density at both trip ends has a negative impact on the probability of choosing automobile mode. This is related to the result of individual mode choice models.

2) *Dissimilarity measure causally influences the likelihood of choosing automobile mode for home-based other trips*. Land use mix measures at both trip ends have negative effects on the probability of automobile choice.

3) *Employment density and design measures (connectivity and road length) at destination directly cause increases in travel time differential (walking time – driving time)*. At trip destination, automobile access is improved and preferred as employment density and network connectivity and roadway miles increase.

4) Land use measures (population and employment density, dissimilarity index and connectivity and roadway length measures) at origin are direct causes of a reduction in travel time differential. It is maintained that land use measures indirectly cause the likelihood of choosing automobile mode through travel prices. Land use measures generally decrease travel time differential at trip origin.

3.3.2.2 Household Automobile Trip Models

1) *Land use measures at origin are direct causes of household automobile trip frequency.* It is expected that land use measures have direct and negative causal relationship with automobile trip rates.

2) *Land use measures at origin are direct causes of reducing travel cost per trip*. Land use measures are assumed to be direct and negative causes of travel cost.

3.3.2.3 Household VMT Models

1) *Land use measures at origin are direct causes of household total VMT.* They are expected to have negative causal relationships with household total VMT for all travel purposes. But there are several variations: population density is not a cause of household VMT for work trips; and employment density is not a cause of household VMT for other trips.

2) *Land use measures at origin are direct causes of travel cost per mile.* They have negative causal impacts on travel cost for all travel purposes.

CHAPTER IV

MEASUREMENT AND METHODOLOGY

This chapter introduces the study area and data sources, and discusses how the variables of interest are measured and applied for the study. The methods are explained for measuring travel behavior outcomes, travel price variables, socioeconomic characteristics and land use factors. In particular, attention is focused on the measurement of both travel outcomes and land use variables. In addition, research methodologies for analyzing the data are discussed. They consist of negative binomial and multinomial logit model for modeling travel demand, and structural equation modeling (SEM) and directed acyclic graphs (DAGs) for clarifying causal structure between land use and travel behavior.

4.1 Study Area and Data Sources

4.1.1 Study Area

Houston-Galveston Area Council (HGAC) region currently consist of 13 counties containing about 5.87 million residents in 2008, and 145 cities and municipalities. The metropolitan area covers about 12,500 square miles in which urban area totals 1,745 square miles.[7] Historically, 8 counties (Brazoria, Chamber, Fort Bend, Galveston, Harris, Liberty, Montgomery and Waller) have been of main interest in terms

[7] Because the area calculations are based on GIS data that contains total area of each object such as town, city and county, they would be a little different from other sources of information.

of regional transportation planning as 5 counties were joined recently in the region (HGAC 2009a).

This study only focuses on 6 counties in traditional regional transportation planning region: Brazoria, Fort Bend, Galveston, Harris, Montgomery and Waller. Not only is detailed land use data available only for 6 counties, but they also have played significant role in regional planning compared with other two counties (Liberty and Chambers). Figure 4.1 represents the HGAC region and 6 counties of interest. The study area comprises 113 cities including 8 towns and 9 villages and 2,829 traffic analysis zones (TAZs) in total. It totals about 6,729 square miles where urban area amounts to 1,558 square miles (23 percent). About 5.40 million people and 1.95 million households are estimated in the study area where 2.67 million people are working in 2008, while 5.52 million residents, 1.99 million households and 2.70 million jobs are estimated in the 8 counties (HGAC 2009b). Regional economic activities are mainly concentrated in both the City of Houston and the City of Galveston.

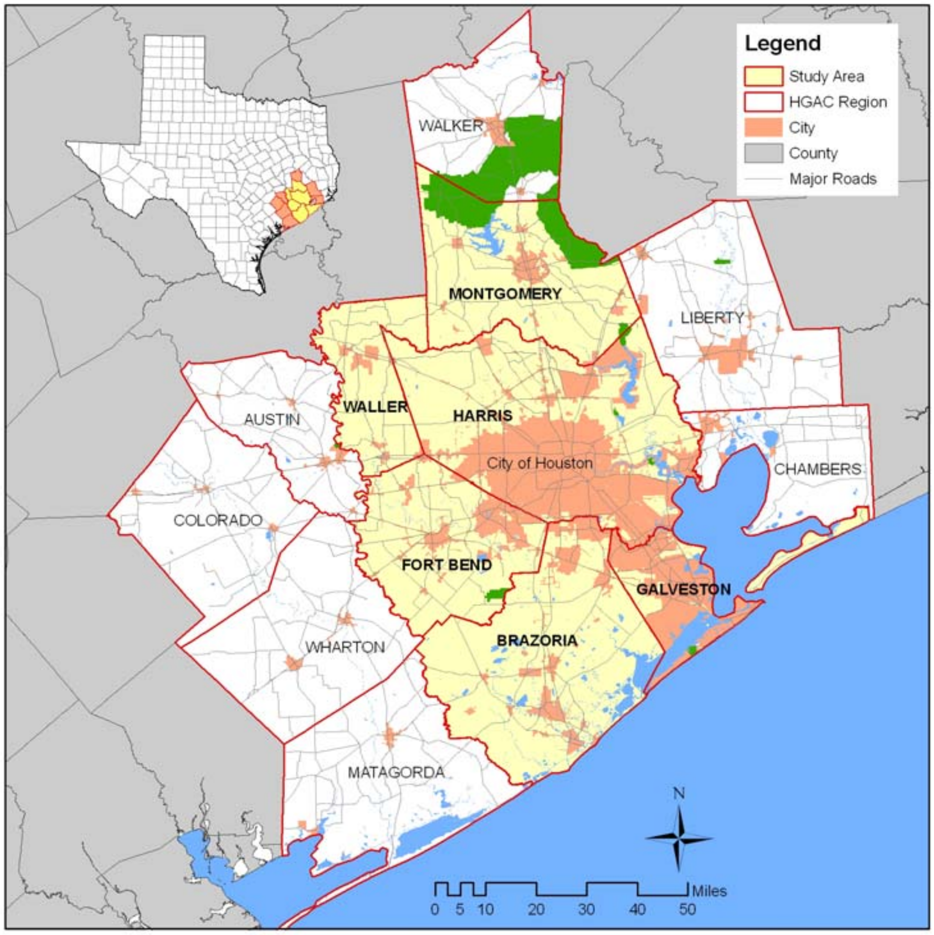


Figure 4.1 Map of the HGAC and the Study Area.

4.1.2 Data Sources

Five different data sources are incorporated in the study as summarized in Table 4.1. The 2007 HGAC Regional Household Activity and Travel Survey data was provided by the Texas Department of Transportation (TxDOT) and the Texas Transportation Institute (TTI). Other data were obtained from the HGAC, the regional association of local governments in the Gulf Coast Planning region of Texas.

Table 4.1 Data Sources and Applications.

Data Source	Applied Measures	Characteristics
2007 HGAC Regional Household Activity and Travel Survey	- Travel outcomes: trip mode, trip frequency and VMT - Individual and household characteristics: sex, age, household size, income residential type, vehicle ownership, bike use - Vehicle information: travel cost	- 84% initially planned total samples for the HGAC survey area - 54,672 trips in 61,731 trip records obtained from 4,775 sampled households - 47,834 trips from 4,367 households collected for 6 counties in total
2007 Land Use GIS Dataset	- Land use measures in 3Ds: population density, employment density, entropy, dissimilarity, connectivity and road length measure	- Parcel-based GIS data prepared by County Appraisal District - 2,074,341 parcels in 6,732 square miles in 6 counties - 66 land use types in 7 major groups
2007 HGAC Regional Travel Model Data	- Travel time by modes - Travel cost by modes	- Travel time, distance and transit fare - Available for travel modes and time of day between TAZs
2008 Population and Employment Forecasts	- Population and employment density	- Forecasts by many spatial units such as city, zip code and census tract - Available from 2005 to 2040 on a yearly basis
2007 STAR Map	- Design measures: connectivity and road length per 1,000 ft^2	- Trademark for the Southeast Texas Addressing and Referencing Map - Including addresses, street name and types and spatial information

First of all, the HGAC Regional Household Activity and Travel Survey intended to obtain the information on both individual and household travel characteristics in the metropolitan region. The survey was conducted for conventional 8 counties of the HGAC region including Brazoria, Chamber, Fort Bend, Galveston, Harris, Liberty, Montgomery and Waller counties. The random stratified household sampling method was introduced for the survey. The survey implementation was composed of three stages. First, randomly selected households were asked to participate in the survey by telephone. If they agree, a packet of household activity and travel diary was sent to the household via mail. When the survey was collected after it was recorded for all household members, the survey data were retrieved by telephone.

Overall, four types of information were collected from the survey: household and individual characteristics, the information on vehicles owned by each household, and the information on every trip and activity made by each individual over 5 years old. The number of sampled households required for the HGAC survey area was 5,700 in total, which were randomly stratified by household size, the number of workers and household income. Due to incomplete survey at the time of the study, the information on 4,775 households (84%), 13,893 people and 54,672 trips was collected. Table 4.2 presents how the survey data has been processed while taking the study area and objectives, and trip purposes into consideration. Two datasets are finally prepared for different travel models. One is for individual mode choice models for 6,239 HBW trips and 10,413 HBO trips. Another dataset is prepared for both household trip generation and VMT models. They are estimated for 6,156 HBW trips by 2,539 households, 14,305 HBO trips by 3,461 households, and 29,858 total trips by 3,976 households.

Note in Table 4.2 that the number of trips and households are different between the mode choice models and other household travel models. The mode choice models are estimated only for HBW and HBO trips. They include return trips to home whose modes are not optional but significantly depend on the mode choice of departure trips from home. In terms of the theory of economic behavior, the inclusion of return trips results in biased estimation therefore, they are all removed from the final datasets. Many distinct trip purposes are included in the HBO trips including school, pick-up and drop-off, personal, social and recreation, and shopping. The study pays main attention to personal, social and recreation and shopping trips because other trips are required for specific times at specific locations which are not appropriate for the study objectives.

Table 4.2 Result of Processing the Travel Survey Data for the Study Area.

Trip Purpose	Surveyed Data		Arranged Data		Studied Data	
	Trips	Households	Trips	Households	Trips	Households
Total trips	47,834	4,367	42,275	4,170	29,729 (29,858) [2)]	4,093 (3,976)
HBW trips [1)]	7,115	2,817	6,558	2,614	6,239 (6,156)	2,614 (2,539)
HBO trips [1)]	25,796	3,917	22,640	3,665	10,413 (14,305)	3,200 (3,461)
NHB trips [1) 3)]	14,923	3,189	13,077	3,010	13,077 (9,397)	3,010 (2,778)

Note: 1) Home-based work (HBW); home-based other (HBO), and non-home based (NHB) trips.
2) Values are related to automobile modes (driving-alone and shared-ride). They are used in the household automobile trip generation model and household VMT model.
3) NHB trips are not analyzed in this study. So studied data are same as arranged data.

The 2007 land use GIS data which are made up of lots of parcels was obtained from the HGAC. The parcel-based land use GIS data has been prepared by each county appraisal distract for assessing property tax. They have been incorporated by the HGAC on a yearly basis, and used for forecasting socio-demographic data and providing regional GIS services. For 6 counties of the study area, over 2 million parcels in 6,729 square miles of total area are available. They are classified into 66 specific land use types along with specific spatial information. The data is useful for measuring dimensions of land use characteristics within specific boundaries.

Travel time and distance data extracted from the 2007 HGAC regional travel model were also obtained from the HGAC. The datasets were introduced to compute travel time and automobile operating cost between each pair of 2829 TAZs by different travel modes.The 2008 population and employment forecast data were downloaded from the HGAC website (HGAC 2009b). The number of population, employment and households were estimated in different spatial levels including counties, TAZs and census tracts. TAZ-level forecast data was used to measure the population and employment density in a quarter-mile boundary of every trip end. In addition, the 2007 STAR Map in GIS format was employed to measure land use design variables. It contains over 1.7 million address points, roadways, street names and types and other information (HGAC 2009c).

4.2 Variable Measurement

4.2.1 Travel Behavior Outcomes

The study employs three general measures of travel behavior: individual mode choice, household automobile trip generation and household total VMT. Table 4.3 presents travel outcome, data type and operational definition for this study.

Four choice options are taken into the multinomial logit (MNL) choice models: driving-alone, shred-ride, transit and walk and bike. According to the trip records in the HGAC household travel survey, driving-alone mode is defined as private vehicle and motorcycle drivers with no passenger. Shared-ride mode is determined as automobile and motorcycle drivers with two and more people, and auto and motorcycle passengers.

These modes exclude commercial vehicle drivers and passengers. Bus and school bus passengers are identified as transit users. Walk and bike modes indicate walk and bike choice in the choice set formation. Only two choice options, automobile vs. non-automobile are considered in the causal modeling.

Table 4.3 Defining Travel Behavior Outcomes.

Travel Outcome	Data Type	Operational Definition
Travel mode	Discrete	Four mode options considered for mode choice models - Driving-alone: auto/van/truck driver and no. people = 1 motorcycle and no. people = 1 - Shared-ride: auto/van/truck driver and no. people ≥ 2 auto/van/truck passenger carpool driver/passenger vanpool driver/passenger motorcycle and no. people ≥ 2 - Transit: bus and school bus passenger - Walk/bike: walk and bicycle mode users
Automobile trip frequency	Count	Total number of automobile trips made by each household regardless of the number of people in the vehicle Automobile trips include driving-alone and shared-ride trips
VMT	Continuous	Total vehicle miles traveled by each household Sum of vehicle miles traveled of every trips made by every member in a household regardless of the number of people in the vehicle

Individual trips are grouped into three different trip purposes: home-based work (HBW), home-based other (HBO) and non-home based (NHB) trips. Trip distribution by travel mode and purpose is shown in Table 4.4 based on the survey result. The table indicates that driving-alone is the dominant choice for HBW trips, and automobile mode covers about 97 % and 93 % for HBO and NHB trips, respectively.

Table 4.4 Trip Distribution by Travel Mode and Purpose in the Study Area.

Travel mode	HBW trips	HBO trips	NHB trips
Driving-alone	5,538 (88.8) [1)]	4,859 (46.7)	6,239 (48.4)
Shared-ride	639 (10.2)	5,293 (50.8)	5,819 (44.5)
Transit	21 (0.3)	34 (0.3)	578 (4.4)
Walk/bike	41 (0.7)	227 (2.2)	351 (2.7)
Total trips	6,239 (100)	10,413 (100)	13,077 (100)

Note: 1) Values are the percentage of trips.

Household automobile trip frequency focuses on trips made by vehicles not by trip-makers. For example, the number of automobile trips is equal to one if two people share a ride in an automobile. It is necessary to remove duplicate trips to count household automobile trips. Household total VMT is calculated as the sum of every automobile trip distance between origin and destination of each trip in a household. Travel distance indicates the shortest network distance along major thoroughfares. It is available from the matrix skim data of the 2007 HGAC regional travel model. It is also essential to eliminate duplicate trips to measure household VMT. Both travel outcomes are summarized by travel purposes in Table 4.5.

Table 4.5 Household Trip Frequency and Total VMT by Trip Purpose in the Study Area.

Travel mode	Total trips	Total home-based trips	HBW trips	HBO trips
Number of households	3,976	3,973	2,539	3,461
Number of auto trips	29,858 (7.5) [1)]	20,461 (5.2)	6,156 (2.4)	14,305 (4.1)
Total VMT	372,321 (93.6) [1)]	273,790 (68.9)	137,843 (54.3)	135,947 (39.3)

Note: 1) Automobile trips per household and VMT per household.

4.2.2 Travel Time and Cost

Travel prices including travel time and cost are very important factors in explaining individual and household travel behavior. Travel times between TAZs for automobile mode are available directly from the HGAC regional travel model skim data in the matrix format. They are organized into two different times of day, i.e. peak periods and off-peak periods. The peak periods are designated as 6:00 to 9:00 A.M. and 4:00 to 7:00 P.M. (Meyer and Miller 2001). While considering both free (no toll) and paid (toll eligible) travel times, travel times are arranged by many specific modes, i.e. driving-alone, two people shared-ride, three people shared-ride and four and more people shared ride. An example of mode specific travel time is driving-alone paid travel time between traffic zones for peak period. It should be noted that travel distances between traffic zones are also arranged in the same way.

Travel times between TAZs for transit mode can be obtained from the skim data of the HGAC regional travel model. Similar to the automobile travel times, they are arranged by two different times of day. However, they consider two specific modes (local and premium buses) and two access modes (drive and walk access to bus). For instance, travel time skim data for drive to local bus for peak period is available for transit mode. Because two specific buses and access modes cannot be clearly identified based on the travel survey, local bus and driving access mode are preferred to premium bus and walk access mode, respectively. It should be also mentioned that transit travel times are estimated using the association between existing times and distances if they are not available in the skim data due to the absence of transit routes. This process was done for generating alternative specific transit time variable for the mode choice models. In addition, travel times for walk and bike mode are computed using Equation 4.1 proposed by Zhang (2004) as follows.

$$WBT_i = 1/[1 + |Age_i - 30|/30] \times Distance_i / Speed \qquad (4.1)$$

where WBT_i indicates walk and bike travel time for a trip *i* and Age_i and $Distance_i$ are age of an individual making trip *i* and travel distance of trip *i*, respectively. The speed is assumed to 3 miles and 9 miles per hour for walk and bike modes, respectively.

Contrary to travel times between traffic zones, automobile travel costs are not available. The study employs the equation as follows for calculation automobile travel costs.[8]

$$Cost_i = [Gas\ price + Mainntence + Tires] \times Distance_i \qquad (4.2)$$

where $Cost_i$ indicates cost spent for a trip *i*. *Gas price* ($ per mile) is calculated as price per gallon[9] multiplied by the inverse of miles per gallon (MPG). Vehicle fuel efficiency or MPG considers a variety of vehicle years, models and makers and types since 1984 (U.S. DOE and U.S. EPA 2009). *Maintenance* and *Tires* are maintenance and tire depreciation costs.[10]

It should be noted that some shortcomings lie in the measurement of the travel time and cost. First, vehicle travel times between TAZs are ideally based on the amount

[8] Travel cost examined in this study represents short-term operating costs. Generalized cost is a broader concept of travel cost measurement. For instance, it is defined in the Dallas-Fort Worth Regional Travel Model as following: automobile operating cost + (value of time × travel time) (Lee 2006).

[9] 2008 average gas price per gallon in Texas is $ 3.283 per gallon for regular gasoline and $ 3.942 per gallon for diesel (AAA 2008).

[10] 2008 average maintenance and tires costs are 4.57 cents per mile and 0.72 cents per mile, respectively (AAA 2008).

of average time spent on the major roadways. Although the estimated times represents the reality, they do not capture actual travel times not only because they fail to consider minor and local network, but also because they do not include access time and frequent stops on the way. Even travel times within zones are not examined, while assuming zero. In addition, transit travel times are estimated in some cases due to unavailable operations of the public transportation.

Why are interzonal travel times used instead of reported travel times in the travel survey? It is because modal attributes are so different from each other that it is not reasonable to assume an identical coefficient for four different mode options. Similarly, the alternative specific model specification is prioritized in the travel demand model, which needs to compute the travel times of different mode choices for every trip. Similar weaknesses are also inherent in the travel cost and distance measurement. Travel costs stand for only short-term operating costs which do not take parking costs and tolls into account as the information is neither available nor reliable in the travel survey. The issues need to be addressed to collect more reliable information in the household travel survey.

4.2.3 Socioeconomic Characteristics

Socioeconomic characteristics play a major role in individual and household travel models. Seven socioeconomic factors are explored in the study: sex, age, bike use, residential type, household size, vehicle ownership and household income. Details of the variable measurement are presented in Table 4.6. Note that both age and household income that are thought to be continuous are grouped in interval scales. Bike use, household size and vehicle ownership are aggregated at the upper ends so that they have enough observations in the category.

Table 4.6 Measurement of Socioeconomic Characteristics.

Variable	Measurement
Sex	Female = 1, male = 0
Age	Below 10 = 0, teens = 1, twenties = 2, thirties = 3, forties = 4, fifties = 5, sixties = 6, 70 and more = 7
Bike use	Number of days person rode bike in last seven days No bike use = 0, 1-2 uses = 1, over 2 uses = 2
Residential type	Detached single-family residence = 1, others = 0
Household size	Number of people living in each household 1 person = 1, 2 people = 2, ···, 6 and more people = 6
Vehicle ownership	Number of vehicles available for each household 0 vehicle = 0, 1 vehicle = 1, ···, 4 and more vehicles = 4
Household income	Combined annual income of all household members Fifteen income brackets

4.2.4 Land Use Measures

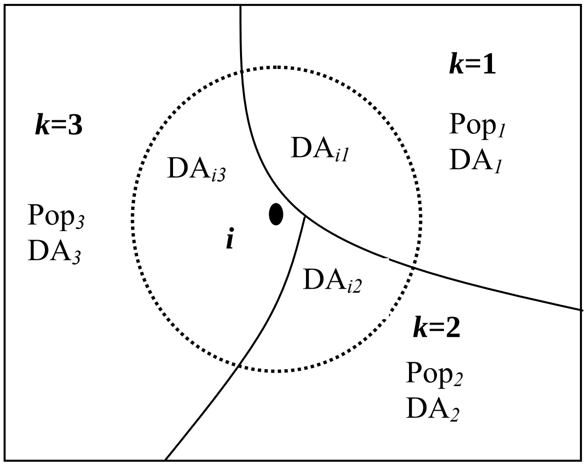
Three dimensions of land use characteristics are examined: density, diversity and design. Two land use measures in each dimension are computed at every trip origin and destination. Density includes population and employment density; diversity encompasses entropy and dissimilarity index; design covers connectivity and road length measures. Thanks to the detailed parcel based land use data and advanced geographic information system (GIS), it is possible to measure land use variables in a quarter-mile radius of trip ends.

4.2.4.1 Density Measures

Population density is generally defined as the number of people living in a certain area divided by total area. The study employs the concepts of net population density as expressed in Equation 4.3. The process of computing the population density is as follows. First, each buffer area of either trip origin or destination is divided into seven different land use types: 1) residential, 2) commercial and industrial, 3) agricultural and farm ranch, 4) park and recreational, 5) school and public, 6) road and transportation, and 7) vacant and other uses. Then, total area of each land use types within a buffer area is computed. Third, the number of people in each buffer area is estimated using TAZ-level population and buffer-level developed area. The basic idea is that the population is homogenously distributed throughout the residential area within each TAZ. Hence, a buffer area spreading over several TAZs takes the number of residents from each TAZ according to the ratio of the residential area dissected by the buffer. Last, the estimated population is normalized by total developed area (residential, commercial and industrial, school and public, and roads and transportation area) in the buffer. An example is illustrated in Figure 4.2.

$$PDR_i = \sum_{k=1}^{n}\left(Pop_k \times \frac{RA_{ik}}{RA_k}\right) \Bigg/ \sum_{k=1}^{n} DA_k \tag{4.3}$$

where PDR_i indicates population density (per acre) within a quarter-mile buffer i, and k is a set of TAZs dissected by the buffer i. Pop_k is total population of zone k, and RA_k, RA_{ik} and DA_k represent total residential area of zone k, and total residential area of zone k included within a buffer i, and total developed area of zone k, respectively.



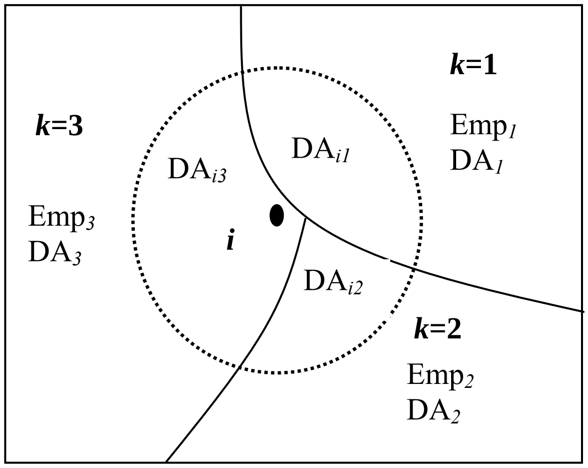
$$PDR_i = \frac{Pop_1 \frac{DA_{i1}}{DA_1} + Pop_2 \frac{DA_{i2}}{DA_2} + Pop_3 \frac{DA_{i3}}{DA_3}}{(DA_1 + DA_2 + DA_3)}$$

Figure 4.2 Measurement of Population Density.

Employment density commonly indicates the number of people working in a certain spatial extent divided by total area. The process of computing the employment density is similar to that of the population density. Equation 4.4 and Figure 4.3 show how the employment density in a buffer area is computed in the study.

$$EDR_i = \sum_{k=1}^{n} \left(Emp_k \times \frac{EA_{ik}}{EA_k} \right) \Bigg/ \sum_{k=1}^{n} DA_k \qquad (4.4)$$

where EDR_i is the employment density (per acre) within a quarter-mile buffer area i, and k is a set of TAZs dissected by the buffer i. Emp_k is the number of employment of zone k, and EA_k, EA_{ik} and DA_k are total commercial and industrial area of zone k, total commercial and industrial area of zone k contained in a buffer area i, and total developed area of zone k, respectively.



$$EDR_i = \frac{Emp_1 \frac{DA_{i1}}{DA_1} + Emp_2 \frac{DA_{i2}}{DA_2} + Emp_3 \frac{DA_{i3}}{DA_3}}{(DA_1 + DA_2 + DA_3)}$$

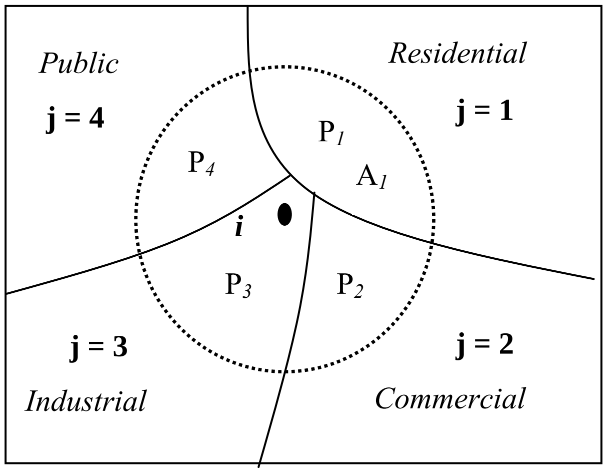
Figure 4.3 Measurement of Employment Density.

4.2.4.2 Diversity Measures

The entropy measure was originally developed for quantifying the energy state in a system and used for gauging how different gases in a system are mixed (Kockelman 1997). This index measures the degree of land use balance or land use heterogeneity (Cervero ana Kockelman 1997). Equation 4.5 presents how the entropy index is computed in which it is normalized with the natural logarithm of the number of land use types. Seven land use types are considered including both developed and undeveloped land uses: residential, commercial and industrial, agricultural and farm ranch, park and recreational, school and public, road and transportation, and vacant and other uses. The values vary between 0 and 1 in which 1 represents perfect balance among different land uses. An example in Figure 4.4 illustrates how to calculate the entropy index in a quarter-mile buffer.

$$EI_i = -\left(\sum_{j=1}^{J} (P_j \cdot \ln P_j) \right) / \ln J \qquad (4.5)$$

where EI_i indicates the entropy index within a quarter-mile buffer area i. P_j is the proportion of a type of land use j, and J is total number of land use types considered.



Entropy index of buffer i (EI_i)

$$= \frac{-\left(P_1 \ln P_1 + P_2 \ln P_2 + P_3 \ln P_3 + P_4 \ln P_4\right)}{\ln 4}$$

conditional on $P_1 + P_2 + P_3 + P_4 = 1$.

Figure 4.4 Measurement of Entropy Index.

The dissimilarity index quantifies how well a place is mixed with its neighboring land uses within a certain area. It focuses on measuring the degree of land use mix in an area. Mean dissimilarity index is introduced in the study as illustrated in Equation 4.6 and Figure 4.5. Seven types of land uses are considered for computing the index. The HGAC region is divided by grids that are 100 feet high and 100 feet wide, resulting in over 80 million cells throughout the region.

A GIS model is created to compute the dissimilarity index (see A1 in Appendix). The computational process is as follows. First, the raster dataset for the HGAC region is prepared as described before. The data is reclassified according to each type of land use, resulting in seven different raster datasets. Each cell in a dataset has 1 value if its land use is same as designated land use type of the dataset. For example, if one out of seven different raster datasets is designated as residential use, then value 1 is assigned to a cell in the dataset if it is residential use; otherwise, value 0 is assigned. Then, the neighborhood function for computing focal statistics is working as illustrated in Figure 4.5. A neighborhood window with five times five cells moves throughout the HGAC region to calculate central grid values. As a result, specific values are assigned to every grid in the raster dataset. After this operation is repeated for seven raster datasets created before, the neighborhood function works for summing up the values in the same cells over six datasets except for a dataset having same designated value in the second step. For instance, if the residential use is designated for a raster dataset, the neighborhood function works only for other six raster datasets to sum up the cell values in the same position. Finally, mean dissimilarity index is computed in a quarter-mile buffer area as done in Figure 4.5.

$$DI_i = \frac{1}{K}\sum_{k=1}^{K}\sum_{j=1}^{J}\frac{D_{jk}}{J} \tag{4.6}$$

where DI_i represents mean dissimilarity index. K is the number of grids in a quarter-mile buffer area, and J is a constant, 24 in this analysis indicating the number of adjacent cells within the neighborhood window (5 by 5 cells). In addition, D_{jk} is a dummy variable for the central grid j within a group of cells k. Its value is 1 if the land use type of the central cell j is different from that of an adjacent cell and 0 otherwise.

C	C	C	R	R	R	R
C	C	R	R	R	C	R
R	R	**C**	**R**	**R**	C	C
I	I	**C**	**R**	**C**	C	C
I	I	**I**	**C**	**R**	R	R
I	I	I	C	C	R	R
P	P	I	C	C	R	R

- Central grid within a square consisting of 3 × 3 grids has 5/8 point because five cells out of eight adjacent to the center are different.
- Mean dissimilarity index of the 3 × 3 square is 0.583.
- Mean dissimilarity index
  = sum of individual index divided by the number of grids within a spatial boundary
  = $\{(6 + 3 + 4 + 6 + 5 + 5 + 4 + 4 + 5) / 8\} / 9 = 0.583$

Figure 4.5 Measurement of Dissimilarity Index.

4.2.4.3 Design Measures

In the literature of the connection between land use and travel behavior, land use design measures generally pay attention to neighborhood street patterns, site development patterns and provision of non-motorized transportation facilities. Two design measures related to street pattern in a quarter-mile buffer area are considered in this study: connectivity and road length measure. Because road network and other amenities for bikers and pedestrians are not available throughout the region, design measures for alternative travel modes are not considered.

Connectivity measure, also called internal connectivity is defined as the number of intersections divided by total number of intersections and dead ends within a certain spatial boundary (Knaap et al. 2007; Song and Knaap 2004). The intersection encompasses 3-way junction, 4-way junction and other types of crossroads. In addition, road length measure examines how long the road network is spread over a buffer area. It is quantified by the sum of roadway miles divided by total area in which roadway length is normalized with buffer area..

Descriptive statistics of explanatory variables for individual trips and for household trips are summarized in Table 4.7 and Table 4.8, respectively.

Table 4.7 Descriptive Statistics for Individual Trips by Trip Purpose.

	HBW Trips		HBO Trips	
Explanatory variable	Mean	Std. Dev.	Mean	Std. Dev.
Travel time (driving-alone)	13.641	11.809	6.996	8.440
Travel time (shared-ride)	24.469	17.132	12.148	11.914
Travel time (transit)	80.444	50.438	50.672	41.532
Travel time (walk/bike)	95.522	75.998	36.637	42.833
Travel cost (driving-alone)	4.956	3.785	2.122	2.371
Travel cost (shared-ride)	2.467	1.893	0.986	1.142
Travel time differential	81.881	65.611	29.642	35.542
Sex (female = 1)	0.436	0.496	0.559	0.497
Age (8 categories)	4.220	1.453	4.344	2.102
Bike use (3 categories)	0.101	0.372	0.185	0.517
Residential type (single-family = 1)	0.916	0.277	0.927	0.261
Household size (6 categories)	3.365	1.335	3.314	1.371
Vehicle ownership (5 categories)	2.558	0.919	2.356	0.900
Total household income (15 groups)	9.329	3.243	8.753	3.547
Population density (per acre) at O	8.516	4.978	8.272	4.897
Population density (per acre) at D	4.817	5.151	5.195	4.828
Employment density (per acre) at O	1.747	3.940	1.777	4.351
Employment density (per acre) at D	27.054	82.382	9.324	27.457
Entropy index at O	0.600	0.138	0.597	0.134
Entropy index at D	0.653	0.140	0.672	0.127
Dissimilarity index at O	0.579	0.129	0.577	0.125
Dissimilarity index at D	0.657	0.129	0.676	0.110
Connectivity at O	0.740	0.169	0.740	0.168
Connectivity at D	0.788	0.174	0.789	0.161
Road length per 1000 ft^2 at O	3.413	1.223	3.321	1.203
Road length per 1000 ft^2 at D	3.880	1.722	3.705	1.587
Sample size	6,239		10413	

Table 4.8 Descriptive Statistics for Household Trips by Trip Purpose.

	Total Trips		HBW Trips		HBO Trips	
Explanatory variable	Mean	SD	Mean	SD	Mean	SD
Travel cost per trip	3.074	2.270	5.188	3.533	2.313	2.136
Travel cost per mile	0.223	0.041	0.226	0.057	0.227	0.067
Household automobile trips	7.510	4.831	2.425	1.336	4.133	2.713
Residential type (single-family = 1)	0.903	0.296	0.906	0.292	0.912	0.284
Household size (6 categories)	2.979	1.348	3.181	1.350	3.047	1.348
Vehicle ownership (5 categories)	2.231	0.880	2.410	0.881	2.262	0.878
Total household income (15 groups)	8.499	3.479	9.167	3.298	8.570	3.491
Population density (per acre) at O	7.465	4.429	7.368	4.300	7.396	4.344
Employment density (per acre) at O	1.978	4.352	1.854	3.965	1.987	4.361
Entropy index at O	0.601	0.138	0.599	0.139	0.599	0.137
Connectivity at O	0.742	0.170	0.738	0.172	0.741	0.170
Road length per 1000 ft^2 at O	3.390	1.240	3.398	1.238	3.371	1.230
Sample size	3976		2539		3461	

4.3 Research Methodology

4.3.1 Negative Binomial Model

The negative binomial regression model is introduced to examine the relationship between land use and household automobile trip frequency, while controlling for trip cost and household socioeconomic variables. In general, the number of automobile trips is skewed to the right; therefore, the Poisson and the negative binomial models are generally employed to estimate the count data.[11] The Poisson model assumes the dependent variable shows Poisson distribution. A popular link function is the log link, which becomes the Poisson loglinear model (Agresti 2007). The probability function for the dependent variable is expressed as follows (Simonoff 2003; Cao et al. 2006).

$$P(Y = y) = \frac{e^{-\lambda}\lambda^{y}}{y!}, \quad y = 0,1,2,\cdots \tag{4.7}$$

[11] Other models such as the zero-inflated Poisson model are also applicable to the count data according to the assumptions and the observational attributes.

where λ is the mean or expected number of frequency of the dependent variable. The mean is expressed in the Poisson loglinear model while satisfying an exponential relationship.

$$\begin{aligned} \log\lambda &= \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n, \\ \lambda &= \exp(\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n) \end{aligned} \tag{4.8}$$

The Poisson model assumes equal mean and variance of the dependent variable. However, there are often heterogeneous Poisson distributions in the population. It causes the variance to be larger than the mean, termed overdispersion. The negative binomial model is proposed to address the overdispersion problem. An unobserved effect is included as follows (Cao et al. 2006).

$$\lambda = \exp(\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + \varepsilon) \tag{4.9}$$

where ε is assumed to have a one-parameter gamma distribution with mean and variance are 1 and α, respectively. The probability function can be described as follows.

$$P(Y = y) = \frac{\Gamma(y+\nu)}{y!\Gamma(\nu)}\left(\frac{\nu}{\nu+\lambda}\right)^{\nu}\left(\frac{\lambda}{\nu+\lambda}\right)^{y}, \quad y = 0, 1, 2, \cdots \tag{4.10}$$

where ν is defined as α^{-1} and Γ indicates the gamma function.

An important attribute of the negative binomial model is that the mean of the dependent variable is λ and the variance is equal to $\lambda(1+\alpha\lambda)$ where α is a dispersion parameter. This type of model with a quadratic influence of variance is called a type 2 negative binomial model. The Poisson model is a special case of the negative binomial model in which the variance is close to λ as the dispersion parameter becomes zero (Simonoff 2003; Cao et al. 2006; Agresti 2007).

4.3.2 Multinomial Logit Model

Regional travel demand is represented as the aggregation of individual travel decisions in a regional context. In terms of individual choice process, Ben-Akiva and Lerman (1985) proposed a series of decision-making steps. Fist, the choice problem is defined, and a set of available choice options are determined; the characteristics of the alternatives are compared and assessed; a choice of an alternative is finally made based on a decision rule. However, the course of actions may not be applied to all decision-making procedure.

It is necessary to define four basic elements to address the individual choice process: the decision maker, the alternatives, the attributes of alternatives and the decision rule (Ben-Akiva and Lerman 1985). The decision makers encounter different choice situation and have different tastes if their characteristics are different. It suggests that the choice model be estimated at individual level and take their differences into account. A set of alternatives considered by the decision maker's environment is called the consideration choice set. The attractiveness of a choice option is assessed by each trip maker with its attribute values where the uncertainty of the attributes can also be considered. They are either generic or alternative specific. It is important to identify policy-related variables because the choice models are basically intended to appraise the impact of policy changes. It is assumed the decision makers are rational, and their decision-making process is both consistent and transitive. Among decision rules, attention is focused on the utility maximization rule in the study. It implies that an individual has a utility function and make trade-offs among the characteristics of alternatives, and make a choice satisfying the highest utility (Ben-Akiva and Lerman 1985; Koppelman and Bhat 2006).

Due to lack of information on the internal decision process and the perception of alternative, individual choice behavior can be explained based on random utility or probabilistic choice theory. According to the theory, the utility function of an alternative for each decision maker is composed of two parts. One is observed component of the utility, called the deterministic or systematic portion; another is unknown component of the utility, called random portion. The utility function is represented as

$$U_{it} = V_{it} + \varepsilon_{it} \tag{4.11}$$

where U_{it} is the true utility of an alternative i for a decision make t, V_{it} and ε_{ij} indicate the deterministic or systematic portion and the random or error portion of utility, respectively.

It is agreed that the deterministic utility is determined by a function of the attributes of the alternative and the chooser's characteristics. This study employs modal attributes such as travel time and cost, individual and household characteristics, and land use attributes surrounding trip origin and destination. The assumptions of the distribution of the error terms determine different mathematical formulation of the choice models. If the error terms are assumed to be normally distributed, for example, the multinomial probit (MNP) choice model is formulated.

Error terms in the utility function are unobserved and random. Two general assumptions of the error component lead to the multinomial logit (MNL) model. One is that the error terms show the extreme value type I or Gumbel distribution. It shapes the mathematical form of the MNL model. Another assumption is that the error terms are identically and independently distributed (IID) across choice alternatives and observations. It implies that there should not be any correlation between the error components of choice options and individuals in the model (Ben-Akiva and Lerman 1985; Meyer and Miller 2001; Koppelman and Bhat 2006; Lee 2006). The MNL model enables us to compute the probability of choosing each alternative using a function of the systematic component (Ben-Akiva and Lerman 1985; Koppelman and Bhat 2006). The probability of choosing an alternative is represented in the MNL model as

$$P(i) = \frac{e^{V_i}}{\sum_{j=1}^{J} e^{V_j}} \tag{4.12}$$

where *P(i)* is the probability of choosing an alternative *i*, and V_j indicates the systematic portion of the utility of alternative *j*.

The probability of choosing an alternative has the S shape as a function of its utility while other utilities are kept constant. It means when the utility of an option is similar to the combined utility of others, the probability of choosing the alternative increases largely with a small increase in its utility. The probabilities of alternative choices rely not on the actual utility values but on the differences in the deterministic utilities of the alternatives. On the other hand, the MNL model has the fundamental property of the independence from irrelevant alternatives (IIA). It indicates that the ratio of the probabilities of making two choices depend only on their attributes; it is independent of the existence of any other alternative. It leads to overestimating the probabilities of choosing similar alternatives and also underestimating the chances of choosing the distinct options. Despite the advantages in model formulation and application, the MNL model is exposed to criticism as it may not appropriately explain the choice behavior (Ben-Akiva and Lerman 1985; Meyer and Miller 2001; Koppelman and Bhat 2006; Lee 2006).

4.3.3 Structural Equation Modeling

Structural equation modeling (SEM) is a class of statistical methodologies incorporating regression analysis, path analysis, confirmatory factor analysis and full scale models including both measurement and structural components. It characterizes hypotheses about the relationships between variables in the structural equation models (SEMs). The SEMs are applicable for experimental data as well as observational data including longitudinal data (Kline 2005).

The SEMs typically consist of two components: the measurement part and the structural part. The measurement model relates latent variables to observed or manifest measures using a confirmatory factor analysis (CFA). The structural model, on the other hand, regresses endogenous or dependent variables with exogenous or independent variables (Thompson 2000; Lee 2007; Kaplan 2009). The CFA method requires a priori measurement structure specifying both the number of latent factors and the relationships between observed variables and latent factors. The path analysis (PA) model can be specified in case that only a measure is available for each measurement part, and the causal relationships among the variables are established based on the relevant theory (Kline 2005; Kaplan 2009).

The SEM approach is effective for reducing the problems related with measurement error and thus achieving better parameter estimates. However, the benefit caused by incorporating both measurement and structural models is obtained at the cost of significant increases in degrees of freedom for testing model fit. It is thus probable that well established path model is rejected in terms of goodness-of-fit indices because of problems in the CFA model (Kaplan and Wenger 1993; Kaplan, 2009).

4.3.3.1 Assumptions

Many assumptions pertinent to the data and the estimation method are required to achieve reasonable estimation result. They cover multivariate normal distribution, complete random missing data, enough sample size, and correct model specification.

First, multivariate normality assumption indicates that observations should be continuous and normally distributed. It becomes relatively loose for categorical data and especially for maximum likelihood (ML) estimation. Nonnormal distribution does not influence coefficients; rather, it makes standard errors to be underestimated (Muthén and Kaplan 1992; Kaplan 2009). DiStefano (2002) also maintained that ML parameter estimates and standard errors were very low by introducing categorical indicators into the model.

The missing data mechanism becomes serious and influences the estimation result significantly if the data are neither missing at random (MAR) nor observed at random (OAR). Two ways to handling missing data are typically applicable: the listwise present approach (LPA) and pairwise present approach (PPA). The LPA utilizes listwise available data for all cases, and the PPA employs pairwise available data focusing on pairwise statistics (Kaplan 2009). Additional model-based approaches deal with missing data by modeling the mechanism that causes missing values (Kline 2005; Kaplan 2009). They are more flexible than conventional methods; moreover, standard errors can be computed while considering missing data, which is important in terms of model estimation and evaluation (Little and Rubin 2002).

Large number of samples should be obtained in order to lower sampling errors. Model complexity needs to be considered to decide sample size because more complicated models require larger number of samples than simpler models for obtaining stable parameter estimates. In addition, it is assumed that there are no model specification errors caused by the omission of relevant variables in any part of SEMs. The SEMs plagued with specification errors produce substantially biased parameter estimates. Some studies also showed that specification error in one part can be reproduced in other parts of the SEMs (Kaplan 1988; Kaplan 2009).

4.3.3.2 Model Specification

As the SEMs include many distinct statistical methodologies, there is no standard way of specifying the SEMs in mathematically and graphically. An effective way is to present different model specifications including PA model, CFA model and full structural model as follows.

Let p and q be the number of endogenous and exogenous variables, respectively. The system of structural part or path model in SEMs can be briefly expressed as

$$y = \alpha + \mathrm{B}y + \Gamma x + \zeta \tag{4.13}$$

where y is a $p\times1$ vector of observed endogenous variables, x is a $q\times1$ vector of observed exogenous variables, α is a $p\times1$ vector of structural intercepts, B is a $p\times p$ coefficient matrix linking endogenous variables, Γ is a $p\times q$ coefficient matrix linking endogenous variables to exogenous variables, and ζ is a $p\times1$ vector of disturbance terms. In the notation, the variance of disturbance terms (Ψ) is a $p\times p$ covariance matrix of the terms. Also, variance of exogenous variables (Φ) is a $q\times q$ covariance matrix for the variables (Kaplan 2009).

Another component of the SEMs, measurement model or CFA model in the form of the linear factor analysis can be specified as

$$x = \Lambda_x \xi + \delta \tag{4.14}$$

where x is a $q\times1$ vector of observed indicators, Λ_x is a $q\times k$ matrix of factor loadings, ξ is a $k\times1$ vector of common factors, and δ is a $q\times1$ vector of unique variance containing measurement error variance and specific variance (Kaplan 2009).

The general SEMs that incorporate both PA and CFA models for continuous latent variables can be represented as

$$\eta = \mathrm{B}\eta + \Gamma\xi + \zeta \tag{4.15}$$

where η is an $m\times1$ vector of endogenous latent variables, ξ is a $k\times1$ vector of exogenous latent variables, B is an $m\times m$ regression coefficient matrix which links the latent endogenous variables. Γ is an $m\times k$ regression coefficient matrix that relates endogenous variables to exogenous variables, and ζ is an $m\times1$ vector of disturbance terms. The latent variables are connected with observed variables via measurement model for both endogenous and exogenous variables. This model can be written as

$$y = \Lambda_y \eta + \varepsilon\,; \quad x = \Lambda_x \xi + \delta \qquad (4.16)$$

where Λ_y and Λ_x indicate $p \times m$ and $q \times k$ vectors of factor loadings, respectively. ε and δ are $p \times 1$ and $q \times 1$ matrices of unique variance, respectively. The parameter vector Ω for the full model consists of nine matrices, Λ_y, Λ_x, Θ_ε, Θ_δ, Φ, B, Γ, Ψ and $\Theta_{\delta\varepsilon}$ as in Figure 4.6 (Kaplan 2009).

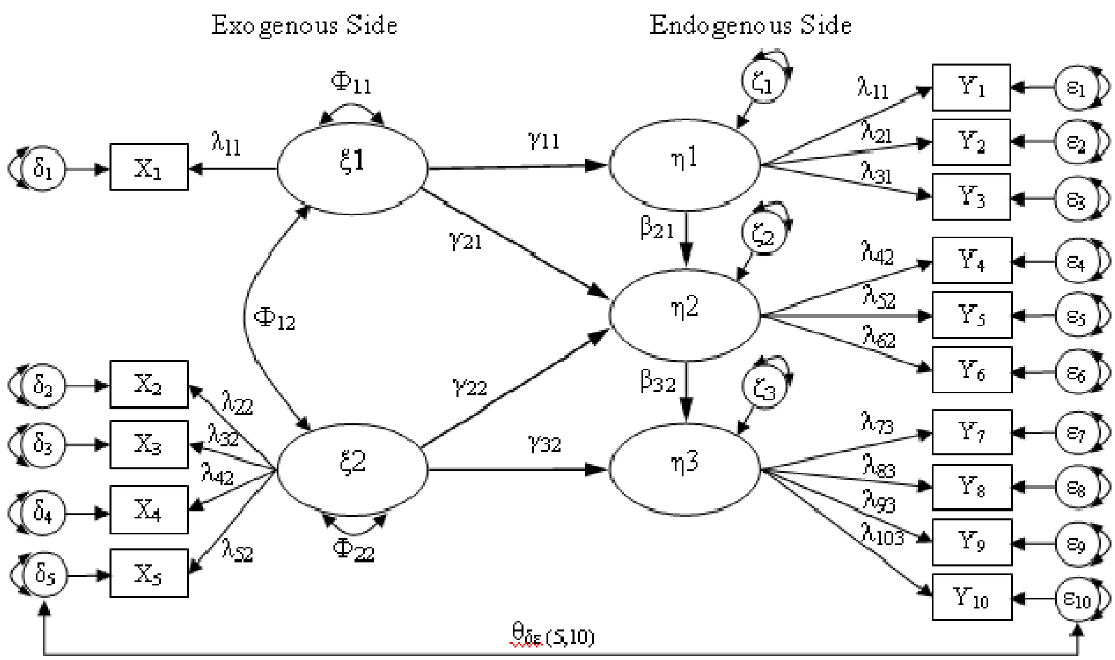


Figure 4.6 Nine Matrices and Four Vectors of General SEMs.

Note: Nine matrices are Φ (PH), B (BE), Γ (GA), Ψ (PS), Λ_X (LX), Λ_Y(LX), Θ_ε (TE), Θ_δ (TD), and $\Theta_{\delta\varepsilon}$ (TH); four vectors are κ (KA), τ_X (TX), τ_Y (TY), and α (AL).

4.3.3.3 Modeling Approach

Conventional SEM approach is characterized as illustrated in Figure 4.7. First, the structural models are established and specified based on a theoretical framework in which presumed relations among observed and latent variables are examined. The process can be accomplished by employing the path diagram method so that the relevant theories are consistent with the diagram, and vice versa. Instead of graphical representations, a series of equations can be established for investigating the assumed connections among variables (Kaplan 2009). Certain requirements have to be met in order for the models to be identified (Kline 2005).

In the next step, a set of variables are measured to be included in the model estimation process. It is crucial in the measurement model to incorporate multiple indicators for making underlying constructs clear. When it comes to measurement model, attention should be paid to both reliability and validity issues (Kline 2005; Kaplan 2009). Reliability indicates to what degree the responses in a sample are consistent across the items without random measurement error. Validity examines the soundness of the inference based on the responses while concerning if underlying constructs are appropriately measured in a sample (Thompson 2003; Kline 2005)

The following is the process of model estimation indicating that parameter estimates of the specified model are obtained using an estimation method. The estimation methods aim to minimize a fit function or a discrepancy function consisting of the sample or observed covariance matrix and model-implied or fitted covariance matrix. They include maximum likelihood (ML), generalized lease squares (GLS) and weighted least squares (WLS).[12] Recently, the WLS based methods for the continuous and categorical observations under nonnormality were developed (Muthén and Muthén 2006).

The estimated model is then evaluated and modified. Model evaluation intends to assess how well the estimated model is fit to the data. Many fit indices and model comparison indices are available especially for the SEMs with continuous variables. Fit indices encompass model chi-square (likelihood ratio chi-square) for exact test, and goodness of fit measures such as the root mean square error of approximation (RMSEA) (Steiger and Lind 1980), the comparative fit index (CFI) (Bentler 1990), and the standardized root mean square residual (SRMR). Model comparison indices are chi-square difference statistic for nested models, Akaike information criterion (AIC) and Bayesian information criterion (BIC) (Kline 2005; Kaplan 2009).

The main reasons of rejecting the model in terms of model fit indices are infringement on the assumptions, incorrect model specification, and insufficient number of samples. The model needs to be modified to better fit the data, which is called model modification. Both modification index (MI) and expected parameter change (EPC) are available commonly for continuous variables. The MI computes the expected decrease in the overall chi-square statistic resulting from freeing the restriction on a parameter estimate, while other constraints kept constant. The EPC gauges the change of a parameter estimate by relaxing the restriction on the parameter. As shown in Figure 4.7,

[12] Refer to Kaplan (2009) and Kline (2005) for detailed explanations of model estimation methods.

the estimated model is assessed and modified until its statistics meet some standards of fit indices (Kaplan 2009).

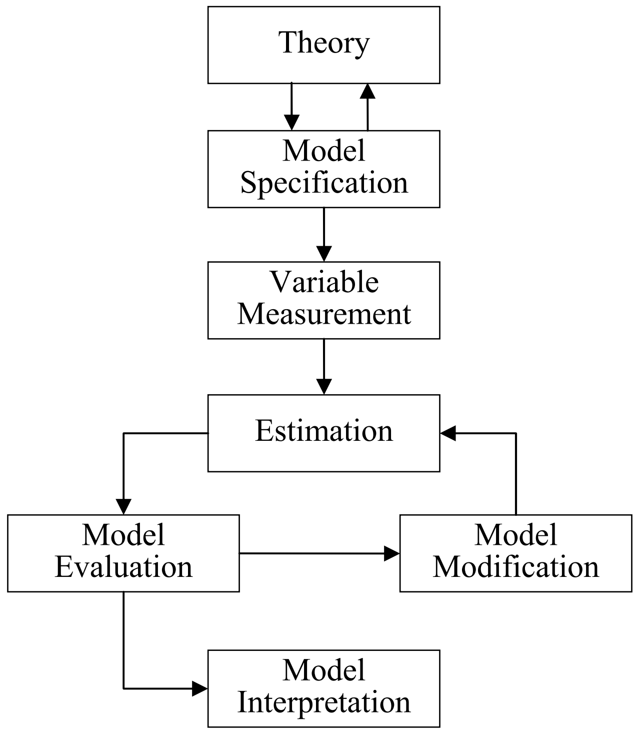


Figure 4.7 Conventional Structural Equation Modeling Approach.
Source: Kaplan (2009)

4.3.4 Direct Acyclic Graphs

In many cases, modeling practices for explanation and prediction have paid much more attention to associations than causalities among variables. Researchers often rely on the theories that support the causal directions among the variables. Even in the case, the theories commonly assume the ceteris paribus situation to clarify the causal relationships in an experimental system. The experiments may work if scientists suppose that one or more variables are functioning in the true system even though the whole system is not yet known. Only if the data are obtained from the randomized experimental design for controlling the system, the causation among variables can be clarified. However, data are not always obtainable in a well controlled system; rather, they are commonly observational.

There are two questions with which we are confronted. How can the causal connections be clarified using observational or non-experimental data in reliable and consistent ways? How the causal structure that is established helps manipulate and predict the system? Studies intend to not only elucidate the causal connections among variables but also forecast the change of the effect by modifying the cause (Spirtes et al. 2000). Causality appears to be linked with intervention and manipulation (Hausman, 1998).

A variety of studies have been conducted during the last several decades to conceptualize causal notions and analyze the causal structure based on graphical representation. A directed graph illustrates the causal flow among a group of variables as a picture. In this way, a great advance of the graphical causal modeling methods has been made based on observational data and nonparametric analyses (Pearl 2000; Spirtes et al. 2000).

4.3.4.1 Elements and Concepts

The directed acyclic graph (DAG) does not consider inference based on a cyclic system of causal flow, i.e. a system of a variable flowing through other variables and finally returning to itself. A graph consists of an ordered three components: **V, M**, and **E**

(Cooper 1999; Roh and Bessler 1999; Roh et al. 1999; Spirtes et al. 2000). **V** is a non-empty set of vertices or variables; **M** is a non-empty set of marks or symbols at the ends of vertices. And **E** is a set of ordered pairs of vertices and marks of which a member is called an edge. Variables which have causal relations are connected by edges, and two variables linked with an edge are adjacent. If a direct edge comes from V_1 to V_2, then V_1 is a parent of V_2 and V_2 is a child of V_1 (Spirtes et al. 2000).

The DAG is an illustration for constructing conditional independence based on a probability theory. In other words, it intends to handle the independence relations among variables in the system resulting from its application under the causal Markov condition. D-separation is a relation between three disjoint subsets of variables, X, Y and Z in a DAG. The concept is to check if a subset of variables in Y blocks any types of causal connections between a set of variables in X and Z. If V_1, V_2 and V_3 belong to a set of variables, then the correlation between V_1 and V_2 conditional on V_3 is zero if and only if V_1 and V_2 are d-separated given V_3 in a DAG (Cooper 1999; Pearl 2000; Spirtes et al. 2000)

In general, inferences on causal relationship in a DAG are structured by asymmetries among causal chains, causal forks, and causal inverted folks (Pearl 2000). The three different types of causal relations among threesomes, V_1, V_2 and V_3 help clarify the inferences behind the concept of d-separation.

First of all, a causal chain can be represented as following if V_1 causes V_2, and V_2 then causes V_3.

$$V_1 \longrightarrow V_2 \longrightarrow V_3$$

In the causal chain, V_1 is correlated with V_2 unconditional on V_3 ($\rho_{v_1,v_2} \neq 0$), i.e. d-connected as unconditional correlation. Also, V_2 is associated with V_3 unconditional on V_1 ($\rho_{v_2,v_3} \neq 0$), namely d-connected as unconditional association. V_1 and V_3 are d-connected as unconditional correlation ($\rho_{v_1,v_3} \neq 0$). However, the correlation between V_1 and V_3 conditional on V_2 is zero ($\rho_{v_1,v_3|v_2} = 0$), i.e. d-separated as conditional correlation.

The causal relation in land use and travel behavior interaction can be exemplified as follows. The number of workers in a household (V_1) is d-connected with total income in the household (V_2) as unconditional association, which then causally affect the number of automobiles in the household (V_3). However, the number of workers (V_1) and car ownership in the household (V_3) are d-separated conditional on total household income (V_2). Among households within same bracket of annual household income, the number of automobiles is not significantly different even though the number of workers increases. It is because the number of workers is no longer an important factor in determining auto ownership for the households earning same amount of annual income.

Another type of causal relation is called a causal fork in which all information is originated from a common cause (V_2). It can be illustrated as follows.

$$V_1 \longleftarrow V_2 \longrightarrow V_3$$

Both the causal fork and the causal chain are defined to be observationally identical because the association structure between vertices of the causal fork is equivalent to that of the causal chain. Any pair of variables among V_1, V_2, and V_3 are correlated or d-connected as unconditional correlation, which can be expressed as

$\rho_{v_1,v_2} \neq 0$, $\rho_{v_2,v_3} \neq 0$, and $\rho_{v_1,v_3} \neq 0$. However, the association between V_1 and V_3 conditional on a common cause (V_2) becomes zero, i.e. d-separated as conditional association ($\rho_{v_1,v_3|v_2} = 0$). In other words, the knowledge of a common cause (V_2) is screening off the relationship between its joint effects (V_1 and V_3).

For instance, car ownership in a household as a common cause causally affect household total vehicle miles of travel (VMT) (V_1) as well as the probability of automobile mode choice for a trip (V_3). If conditional on household car ownership (V_2), household total VMT (V_1) and the likelihood to choose automobile mode (V_3) become d-separated. If households have same number of automobiles available, both total VMT of each household and automobile choice probability appear to be constant across the households of interest.

Third type of causal relation is called a causal inverted fork in which a common effect (V_2) takes all information flowing from different adjacent causes (V_1 and V_3), but is not open to any other variable. The variable V_2 is defined as a collider because causal impacts of different causes converge or collide on it as follows.

$$\boldsymbol{V_1} \longrightarrow \boldsymbol{V_2} \longleftarrow \boldsymbol{V_3}$$

The variables V_1 and V_3 in the causal inverted fork are associated with the collider V_2, respectively unconditional on remaining variable, i.e. d-connected as unconditional correlation ($\rho_{v_1,v_2} \neq 0$ and $\rho_{v_2,v_3} \neq 0$). But the unconditional association between V_1 and V_3 is zero, indicating that V_1 and V_3 are d-separated in the directed graph. On the other hand, the correlation between V_1 and V_3 conditional on the common effect (V_2) becomes significantly different from zero, i.e. d-connected as conditional association ($\rho_{v_1,v_3|v_2} \neq 0$). It implies that the information of a common effect V_2 does not block or screen off the association between its joint causes V_1 and V_3.

For example, increase in diversity of land use pattern (V_1) and improved pedestrian connectivity (V_3) causally influence decrease in the probability of choosing automobile modes for a trip (V_2). The two causes, land use mix (V_1) and pedestrian connectivity (V_3) do not seem to be correlated unconditional on the common effect, automobile choice probability (V_2). However, given the common effect, the relationship between land use diversity and pedestrian connectivity becomes significant. If a group of individual trip-makers are highly dependent on automobile choice for every trip, it is more likely that not only are land uses not well mixed among different uses such as residential, commercial and recreational uses, but pedestrian and bike road network around their origins or destinations is poorly prepared and connected.

4.3.4.2 Assumptions

Causation is assumed to be transitive, irreflexive and antisymmetric. If V_1 is a cause of V_2, and V_2 is a cause of V_3, the V_1 is a cause of V_3 (transitive). Any event (V_1, V_2 or V_3) cannot cause itself (irreflexive). If V_1 causes V_2, then V_2 cannot cause V_1 (antisymmetric) (Sprites et al. 2000).

Three assumptions are generally considered on which probability distributions are connected with the DAGs: causal Markov condition, faithfulness condition and causal sufficiency. They are not independent, but connected with each other.

First, the causal Markov condition indicates that all information on the probability distribution of a variable must be carried from its parents. It can be formally defined as follows (Spirtes et al. 2000).

> Let G be a causal graph with vertex set V and P be a probability distribution over the vertices in V generated by the causal structure represented by G. G and P satisfy the Causal Markov Condition if and only if for every W in V, W is independent of V\(Descendants(W) ∪ Parents(W)) given Parents(W).

As implied in the definition, causality based on the Markov condition is local in time and space, so direct cause screen off remote or indirect causes (Cooper 1999). The condition allows us to have two intuitions. First, variables are independent of their indirect causes conditional on their parents or direct causes. Another principle is that a variable is independent of others conditional on its common causes (Scheines et al. 1996).

Based on the condition, a class of probability distributions can be determined, and the probability is represented with a recursive product (Spirtes et al. 2000).

$$\mathrm{P}(v_1, v_2, \ldots, v_n) = \prod_{i=1}^{n} \mathrm{P}(v_i \,|\, Parents\,(v_i)) \qquad (4.17)$$

where P is the joint probability of vertices or variables V_1, V_2,…, V_n, and Parents(v_i) represents direct causes of a variable v_i. Π indicates the functional product operation. The equation is represented as d-separation, a generalized graphical relation proposed by Pearl (1995, 2000).

Another assumption is faithfulness condition which concerns conditional independence relations presented both by a probability distribution and by the causal Markov condition. It is specified as follows (Spirtes et al. 2000).

> Let G be a causal graph and P a probability distribution generated by G. <G, P> satisfied the Faithfulness Condition if and only if every conditional independence relation true in P is entailed by the Causal Markov Condition applied to G.

In some cases, independence relations that are not generated based on the Markov condition could be existed in a probability distribution on a DAG in which the Markov condition is met. The faithfulness condition is important for figuring out causal structure because it pays attention to the relationship between probability distributions and causal connections (Cooper 1999; Spirtes et al. 2000). The causal Markov condition connects causal structure on a causal graph with independence relationships shown in a probability distribution. The faithfulness condition, on the other hand, link causal structure with dependence relations represented in a probability distribution (Cooper 1999).

Third, the causal sufficiency should be satisfied for constructing a directed graph. It indicates that a set of variables is said to be causally sufficient if the group of variables contains the variables which causes two or more other variables in the group. Therefore, causal sufficiency assumption ensures that no variable should be omitted in the investigation if the variable is a common cause of other variables (Scheines et al. 1996).

4.3.4.3 PC Algorithm

Some computing algorithms based on the idea of d-separation have been created and developed.[13] A series of procedures for producing the DAGs has been integrated into each algorithm (Spirtes, et al. 2000). PC algorithm has also been designed for the purpose that is conducted by the serial versions of TETRAD programs (Scheines, et al. 1996; Spirtes, et al. 2000).

The PC algorithm is composed of a series of ordered computing commands (Spirtes et al. 2000). First, a complete undirected graph is constructed on all pairs of variables. The undirected graph represents a group of undirected edges between every pair of variables in the analytical system. Then, tests for an ordered pair of variables are performed consecutively to check out if unconditional correlation between the pair of vertices is statistically equal to zero, i.e. $\rho_{v_1,v_2} = 0$. The edge is taken away from the graph if it is not significantly different from zero. Furthermore, edges which are still connected in the undirected graph are tested if conditional correlation between each pair of vertices is equal to zero in an orderly manner. If the partial correlation is equal to zero, then the edge is removed. In terms of statistical decisions, Fisher's z statistic is employed to conduct the tests of conditional correlation.[14] It is described as follows.

$$z(\rho_{i,j|k}) = \frac{\sqrt{(n-|k|-3)}}{2} \times \ln\left(\frac{1+\rho_{i,j|k}}{1-\rho_{i,j|k}}\right) \tag{4.18}$$

[13] Many applicable algorithms except PC algorithm are presented in Spirtes et al. (2000) including the Wermuth-Lauritzen algorithm, the SGS algorithm, Modified PC algorithm, Causal Inference algorithm, and Fast Causal Inference algorithm.

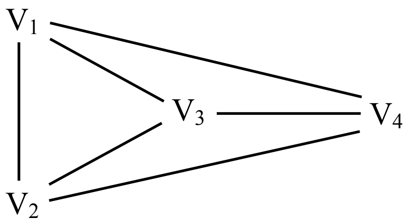
[14] The z test is only applicable for continuous variables. For the discrete case, PC algorithm conducts tests for independence using G^2 that is defined as: $G^2 = 2\sum(Observed\,Value)\ln(Observed/Expected)$ (Spirtes et al. 2000).

where $\rho_{i,j|k}$ is population conditional correlation between a series of variables i and j conditional on k. $|k|$ is the number of variables in k, and n is the number of observations.

Finally, remaining edges are connected based on separation set or sepset and the away-from-a collider test. The sepset is defined as the subset of conditioning variable(s) on removed edges between two vertices after partial correlation tests in series. In a DAG consisting of three variables, X, Y and Z, then Z is the sepset of the edge between X and Y if this edge is removed by conditioning on Z, i.e. $\rho_{X,Y|Z} = 0$. But, Z becomes a collider if the edge cannot be removed conditional on Z, i.e. $\rho_{X,Y|Z} \neq 0$.

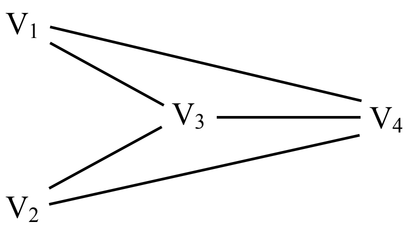
An example of how the PC algorithm in a directed graph works is presented in Figure 4.8. Three assumptions, causal Markov condition, faithfulness condition and causal sufficiency suffice for the exemplary application. It is also assumed that the true structure of the directed graph that generated the data is illustrated in the last stage of Figure 4.8. It is called a pattern representing a set of directed causal graphs as they entail the same conditional independence relations and are consistent with the knowledge of causal structure (Verma and Pearl 1990).

To begin with, the algorithm automatically builds a complete undirected graph as presented in stage (i) in which every pair of four variables (V_1, V_2, V_3 and V_4) is connected without causal direction. Then, unconditional or zero-order partial correlation test is conducted for every pair of vertices. The undirected link between V_1 and V_2 that is not significantly different from zero in the test is removed as illustrated in the stage (ii). In the third step, first order partial correlation test on each pair of variables given one of other variables. This procedure continues until the tests for k^{th} order partial correlation conditional on all other remaining k variables are completed. Any undirected edge that is conditionally independent is eliminated in the graph. As shown in the stage (iii), first order partial correlation tests for V_1 and V_4 and for V_2 and V_4 conditional on V_3 are conducted consecutively. As a consequence, their links are taken away from the graph because they are proved to be conditionally independent, i.e. $\rho_{v_1,v_4|v_3} = 0$ and $\rho_{v_2,v_4|v_3} = 0$. The causal relation among V_1, V_2 and V_3 is the type of a causal inverted fork because of their unconditional independence in the step (ii) and their conditional association ($\rho_{v_1,v_2|v_3} \neq 0$). Thus, all information from V_1 and V_2 flows toward V_3, a common effect as represented in the step (iv). Last step determines the causal direction between V_3 and V_4 based on the fact that the associations between V_1 and V_4 and V_2 and V_4 are screened off by V_3 in the step (iii). The fact in the step (iii) indicates that both the causal connections among V_1, V_3 and V_4, and among V_2, V_3 and V_4 are either causal chains or causal forks. When the causal relationships between V_1, V_3 and V_2, V_3 in the step (iv) are considered, they should be causal chains.



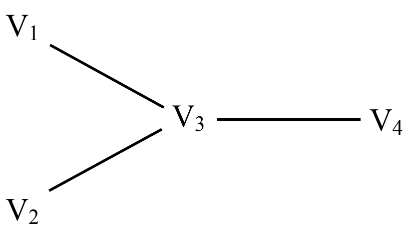
(i)

Step 1:
Build a complete undirected graph between each pair of variables.



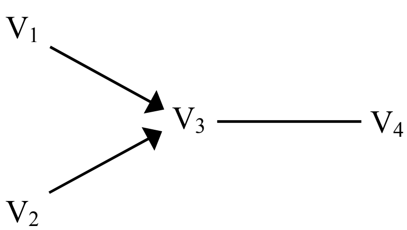
(ii)

Step 2:
An undirected edge between of V_1 and V_2 is removed as they are marginally independent, i.e. $\rho_{v_1,v_2} = 0$.



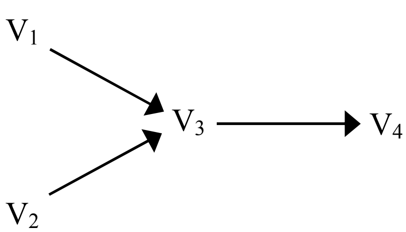
(iii)

Step 3:
Undirected links between of V_1 and V_4 and V_2 and V_4 conditional on V_3 are eliminated because they are conditionally independent, i.e. $\rho_{v_1,v_4|v_3} = 0$ and $\rho_{v_2,v_4|v_3} = 0$.



(iv)

Step 4:
V_3 is a common effect as V_1 and V_2 are unconditionally independent ($\rho_{v_1,v_2} = 0$) and conditionally dependent ($\rho_{v_1,v_2|v_3} \neq 0$).



(v)

Step 5:
Based on the facts found in the previous steps, V_3 should be a cause of V_4. A directed acyclic graph is finally accomplished.

Figure 4.8 An Example of How the PC Algorithm Works.

Source: Cooper (1999); Druzdzel and Glymour (1999); Lee (2006)

CHAPTER V

RESULTS

This chapter exhibits overall household travel pattern and land use characteristics based on the 2007 HGAC Regional Household Activity and Travel Survey and 2007 parcel-based land use datasets, respectively. Empirical results are presented and interpreted with tables and figures representing estimated models of individual mode choice, household auto trip generation, and household total VMT. They are specified for different travel purposes from travel demand and causal relationship approaches. The results are summarized for each travel behavior outcome of interest and some related issues are further discussed.

5.1 Household Travel and Land Use Characteristics

5.1.1 Household Travel Pattern[15]

Based on partial data of the 2007 HGAC household travel survey, 42,275 trips are made by 4,170 sampled households in total. The household travel survey collected household and individual socioeconomic characteristics, vehicle information and trip and activity information. Total 42,275 trips are classified into 6,558 HBW trips, 22,640 HBO trips and 13,077 NHB trips. Table 5.1 presents the distribution of automobile travel time by trip purpose. Each household is estimated to drive an average 8.4 minutes per trip in the HGAC region every weekday. Commuters averaged about 13.3 minutes

[15] As described in the previous chapter, only 84% of the 2007 HGAC Regional Household Activity and Travel Survey data are used in the study because the survey was not completed then.

per trip for HBW trips that is more than average trip times for other purposes (6.1 and 9.9 minutes per trip for HBO and NHB trips, respectively). In terms of travel time distribution, more than 90% of total trips are made within 20 minutes. This pattern is similar for both HBO and NHB trips. But, trip times are spread more widely for HBW trips, while only about 77% of trips are made within 20 minutes.

Table 5.1 Automobile Travel Time Distribution by Trip Purpose.

Travel Time (min) [1)]	Total Trips		HBW Trips		HBO Trips		NHB Trips	
	Trips	%	Trips	%	Trips	%	Trips	%
0 – 10	32,392	76.6	3,307	50.4	18,767	82.9	10,318	78.9
10 – 20	6,199	14.7	1,765	26.9	2,680	11.8	1,754	13.4
20 – 30	2,211	5.2	880	13.4	684	3.0	647	4.9
30 – 40	902	2.1	356	5.4	314	1.4	232	1.8
40 – 50	344	0.8	154	2.3	117	0.5	73	0.6
50 – 60	143	0.3	70	1.1	37	0.2	36	0.3
Over 60	84	0.2	26	0.4	41	0.2	17	0.1
Total	42,275	100.0	6,558	100.0	22,640	100.0	13,077	100.0

Note: 1) Travel time is driving-alone (DA) travel time based on 2007 transportation skim data obtained from the HGAC Transportation Department.

The observed patterns of travel time by different mode options are similar as shown in Figure 5.1.[16] A great number of trips are made within 20 minutes, and they generally decreases as travel time increases. However, the percentage of total trips within 20 minutes is relatively small except for walk and bike mode when compared with the previous distribution. Travel time distribution of transit mode is different from other modes.

[16] Travel time is reported on the 2007 HGAC Regional Household Activity and Travel Survey.

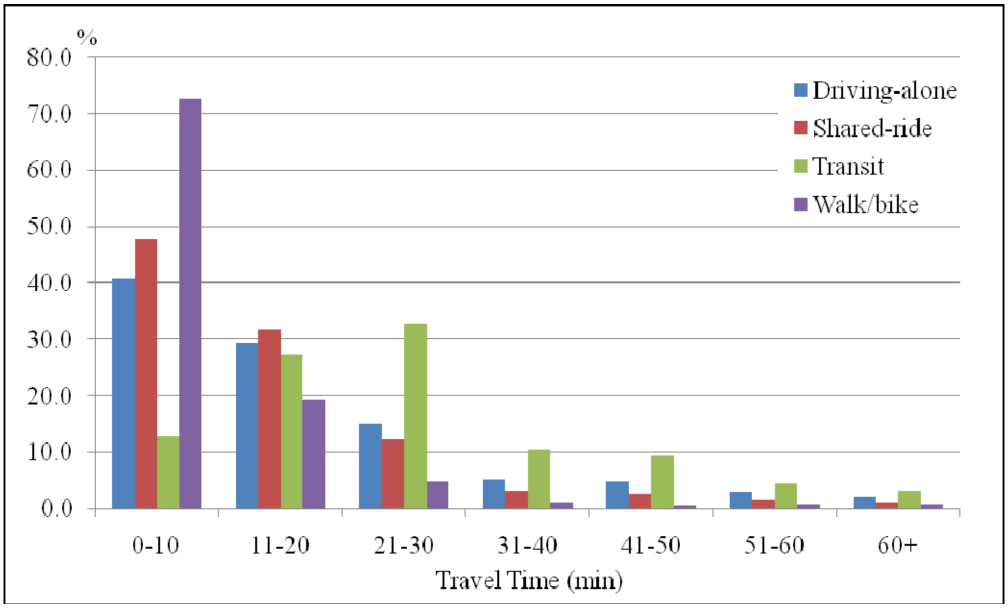


Figure 5.1 Automobile Travel Time Distribution by Trip Mode Choice.

It is also examined how household income levels are associated with travel cost by travel purposes. Table 5.2 shows the number of trips and average short-term auto travel cost ($/trip) calculated by household income levels and trip purposes. An increase in household income is associated with more number of trips generated despite a few variations. It is also evident that a household spends more in automobile trips as it makes more money. This pattern is similar for every travel purpose. An average household spends 2.05 dollar per trip. For HBW trips, over two times more than the average cost is needed, and it is more expensive than the average travel costs for any other travel purposes at all income brackets. These characteristics seem to be closely related to more travel time and distance as well as more disposable income for higher income groups.

Table 5.2 Automobile Travel Cost by Household Income Level.

Household Income	Total Trips		HBW Trips		HBO Trips		NHB Trips	
	Trips	Mean [1)]	Trips	Mean	Trips	Mean	Trips	Mean
Below 10K	1,164	1.33	106	2.90	724	1.18	334	1.17
10K – 20K	3,189	1.76	411	3.93	1,846	1.41	932	1.49
20K – 30K	5,966	1.89	841	4.40	3,292	1.45	1,833	1.53
30K – 40K	8,886	2.03	1,486	4.49	4,695	1.41	2,705	1.76
40K – 60K	5,519	2.18	888	4.96	2,859	1.53	1,772	1.83
60K – 100K	10,243	2.17	1,691	5.00	5,392	1.42	3,160	1.92
Over 100K	7,308	2.16	1,135	5.17	3,832	1.47	2,341	1.83
Total	42,275	2.05	6,558	4.73	22,640	1.44	13,077	1.76

Note: 1) Mean trip cost is short-term cost including operation and maintenance expenses for driving-alone (DA) mode. For details, see the measurement section of chapter IV.

In terms of mode shares, automobile modes including driving-alone and shared-ride are dominant as illustrated in Figure 5.2. The share of driving-alone mode is slightly larger than that of shared-ride for total trips. However, the proportion of driving-alone mode is dominant for HBW trips. Trip modal splits by activities indicate that trip-makers drive alone more from home for working, shopping and social and recreational activities.

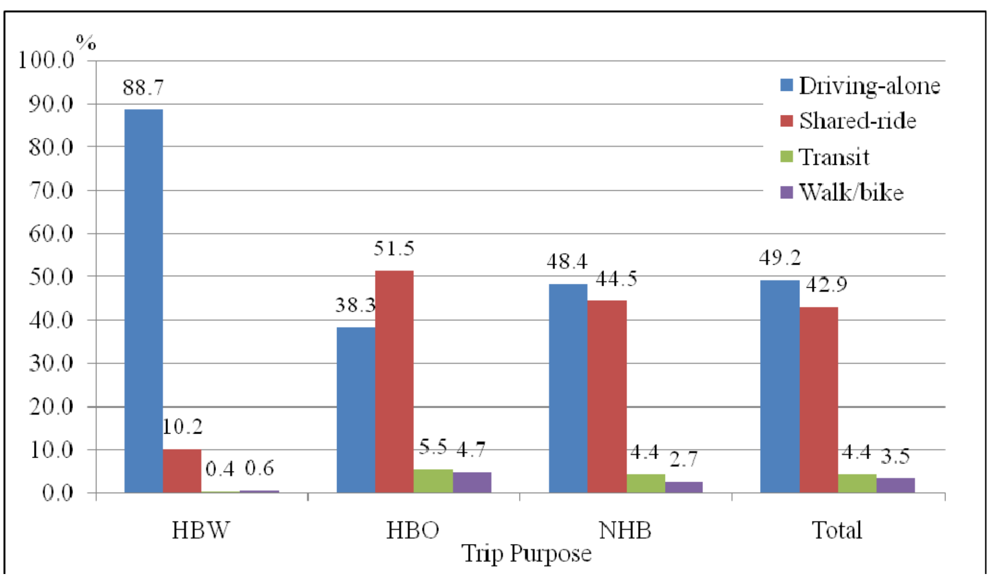


Figure 5.2 Trip Mode Shares by Trip Purpose.

Table 5.3 shows trip distribution by household size and travel purposes. It is found that total number of trips increases with household size. The number of HBW trips increases until three of household size, and decrease after then. The pattern of total trips is attributed to that of HBO trips. The number of bike uses for last seven days is presented in Figure 5.3. As expected, the greater the household size is, the more likely bike uses are. This tendency becomes intensified for HBO trips; however, it is not consistent for HBW trips.

Table 5.3 Trip Distribution by Household Size.

Household Size	Total Trips		HBW Trips		HBO Trips		NHB Trips	
	Trips	%	Trips	%	Trips	%	Trips	%
1	2,367	5.6	463	7.1	981	4.3	923	7.1
2	9,120	21.6	1,359	20.7	4,559	20.1	3,202	24.5
3	9,502	22.5	1,920	29.3	4,694	20.7	2,888	22.1
4	10,599	25.1	1,504	22.9	5,995	26.5	3,100	23.7
5+	10,687	25.3	1,312	20.0	6,411	28.3	2,964	22.7
Total	42,275	100.0	6,558	100.0	22,640	100.0	13,077	100.0

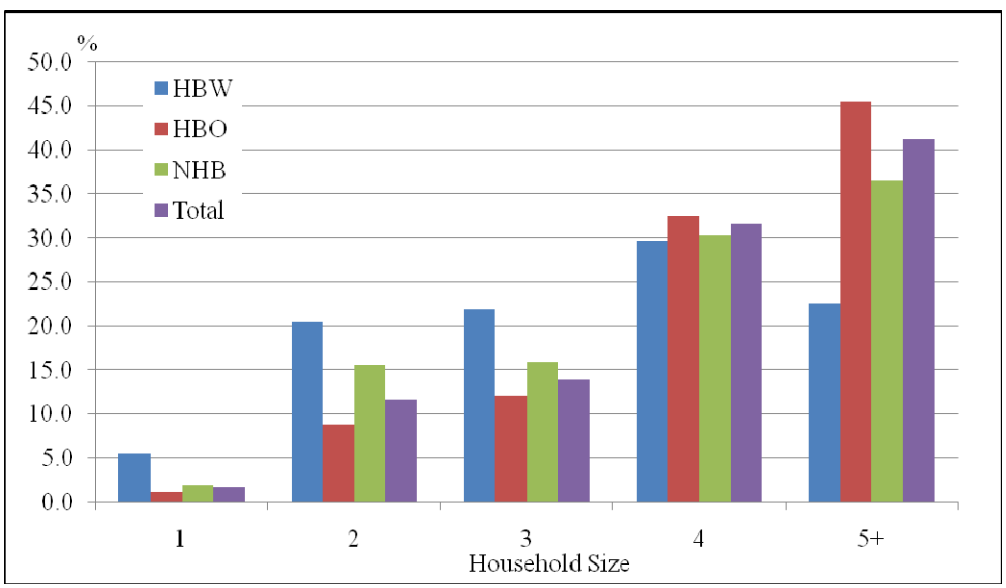


Figure 5.3 Bike Use by Household Size.

Average VMT is arranged by different residential types in Table 5.4. On average, VMT for total trips is 12.12 and 9.53 miles for driving-alone and automobile mode, respectively. Several similarities are observed. First, the average VMT for driving-alone is greater than that for automobile for all purposes. Second, people living in single-family homes make longer trip than those in multi-family homes and apartments. Last, the average VMT for HBW trips are longer than that for both HBO and NHB trips. These characteristics seem to be associated with the distribution of travel time and cost presented in Table 5.1 and Table 5.2.

Table 5.4 Average VMT by Residential Type.

Residential Type	Total Trips		HBW Trips		HBO Trips		NHB Trips	
	DA	Auto [1)]	DA	Auto	DA	Auto	DA	Auto
Single-family	14.79	11.00	22.95	21.77	11.24	8.46	12.18	9.54
Multi-family	9.59	7.88	18.36	17.74	9.21	7.21	9.59	7.88
Apartment	11.64	8.93	18.73	17.45	10.15	7.70	11.64	8.93
Other	13.30	11.58	21.67	20.34	12.27	11.16	13.30	11.58
Total Mean	12.12	9.53	22.70	21.51	11.19	8.48	12.12	9.53

Note: 1) Auto includes driving-alone (DA) and shared-ride (SR).

5.1.2 Land Use and Development Characteristics

It is pivotal in this study to understand and measure land use and development characteristics of the HGAC region. As explained in the measurement part of the previous chapter, this study conducts land use measurement in a quarter-mile boundary for every trip ends. However, this section examines overall land use patterns measured at the level of traffic analysis zones (TAZs). The zones are spread over the entire HGAC region without any overlap between them. Sometimes they reflect the spatial extents in which various activities take place.

Table 5.5 summarizes overall land use and development patterns of six counties in the HGAC region. An area of 1,500 out of 6,730 square miles was developed for sustaining human activities. Residential use area (55%) shows the largest share of total developed area. Among the residential shares, the area of single-family houses (50%) is far larger than that of multi-family (2%) and condo and apartment (0.3%). This land use pattern is consistent throughout the region from 86 percent for single-family residential area in Montgomery to 34 percent in Brazoria. Harris County has the largest single-family residential area covering 353 square miles. Dominant area of single-family residential use may encourage people to choose automobile mode, make more auto trips and drive farther to meet their travel demands.

Table 5.5 Land Use Pattern of Developed Area in the HGAC Region.

Land Use	Total	Brazoria	Fort Bend	Galveston	Harris	Montgo-mery	Waller
Residential	823.78 (54.82) 1)	103.99 (35.48)	96.23 (66.39)	64.24 (64.51)	397.67 (51.62)	154.05 (88.68)	7.61 (36.26)
Single-family	758.54 (50.48)	100.24 (34.20)	86.79 (59.88)	60.66 (60.93)	353.28 (45.86)	150.07 (86.39)	7.50 (35.73)
Multi-family	33.37 (2.22)	1.23 (0.42)	1.29 (0.89)	1.93 (1.94)	27.68 (3.59)	1.13 (0.65)	0.11 (0.53)
Condo/Apt	4.83 (0.32)	0.00 (0.00)	0.19 (0.13)	0.36 (0.36)	4.05 (0.53)	0.23 (0.13)	0.00 (0.00)
Others	27.05 (1.80)	2.51 (0.86)	7.96 (5.49)	1.29 (1.29)	12.66 (1.64)	2.63 (1.51)	0.00 (0.00)
Commercial /Industrial	354.06 (23.56)	25.46 (8.69)	48.58 (33.52)	33.50 (33.64)	217.37 (28.21)	19.66 (11.32)	9.48 (45.20)
Commercial	244.24 (16.25)	14.96 (5.10)	32.34 (22.31)	25.77 (25.88)	146.71 (19.04)	19.08 (10.99)	5.38 (25.63)
Industrial	77.13 (5.13)	8.72 (2.97)	7.29 (5.03)	5.68 (5.70)	52.61 (6.83)	0.58 (0.33)	2.26 (10.77)
Others	32.69 (2.18)	1.79 (0.61)	8.95 (6.18)	2.05 (2.06)	18.05 (2.34)	0.00 (0.00)	1.85 (8.80)
School/Public	324.89 (21.62)	163.64 (55.83)	0.14 (0.10)	1.84 (1.85)	155.38 (20.17)	0.00 (0.00)	3.89 (18.54)
Total developed 2)	1502.73 (22.33)2)	293.09 (18.35)	144.95 (16.37)	99.57 (11.37)	770.42 (43.35)	173.71 (16.15)	20.98 (4.05)
Total area	6,729.44	1,597.31	885.64	875.75	1,777.32	1,075.81	517.62

Note: 1) Values in parenthesis in land use types are the percentage of total developed.
2) Total developed is the sum of residential, commercial/industrial, and school/public. Values in parenthesis are the percentage of total area.

Table 5.6 and Figure A2-1 and A2-2 in the appendix show the distribution of population and employment density. Regional median population and employment density are 4.2 and 1.7 per acre, respectively. About one third of total TAZs have less than 2 residents per acre, and one fifth of total zones have over 10 people per acre. Harris County shows the highest population density, but Waller County exhibits the lowest. More than 86 percent of total TAZs have less than 2 people per acre in Waller

County. These patterns are similar to those of employment density, which shows the extremes of lower and higher density.

Figure A2-1 helps understand the spatial distribution of population density in depth. Inner areas of Harris County and the City of Houston show relatively higher density; however, the outskirts of the region reveal very low density. Figure A2-2 also supports the fact that the employment density is polarized into two extremes more than population density. It is very high in the central area of the City of Houston, but becomes very low out of the area. It is also found that residential and commercial and industrial uses are highly segregated throughout the region.

Table 5.6 Distribution of Land Use Density Measures of TAZs in the HGAC Region.

Density (per acre)	Population Density				Employment Density			
	Total	Harris	Galveston	Waller	Total	Harris	Galveston	Waller
0	6.5 [2)]	9.4	0.9		5.1 [2)]	5.1	5.4	8.6
0 - 1	17.7	10.1	15.2	65.5	34.9	21.4	49.6	74.1
1 - 2	10.9	7.5	15.2	20.7	14.1	15.1	14.7	12.1
2 - 3	7.8	6.0	14.3	10.3	8.0	8.7	7.1	1.7
3 - 4	5.8	4.8	8.0	1.7	6.3	7.9	5.8	
4 - 5	6.5	6.4	12.5		4.4	5.8	2.2	
5 - 6	5.6	5.7	7.6		2.9	3.4	1.8	1.7
6 - 7	5.3	6.0	4.9	1.7	2.7	2.9	1.3	1.7
7 - 8	4.8	5.5	3.6		2.1	2.5	2.7	
8 - 9	5.0	6.1	3.6		1.5	1.8	1.3	
9 - 10	4.2	5.1	2.2		1.0	1.2	0.0	
10+	20.0	27.2	12.1		17.1	24.3	8.0	
Total [1)]	2,829	1,846	224	58	2,829	1,846	224	58
Mean	5.94	7.20	4.71	1.06	24.50	36.51	3.36	0.54
Median	4.22	5.99	3.53	0.62	1.68	2.94	0.79	0.12

Note: 1) Total number of traffic analysis zones (TAZs).
2) Individual density values except for total, mean and median are percentage of total TAZs.

Entropy index measures the degree of balance among different uses within an area. Dissimilarity index shows how well different land uses in a place are mixed with its neighbors. The distribution of both diversity measures is summarized along with three counties in Table 5.7. Both entropy and dissimilarity measures average 0.62 and 0.55 in total, respectively. It is clearer in Harris County that higher county average and portion over the regional average are observed. Waller County, however, is in the opposite direction where the majority of the zones are placed below the regional average. These patterns are comparable with those of land use mix.

The spatial distributions of both measures are also illustrated in Figure A2-3 and A2-4. The map of entropy index distribution indicates that the value becomes larger as the distance from the central district of the City of Houston surrounded by I-610 increases. Both northern and southern areas including the border areas of Harris County shows higher level of land use balance. The spatial distribution of dissimilarity index is also similar except that it shows lower mean and median values, and the central district exhibits higher rates.

Table 5.7 Distribution of Land Use Diversity Measures of TAZs in the HGAC Region.

Value	Entropy Index				Dissimilarity Index			
	Total	Harris	Galveston	Waller	Total	Harris	Galveston	Waller
0.0 - 0.1	1.1 [2)]	0.7	0.9	3.4	1.6 [2)]	0.7	2.2	6.9
0.1 -0.2	1.6	0.8	1.3	17.2	4.3	1.4	3.6	56.9
0.2 - 0.3	2.9	1.8	2.2	25.9	5.2	2.0	6.7	19.0
0.3 - 0.4	10.0	11.7	6.3	20.7	6.4	4.1	7.6	8.6
0.4 - 0.5	8.3	6.4	12.1	19.0	9.3	7.5	10.7	5.2
0.5 - 0.6	15.1	14.0	19.6	12.1	21.0	21.8	19.2	1.7
0.6 - 0.7	21.3	19.2	28.6		38.2	44.5	35.3	1.7
0.7 - 0.8	24.1	26.4	19.6	1.7	13.4	17.2	14.7	
0.8 - 0.9	14.2	16.8	8.5		0.5	0.8		
0.9 - 1.0	1.5	2.1	0.9					
Total [1)]	2,829	1,846	224	58	2,829	1,846	224	58
Mean	0.620	0.639	0.606	0.326	0.553	0.600	0.541	0.204
Median	0.658	0.681	0.623	0.318	0.605	0.628	0.599	0.152

Note: 1) Total number of traffic analysis zones (TAZs).
2) Individual diversity values except for total, mean and median are percentage of total TAZs.

Table 5.8 presents the distribution of connectivity measure along with six counties in the HGAC region. Connectivity indicates how well the road network is connected in an area. Regional average is 0.26, and around two thirds of total TAZs have less than 0.2 in the region. Both Harris and Galveston counties show less than 60 percent in the number of TAZs with under 0.2. In Harris County, 12 percent of TAZs have more than 0.9 of connectivity measure. On the other hand, over 80 percent of TAZs in Fort Bend and Montgomery show below 0.2.

A map is also prepared to examine the spatial distribution of connectivity throughout the region as shown in Figure A2-5. It is observed that the central district of the City of Houston is higher in value. However connectivity generally diminishes as it becomes distant from the center.

Table 5.8 Distribution of Connectivity Measure of TAZs in the HGAC Region.

Value	Total	Brazoria	Fort Bend	Galveston	Harris	Montgomery	Waller
0.0 - 0.1	29.8 [2)]	48.1	48.4	34.8	20.9	50.0	63.8
0.1 - 0.2	33.2	27.9	37.5	24.6	35.0	33.6	15.5
0.2 - 0.3	11.7	9.5	9.4	13.4	12.7	8.0	5.2
0.3 - 0.4	5.7	8.8	2.1	7.1	5.8	3.5	1.7
0.4 - 0.5	3.4	0.7	0.5	4.0	4.2	1.8	3.4
0.5 - 0.6	3.2	2.5	0.5	3.6	3.6	1.8	5.2
0.6 - 0.7	2.4	1.1	1.0	3.1	2.8	0.4	5.2
0.7 - 0.8	1.5	1.1	0.5	1.8	1.8	0.4	0.0
0.8 - 0.9	0.9	0.4	0.0	4.5	0.8	0.0	0.0
0.9 - 1.0	8.3	0.0	0.0	3.1	12.2	0.4	0.0
Total [1)]	2,829	283	192	224	1,846	226	58
Mean	0.260	0.151	0.131	0.248	0.311	0.132	0.137
Median	0.147	0.107	0.101	0.146	0.173	0.099	0.065

Note: 1) Total number of traffic analysis zones (TAZs).
2) Individual values except for total, mean and median are percentage of total TAZs.

5.2 Individual Mode Choice Models

5.2.1 Results of Multinomial Logit Models

The multinomial logit (MNL) model is employed as a conventional travel demand model for analyzing individual mode choice behavior. Four choice options are taken into consideration: driving-alone (DA), shared-ride (SR), transit (TR), and walk and bike (WB). Driving-alone is chosen as the reference mode; each estimated constant term on the utility function, therefore, has to be explained in consideration of the reference. In addition, MNL specification introduces alternative specific variables instead of generic variables because trip-makers are affected by the attributes of different modes in different ways. Results of the MNL models for HBW and HBO trips are presented in Table 5.9 and Table 5.10. Base models for different travel purposes are also estimated to be compared with extended models or full models.

The MNL extended model for HBW trips in Table 5.9 shows same patterns as the base model for travel time variables and many socioeconomic characteristics. Four alternative specific travel time variables are all negatively significant at 1% level as suggested by the theory. They imply that an increase in travel time for each travel mode reduces the probability of choosing the mode, which is consistent for all choice modes. Travel time and cost measures play an important role in making choice decisions for commute trips.

Socioeconomic attributes have positive and significant impacts on the likelihood of specific mode choices, which agrees with both the theory and the results of previous studies. In specific, personal attributes such as gender and age are of significance: females are more likely to drive to work, and the older are also likely to drive alone for commute trips. Commuters who have used a bike mode tend to take more transit and alternative modes (walk and bike). Household socioeconomics including household size, vehicle ownership and total income have meaningful effects on the likelihood of specific mode choice. Larger households are more inclined to use shared-ride, transit and walk and bike rather than to drive alone. Individuals having more income and vehicles depend more on driving-alone mode.

Five out of eight land use measures have significant effects on mode choice probability for HBW trips. Density measures including population and employment density are worthy to be focused. At trip destination, two density measures are negatively associated with the probability of choosing automobile modes (driving-alone and shared-ride). On the contrary, increases in population and employment density at origin encourage individual travelers to make driving-alone and automobile choice, respectively. Also, improved connectivity at destination is significantly correlated with more chances of automobile mode choice. Dissimilarity variables at both trip ends, however, are not significant in the HGAC area.

The goodness-of-fit indices and the model improvement test confirm that the extended model works better than the base model for HBW trips. This evidence supports that land use measures play a significant role in influencing individual mode choice behaviors for HBW trips.

Table 5.9 MNL Model of Mode Choice for Home-based Work Trips.

Variables	Base Model			Extended Model		
	Estimate	Std. err.	p-value	Estimate	Std. err.	p-value
Constant (SR) [1)]	-0.4028	0.231	0.082	-0.2521	0.252	0.317
Constant (TR)	-2.4163	0.572	0.000	-2.8180	1.160	0.015
Constant (WB)	-0.9942	0.525	0.058	-1.4531	1.140	0.202
Travel time (DA)	-0.0854	0.014	0.000	-0.0814	0.014	0.000
Travel time (SR)	-0.0766	0.010	0.000	-0.0745	0.010	0.000
Travel time (TR)	-0.0388	0.007	0.000	-0.0382	0.007	0.000
Travel time (WB)	-0.0578	0.008	0.000	-0.0554	0.008	0.000
Sex (DA, SR)	0.5954	0.275	0.030	0.6746	0.286	0.018
Age (DA)	0.1748	0.031	0.000	0.1738	0.031	0.000
Bike use (TR,WB)	0.9157	0.197	0.000	0.9903	0.205	0.000
Household size (SR,TR,WB)	0.4230	0.036	0.000	0.4241	0.036	0.000
Vehicles in household (DA)	0.5396	0.054	0.000	0.5344	0.054	0.000
Household income (DA)	0.0651	0.014	0.000	0.0679	0.014	0.000
Population density at O (DA) [1) 2)]				0.0196	0.010	0.058
Population density at D (DA,SR)				-0.0502	0.025	0.046
Employment density at O (TR,WB)				0.0624	0.013	0.000
Employment density at D (DA,SR)				-0.0049	0.001	0.000
Dissimilarity index at O (TR,WB)				1.7535	1.188	0.140
Dissimilarity index at D (DA,SR)				1.4888	1.096	0.174
Connectivity at O (TR,WB)				0.5777	0.672	0.390
Connectivity at D (DA,SR)				1.3564	0.788	0.085
Sample size	6239			6239		
Log Likelihood ($\mathcal{L}$) at converge	-2181.98			-2157.18		
Goodness-of-fit index: $\rho^2, \bar{\rho}^2$	0.7477, 0.7462			0.7506, 0.7482		
Model improvement test: -2[$\mathcal{L}(\mathcal{B}) - \mathcal{L}(\mathcal{E})$]	χ^2 = 49.594, df = 8, Prob. <0.001					

Note: 1) DA = driving-alone, SR = shared-ride, TR = transit, WB = walk/bike. Parenthesis indicates the modes to which the variable is specified.
2) O = trip origin, D = trip destination

Table 5.10 exhibits the results of the MNL choice model for HBO trips where the base and extended model have same patterns for travel times and socioeconomic attributes. Contrary to the model for HBW trips, two travel costs specific to automobile modes and two travel times specific to transit and alternative modes are introduced. It is because cost variables seem to be more important than travel times specific to automobile modes mainly for shopping and recreational trips. As a result, four mode attributes prove all negatively associated with each specific mode, which is consistent with the theory. The effects of socioeconomics are compatible with those for HBW trips except for gender. Females are no more likely to drive for HBO trips.

Many differences are observed in the association of land use measures at trip ends with individual mode choice behavior for HBO trips. Six land use measures have meaningful impacts. More land use diversity and design factors than density measures become significant. Among density variables, only population density at origin shows significance. Dissimilarity and road length at both trip ends and connectivity at origin are significantly associated with the probability of choosing specific modes. In specific, an increase in dissimilarity index enhances the likelihood to choose non-automobile modes at origin and automobile modes at destination at the same time. In addition, two design measures, connectivity and road length at origin promote travelers to use non-automobile modes. However, more road length at destination is significantly associated with more chances of driving automobiles. In summary, land use measures obviously contribute to model improvement in terms of χ^2 model improvement test. It implies that the MNL model for HBO trips can be significantly enhanced with full considerations of land use attributes.

Table 5.10 MNL Model of Mode Choice for Home-based Other Trips.

Variables	Base Model			Extended Model		
	Estimate	Std. err.	p-value	Estimate	Std. err.	p-value
Constant (SR) [1)]	1.4982	0.124	0.000	1.9281	0.138	0.000
Constant (TR)	-2.5469	0.328	0.000	-2.2468	0.704	0.001
Constant (WB)	-0.5214	0.269	0.052	-0.1985	0.689	0.699
Travel cost (DA)	-1.6344	0.093	0.000	-1.6056	0.094	0.000
Travel cost (SR)	-3.7538	0.186	0.000	-3.7295	0.188	0.000
Travel time (TR)	-0.1593	0.012	0.000	-0.1564	0.012	0.000
Travel time (WB)	-0.1836	0.010	0.000	-0.1822	0.010	0.000
Sex (DA, SR)	0.0769	0.131	0.558	0.0855	0.132	0.531
Age (DA)	0.1713	0.013	0.000	0.1728	0.013	0.000
Bike use (TR,WB)	0.8122	0.082	0.000	0.8119	0.084	0.000
Household size (SR,TR,WB)	0.3301	0.022	0.000	0.3164	0.022	0.000
Vehicles in household (DA)	0.3130	0.030	0.000	0.3354	0.030	0.000
Household income (DA)	0.0771	0.007	0.000	0.0764	0.007	0.000
Population density at O (DA) [1) 2)]				0.0412	0.006	0.000
Population density at D (DA,SR)				-0.0146	0.017	0.395
Employment density at O (TR,WB)				-0.0149	0.017	0.388
Employment density at D (DA,SR)				-0.0013	0.003	0.675
Dissimilarity index at O (TR,WB)				2.3898	0.584	0.000
Dissimilarity index at D (DA,SR)				3.3233	0.535	0.000
Connectivity at O (TR,WB)				1.4859	0.504	0.003
Connectivity at D (DA,SR)				0.3848	0.432	0.373
Road length at O (TR,WB)				0.1386	0.065	0.034
Road length at D (DA,SR)				0.1345	0.052	0.009
Sample size	10413			10413		
Log Likelihood ($\mathcal{L}$) at converge	-6981.78			-6921.83		
Goodness-of-fit index: $\rho^2, \bar{\rho}^2$	0.5163, 0.5154			0.5205, 0.5189		
Model improvement test: $-2[\mathcal{L}(\mathcal{B}) - \mathcal{L}(\mathcal{E})]$	χ^2 = 119.901, df = 10, Prob. <0.001					

Note: 1) DA = driving-alone, SR = shared-ride, TR = transit, WB = walk/bike. Parenthesis indicates the modes to which the variable is specified.

2) O = trip origin, D = trip destination

5.2.2 Results of Structural Equation Models

As expressed in research design section of Chapter III, only two mode choices are considered in the specification process of the causal relationship models. It is mainly because the directed acyclic graph (DAG) has a methodological limitation in handling multiple choice data; rather, binary data with alterative specific specification can be working properly with DAGs. Therefore, two mode options, automobile vs. non-automobile are taken into account for specifying causal models. Automobile mode includes both driving-alone and shared-ride.

The estimation results of binomial logit models for HBW and HBO trips are presented in Table 5.11. They are to set up the basis for comparing the results with the outcomes of following causal relationship models. Mode attribute (travel time differential) is positively significant for both trip purposes, which indicates that the bigger the difference between driving time and walk time from home to destinations, the more likely trip-makers to use automobile mode.

Socioeconomic characteristics are significantly associated with the probability of automobile choice except household income for HBW trips. Their signs and effects are generally congruous with the theory and the arguments of previous studies. However, there are some variations in the impacts of land use measures on automobile choice probability. For HBW trips, only two employment density variables out of ten measures are significant. Employment densities at both trip ends are negatively correlated for HBW trips. On the contrary, diversity and design measures become significant for HBO trips, and population density at origin shows significant relationship. In addition, several statistics suggest that the models for both trip purposes are significantly improved by the inclusion of various land use measures.

Table 5.11 Binomial Logit Models for Home-based Trips.

	Home-based Work Trips			Home-based Other Trips		
Variables	Estimate	Std. err.	p-value	Estimate	Std. err.	p-value
Constant	2.4195	1.528	0.113	1.5878	0.737	0.031
Travel time differential [1)]	0.0512	0.008	0.000	0.0411	0.006	0.000
Household size	-0.3765	0.107	0.000	-0.3479	0.050	0.000
Vehicles in household	0.5351	0.185	0.004	0.6094	0.091	0.000
Household income	-0.0174	0.046	0.708	0.0685	0.022	0.002
Bike use	-1.1163	0.215	0.000	-0.7652	0.086	0.000
Single-family housing	1.0968	0.362	0.002	0.5460	0.206	0.008
Population density at O [2)]	-0.0233	0.039	0.547	-0.0462	0.019	0.015
Population density at D	-0.0416	0.030	0.168	0.0060	0.017	0.731
Employment density at O	-0.0489	0.015	0.001	0.0219	0.018	0.235
Employment density at D	-0.0036	0.001	0.002	0.0004	0.003	0.909
Dissimilarity index at O	-0.9035	1.214	0.457	-1.3833	0.587	0.018
Dissimilarity index at D	1.5594	1.070	0.145	3.3668	0.521	0.000
Connectivity at O	-0.3965	1.026	0.699	-1.3363	0.498	0.007
Connectivity at D	-0.4930	0.902	0.585	0.2952	0.408	0.469
Road length at O	0.1220	0.152	0.422	-0.0532	0.070	0.447
Road length at D	0.1101	0.109	0.310	0.1017	0.049	0.039
Sample size	6239			10413		
Goodness-of-fit index: $\rho^2, \bar{\rho}^2$	0.9443, 0.9404			0.8576, 0.8553		
Model improvement test: $-2[\mathcal{L}(\mathcal{B}) - \mathcal{L}(\mathcal{E})]$	χ^2 = 27.14, df = 10, Prob. = 0.0025			χ^2 = 57.10, df = 10, Prob. <0.001		

Note: 1) Travel time differential = walk time – driving time
2) O = trip origin, D = trip destination

The structural equation models (SEMs) includes same groups of variables as used in the binomial logit models. Table 5.12 presents the estimated results for both HBW and HBO trips. Each structural model consists of two main parts in addition to the intercept part: automobile choice (*Automobile ON*) and travel time part (*Travel time* *differential ON*). The automobile choice part shows the estimated results of binomial logit regression where non-automobile mode is the reference choice. Therefore, the estimated outcomes are very similar to those of the binomial logit models shown in Table 5.11. In summary, land use and development patterns are causally associated with the probability of automobile choice in terms of the causality based on the SEM approach. Employment density measures are significant for HBW trips; diversity and design measures as well as population density at origin have significant causal effects on the likelihood of automobile choice.

Attention is paid to interpreting the travel time part for both trip purposes. It is assumed that household income and various dimensions of land use measures are causally connected with travel time differential. They are also expected to have indirect relationship with automobile choice probability through the travel time. The results exhibit same patterns for HBW and HBO trips except that the magnitudes of the parameters for HBW trips are generally larger than their counterparts for HBO trips. Specifically, household income positively affects the travel time differential. Most land use measures at trip origin (population and employment density, connectivity and road length measure) reduce the travel time differential as they increase. Increases in employment density and roadway length at destination widen the differential. Dissimilarity indices at both trip ends have opposite impacts.

Table 5.12 Structural Equation Models of Binary Mode Choice for Home-based Trips.

Variables	Home-based Work Trips			Home-based Other Trips		
	Estimates	Std. err.	p-value	Estimates	Std. err.	p-value
Automobile ON						
Travel time differential [1)]	0.051	0.008	0.000	0.041	0.006	0.000
Household size	-0.376	0.107	0.000	-0.348	0.050	0.000
Vehicles in household	0.535	0.185	0.004	0.609	0.091	0.000
Household income	-0.017	0.046	0.708	0.068	0.022	0.002
Bike use	-1.116	0.215	0.000	-0.765	0.086	0.000
Single-family housing	1.097	0.362	0.002	0.546	0.206	0.008
Population density at O [2)]	-0.023	0.039	0.547	-0.046	0.019	0.015
Population density at D	-0.042	0.030	0.168	0.006	0.017	0.731
Employment density at O	-0.049	0.015	0.001	0.022	0.018	0.235
Employment density at D	-0.004	0.001	0.002	0.000	0.003	0.909
Dissimilarity index at O	-0.904	1.214	0.457	-1.383	0.587	0.018
Dissimilarity index at D	1.560	1.070	0.145	3.367	0.521	0.000
Connectivity at O	-0.394	1.026	0.701	-1.336	0.498	0.007
Connectivity at D	-0.494	0.902	0.584	0.293	0.408	0.472
Road length at O	0.122	0.152	0.423	-0.053	0.070	0.447
Road length at D	0.110	0.109	0.310	0.102	0.049	0.039
Travel time differential ON						
Household income	1.742	0.252	0.000	0.669	0.097	0.000
Population density at O	-1.165	0.237	0.000	-0.573	0.106	0.000
Population density at D	0.119	0.188	0.525	0.063	0.087	0.472
Employment density at O	-1.376	0.236	0.000	-0.270	0.090	0.003
Employment density at D	0.048	0.010	0.000	0.126	0.011	0.000
Dissimilarity index at O	23.656	6.578	0.000	8.816	2.817	0.002
Dissimilarity index at D	-42.636	6.328	0.000	-15.673	3.125	0.000
Connectivity at O	-31.634	5.302	0.000	-4.842	2.250	0.031
Connectivity at D	0.401	5.056	0.937	-1.781	2.236	0.426
Road length at O	-5.221	0.869	0.000	-2.400	0.376	0.000
Road length at D	5.712	0.549	0.000	1.144	0.245	0.000
Intercept						
Travel time differential	106.652	8.257	0.000	38.208	3.572	0.000
Sample size	6239			10413		
Log Likelihood (H_0 value)	-57204.386			-90881.534		
Information Criteria						
No. of free parameters	42			42		
Akaike (AIC)	114492.772			181847.067		
Bayesian (BIC)	114775.793			182151.601		

Note: 1) Travel time differential = walk time – driving time
2) O = trip origin, D = trip destination

5.2.3 Results of Directed Acyclic Graphs

The directed acyclic graphs (DAGs) depend upon the multivariate distribution of variables from the observational data. As clarified in the previous chapter, some assumptions are required to apply this method including causal Markov condition, faithfulness and causal sufficiency. In order to make an analysis of DAGs for each trip purpose, a lower triangular correlation matrix should be computed. This input of the unconditional correlation matrix between pairs of variables is the starting point for estimating causal graphs in the TETRAD III algorithm. Then, it explores both conditional and unconditional independence relations among input variables. This study employs 17 variables as considered in the previous SEMs: one binary choice variable (auto choice), one mode attribute (travel time differential), five socioeconomic characteristics (household size, vehicle ownership, income, bike use and single-family residence), and ten land use measures at trip origin and destination (population density, employment density, dissimilarity index, connectivity and road length at both trip ends).

To obtain reasonable results, three constraints are imposed in the estimation process. One is that four socioeconomic variables except bike use precede travel time, bike use and land use measures. It implies that these socioeconomics cannot be effects of others. Another constraint is that land use variables at origin do not cause those at destination, and vice versa. Last one is that land use measures can only be causes of travel time differential, bike use and auto choice variables. It suggests that opposite causation from the latter variables to land use patterns is a long-term process; moreover, it is beyond the scope of this study. A 1% significance level is applied to produce the directed graphs as suggested by Spirtes et al. (2000).

The result of estimated DAG for HBW trips is illustrated in Figure 5.4. The direct graph indicates that automobile choice for HBW trips is causally affected by five factors: travel time differential (TRAVEL TIME: +), bike use (BIKE USE: –), number of vehicles (NOVEHICLE: +), single-family residence (SF RESID: +), and employment density at origin (O_EMPDEN: –). Increased difference between driving time and walk time from an origin to a destination promotes individual trip-makers to drive. More vehicles available and single-family residence causally affect the increase in the chances of making automobile choices. On the contrary, an increase in bike use experiences discourages travelers to use an automobile. In particular, only one land use measure shows significant causal connection to automobile choice probability. Employment density at trip origin has a negative causal impact on the likelihood of automobile mode choice for HBW trips. These results are quite consistent with those of the SEMs for HBW trips except that household size and employment density at destination are not causally connected with automobile choice in the DAGs.

Additional attention needs to be paid to travel time differential. According to the estimated direct graphs, it is causally influenced by two socioeconomic attributes and many land use variables: household size (HHSIZE: +), household income (INCOME: +), population density at origin (O_POPDEN: –), employment density at both origin (O_EMPDEN: –) and destination (D_EMPDEN: +), road length at both origin (O_ROADMI: –) and destination (D_ROADMI: +), connectivity at origin

(O_CONNECT: –), and dissimilarity at destination (D_DISSINDEX: –). These causal connections are similar to the results of SEM estimation except that dissimilarity index at destination is no more direct cause in the directed graphs. The result suggests these variables indirectly affect automobile choice probability through travel time differential. For instance, increases in land use variables at origin reduce travel time differential; decreased time differential then lowers the chances of automobile mode choice.

There are two colliders, population density at destination and dissimilarity index at origin at which causal information flowing from other variables comes into collision. Bi-directed or double-headed edges are also observed between land use variables at both trip ends. They suggest that there should be an unmeasured common cause or a latent variable between two variables. For example, roadway development and improvement can be a common cause between connectivity and road length at trip origin. There are undirected edges between socioeconomic factors. Personal judgment and the arguments of relevant studies are introduced to provide causal orientation for each pair of variables.

Figure 5.5 displays the result of estimated directed graphs for HBO trips. It is found that five variables causally influence the likelihood to drive for HBO trips: travel time differential (TRAVEL TIME: +), bike use (BIKE USE: –), number of vehicles (NOVEHICLE: +), single-family residence (SF RESID: +), and dissimilarity index at destination (D_DISINDEX: +). Travel time, vehicle ownership and single-family residential type have positive causal relationship with the chances of automobile choice for HBO trips. In addition, more experiences of using a bike reduce the likelihood that an individual drives for shopping and recreational trips. When compared with the result for HBW trips, they are quite similar except that dissimilarity at destination instead of employment density at origin becomes direct cause. The estimated outcomes seem to be little consistent with those of the SEM estimation for HBO trips because the latter claims that five land use measures are causally connected with driving probability. Furthermore, the positive sign of dissimilarity index implies that higher level of land use mix at destination encourages people to drive more. Details in the issues will be discussed later.

From the perspective of the causes of travel time differential for HBO trips, it is causally explained by household income (INCOME: +), number of vehicles (NOVEHICLE: +), population density (O_POPDEN: –) and roadway length (O_ROADMI: –) at trip origin, and employment density (D_EMPDEN: +) and dissimilarity (D_DISSINDEX: –) at destination. Although it is argued that three dimensions of land use patterns all causally influence travel time differential, only four land use measures as direct causes are relatively fewer than eight land use variables in the SEM results. Overall, it is confirmed that land use variables have indirect impacts on automobile choice probability through travel time differential.

Four colliders are observed in the directed graphs: employment density, dissimilarity index and connectivity at trip origin, and population density at destination. A number of bi-directed edges between land use measures suggest the existence of unmeasured common causes between them. Improved facilities for alternative transportation, for instance, can be a common cause between travel time differential and bike use. Same approach as used for HBW trips is applied to construct causal connections between socioeconomic variables.

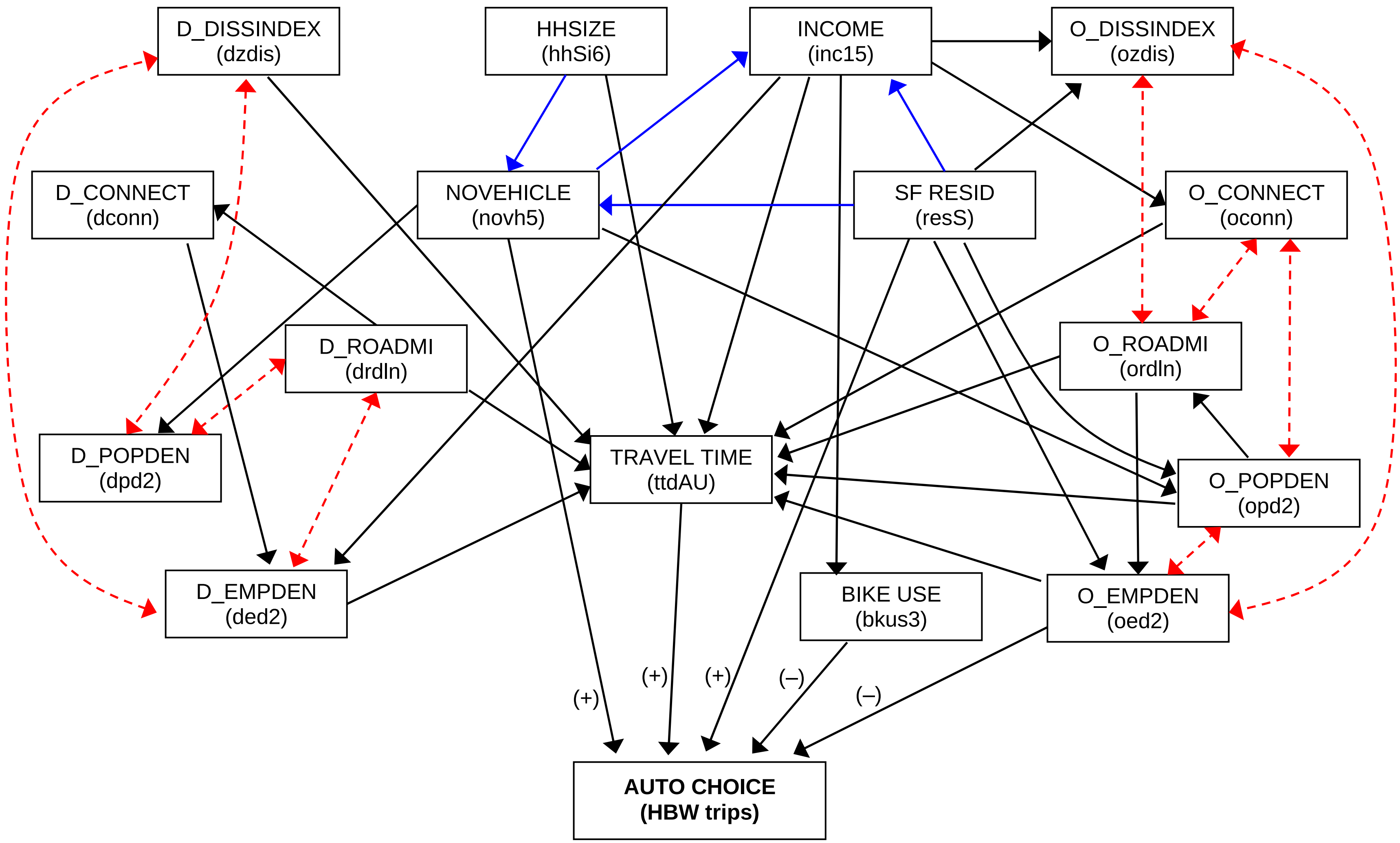


Figure 5.4 Directed Acyclic Graphs (DAGs) on Binary Mode Choice for Home-based Work Trips (1% significance level).
Note: Double-headed or bi-directed edges, x1↔x2 in a pattern suggest that there is a latent common cause between two variables.
Names in parentheses indicate variable names that are used in the analytical process.

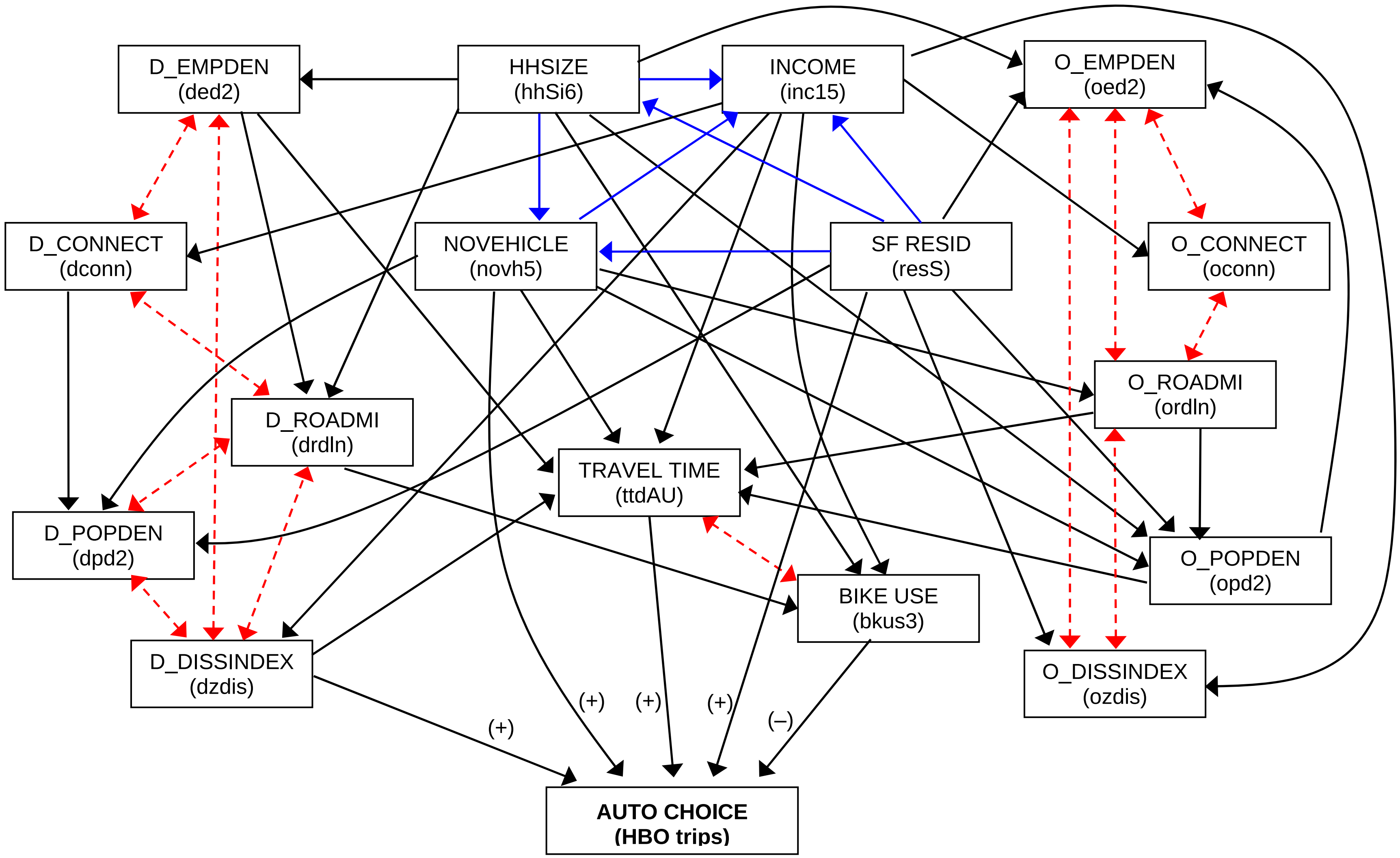


Figure 5.5 Directed Acyclic Graphs (DAGs) on Binary Mode Choice for Home-based Other Trips (1% significance level).
Note: Double-headed or bi-directed edges, x1↔x2 in a pattern suggest that there is a latent common cause between two variables.
Names in parentheses indicate variable names that are used in the analytical process.

5.3 Household Automobile Trip Generation Models

5.3.1 Results of Negative Binomial Models

The negative binomial model is introduced as a travel demand model for household automobile trip generation. Based on the travel demand theory, important sets of explanatory variables should be taken into the modeling process. They contain travel cost ($/trip), household socioeconomic characteristics (household size, vehicle ownership, total income, income squared and single-family residence), and land use measures at trip origin (population density, employment density, entropy index, connectivity and road length measure). In order to understand their effects on household automobile trip rates in depth, the travel demand models are estimated for different purposes: total trips, total home-based trips, HBW trips and HBO trips. A comparison is made between a base model and an extended model for each trip purpose to examine if land use measures are collectively significant in improving the travel demand model.

The estimation results of the negative binomial models for both total and total home-based trips are shown in Table 5.13. The patterns of travel cost and household socioeconomic characteristics in the extended models are same as those in the base models. The results of two extended models are quite similar to each other. Travel cost variables which include operation and maintenance costs are negatively associated with total household auto trips.

Household socioeconomic attributes have positively significant effects on automobile trip frequency. But household income squared is exceptional, which is inconsistent with the demand theory of household trip rates established by Boarnet and Sarmiento (1998) and Boarnet and Crane (2001a). Specifically, a household tends to make more auto trips as household members and vehicles increase. Higher household income promotes trip-makers to depend more on automobiles, which results in an increase in household trip frequency. Single-family households are more likely to make auto trips than multi-family households.

One notable feature of the models is that no land use measure at trip origin is significant in estimating household automobile trip rates. As shown in the table, three dimensions of land use characteristics at origin do not have meaningful impacts on household automobile trip generation. As a result, likelihood ratio tests show a group of land use measures do not significantly contribute to model improvement for both total and total home-based trips.

Table 5.13 Household Auto Trip Generation Models for Total Trips.

	Total Trips				Total Home-based Trips			
	Base Model		Extended Model		Base Model		Extended Model	
Variables	Est.	χ^2	Est.	χ^2	Est.	χ^2	Est.	χ^2
Constant	1.081	**420.17**	1.196	**222.22**	0.606	**141.25**	0.628	**68.52**
Travel cost ($ / trip)	-0.088	**448.98**	-0.089	**441.77**	-0.055	**250.05**	-0.054	**230.43**
Household size	0.124	**327.03**	0.124	**321.43**	0.141	**521.74**	0.142	**522.51**
Vehicles in household	0.156	**197.81**	0.155	**194.13**	0.168	**282.36**	0.168	**282.66**
Household income	0.053	**20.87**	0.052	**20.47**	0.043	**15.26**	0.043	**15.35**
Income squared	-0.001	2.21	-0.001	2.21	-0.001	1.85	-0.001	2.01
Single-family housing	0.060	**3.85**	0.054	***2.90***	0.065	**4.85**	0.064	**4.30**
Pop. density at O [1)]			0.001	0.05			-0.001	0.06
Emp. density at O			0.000	0.01			0.001	0.09
Entropy index at O			-0.070	1.28			-0.082	2.07
Connectivity at O			-0.048	0.79			-0.008	0.02
Road length at O			-0.007	0.68			0.009	1.29
Dispersion	0.123 (p<0.001)				0.028 (p<0.001)			
Sample size	3976				3973			
Log Likelihood ($\mathcal{L}$)	32826.17		32828.13		14377.50		14379.54	
Model improvement: -2[$\mathcal{L}(\mathcal{B}) - \mathcal{L}(\mathcal{E})$]	χ^2 = 3.928, df = 5, Prob. = 0.560				χ^2 = 4.073, df = 5, Prob. = 0.539			

Note: 1) O = trip origin
2) Estimates in bold are significant at 5% level; estimates in italic bold are significant at 10% level (two-tailed test).

Household automobile trip generation models for HBW and HBO trips are presented in Table 5.14. There are similarities between the outcomes of two extended models; on the other hand, differences clearly exist between total trip and home-based trip purposes. Travel costs have negative relationship with household auto trip frequency for both travel purposes. Based on the magnitude of the coefficients, HBO trips are more sensitive to travel cost than HBW trips.

Major dissimilarities lie in household socioeconomic variables when compared with total trip purposes. Household size and vehicle ownership still have significantly positive effects. But both household income and income squared appear not to play an important role in estimating household auto trip rates. Single-family households also do not make more auto trips for commuting purpose; however, they rely more on automobile trips for other trips.

A similarity exists in the role of land use measures in household automobile trip generation models. No land use measure is meaningful for explaining household automobile trip rates. Thus, there is no significance of model improvement tests.

Table 5.14 Household Auto Trip Generation Models for Home-based Trips.

	Home-based Work Trips				Home-based Other Trips			
	Base Model		Extended Model		Base Model		Extended Model	
Variables	Est.	χ^2	Est.	χ^2	Est.	χ^2	Est.	χ^2
Constant	0.221	**6.59**	0.204	***2.85***	0.723	**130.59**	0.679	**49.33**
Travel cost ($ / trip)	-0.018	**28.73**	-0.017	**23.45**	-0.048	**84.59**	-0.047	**76.64**
Household size	0.015	2.47	0.016	***2.80***	0.149	**348.5**	0.151	**347.90**
Vehicles in household	0.208	**203.00**	0.210	**204.78**	0.064	**23.88**	0.064	**24.21**
Household income	0.024	1.76	0.024	1.75	0.005	0.14	0.005	0.13
Income squared	-0.001	1.15	-0.001	1.20	0.001	0.89	0.001	0.92
Single-family housing	0.071	2.49	0.061	1.73	0.070	***3.38***	0.079	**3.95**
Pop. density at O [1)]			0.004	1.12			0.001	0.08
Emp. density at O			-0.005	1.61			0.002	0.53
Entropy index at O			-0.072	0.71			-0.022	0.08
Connectivity at O			-0.026	0.12			0.038	0.36
Road length at O			0.016	1.79			0.001	0.01
Dispersion	-0.083 ($p<0.001$)				0.078 ($p<0.001$)			
Sample size	2539				3461			
Log Likelihood ($\mathcal{L}$)	-500.61		-496.69		6656.22		6657.09	
Model improvement: $-2[\mathcal{L}(\mathcal{B}) - \mathcal{L}(\mathcal{E})]$	$\chi^2 = 7.845$, df = 5, Prob. = 0.165				$\chi^2 = 1.739$, df = 5, Prob. = 0.884			

Note: 1) O = trip origin

2) Estimates in bold are significant at 5% level; estimates in italic bold are significant at 10% level (two-tailed test).

5.3.2 Results of Structural Equation Models

Structural models of household automobile trip generation for different travel purposes have same modeling structure consisting of three major parts: household automobile trip frequency part (*Auto trips ON*), travel cost ($/trip) part (*Travel Cost ON*) and household income part (*Household income ON*). Table 5.15 and 5.16 reveal that the estimation results of the automobile trip frequency part are similar to the travel demand model outcomes for same travel purpose. Land use measures at origin do not significantly affect household automobile trip rates when other variables are kept constant.

Attention needs to be focused on travel cost model to investigate whether land use attributes around home places causally influence travel cost based on assumed causality. The results in Table 5.15 for both total and total home-based trips indicate that household income and land use measures except entropy are statistically significant. Higher household income affects higher travel cost. Land use factors at origin negatively affect travel cost, but entropy index is insignificant. In other words, increased density and improved neighborhood design lead to the reduction in travel cost for total trips.

The results of travel cost model for both total trip purposes also work for HBW trips; however, several distinctions are observed in the results for HBO trips as shown in Table 5.16. Household income is no longer significant in the effect on travel cost. Travel cost per trip is affected by household economic status for commute trips, but not for shopping and recreational trips. Entropy index becomes negatively significant; however, employment density has no significant relationship with HBO trip frequency.

Furthermore, household income model displays significant impacts of socioeconomic characteristics on household income. As expected, household size, vehicle availability and single-family residence are all positively affect total household income. These connections are in effect for all travel purposes.

5.3.3 Results of Directed Acyclic Graphs

Eleven variables are taken into TETRAD III algorithm in the form of lower triangular correlation matrix to make causal graphs for different purposes. They are composed of four groups: one travel behavior outcome (automobile trip frequency), one travel cost ($/trip), four household socioeconomic characteristics (household size, vehicle ownership, total income, and single-family residence), and five land use measures at trip origin (population and employment density, entropy index, connectivity and road length measures). Same three restrictions are established on the estimation process. Also, undirected edges between household socioeconomic factors are causally oriented based on personal reasoning and evidence of related research.

Table 5.15 Structural Equation Models of Household Auto Trip Generation for Total Trips.

	Total Trips			Total Home-based Trips		
Variables	Estimates	Std. err.	p-value	Estimates	Std. err.	p-value
Auto trips ON						
Travel cost ($ / trip)	-0.089	0.004	0.000	-0.054	0.003	0.000
Household size	0.123	0.007	0.000	0.142	0.006	0.000
Vehicles in household	0.155	0.011	0.000	0.168	0.010	0.000
Household income	0.052	0.012	0.000	0.043	0.011	0.000
Household income squared	-0.001	0.001	0.153	-0.001	0.001	0.153
Single-family housing	0.054	0.034	0.115	0.064	0.030	0.035
Population density at O [1)]	0.001	0.002	0.817	-0.001	0.002	0.802
Employment density at O	0.000	0.002	0.904	0.001	0.002	0.743
Entropy index at O	-0.070	0.061	0.247	-0.082	0.056	0.139
Connectivity at O	-0.048	0.058	0.403	-0.008	0.052	0.884
Road length at O	-0.007	0.009	0.432	0.009	0.008	0.256
Travel cost ON						
Household income	0.050	0.010	0.000	0.064	0.011	0.000
Population density at O	-0.029	0.010	0.002	-0.039	0.010	0.000
Employment density at O	-0.030	0.006	0.000	-0.027	0.007	0.000
Entropy index at O	-0.213	0.282	0.450	-0.281	0.303	0.353
Connectivity at O	-0.863	0.293	0.003	-1.089	0.300	0.000
Road length at O	-0.206	0.038	0.000	-0.242	0.041	0.000
Household income ON						
Household size	0.265	0.043	0.000	0.265	0.043	0.000
Vehicles in household	1.223	0.064	0.000	1.223	0.064	0.000
Single-family housing	1.075	0.177	0.000	1.076	0.178	0.000
Intercept						
Auto trips	1.319	0.083	0.000	0.770	0.077	0.000
Travel cost	4.398	0.332	0.000	4.803	0.344	0.000
Household income	4.275	0.187	0.000	4.277	0.187	0.000
Dispersion	0.123	0.006	0.000	0.028	0.005	0.000
Sample size		3976			3973	
Log Likelihood (H_0 value)		-42316.490			-40793.704	
Information Criteria						
Free parameters		36			36	
Akaike (AIC)		84704.979			81659.407	
Bayesian (BIC)		84931.349			81885.749	

Note: 1) O = trip origin

Table 5.16 Structural Equation Models of Household Auto Trip Generation for Home-based Trips.

Variables	Home-based Work Trips			Home-based Other Trips		
	Estimates	Std. err.	p-value	Estimates	Std. err.	p-value
Auto trips ON						
Travel cost ($ / trip)	-0.017	0.003	0.000	-0.047	0.005	0.000
Household size	0.015	0.008	0.068	0.151	0.008	0.000
Vehicles in household	0.211	0.013	0.000	0.064	0.013	0.000
Household income	0.023	0.014	0.107	0.005	0.013	0.708
Household income squared	-0.001	0.001	0.214	0.001	0.001	0.331
Single-family housing	0.059	0.032	0.063	0.079	0.040	0.046
Population density at O [1)]	0.003	0.003	0.215	0.001	0.003	0.771
Employment density at O	-0.005	0.003	0.098	0.002	0.003	0.473
Entropy index at O	-0.072	0.075	0.336	-0.022	0.073	0.767
Connectivity at O	-0.027	0.067	0.688	0.038	0.067	0.567
Road length at O	0.015	0.010	0.143	0.001	0.011	0.938
Travel cost ON						
Household income	0.092	0.021	0.000	0.011	0.010	0.298
Population density at O	-0.063	0.018	0.001	-0.041	0.010	0.000
Employment density at O	-0.061	0.017	0.000	-0.002	0.007	0.787
Entropy index at O	0.285	0.540	0.597	-0.861	0.304	0.005
Connectivity at O	-2.236	0.525	0.000	-0.654	0.271	0.016
Road length at O	-0.322	0.071	0.000	-0.227	0.039	0.000
Household income ON						
Household size	0.200	0.052	0.000	0.325	0.045	0.000
Vehicles in household	0.958	0.079	0.000	1.213	0.068	0.000
Single-family housing	1.293	0.226	0.000	1.262	0.199	0.000
Intercept						
Auto trips	0.448	0.105	0.000	0.829	0.092	0.000
Travel cost	7.499	0.632	0.000	4.293	0.336	0.000
Household income	6.208	0.220	0.000	4.010	0.210	0.000
Dispersion	-0.083	0.004	0.000	0.078	0.007	0.000
Sample size		2539			3461	
Log Likelihood (H_0 value)		- 25471.792			- 34864.910	
Information Criteria						
Free parameters		34			36	
Akaike (AIC)		51011.585			69801.820	
Bayesian (BIC)		51210.129			70023.195	

Note: 1) O = trip origin

Figure 5.6 and 5.7 display the results of estimated directed graphs at 1 % significance level for total and total home-based trips. They show that the number of household automobile trips for total trips is causally connected with four factors: travel cost (TRAVEL COST: –), household size (HHSIZE: +), number of vehicles (NOVEHICLE: +), and household income (INCOME: +). Household members tend to reduce their auto trips as travel cost gets higher. Also, bigger household size, and more available vehicles and income in a household are causally connected with more automobile trips for both total and total home-based trips. It should be noted that no land use measure is a direct cause of total automobile trips. These causal connections are identical to the results of the structural models shown in Table 5.15.

According to the figures, travel cost is causally affected by many factors: vehicle ownership (NOVEHICLE: +), population density (O_POPDEN: –), employment density (O_EMPDEN: –), connectivity (O_CONNECT: –) and road length (O_ROADMI: –). Household income (INCOME: +) is added for total home-based trips (see Figure 5.7). Household automobile trips increase as vehicles and income in a household increases. It is noteworthy that land use measures except entropy index have negative causal impacts on travel cost. Therefore, it is argued that land use measures are not direct but indirect causes of household automobile trip frequency through travel cost. These causal effects are congruous with the outcomes of the SEMs for total trip purposes (see Table 5.15).

The DAGs have a collider, entropy index taking up causal information from the precedents, but blocks its flow into others. In addition, two bi-directed edges are recognized between employment density and entropy and road length.

The directed graphs for HBW and HBO trips are presented in Figure 5.8 and 5.9, respectively. For HBW trips, only two variables serve as direct causes of household auto trip generation: travel cost (TRAVEL COST: –) and vehicle ownership (NOVEHICLE: +). It makes sense that household size and income are no longer direct causes for commute trips. However, household auto trip rates for HBO trips are causally influenced by same variables as shown for total trip purposes. They are travel cost, household size, vehicle ownership and household income as illustrated in Figure 5.9.

When it comes to the causal relationship with travel cost, some differences between different travel purposes are observable. For HBW trips, positive household income is the only socioeconomic variable that causally affects travel cost. Land use measures except entropy show same patterns as in the SEMs for HBW trips. For HBO trips, both household size and vehicle ownership are positive causes of travel cost. Population density, connectivity and road length also affect travel cost, but employment density becomes insignificant by the nature of HBO trips.

One collider (entropy index) for HBW trips and two colliders (employment density and entropy index) for HBO trips are identified. A bi-directed edge between employment density and entropy index is observed for both trip purposes.

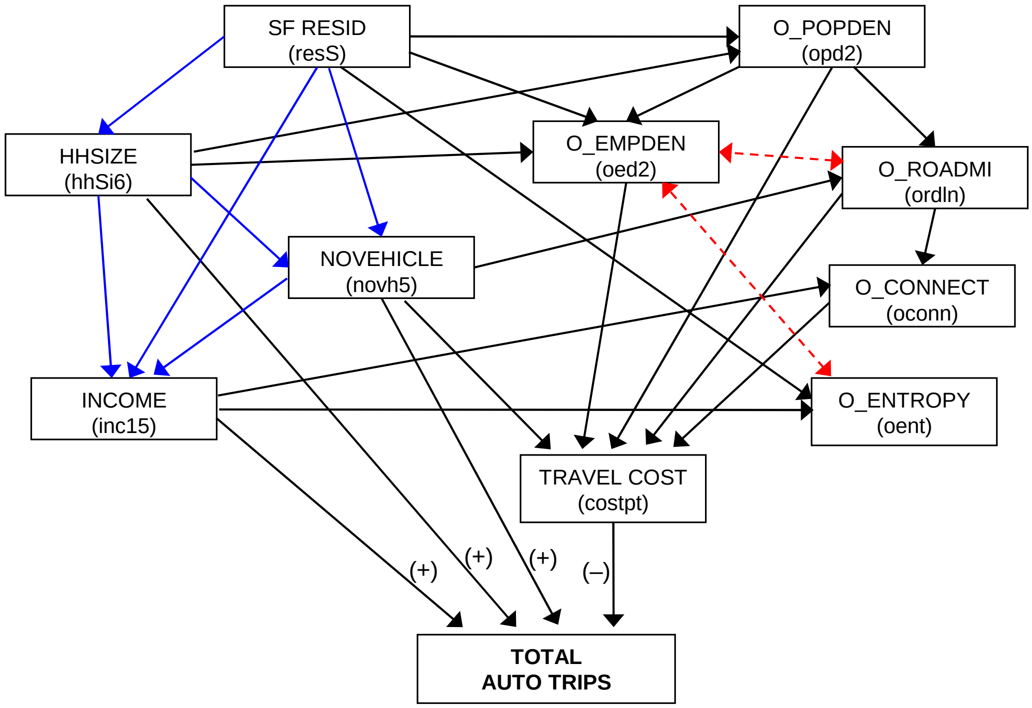


Figure 5.6 DAGs on Household Auto Trip Generation for Total Trips.

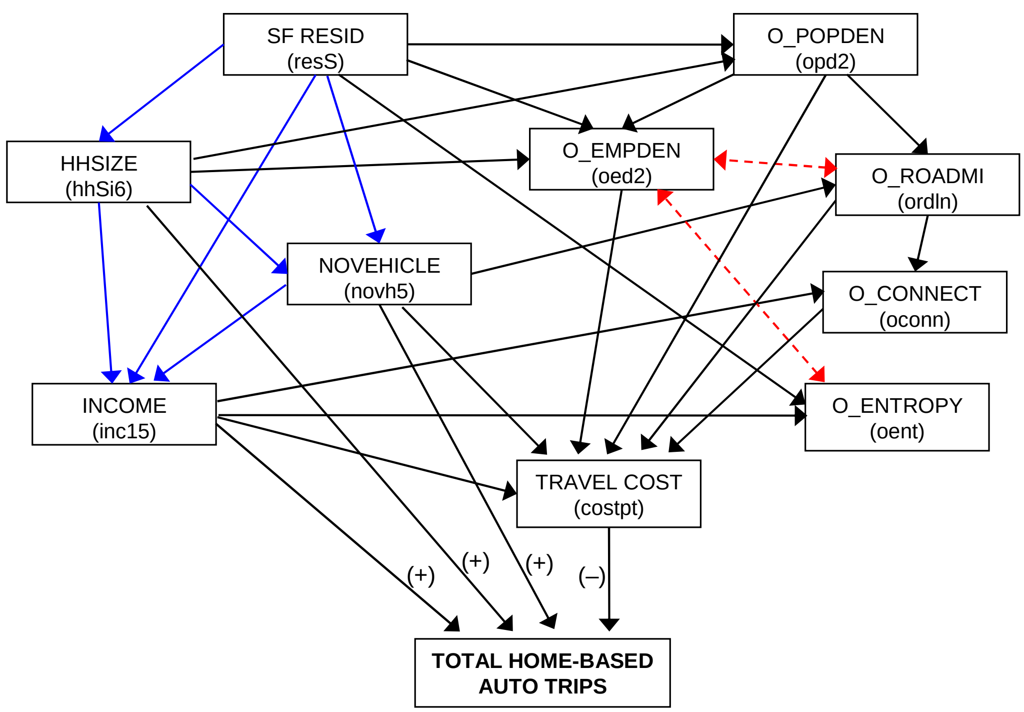


Figure 5.7 DAGs on Household Auto Trip Generation for Total Home-based Trips.
Note: Double-headed edges in a pattern suggest a latent common cause between two variables.

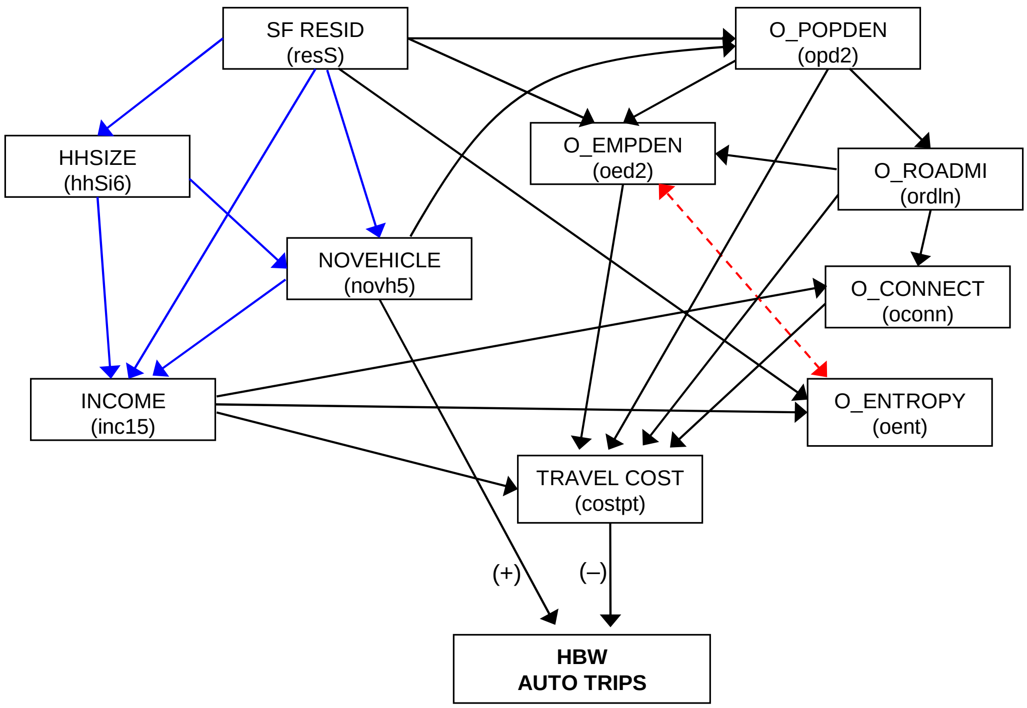


Figure 5.8 DAGs on Household Auto Trip Generation for Home-based Work Trips.

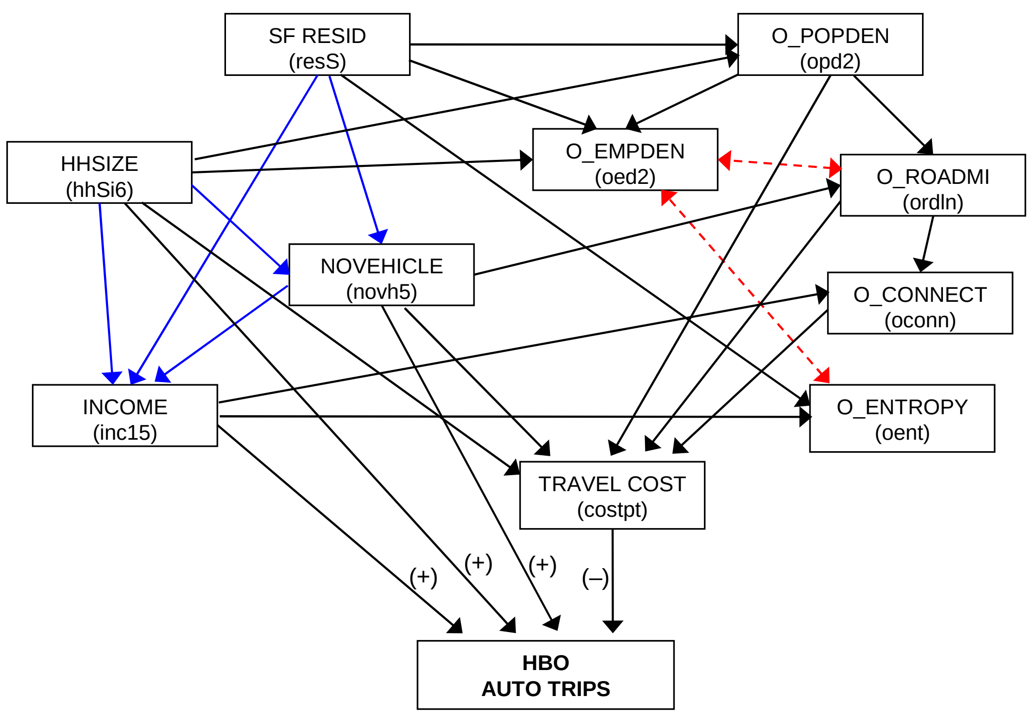


Figure 5.9 DAGs on Household Auto Trip Generation for Home-based Other Trips.
Note: Double-headed edges in a pattern suggest a latent common cause between two variables.

5.4 Household Total VMT Models

5.4.1 Results of OLS Regression Models

To estimate household total vehicle miles traveled (VMT), the ordinary least squares (OLS) regression model is specified where the dependent variable, household total VMT is regressed on a set of explanatory variables. Logarithmic transformation of household total VMT is implemented to have better estimation results. Independent variables consist of travel attributes (travel cost per mile and automobile trip frequency), household socioeconomic characteristics (household size, vehicle ownership, total income and income squared), and land use measures at trip origin (population density, employment density, entropy index, connectivity and road length measure). Both a base and an extended model for each travel purpose are estimated for each travel purpose to examine the role of land use variables in model improvement.

Table 5.17 presents the estimation results of household VMT models for total and total home-based trips. The base models and extended models for both travel purposes show same patterns in terms of the effects of travel attributes and socioeconomic factors. Higher travel cost per mile and fewer auto trips are associated with the reduction of household VMT. Vehicle ownership is positively connected with household VMT. Both household income and income squared suggest that household VMT increases with household income; the intensity of the income effect, however, diminishes as income increases.

According to the extended models, land use measures except entropy index are all significant in explaining household VMT. Population and employment density, connectivity and road length measures have negative effects on household VMT. When explanatory variables are controlled, the extended models explain 41.7% and 37.0% of variations for both total and total home-based trips, respectively. It is also confirmed that land use measures significantly contribute to improving household VMT models for two total trip purposes.

Table 5.17 Household Total VMT Models for Total Trips.

	Total Trips				Total Home-based Trips			
	Base Model		Extended Model		Base Model		Extended Model	
Variables	Est.	*t*	Est.	*t*	Est.	*t*	Est.	*t*
Constant	2.675	**29.58**	3.217	**26.38**	2.333	**23.95**	2.934	**22.33**
Travel cost ($ / mile)	-0.921	**-3.16**	-1.093	**-3.81**	-0.704	**-2.27**	-0.919	**-3.01**
Auto trips	0.090	**32.09**	0.091	**33.18**	0.127	**25.58**	0.131	**26.93**
Household size	0.015	1.47	0.003	0.32	0.001	0.13	-0.014	-1.23
Vehicles in household	0.204	**12.63**	0.182	**11.34**	0.238	**13.49**	0.211	**12.07**
Household income	0.106	**6.61**	0.105	**6.70**	0.090	**5.20**	0.090	**5.30**
Income squared	-0.004	**-4.58**	-0.004	**-4.69**	-0.003	**-3.33**	-0.003	**-3.43**
Pop. density at O [1)]			-0.011	**-3.42**			-0.015	**-4.23**
Emp. density at O			-0.011	**-3.89**			-0.011	**-3.34**
Entropy index at O			0.102	1.18			0.144	1.54
Connectivity at O			-0.310	**-4.07**			-0.329	**-4.01**
Road length at O			-0.044	**-3.55**			-0.051	**-3.81**
Sample size	3976				3973			
$R^2, \overline{R}^2$	0.395, 0.394		0.417, 0.416		0.343, 0.342		0.370, 0.368	
Model improvement: *F-test*	F=29.917, df_1=5, df_2=3964 Prob.< 0.001				F=33.951, df_1=5, df_2=3961 Prob.< 0.001			

Note: 1) O = trip origin
2) Estimates in bold are significant at 5% level; estimates in italic bold are significant at 10% level (two-tailed test).

Household VMT models for HBW and HBO trips are estimated and presented in Table 5.18. Similar to the results for total trips in Table 5.17, both travel cost and automobile trips are significant with negative and positive impacts, respectively. Socioeconomic variables except household size are significant for HBW trips. Contrary to other models, however, household size and vehicle ownership are only meaningful in explaining household VMT for HBO trips. Household income and income squared become insignificant.

Some differences are observed between household VMT models for HBW and HBO trips with regard to the role of land use. For HBW trips, population density has no significant impact; however, entropy index is positively significant. Employment density as well as entropy index becomes unimportant for explaining household VMT for HBO trips. When other factors kept constant, 29.2% and 33.7% of total variations in household VMT are explained for HBW and HBO trips, respectively. Model improvement tests support that household VMT models are significantly refined when land use measures are considered in the modeling process.

Table 5.18 Household Total VMT Models for Home-based Trips.

Variables	Home-based Work Trips				Home-based Other Trips			
	Base Model		Extended Model		Base Model		Extended Model	
	Est.	*t*	Est.	*t*	Est.	*t*	Est.	*t*
Constant	2.216	**17.69**	2.771	**16.53**	2.355	**23.88**	3.134	**21.99**
Travel cost ($ / mile)	-0.731	**-2.54**	-0.841	**-2.98**	-1.202	**-5.16**	-1.289	**-5.67**
Auto trips	0.304	**23.17**	0.310	**24.09**	0.203	**32.65**	0.206	**34.06**
Household size	0.033	**2.41**	0.018	1.33	-0.080	**-5.82**	-0.094	**-7.00**
Vehicles in household	0.105	**4.58**	0.084	**3.73**	0.203	**9.80**	0.169	**8.27**
Household income	0.101	**4.06**	0.091	**3.73**	0.013	0.61	0.014	0.71
Income squared	-0.005	**-3.40**	-0.004	**-3.09**	0.000	-0.32	-0.001	-0.45
Pop. density at O [1)]			-0.006	-1.25			-0.021	**-4.86**
Emp. density at O			-0.019	**-4.32**			-0.005	-1.42
Entropy index at O			0.355	**2.98**			-0.027	-0.24
Connectivity at O			-0.574	**-5.51**			-0.214	**-2.18**
Road length at O			-0.033	***-1.95***			-0.093	**-5.78**
Sample size	2539				3461			
$R^2, \overline{R}^2$	0.259, 0.258		0.292, 0.289		0.300, 0.299		0.337, 0.334	
Model improvement: *F-test*	F=23.505, df_1=5, df_2=2527 Prob.< 0.001				F=37.631, df_1=5, df_2=3449 Prob.< 0.001			

Note: 1) O = trip origin
2) Estimates in bold are significant at 5% level; estimates in italic bold are significant at 10% level (two-tailed test).

5.4.2 Results of Structural Equation Models

There are four major models in the results of SEM estimation of household total VMT for different travel purposes: household VMT (*VMT ON*), travel cost (*Travel cost ON*), household automobile trip frequency (*Auto trips ON*), and total household income (*Household income ON*). The structural models estimated for different trip purposes are presented in Table 5.19 and 5.20. The outcomes of household VMT part in the SEMs for different trip purposes are comparable to those of household VMT models in Table 5.17 and 5.18, respectively. To summarize, land use measures at trip origin have negatively significant relationships with household total VMT while other variables are controlled.

Travel cost models for different trip purposes deserve attention to examine if the assumed causal connections of travel cost with land use measures at origin are still valid. Based on the outcomes of both Table 5.19 and 5.20, it is confirmed that two land use variables, population density and entropy index are negatively affect travel cost. Increases in population density and entropy around residential locations are causally connected with the reduction of travel cost per mile. These causal relationships remain unchanged for different travel purposes. Household income is also positively significant in its linkage with travel cost except for HBW trips. It implies there is no significant difference in commuting trip cost among income brackets.

The model specification of the automobile trip frequency part is similar to the automobile trip generation models, but the former excludes land use measures. Socioeconomic factors in addition to travel cost per trip have significant effects on household automobile trip rates. But household income and income squared are not significant for both HBW and HBO trips, and income squared is not for total trip purposes.

Table 5.19 Structural Equation Models of Household Total VMT for Total Trips.

	Total Trips			Total Home-based Trips		
Variables	Estimates	Std. err.	p-value	Estimates	Std. err.	p-value
VMT ON						
Travel cost ($ / mile)	-1.088	0.409	0.008	-0.913	0.404	0.024
Auto trips	0.091	0.003	0.000	0.131	0.005	0.000
Household size	0.003	0.011	0.754	-0.014	0.012	0.236
Vehicles in household	0.182	0.015	0.000	0.211	0.017	0.000
Household income	0.105	0.018	0.000	0.090	0.018	0.000
Household income squared	-0.004	0.001	0.000	-0.003	0.001	0.001
Population density at O [1)]	-0.011	0.003	0.001	-0.015	0.004	0.000
Employment density at O	-0.011	0.003	0.000	-0.011	0.003	0.001
Entropy index at O	0.102	0.089	0.253	0.144	0.096	0.136
Connectivity at O	-0.309	0.076	0.000	-0.329	0.081	0.000
Road length at O	-0.044	0.013	0.000	-0.051	0.013	0.000
Travel cost ON						
Household income	0.001	0.000	0.000	0.001	0.000	0.000
Population density at O	-0.001	0.000	0.000	-0.001	0.000	0.000
Employment density at O	0.000	0.000	0.445	0.000	0.000	0.561
Entropy index at O	-0.016	0.005	0.001	-0.015	0.005	0.001
Connectivity at O	-0.001	0.004	0.889	-0.001	0.004	0.787
Road length at O	-0.001	0.001	0.255	-0.001	0.001	0.200
Auto trips ON						
Travel cost ($ / trip)	-0.088	0.004	0.000	-0.055	0.003	0.000
Household size	0.124	0.007	0.000	0.141	0.006	0.000
Vehicles in household	0.156	0.011	0.000	0.168	0.010	0.000
Single-family housing	0.060	0.033	0.069	0.065	0.029	0.024
Household income	0.053	0.012	0.000	0.043	0.011	0.000
Household income squared	-0.001	0.001	0.153	-0.001	0.001	0.171
Household income ON						
Household size	0.265	0.043	0.000	0.265	0.043	0.000
Vehicles in household	1.223	0.064	0.000	1.223	0.064	0.000
Single-family housing	1.075	0.177	0.000	1.076	0.178	0.000
Intercept						
VMT	3.219	0.145	0.000	2.918	0.152	0.000
Travel cost	0.234	0.004	0.000	0.235	0.005	0.000
Auto trips	1.205	0.057	0.000	0.747	0.050	0.000
Household income	4.275	0.187	0.000	4.277	0.187	0.000
Sample size	3976			3973		
Log Likelihood (H_0 value)	-30798.264			-29370.513		
Information Criteria						
Free parameters	44			44		
Akaike (AIC)	61684.528			58829.026		
Bayesian (BIC)	61961.201			59105.666		

Note: 1) O = trip origin

Table 5.20 Structural Equation Models of Household Total VMT for Home-based Trips.

	Home-based Work Trips			Home-based Other Trips		
Variables	Estimates	Std. err.	p-value	Estimates	Std. err.	p-value
VMT ON						
Travel cost ($ / mile)	-0.842	0.507	0.097	-1.285	0.219	0.000
Auto trips	0.310	0.015	0.000	0.206	0.007	0.000
Household size	0.018	0.014	0.185	-0.094	0.014	0.000
Vehicles in household	0.084	0.023	0.000	0.169	0.019	0.000
Household income	0.091	0.026	0.000	0.014	0.021	0.504
Household income squared	-0.004	0.001	0.003	-0.001	0.001	0.672
Population density at O [1)]	-0.006	0.004	0.195	-0.021	0.005	0.000
Employment density at O	-0.019	0.004	0.000	-0.005	0.004	0.146
Entropy index at O	0.354	0.123	0.004	-0.027	0.115	0.814
Connectivity at O	-0.574	0.108	0.000	-0.213	0.095	0.024
Road length at O	-0.033	0.016	0.041	-0.092	0.017	0.000
Travel cost ON						
Household income	0.000	0.000	0.957	0.001	0.000	0.000
Population density at O	-0.001	0.000	0.001	-0.001	0.000	0.054
Employment density at O	0.000	0.000	0.912	0.000	0.000	0.350
Entropy index at O	-0.010	0.006	0.070	-0.022	0.008	0.009
Connectivity at O	0.005	0.009	0.556	0.000	0.007	0.987
Road length at O	-0.001	0.001	0.150	-0.001	0.001	0.497
Auto trips ON						
Travel cost ($ / trip)	-0.018	0.003	0.000	-0.048	0.004	0.000
Household size	0.014	0.008	0.089	0.149	0.008	0.000
Vehicles in household	0.208	0.013	0.000	0.064	0.013	0.000
Single-family housing	0.069	0.030	0.021	0.070	0.038	0.068
Household income	0.023	0.014	0.115	0.005	0.013	0.695
Household income squared	-0.001	0.001	0.238	0.001	0.001	0.341
Household income ON						
Household size	0.200	0.052	0.000	0.325	0.045	0.000
Vehicles in household	0.958	0.079	0.000	1.213	0.068	0.000
Single-family housing	1.293	0.226	0.000	1.262	0.199	0.000
Intercept						
VMT	2.874	0.202	0.000	3.038	0.146	0.000
Travel cost	0.239	0.006	0.000	0.235	0.007	0.000
Auto trips	0.456	0.061	0.000	0.872	0.059	0.000
Household income	6.208	0.220	0.000	4.010	0.210	0.000
Sample size		2539			3461	
Log Likelihood (H_0 value)		-17834.592			-27394.273	
Information Criteria						
Free parameters		42			44	
Akaike (AIC)		35753.184			54876.545	
Bayesian (BIC)		35998.444			55147.115	

Note: 1) O = trip origin

5.4.3 Results of Directed Acyclic Graphs

Eleven variables in four broad categories are introduced in the estimation of DAGs for different trip purposes: one travel behavior measure of interest (household VMT), two travel attributes (travel cost and auto trip frequency), three household socioeconomic factors (household size, vehicle ownership and total income), and land use measures at trip origin (population and employment density, entropy index, connectivity and road length measure). Three constraints imposed for both binary choice and household auto trip generation are still in effect. Undirected edges between socioeconomic factors are causally oriented based on personal judgment and academic evidence.

The directed graphs for both total VMT and total home-based VMT are illustrated in Figure 5.10 and 5.11. They produce exactly same results of causal connections. Household total VMT (TOTAL VMT), according to the DAGs, is causally influenced by six measures: automobile trips (AUTO TRIPS: +), vehicle ownership (NOVEHICLE: +), household income (INCOME: +), population density (O_POPDEN: –), employment density (O_EMPDEN: –), and connectivity (O_CONNECT: –). More automobile trips, more vehicles available and higher income lead to longer household automobile trip distance. On the other hand, higher population and employment density and improved connectivity discourage people from driving longer. Contrary to the SEM results, both travel cost (TRAVEL COST) and road length (O_ROADMI) variables are not direct causes of household total VMT.

Interestingly travel cost is not significant in the causality with household VMT. According to the results in Figure 5.10 and 5.11, travel cost has causal relationships with two household socioeconomic factors and two land use measures: household size (HHSIZE: +), household income (INCOME: +), population density (O_POPDEN: –), and entropy index (O_ENTROPY: –). These connections are generally congruous with those clarified in the SEM results. However, it is identified as a collider in the directed graphs. As explained before, travel cost as a collider receives information from the parent variables, but prevents the information from flowing to others. In terms of automobile trip frequency, all socioeconomic variables are direct causes, but no land use measure is. It is quite consistent with the findings in other models for total trips (see Table 5.19 and Figure 5.6 and 5.7).

The estimated directed graphs for both HBW and HBO trips are displayed in Figure 5.12 and 5.13. For HBW trips, household VMT has various causal factors: automobile trips (AUTO TRIPS: +), vehicle ownership (NOVEHICLE: +), household income (INCOME: +), employment density (O_EMPDEN: –), and connectivity (O_CONNECT: –). For HBO trips, household VMT is also affected by many variables: travel cost per mile (TRAVEL COST: –), automobile trips (AUTO TRIPS: +), vehicle ownership (NOVEHICLE: +), population density (O_POPDEN: –), and road length (O_ROADMI: –). Several differences are found when they are compared with the results total trips. First, population density is not significant for HBW trips; neither employment density is for HBO trips. Moreover, road length measure instead of connectivity becomes direct cause for HBO trips. Another difference lies in the role of travel cost: travel cost negatively cause household total VMT for HBO trips.

It should be also noted that no land use measure causally affects travel cost for both HBW and HBO trips; thus, travel cost is only determined by household socioeconomic characteristics. There are two colliders, travel cost and entropy index for HBW trips; on the other hand, there is only one collider, entropy index for HBO trips. They are similar in that entropy index is a collider which is causally independent of household VMT for both travel purposes.

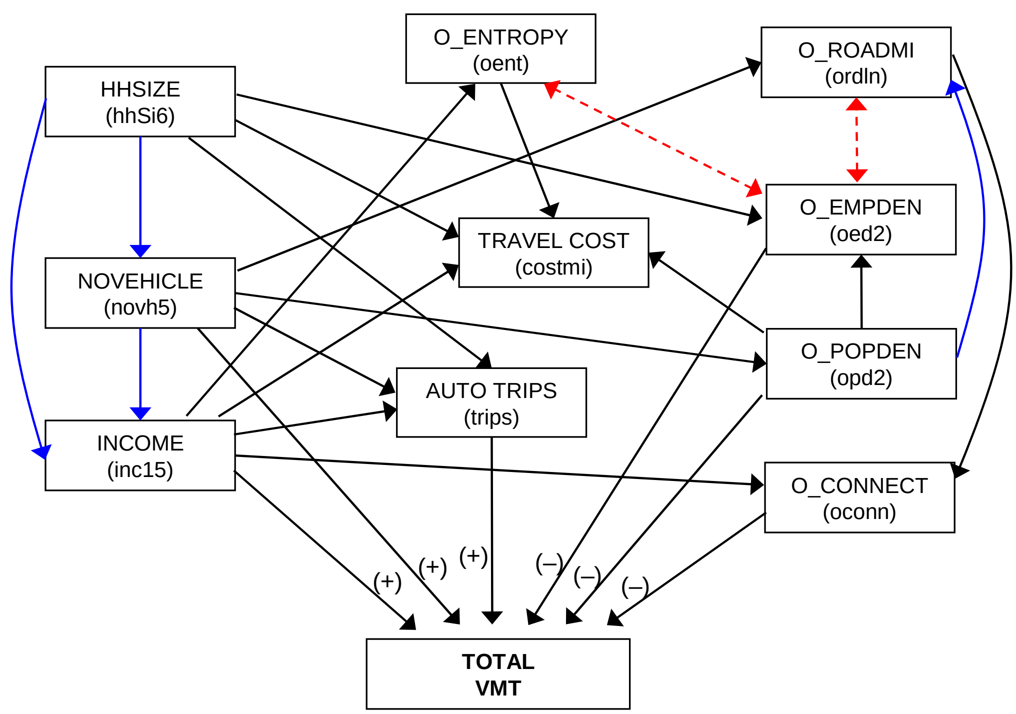


Figure 5.10 DAGs on Household VMT for Total Trips (1% sig. level).
Note: Double-headed edges in a pattern suggest a latent common cause between two variables.

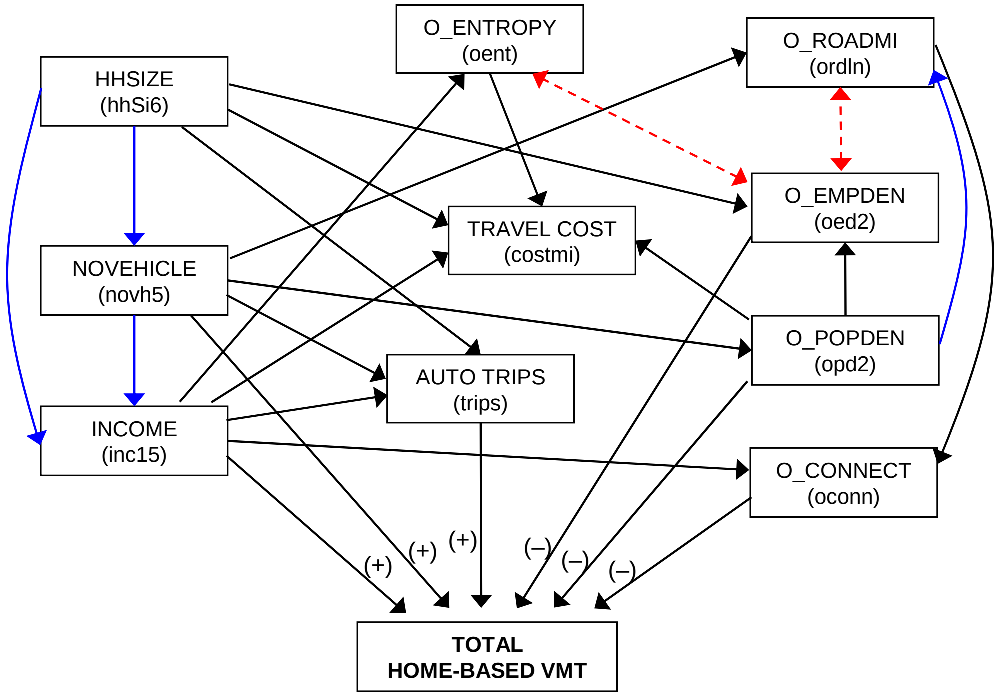


Figure 5.11 DAGs on Household VMT for Total Home-based Trips (1% sig. level).

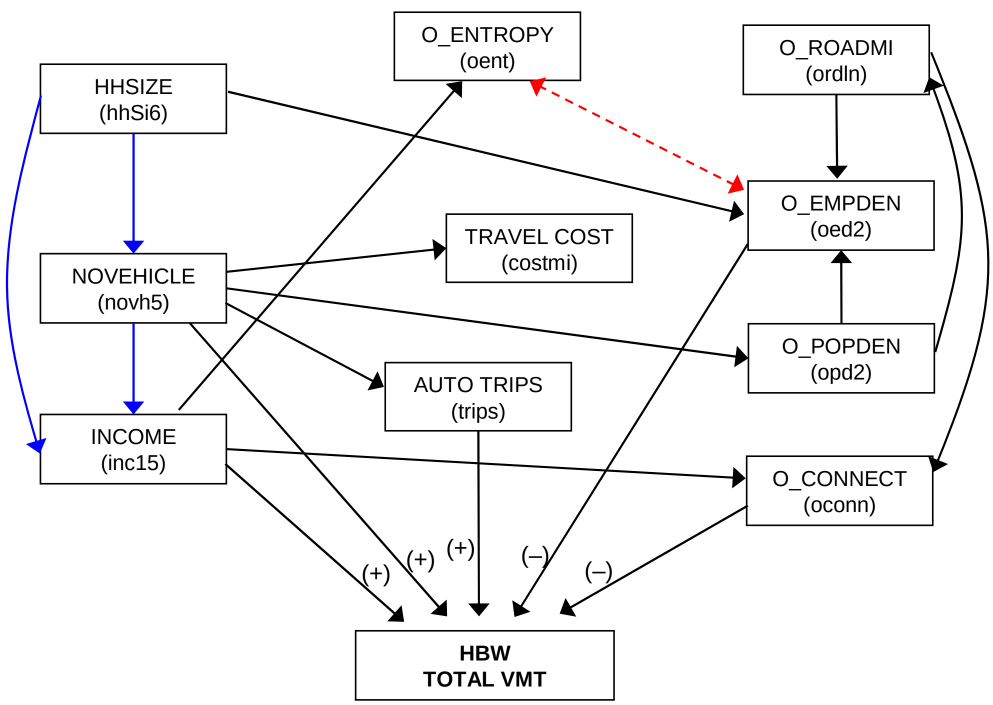


Figure 5.12 DAGs on Household VMT for Home-based Work Trips (1% sig. level).
Note: Double-headed edges in a pattern suggest a latent common cause between two variables.

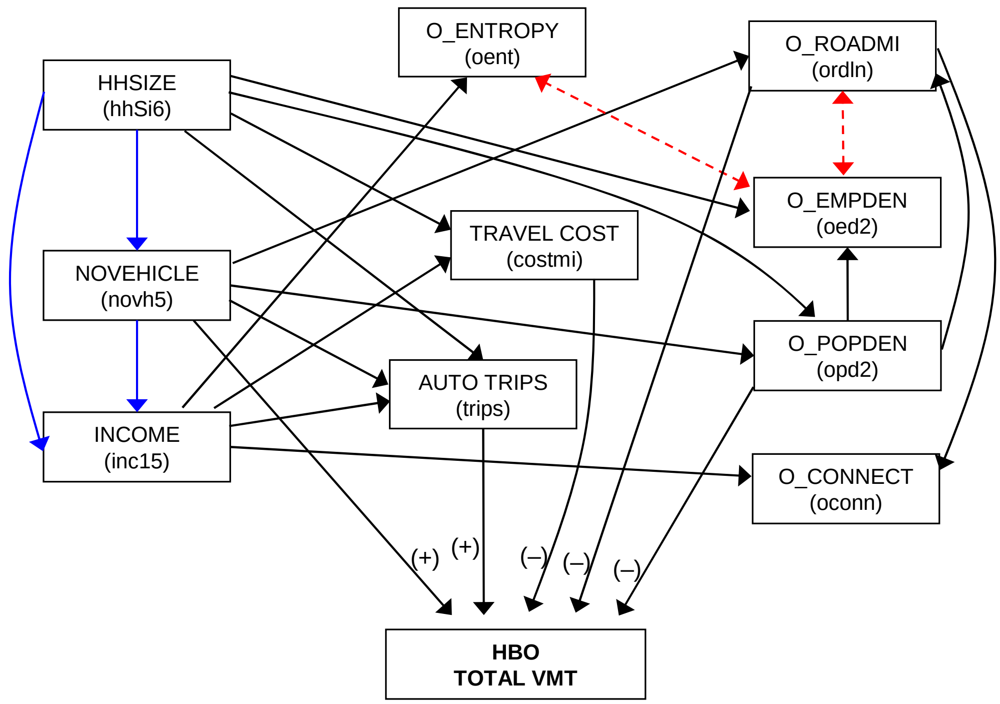


Figure 5.13 DAGs on Household VMT for Home-based Other Trips (1% sig. level).
Note: Double-headed edges in a pattern suggest a latent common cause between two variables.

5.5 Summary and Discussion

There are some findings observed from the patterns of household travel and land use and development in the HGAC region. According to the 2007 HGAC household travel survey, people in the metropolitan region are highly depending on automobiles in terms of mode choice, trip rates and travel distance. First of all, 92 percent of total trips in the region are made by automobile modes including driving-alone and shared-ride options. 99 percent of HBW trips are made by automobile modes, and driving-alone trips account for 89 percent of the total. Average VMT (miles per trip) by automobile and driving-alone modes are 9.53 and 12.12 for regional total, and 21.51 and 22.70 for HBW trips, respectively. It indicates that people drive longer for commuting purpose than for other trip purposes. In addition, regional driving time and cost average 8.4 minutes and 2.05 dollars per trip; for HBW trips, they are 13.3 minutes and 4.73 dollars per trip, respectively. High level of automobile dependence, longer driving miles and time, and higher driving cost for commuting trips suggest that residential, commercial and industrial areas are not only developed with low-density, but they are also much segregated rather than well mixed and balanced even in urban areas.

Land use and development patterns in the HGAC region indicate that single-family residential use is prevailing among various types of residential uses covering 50 percent of total developed area and 92 percent of total residential area. Land use density measures including population and employment density show that they are relatively high in urban areas, especially in the City of Houston, but very low in other areas. Employment density tends to be more concentrated on the central area. Land use diversity and design measures also reveal that land uses are well mixed and connected with each other in the central area, implying that these patterns of land use lead to high level of automobile dependence throughout the region. The analysis of average VMT by residential type suggests that single-family households make longer trips than others including multi-family households for all travel purposes. It is inferred that land use patterns can significantly influence travel behavior in many ways.

5.5.1 Individual Mode Choice

Mode choice models center on whether and how various dimensions of land use and development affect individual mode choice behavior in the HGAC region. The models also examine how land use measures are causally connected with travel time. A summary of the major findings based on MNL models, SEMs and DAGs is presented in Table 5.21.

Table 5.21 Summary of Mode Choice Models.

Variable of Interest	Mode Choice [3) 4)]						Travel Time			
Trip purpose	HBW			HBO			HBW		HBO	
Variables	MNL [2)]	SEM	DAG	MNL	SEM	DAG	SEM	DAG	SEM	DAG
Travel time	–	+	+	–	+	+				
Household income	+ (D)			+ (D)	+		+	+	+	+
Pop. density at O [1)]	+ (D)*			+ (D)	–		–	–	–	–
Pop. density at D	– (D,S)									
Emp. density at O	+ (T,W)	–	–				–	–	–	
Emp. density at D	– (D,S)	–					+	+	+	+
Dissim. index at O				+ (T,W)	–		+		+	
Dissim. index at D				+ (D,S)	+	+	–	–	–	–
Connectivity at O				+ (T,W)	–		–	–	–	
Connectivity at D	+ (D,S)*									
Road length at O				+ (T,W)			–	–	–	–
Road length at D				+ (D,S)	+		+	+	+	

Note: 1) O = trip origin, D = trip destination
2) D = driving-alone, S = shared-ride, T = transit, W = walk/bike
3) For MNL models, 4 alternative specific travel times are used for HBW trips; 2 times and 2 costs are for HBO trips. For SEMs and DAGs, travel time differential is used.
4) SEMs and DAGs are estimated with binary choice in which non-automobile is reference.
5) DAGs are estimated at 1% significance level; others are at 5% level except for * at 10% level.

MNL models show that many land use measures have significant impacts on individual mode choice behavior. Not only do they individually influence the probability of specific mode choice, but they collectively contribute to the improvement of multinomial choice models. These facts are consistent for both HBW and HBO trips. As shown in Table 5.21, several differences are noticeable. First, density measures are very important in mode choice for HBW trips; however, both diversity and design measures become significant for HBO trips. They increase the probability of non-automobile choice at origin and of automobile choice at destination at the same time for HBO trips. It is reasonable to argue that both land use mix and design factors are positively associated with the likelihood of non-automobile mode choice at origin. However, it may not be plausible to maintain that increased land use mix at destination enlarges the automobile choice probability. Travel and land use characteristics of the HGAC region give a tenable explanation. As described earlier, automobile dependent travel patterns are pandemic in the region. Land uses are also highly segregated and low-density residential areas are widely spread. Although land uses are well mixed and various activities are accommodated in the destinations, people tend to drive to the places so long as they are not close to their homes. This automobile captive behavior can also explain the positive relationship between population density at origin and driving-alone choice for both travel purposes.

The results of the mode choice models in the SEMs are similar to those of binomial logit models estimated to be compared with causal models. They are also generally congruous with MNL model outcomes: density factors are important for HBW trips; both diversity and design measures become significant for HBO trips; and the models are significantly enhanced by taking various land use measures into account. In short, land use and development factors significantly affect automobile choice behavior based on the assumed causality. Single-family housing as a socioeconomic factor closely related to land use measures positively influence the automobile choice probability.

The outcomes of travel time models in the SEMs for both trip purposes confirm that various measures of land use are causally associated with travel time differential; therefore, they have indirect connections with the automobile choice probability through travel time. Land use measures at origin and dissimilarity at destination negatively affect the travel time differential, which then reduces the likelihood to choose automobile mode. However, employment density and roadway length at destination and dissimilarity at origin are working in the opposite direction. Their positive effects on the travel time differential are inconsistent with the hypothesis and general reasoning. These impacts are partially attributed to both high automobile dependence and segregated and loose land use in the region.

According to the directed acyclic graph (DAG) for HBW trips, automobile choice is causally affected by many factors: travel time differential, bike use, vehicle ownership, single-family residence and employment density at origin. Single-family residence as related to land use attributes is a direct cause of automobile choice. Only employment density at origin among land use measures has a negative causal impact on the automobile choice probability. However, employment density at destination which is significant in the SEMs is not causally connected. The directed graph also shows that land use measures causally influence travel time differential. Dissimilarity at origin that is arguable in the SEM is no longer a direct cause. The result suggests that land use measures indirectly affect the automobile choice behavior through travel time.

Direct causes of the automobile mode choice for HBO trips are travel time differential, bike use, vehicle ownership, single-family residence and dissimilarity index at destination. While single-family residence is still a direct cause of automobile choice for HBO trips, the result is little consistent with the SEMs in which both diversity and design measures are significant. Only dissimilarity index at destination positively affects the probability of driving choice, which seems to be contrary to the hypothesis. As described in the MNL model results, it seems to be due in part to the automobile captive behavior that is caused by high automobile dependence and sprawling land use patterns in the region. In addition, travel time differential is causally associated with four land use measures. Similar to the result for HBW trips, these land use measures have indirect relationships with the automobile choice probability through travel time. Employment density at destination is positively significant for all trip purposes, which is a representation of land use and travel characteristics in the region.

To summarize, conventional travel demand model and causal models support that land use characteristics are directly affect individual mode choice behavior. In addition, both causal models based on different causal notions confirm that land use measures have indirect causal connections with individual mode choice through travel time in the metropolitan region. Although several discrepancies are observed, land use measures at both trip ends, in general, encourage trip-makers to use non-automobile modes as well as discourage them to use automobile modes.

5.5.2 Household Automobile Trip Generation

Household automobile trip generation models investigate whether and how land use characteristics directly influence the number of household automobile trips. Attention is also concentrated upon how land use and development patterns indirectly affect household automobile trip rates through travel cost ($/trip). Major findings based on negative binomial models, SEMs and DAGs are summarized in Table 5.22.

Table 5.22 Summary of Household Trip Generation Models.

Variable of Interest	Household Automobile Trip Generation [3]								
Trip purpose		Total			HBW			HBO	
Variables	NB [2]	SEM	DAG	NB	SEM	DAG	NB	SEM	DAG
Travel cost ($/trip)	–	–	–	–	–	–	–	–	–
Pop. density at O [1]									
Emp. density at O									
Entropy index at O									
Connectivity at O									
Road length at O									
Variable of Interest	Travel Cost ($/trip)								
Trip purpose		Total			HBW			HBO	
Variables		SEM	DAG		SEM	DAG		SEM	DAG
Household income		+			+	+			
Pop. density at O		–	–		–	–		–	–
Emp. density at O		–	–		–	–			
Entropy index at O								–	
Connectivity at O		–	–		–	–		–	–
Road length at O		–	–		–	–		–	–

Note: 1) O = trip origin; 2) Negative binomial model
3) DAGs are estimated at 1% significance level; others are at 5% level except for * at 10% level.

The results of negative binomial models for total, HBW and HBO trips are identical in this regard. Travel cost is negatively associated with household automobile trip rates. However, no land use measures at trip origin appear to be significant when travel cost and socioeconomic factors are kept constant. As a consequence, land use measures do not significantly contribute to the improvement of household automobile trip generation models.

The results of the travel demand models are similar to the estimation outcomes of causal models, the SEMs and the DAGs for different travel purposes. Higher travel cost per trip causally influences the reduction of total number of household automobile trips for all trip purposes. Another similarity exists between the travel demand models and causal models in that no land use measure is a direct cause of household automobile trip generation. In fact, the insignificant role of land use measures in the models has been a subject of academic controversy in the fields of urban and transportation planning. For instance, this result generally agrees with the arguments made by Boarnet and Sarmiento (1998) and Boarnet and Crane (2001a). On the other hand, it is quite contrary to the evidence found in Cervero and Kockelman (1997), Khattak and Rodriguez (2005), and Lee (2006).

The travel cost models in the SEMs are estimated on the assumption that land use measures and household income affect the travel cost per trip. The result indicates that most land use measures have negative impacts on the travel cost. It can be inferred that land use measures are causally connected with household automobile trip generation through the travel cost. There are several differences among travel purposes. For total and HBW trips, both density and design measures are statistically significant, but the entropy measure is not; on the other hand, land use measures except employment density are all meaningful for HBO trips.

The DAG for total trips shows that household automobile trip frequency is causally influenced by travel cost, household size, vehicle ownership, and household income. The result is consistent with that of automobile trip model in the SEMs for total trips. No land use measure at origin is a direct cause of automobile trip rates. Rather, density and design measures are negative causes of travel cost through which they are indirect causes of automobile trip generation. For HBW trips, direct causes of household automobile trip frequency are travel cost and vehicle ownership. Unlike other DAGs, household size and income do not causally influence household automobile trips for commuting trips. Automobile trip generation for HBO trips is causally affected by travel cost, household size, vehicle ownership and household income. Any causal connection between land use measures and automobile trip rates is still insignificant. Density and design measures for both HBW and HBO trips have negative causal relationships with travel cost through which they indirectly affect household automobile trip rates.

It is noteworthy that single-family residence, also a dominant land use pattern in the region is positively significant in the automobile trip generation models for HBO trips and marginally for total trips. In the structural models, it is still meaningful for HBO trips and marginally for HBW trips, but not for total trips. However, it no more directly affects household automobile trip frequency in the DAGs. It has indirect causality with automobile trips through other socioeconomic factors, household income and vehicle ownership.

As clarified in the structural models and directed graphs for different travel purposes, most land use measures have negative causal connections with travel cost as presented in Table 5.22. Due to the negative effect of travel cost on household automobile trip frequency, reduced travel cost by intensified land use and enhanced neighborhood network results indirectly in an increase in household automobile trips. In this regard, a question is raised whether land use policy is effective to manage the travel demand (Gomez-Ibanez 1991; Giuliano and Small 1993; Giuliano 1995). However, this argument may not be valid if other aspects of land use and travel behavior connections are considered. The effects of land use should be assessed from comprehensive standpoint with careful investigations into the land use impacts on mode choice and VMT. Some studies have found the evidence that land use directly affects household automobile trips, which may offset the indirect impact of land use through travel cost (Cervero and Kockelman 1997; Khattak and Rodriguez 2005; Lee 2006).

5.5.3 Household Total VMT

Household VMT models explore whether and how land use measures affect household total driving distance in the HGAC region. It is also investigated whether and how land use measures have causal relationships with travel cost. Table 5.23 gives a summary of household VMT models from ordinary least squares (OLS) regression, SEMs and DAGs.

Table 5.23 Summary of Household VMT Models.

Variable of Interest	Household Total VMT [3]								
Trip purpose		Total			HBW			HBO	
Variables	OLS [2]	SEM	DAG	OLS	SEM	DAG	OLS	SEM	DAG
Travel cost ($/mi)	–	–		–	–		–	–	–
Pop. density at O [1]	–	–	–				–	–	–
Emp. density at O	–	–	–	–	–	–			
Entropy index at O				+	+				
Connectivity at O	–	–	–	–	–	–	–	–	
Road length at O	–	–		–*	–		–	–	–
Variable of Interest	Travel Cost ($/mile)								
Trip purpose		Total			HBW			HBO	
Variables		SEM	DAG		SEM	DAG		SEM	DAG
Household income		+	+					+	+
Pop. density at O		–	–		–			–*	
Emp. density at O									
Entropy index at O		–	–		–*			–	
Connectivity at O									
Road length at O									

Note: 1) O = trip origin; 2) Ordinary least squares regression model
3) DAGs are estimated at 1% significance level; others are at 5% level except for * at 10% level.

Table 5.23 summarizes the land use effects on both household VMT and travel cost for different travel purposes. Both density and design measures at trip origin are negatively associated with household VMT for total trips. For HBW trips, land use measures except population density at origin are significant. Entropy index has a positive relationship, which is not consistent with the theory. But it becomes insignificant in the DAG; it is therefore claimed that the positive impact of entropy index is spurious in terms of causality between land use and household VMT. For HBO trips, population density and design measures have negative association with household VMT. In particular, employment density is not important for HBO trips, which is generally consistent with the findings from other travel demand models. To sum up, land use factors are individually significant in the models. They also collectively refine household VMT models for all trip purposes.

Household VMT models in the SEMs for total, HBW and HBO trips are almost identical to OLS regression models of household VMT. In short, land use measures in general significantly affect household total VMT based on the assumed causality. Travel cost models in the SEMs suggest that two land use measures, population density and entropy index at origin reduce travel cost per mile as they increase. Therefore, it can be stated that land use measures at origin directly affect household VMT as well as indirectly influence household VMT through travel cost. Similar to the results of the automobile trip generation models, the indirect impact of land use measure on household VMT could be debatable between two different viewpoints. However, what is different from the trip generation models is that land use measures have negative direct connections with household VMT. The indirect effects of land use measures disappear according to the results of the directed graphs for all trip purposes.

The DAG for total trips illustrates that household VMT is directly caused by automobile trips, vehicle ownership, household income and three land use measures, population density, entropy index and connectivity. Contrary to the regression model and the SEMs, both travel cost and road length variables are not causally connected with household VMT. Travel cost has causal relationships with population density and entropy index as well as household income, which is similar to the SEM result for total trips. Travel cost is, however, identified as a collider.

According to the directed graph for HBW trips, direct causes of household VMT are auto trips, vehicle ownership, household income and two land use measures, employment density and connectivity. Household VMT for HBO trips is directly affected by travel cost, automobile trips, vehicle ownership, and two land use measures, population density and road length. It is natural that employment density instead of population density is significant for HBW trips; the opposite is also sensible for HBO trips. Travel cost is a direct cause for HBO trips, while it is not for other trip purposes. It is notable that no land use measure is causally connected with travel cost; rather, travel cost is affected by household socioeconomic factors. Travel cost is a collider for HBW trips, and entropy is a collider for both trip purposes.

In short, conventional regression models, structural models and directed graphs consistently maintain that various land use patterns around residential locations have direct relationships with household total VMT. Compact development with high density and improved neighborhood network design significantly contribute to the reduction in household VMT despite several variations. However, land use mix does not play a crucial role in contrast with the academic evidence and expectation. There are some differences between the model results. The structural models indicate that both population density and entropy indirectly affect household VMT through travel cost per mile; however, no land use measure is causally connected with travel cost in the directed graphs for both HBW and HBO trips.

CHAPTER VI

CONCLUSIONS AND IMPLICATIONS

This chapter generalizes conclusions by summarizing major findings of the investigation. Then, policy implications are explored that are related to land use and development to deal with automobile dependence as well as to achieve the objectives of sustainable development in the HGAC region. In addition, the limitations and possible improvements of this research are discussed.

6.1 Conclusions

The study investigates how land use patterns affect individual and household travel behavior in a regional context for reducing automobile dependence and achieving sustainability goals. Previous researches are significant in that they enhanced our understanding of land use effects on travel behavior, suggested land use and development policies for reducing automobile dependence, and provided suggestions for improving travel demand models. Nonetheless, questions are still remaining about land use measurement, theory and framework for travel demand models, and causal connections between land use and travel behavior.

The study focuses on six counties of the Houston-Galveston Area Council (HGAC) regions: Brazoria, Fort Bend, Galveston, Harris, Montgomery and Waller County. Major data sources are the 2007 HGAC regional household activity and travel survey, 2007 parcel-based land use GIS dataset and HGAC regional travel model and forecast data. Three travel behavior measures are considered as principal dependant variables in the model estimation: individual mode choice, household automobile trip generation and household total VMT. Also, three major categories of explanatory variables, i.e., travel time and cost, socioeconomic characteristics and land use measures are taken into account. A variety of land use characteristics are measured using quarter-mile buffers for both trip origins and destinations.

In terms of model estimation strategies, attention is focused on the effects of land use on travel behavior from different modeling perspectives. One is conventional travel demand modeling for exploring the association between land use and travel behavior. Another is causal modeling for clarifying the causal connections between them. The causal modeling approaches include both the structural equation modeling (SEM) and the directed acyclic graphs (DAGs). They are different in that the SEM depends on causal assumptions based on the theory; the DAGs, however, rely not on assumed causality but on causality based on observational data. Both the SEM and the DAGs pay attention to not only direct impacts of land use on travel behavior outcomes, but also indirect impacts of land use through travel cost. Models are estimated for different travel purposes including total, HBW and HBO trips. As a consequence, three travel outcomes and different trip purposes are taken into three modeling strategies with full array of explanatory variables. Conclusions are drawn as follows.

For travel and land use patterns:

- People in the HGAC region are highly dependent on automobiles in terms of mode choice, trip frequency and travel time and distance. The automobile dependence is noticeable for home-based work (HBW) trips.
- Residential, commercial and industrial areas are not only developed with low-density, but they are also much segregated rather than well mixed and connected with each other even in urban areas.
- Land use patterns characterized by low density, less diversity and poor network design may lead to high level of automobile dependence throughout the region. It is inferred that land use patterns significantly influence travel behavior in the region while other factors kept constant.

For individual mode choice:

- Land use measures have significant impacts on individual mode choice behavior based on MNL choice models. Not only do they individually influence specific mode choice probability, but they collectively contribute to improving the choice models for both HBW and HBO trips.
- Land use measures at trip origin and destination are meaningful in explaining automobile mode choice in the SEMs. The results are consistent with those of MNL choice models.
- Travel time models in the SEMs confirm that various land use factors are causally connected with travel time differential (walk time – driving time).

Through the travel time, land use measures are indirectly connected with the automobile choice probability.

- Most land use measures at origin have negative relationship with the travel time differential through which they reduces the probability of automobile choice. Several land use factors at destination are working in opposite direction.
- The DAGs for both HBW and HBO trips show that individual automobile choice are directly caused by travel time, vehicle ownership, bike use, single-family residence and one land use measure for each trip purpose.
- For HBW trips, only employment density at origin has a negative causal impact on the automobile choice probability. For HBO trips, on the other hand, dissimilarity measure at destination positively affects the probability of driving choice.
- In terms of direct land use effects, the results of the DAGs are not consistent with those of the SEMs. It suggests that there is a gap between assumed and data generated causal relationships of land use with automobile choice behavior.
- Many land use measures have indirect causal connections with the automobile choice probability through travel time, which is congruous with the SEM results.
- In short, conventional travel demand model and causal models support that land use measures directly affect individual mode choice behavior in varying degrees. There is also clear evidence from causal models that land use factors indirectly influence it through travel time.

- Land use measures at both trip ends generally encourage trip-makers to use non-automobile modes as well as discourage them to use automobile modes even though several variations are observable.

For household automobile trip generation:

- Based on the results of the negative binomial models, no land use measure in density, land use balance and network design at trip origin is significantly associated with household automobile trip frequency when travel cost and socioeconomic factors are controlled. It is consistent for all trip purposes, i.e. total, HBW and HBO trips
- As a consequence, land use measures do not significantly contribute to the improvement of household automobile trip generation models for all travel purposes.
- Similar to the results of travel demand models, both the SEMs and the DAGs show no evidence to support that land use measures at origin are direct causes of household automobile trip rates.
- Instead, there is strong evidence based on the causal models that land use measures have indirect causal connections with household automobile trip frequency through travel cost per trip. It is valid for all trip purposes.
- As a whole, density and design measures negatively affect travel cost per trip through which they have indirect causal relationships with the number of

household automobile trips. The argument is consistent for two different approaches to causal modeling and for all trip purposes with minor differences.

- The results of two different causal modeling approaches are quite consistent. It implies that the theory for assumed causal relationships between land use and automobile trip rates are well established.
- Due to the negative effect of travel cost on automobile trip frequency, reduced travel cost by high density and improved network design results in an increase in household automobile trip frequency. In this regard, it is questionable whether land use strategies are effective to manage household automobile trip generation in the HGAC region.
- It should be noted that the argument may be plausible only if an aspect of land use and travel behavior connections is considered. The effects of land use should be evaluated with comprehensive investigations into the relationships between land use and various travel behavior measures including mode choice and VMT.

For household total VMT:

- Regression models of household VMT indicate that land use factors are individually significant; they also collectively contribute to refining the household VMT models. Their significance in explaining household total VMT is justifiable for all travel purposes.

- Density and design measures at trip origin are significantly associated with the reduction in household VMT for total, HBW and HBO trips although minor variations exist.
- Household VMT models in the SEMs are congruous with the regression models, supporting the significance of density and design factors based on the assumed causality.
- Travel cost ($/mile) in the SEMs has negative relationship with population density and land use balance measures for all trip purposes. It is claimed that land use measures at origin directly affect household VMT as well as indirectly influence it through travel cost.
- The DAGs exhibit that household VMT is not only positively caused by automobile trips and several socioeconomic factors, but also negatively affected by several land use measures, especially both density and design measures for all travel purposes.
- Similar to the results of both the SEMs and the household automobile trip generation models, land use measures have no direct connections with household automobile trips.
- Contrary to the results of the SEMs, travel cost ($/mile) is not a direct cause of household total VMT except for HBO trips. It is inferred that travel cost is an important factor for determining household VMT for shopping, social and recreational trips, but not for commuting trips.

- No land use measure causally influences travel cost per mile; rather, travel cost is affected by household socioeconomic factors. Hence, there is no evidence that land use measures have indirect causal relationships with household VMT via travel cost, which disagrees with the finding from the structural models.
- In terms of the role of travel cost in household VMT models, there exists lack of consistency between assumed and data generated causal connections of travel cost with household VMT and land use measures.
- To sum up, the results from different modeling strategies confirm that various land use patterns around residential locations are not only significantly associated, but also causally connected with household total VMT. Compact development with high density and improved network design significantly contribute to the reduction in household VMT despite several differences.

6.2 Policy Implications

Consistently growing automobile dependence over past decades has resulted in a number of malign impacts on our economic, social and environmental system although some economic benefits are attributed to it. As a consequence, it has been adverse to our continuous efforts for sustainable development and transportation since late 1980s. Many studies suggested that land use and development pattern is one of main causes of automobile dependence (Newman and Kenworthy 1989b; Raad 1998; Lee 2006; VTPI 2008a). Hence, policies and strategies related to land use and development have been proposed to reduce automobile dependence as well as to accomplish the goals of sustainability.

Note in the study results that two variables in addition to various land use measures attract attention in relation to the public policy: bike use and single-family residence. Bike and walk are regarded as the most sustainable travel modes; therefore, they are strongly encouraged in most cities and metropolitan regions in terms of sustainability and public health. According to the study results, bike use is significantly associated with increased likelihood of choosing non-automobile modes and reduced probability of automobile mode choice. It is a direct cause of the reduction of automobile choice probability for both HBW and HBO trips. It is also presumed based on the theory to be closely connected with the reduction of household automobile trip rates and total VMT. The actual share of non-motorized modes is, however, very low compared with that of automobile modes in the HGAC region.

Currently, some efforts are being made to develop comprehensive pedestrian and bicyclist plans and programs and to prepare the bikeway network. They include livable centers project and transit and land use coordination (HGAC 2007). More collaborative and continuing researches and programs are needed for replacing automobile trips with non-motorized trips, connecting with transit mode, widening travel mode choice options, and enhancing pedestrian and bicyclist's safety. It should be noted that higher land use density, better land use mix and balance and improved network design are a prerequisite for encouraging people to walk and bike to their destinations.

Another factor that deserves attention is single-family residence indicating detached low-density housing for single-family households. Based on the results, it has significantly positive relationships with automobile choice probability and household automobile trips for almost all travel purposes. It also shows a direct causal relationship with automobile mode choice, and an indirect causality with household automobile trip rates and household VMT through household income. Single-family residential neighborhoods have three common characteristics especially located in suburban areas in the U.S (Knaap et al. 2007; Kopits et al. 2009). One is that most of them are detached and segregated from other land uses such as commercial areas and employment centers. Their development density is commonly very low in which the street network is curvilinear with lots of cul-de-sacs thus lack of connectivity. Moreover, the single-family residential area frequently covers the largest proportion out of total developed area. The land use patterns represented by the neighborhoods serve as important indications of urban sprawl. Urban sprawl is considered unsustainable and undesirable attributes of urban land use and development patterns (Knaap et al. 2007). What is worse, there is evidence of self-reinforcing cycle of automobile dependence, auto-oriented planning and sprawling land use (VTPI 2008a).

Based on common awareness of the issues of urban sprawl and growing automobile dependence, two remedies have historically been suggested: the planning- and market-oriented approaches. The former has paid attention to the role of land use planning and regulation to encourage denser, more diverse and pedestrian-friendly land use and development. The latter, however, has focused mainly on economic measures to prevent sprawling land use and growing automobile uses, while arguing that land use and transportation linkage has been weakening. Conventional zoning and other local government land use regulations do not contributed to countering sprawling land use pattern and auto-dependent travels. Rather, it has been claimed that the government interventions led to lower density and separated land uses because they control building heights and uses, lot coverage, parking spaces and roadway width (Kopits et al. 2009; Levine 2006; Litman 2009).

In response, new strategies and policies have been proposed to integrate land use and transportation in planning field: smart growth and new urbanism. Smart growth generally focuses on the policy and planning, and new urbanism tends to focus on specific design practices (Handy et al. 2005; VTPI 2008b, 2008c, 2008d). But they have common objectives in transportation: increase the share of choosing non-automobile modes, decrease the number of automobile trips, and reduce vehicle miles of travel and increase vehicle occupancy (Cervero and Kockelman 1997). Specific land use policies include mixed-use zoning, form-based zoning code, cluster and infill development, brownfield development, transit-oriented development, and bicycle and pedestrian network (Handy et al. 2005; VTPI 2008b, 2008c, 2008d). They should be conducted in cooperation with local governments in the region that have authority for land use regulations and decisions. It should be also noted that some strategies may not be applicable to the cities and municipalities in the HGAC region where zoning codes are not established.

Noteworthy is the market-enabling strategy proposed by Levine (2006). Similar to the land use policies proposed by new planning movements, it is skeptical about the role of zoning and other land use regulations in controlling urban sprawl and auto-oriented trips. However, they are different in that the market-enabling strategy attributes the problems to planning failure, so land use policy reform is essential for overcoming obstacles to high-density, well-mixed and pedestrian-friendly areas. The land use regulatory reform finally results in an increase in individual and household choice in both travel and land use (Levine 2006; Levine and Inam 2004). Three types of policy reform are suggested: unchanging local government's land use regulatory power but promoting compact development; economic incentives from higher-level of governments to encourage municipalities for compact development; and sharing land use authority with higher-level of governments. The approach provides important suggestions of land use policy reform ensuring alternative development for the HGAC region.

Connecting land use and transportation is not only a goal of the 2035 regional transportation plan (RTP), but it is also considered one of the most effective strategies to enhance mobility and accessibility and improve quality of life (HGAC 2007). How to measure and evaluate the current performance and the progress throughout the metropolitan area? Sustainable transportation indicators can help assess the progress and make decisions. Transportation sustainability guides to set up goals and select a set of measures, and indicators determine what should be measured to achieve the goals (Zietsman and Rilett 2002; STI 2008; Ramani et al. 2009). In terms of sustainability goals of integrating land use and transportation, it is necessary for the HGAC to consider various land use measures such as density, land use mix and design measures.

One objective of the study is to convey implications for improving current transportation demand modeling of the HGAC region. This study introduces various land use measures that are computed within walking distance (one quarter-mile radius) of both trip origin and destination to reflect trip-maker's surrounding context. These state-of-the-art methods of land use measurement lead to refining current travel demand forecasting models. They help ameliorate underestimated and biased estimators of the models due to lack of full array of land use factors into consideration. Another issue is that current regional travel demand models do not estimate the effect of neighborhood-level land use and development on transportation demand. As Cervero (2006) suggested, either post-processing using elasticity estimates or direct modeling method is useful to capture the land use effects on travel demand in small-scale projects.

The HGAC region is made up of 13 counties containing 5.4 million residents and 145 municipalities. It implies that proposed policies and programs can be successfully accomplished with close collaboration and elaborate coordination among the local governments and interest groups. The policy implications are still neither complete nor satisfactory for achieving the goals of sustainability in the region.

6.3 Limitations

The study sheds light on the significant land use effects on individual and household travel behavior. New methodology is introduced for analyzing causal relationships between land use and travel behavior. Land use characteristics are measured fully in three dimensions. They are also calibrated in quarter-mile buffers of both trip ends in order to represent trip-maker's environment influencing travel decisions. Moreover, the entire set of explanatory variables is properly included in the model estimation of mode choice, automobile trip generation and VMT.

However, there are several limitations in this investigation. First, the household survey data and land use GIS data are not complete. Only 84% of planned survey total samples, and 6 counties out of 8 surveyed counties are included in the research due to data availability. Another weakness comes from causal modeling methodologies. Many latent variables shown with the bidirected edges mostly between land use variables cannot be clarified in the directed graphs. In addition, the issue of self-selection or unrevealed preferences is not properly addressed in the causal modeling. It is mainly because no questions related to travel attitudes and preferences are included in current regional travel survey.

REFERENCES

Agresti, A. (2007). *An introduction to categorical data analysis*. Hoboken, NJ: Wiley-Interscience.

American Automobile Association (AAA). (2008). *Your driving costs. 2008 edition*. Retrieved April 3, 2009, from http://www.aaaexchange.com/Assets/Files/20084141552360.DrivingCosts2008.pdf

Badoe, D., & Miller, E. (2000). Transportation-land-use interaction: empirical findings in North America, and their implications for modeling. *Transportation Research D*, *5*(4), 235-263.

Bagley, M., & Mokhtarian, P. (2002). The impact of residential neighborhood type on travel behavior: A structural equations modeling approach. *The Annals of Regional Science*, *36*(2), 279-297.

Basar, G.., & Bhat, C. (2004). A parameterized consideration set model for airport choice: An application to the San Francisco Bay Area. *Transportation Research B*, *38*, 889-904.

Beatley, T. (1995). Many meanings of sustainability. *Journal of Planning Literature*, *9*(4), 339-342.

Beimborn, E., Greenwald, M., & Jin, X. (2003). Accessibility, connectivity, and captivity: Impacts on transit choice. *Transportation Research Record*, *1835*, 1-9.

Ben-Akiva, M., & Bierlaire, M. (1999). Discrete choice models with applications to departure time and route choice. In R. Hall (Ed.), *Handbook of transportation science* (pp. 7-37). New York: Springer.

Ben-Akiva, M., & Boccara, B. (1995). Discrete choice models with latent choice sets. *International Journal of Research in Marketing*, *12,* 9-24.

Ben-Akiva, M., & Lerman, S. (1985). *Discrete choice analysis: Theory and application to travel demand*. Cambridge, MA: MIT Press.

Black, W. B. (2005). Sustainable transport: definition and responses. Resource Paper for Conference Proceedings 37. *Integrating Sustainability into the Transportation Planning Process*. 2004 Conference Report. Washington, DC: Transportation Research Board.

Boarnet, M., & Crane, R. (2001a). The influence of land use on travel behavior: Specification and estimation strategies. *Transportation Research A*, *35*(9), 823-845.

Boarnet, M., & Crane, R. (2001b). *Travel by design: The influence of urban form on travel*. New York: Oxford University Press.

Boarnet, M., & Greenwald, M. (2000). Land use, urban design, and nonwork travel. *Transportation Research Record*, *1722*, 27-37

Boarnet, M., & Sarmiento, S. (1998). Can land-use policy really affect travel behaviour? A study of the link between non-work travel. *Urban Studies*, *35*(7), 1155-1169.

Bureau of Transportation Statistics (BTS) (2008). *National transportation statistics 2008*. Washington DC: U.S. Department of Transportation. Retrieved July 17, 2009, from http://www.bts.gov/publications/national_transportation_statistics/2008

Cao, X., Handy, S., & Mokhtarian, P. (2006). The influences of the built environment and residential self-selection on pedestrian behavior: Evidence from Austin, TX. *Transportation*, *33*(1), 1-20.

Center for Sustainable Transportation (CST). (2000). *Sustainable transportation performance indicators project. phase I report*. Retrieved July 17, 2009, from http://cst.uwinnipeg.ca/ completed.html#indicators

Cervero, R. (1991). Land uses and travel at suburban activity centers. *Transportation Quarterly*, *45*, 479-491.

Cervero, R. (1996). Mixed land-uses and commuting: Evidence from the American Housing Survey. *Transportation Research A*, *5*, 361-377.

Cervero, R. (2002). Built environments and mode choice: Toward a normative framework. *Transportation Research D*, *7*(4), 265-284.

Cervero, R. (2006). Alternative approaches to modeling the travel-demand impacts of smart growth. *Journal of the American Planning Association*, *72*(3), 285-295.

Cervero, R., & Gorham, R. (1995). Commuting in transit versus automobile neighborhoods. *Journal of the American Planning Association*, *61*, 210-225.

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research D*, *2*(3), 199-219.

Cervero, R., & Landis, J. (1995). The transportation–land-use connection still matters. *Access*, *7*, 2-10.

Cervero, R., & Radisch, C. (1996). Travel choices in pedestrian versus automobile oriented neighborhoods. *Transport Policy*, *3*(3), 127-141.

Cooper, G. (1999). An overview of the representation and discovery of causal relationships using Bayesian networks. In C. Glymour and G. Cooper (Eds.), *Computation, causation and discovery* (pp. 3-62). Cambridge, MA: The MIT Press.

Crane, R. (1996). On form versus function: will the New Urbanism reduce traffic, or increase it? *Journal of Planning Education and Research*, *15*(2), 117-126.

Crane, R. (2000). The influence of urban form on travel: An interpretive review. *Journal of Planning Literature*, *15*(1), 3-23.

Crane, R., & Crepeau, R. (1998). Does neighborhood design influence travel? A behavioral analysis of travel diary and GIS data. *Transportation Research D*, *4*, 225-238.

DiStefano, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling*, *9*, 327-346.

Domencich, T. A., & McFadden, D. L. (1975). *Urban travel demand: A behavioral analysis.* A Charles River Associates research study. Amsterdam: North-Holland Pub. Co. Reprinted by The Blackstone Company, Mount Pleasant, MI, 1996.

Druzdzel, M., & Glymour, C. (1999). Causal inference from databases: Why university lose students. In C. Glymour and G. Cooper (Eds.), *Computation, causation and discovery* (pp. 521-539). Cambridge, MA: The MIT Press.

Dupuy, G. (1999). From the "Magic Circle" to "Automobile Dependence": Measurements and political implications. *Transport Policy*, *6*, 1-17.

European Council of Ministers of Transport (ECMT). (2004). *Assessment and decision making for sustainable transport*. Paris, France: Organization for Economic Co-operation and Development.

Ewing, R., & Cervero, R. (2001). Travel and the built environment: A synthesis. *Transportation Research Record*, *1780*, 87-114.

Gaudry, M., & Dagenais, M. (1979). The Dogit model. *Transportation Research B*, *13B*, 105-111.

Giuliano, G. (1995). The weakening transportation–land-use connection. *Access*, *6*, 3-11.

Giuliano, G., & Small, K. (1993). Is the journey to work explained by urban structure? *Urban Studies*, *30*, 1485-1500.

Gomez-Ibanez, J. (1991). A global view of automobile dependence. *Journal of the American Planning Association*, *57*, 376-379.

Goodwin, P. (1997). Mobility and car dependence. In T. Rothengatter and E. Vaya (Eds.), *Traffic and transport psychology: Theory and application* (pp. 449-464). Amsterdam, Netherlands: Pergamon.

Gordon, P., & Richardson, H. (1989). Gasoline consumption and cities: A reply. *Journal of the American Planning Association*, *55*, 342-346.

Handy, S. (1996a). Methodologies for exploring the link between urban form and travel behavior. *Transportation Research D*, *2*, 151-165.

Handy, S. (1996b). Urban form and pedestrian choices: study of Austin neighborhoods. *Transportation Research Record*, *1552*, 135-144.

Handy, S. (2002). *Accessibility- vs. mobility-enhancing strategies for addressing automobile dependence.* Paper prepared for the 2002 European Conference of Ministers of Transport. Department of Environmental Science and Policy, University of California at Davis.

Handy, S., Cao, X., & Mokhtarian, P. (2005). Correlation or causality between the built environment and travel behavior? Evidence from Northern California. *Transportation Research D*, *10*(6), 427-444.

Handy, S., Cao, X., & Mokhtarian, P. L. (2006). Self-selection in the relationship between the built environment and walking. *Journal of the American Planning Association*, *72*(1), 55-74.

Handy, S., Clifton, K., & Fisher, J. (1998). *The effectiveness of land-use policies as a strategy for reducing automobile dependence: A study of Austin neighborhoods* (Report SWUTC/98/465650-1). The University of Texas at Austin.

Hausman, D. (1998). *Causal asymmetries*. Cambridge, UK: Cambridge University Press.

Houston-Galveston Area Council (HGAC). (2007). *Bridging our communities: The 2035 Houston-Galveston regional transportation plan*. October 26, 2007. Retrieved September 17, 2009, from http://www.h-gac.com/taq/plan/default.aspx

Houston-Galveston Area Council (HGAC). (2009a). *Transportation and air quality*. Retrieved September 1, 2009, from http://www.h-gac.com/home/

Houston-Galveston Area Council (HGAC). (2009b). *Population and employment forecasts*. Retrieved September 1, 2009, from http://www.h-gac.com/rds/forecasts/default.aspx

Houston-Galveston Area Council (HGAC). (2009c). *STAR*Map*. Retrieved September 1, 2009, from http://www.h-gac.com/rds/gis/starmap/default.aspx

Jacobs, J. (1992). *The death and life of great American cities.* New York: Vintage Books.

Jeon, C. M., & Amekudzi, A. (2005). Addressing sustainability in transportation systems: Definition, indicators, and metrics. *Journal of Infrastructure System, 11*(1). 31-50.

Jeon, C. M., Amekudzi, A., & Vanegas, J. (2006). Transportation system sustainability issues in high-, middle-, and low-income economies: Case studies from Georgia (U.S.), South Korea, Colombia, and Ghana. *Journal of Urban Planning and Development*, *132*(3), 172-186.

Kaplan, D. (1988). The impact of specification error on the estimation, testing, and improvement of structural equation models. *Multivariate Behavioral Research*, *23*, 69-86.

Kaplan, D. (2009). *Structural equation modeling: Foundations and extensions* (2nd ed.), Los Angeles, CA: Sage.

Kaplan, D., & Wenger, R. N. (1993). Asymptomatic independence and separability in covariance structure models: Implications for specification error, power, and model modification. *Multivariate Behavioral Research*, *28*, 483-498.

Khattak, A., & Rodriguez, D. (2005). Travel behavior in neo-traditional neighborhood developments: A case study in USA. *Transportation Research A*, *39*(6), 481-500.

Kitamura, R., Mokhtarian, P., & Daidet, L. (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. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Gilford Press.

Knaap, G., Song, Y., & Nedovic-Budic, Z. (2007). Measuring patterns of urban development: New intelligence for the war on sprawl. *Local Environment*, *12*(3), 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.

Kopits, E., McConnell, V., & Miles, D. (2009). *Lot size, zoning, and household preferences: Impediments to smart growth?* Resources for the Future discussion paper, RFF DP 09-15, April 2009.

Koppelman, F., & Bhat, C. (2006). *A self instructing course in mode choice modeling: Multinomial and nested logit models*. Manual prepared for U.S. Department of Transportation Federal Transit Administration, January 2006.

Krizek, K. (2003). Residential relocation and changes in urban travel: Does neighborhood-scale urban form matter? *Journal of the American Planning Association*, *69*(3), 265-281.

Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy*, *34*, 132-157.

Lee, S. (2006). *The correlational and causal investigation into the land use-transportation relationships: evidence from the Dallas-Fort Worth metropolitan area.* Unpublished doctoral dissertation, Texas A&M University, College Station, TX.

Lee, S. (2007). *Structural equation modeling: A Bayesian approach.* Chichester, England; Wiley.

Levine, J. & Inam, A. (2004). The market for transportation-land use integration: Do developers want smarter growth than regulations allow? *Transportation*, *31*(4), 409-427.

Levine, J. (2006). *Zoned out: Regulation, markets, and choices in transportation and metropolitan land-use*. Washington, DC: Resources for the Future.

Levinson, D., & Kumar, A. (1997). Density and the journey to work. *Growth and Change*, *28*: 147-172.

Litman, T. (2000). Transportation market reforms for sustainability. *Transportation Research Record*, *1702*, 11-20.

Litman, T. (2002). *The cost of automobile dependency and the benefits of balanced transportation.* Victoria Transport Policy Institute. Retrieved October 20, 2004, from http://www.vtpi.org

Litman, T. (2008a). *Well measured: developing indicators for comprehensive and sustainable transport planning*. Victory Transport Policy Institute. Retrieved July 17, 2009, from http://www.vtpi.org/documents/evaluation.php

Litman, T. (2008b). *Land use impacts on transport: How land use factors affect travel behavior.* Victory Transport Policy Institute. Retrieved July 17, 2009, from http://www.vtpi.org/ documents/evaluation.php

Litman, T. (2009) *Smart growth reforms: Changing planning, regulatory and fiscal practices to support more efficient land use*. Victoria Transport Policy Institute. Retrieved September 17, 2009, from http://www.vtpi.org/smart_growth_reforms. pdf.

Litman, T., & Burwell, D. (2006). Issues in sustainable transportation. *International Journal of Global Environmental Issues, 6*(4), 331-347.

Litman, T., & Laube, F. (2002). *Automobile dependency and economic development*. Victoria Transport Policy Institute. Retrieved July 17, 2009, from http://www.vtpi.org/ecodev.pdf

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). New York: Wiley.

Manski. C. (1977). The structure of random utility models. *Theory and Decision*, *8*, 229-254.

McCarthy, P. (2001). *Transportation economic: Theory and practice: A case study approach*. Malden, MS: Blackwell Publishers Inc.

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

McFadden, D. (1976). The multinomial logit model when the population contains "captive" subpopulations. Unpublished memorandum.

Meyer, M., & Miller, E. (2001). *Urban transportation planning: A decision-oriented approach* (2nd ed.). Boston, MA: McGraw-Hill.

Mindali, O., Raveh, A., & Salomon, I. (2004). Urban density and energy consumption: A new look at old statistics. *Transportation Research A*, *38*, 143-162.

Muthén. L., & Muthén, B. (2006). *Mplus: Statistical analysis with latent variables*. Los Angeles, CA: Muthén & Muthén.

Newman, P., & Kenworthy, J. (1989a). *Cities and automobile dependence: An international sourcebook*. Brookfield, VT: Gower Publishing.

Newman, P., & Kenworthy, J. (1989b). Gasoline consumption and cities: A comparison of U.S. cities with a global survey. *Journal of the American Planning Association*, *55*, 24-37.

Newman, P., & Kenworthy, J. (1999). *Sustainability and cities: Overcoming automobile dependence.* Washington, DC: Island Press

Newman, P., Kenworthy, J., & Vintila, P. (1995). Can we overcome automobile dependence? Physical planning in an age of urban cynicism. *Cities*, *12*, 53-65.

Organization for Economic Co-operation and Development (OECD) (1999). *Indicators for the integration of environmental concerns into transport policies*, Environment Directorate, Paris, France. Retrieved July 17, 2009, from http://www.olis.oecd.org/

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, *82*, 669-710.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge, UK: Cambridge University Press.

Raad, T. (1998). *The car in Canada: a study of factors influencing automobile dependence in Canada's seven largest cities, 1961-1991.* Unpublished master's thesis, The University of British Columbia, Vancouver, Canada.

Roh, J., & Bessler, D. (1999). Occupant death: A study with direct graphs. *Applied Economics Letters*, *6*, 303-306.

Roh, J., Bessler, D., & Gilbert, R. (1999). Traffic fatalities, Peltzman's model, and directed graphs. *Accident Analysis and Prevention*, *31*, 55-61.

Scheines R., Spirtes, P., Glymour, C., Meek, C., & Richardson, T. (1996). *TETRAD 3: Tools for causal modeling*. User's manual. Manwah, NJ: Lawrence Erlbaum Associates.

Schipper, L. (2002). Sustainable urban transport in the 21st century: A new agenda. *Transportation Research Record*, *1792*. 12-19.

Schrank, D., & Lomax, T. (2002). *The 2002 urban mobility report*. College Station, TX: Texas Transportation Institute.

Schrank, D., & Lomax, T. (2007). *The 2007 urban mobility report*. College Station, TX: Texas Transportation Institute. Retrieved November 20, 2007, from http://mobility.tamu.edu/

Schwanen, T., & Mokhtarian, P. L. (2005). What affects commute mode choice: Neighborhood physical structure or preferences toward neighborhoods? *Journal of Transport Geography*, *13*(1), 83-99.

Shocker, A., Ben-Akiva, M., Boccara, B., and Nedungadi, P. (1991). Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions. *Marketing Letters*, *2*(3), 181-197.

Simonoff, J. (2003). *Analyzing categorical data*. New York: Springer.

Song, Y., & Knaap, G. (2004). Measuring urban form: Is Portland winning the war on sprawl? *Journal of the American Planning Association*, *70*(2), 210-225.

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction and search* (2nd ed.), Cambridge, MA: The MIT Press.

Steiger, J. H., & Lind, J. M. (1980). *Statistically based tests for the number of common factors*. Paper presented at the Psychometric Society, Iowa City, IA.

Stradling, S. (2001). *Measuring individual car dependence*. Paper presented at the Universities Transport Study Group Annual Conference, Oxford, UK.

Sustainable Transportation Indicators Subcommittee (STI) (2008). *Sustainable transportation indicators: A recommended research for developing sustainable transportation indicators and data* (ADD40 [1]). Paper submitted at the 88th Annual Meeting of the Transportation Research Board, Washington DC: Transportation Research Board. Retrieved July 17, 2009, from http://www.vtpi.org/documents/ evaluation.php

Swait, J. (2001). Choice set generation within the generalized extreme value family of discrete choice models. *Transportation Research B*, *35*, 643-666.

Swait, J., & Ben-Akiva, M. (1986a). Constraint on individual travel behavior in a Brazilian city. *Transportation Research Record*, *1085*, 75-85.

Swait, J., & Ben-Akiva, M. (1986b). Analysis of the effects of captivity on travel time and cost elasticities. In *Behavioural Research for Transport Policy* (pp. 119-134). Utrecht, The Netherlands: VNU Science Press.

Swait, J., & Ben-Akiva, M. (1987a). Incorporating random constraints in discrete models of choice set generation. *Transportation Research B*, *21 (2)*, 91-102.

Swait, J., & Ben-Akiva, M. (1987b). Empirical test of a constrained choice discrete model: Mode choice in Sao Paulo, Brazil. *Transportation Research B*, *21(2)*, 103-115.

Ramani, T., Zietsman, J., Eisele, W., Rosa, D., Spillane, D., & Bochner, B. (2009). *Developing sustainable transportation performance measures for TxDOT's* *strategic plan* (Report FHWA/TX-09-5541-1). College Station, TX: Texas Transportation Institute.

Thompson, B. (2000). Ten commandments of structural equation modeling. In L. G. Grimm, & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics*. (pp. 261-283). Washington, DC: American Psychological Association.

U.S. Census Bureau (2002). Demographic trend in the 20th century, November 2002. Retrieved July 17, 2009, from http://www.census.gov/

U.S. Census Bureau (2007). 2005-2007 American community survey 3-year estimates. Retrieved July 17, 2009, from http://factfinder.census.gov/

U.S. Census Bureau (2008). The 2008 statistical abstract: 2008 edition. Retrieved July 17, 2009, from http://www.census.gov/

U.S. Census Bureau (2009). 2008 population estimates. Retrieved July 17, 2009, from http://factfinder.census.gov/

U.S. Department of Energy (U.S. DOE) & U.S. Environmental Protection Agency (U.S. EPA). (2009). *Fuel economy estimates*. Retrieved April 4, 2009, from http://www.fueleconomy.gov/

Verma, T. & Pearl, J. (1990). Equivalence and synthesis of causal models. *Proceedings of the sixth conference on uncertainty in AI*, Mountain View, CA: Association for Uncertainty in AI.

Victory Transportation Policy Institute (VTPI). (2008a). *Automobile dependency: Transportation and land use patterns that cause high levels of automobile use and reduced transport options.* TDM Encyclopedia: Victory Transportation Policy Institute. Retrieved July 17, 2009, from http://www.vtpi.org/tdm/tdm100.htm

Victory Transportation Policy Institute (VTPI). (2008b). *New urbanism: Clustered, mixed-Use, multi-modal neighborhood*. TDM Encyclopedia: Victory Transportation Policy Institute. Retrieved September 20, 2009, from http://www.vtpi.org/tdm/tdm24.htm

Victory Transportation Policy Institute (VTPI). (2008c). *Smart growth: More efficient land use management*. TDM Encyclopedia: Victory Transportation Policy Institute. Retrieved September 20, 2009, from http://www.vtpi.org/tdm/tdm38.htm

Victory Transportation Policy Institute (VTPI). (2008d). *Transit oriented development: Using public transit to create more accessible and livable neighborhoods*. TDM

Encyclopedia: Victory Transportation Policy Institute. Retrieved September 20, 2009, from http://www.vtpi.org/tdm/tdm45.htm

Wachs, M. (2005) *What are the challenges to creating sustainable transportation?: How can transportation systems become more sustainable?* Resource Paper for Conference Proceedings 37. Integrating Sustainability into the Transportation Planning Process. Washington DC: Transportation Research Board.

World Commission on Environment and Development (WCED). (1987). *Our common future*, Oxford, UK: Oxford University Press.

World Health Organization (WHO). (2000). *Transport, environment and health*. Copenhagen: World Health Organization, Regional Office for Europe.

Zhang, M. (2004). The role of land use in travel mode choice. *Journal of the American Planning Association*, *70*(3), 344-360.

Zhang, M. (2005). Intercity variations in the relationship between urban form and automobile dependence: Disaggregate analyses of Boston, Massachusetts; Portland, Oregon; and Houston, Texas. *Transportation Research Record*. *1902*, 55-62.

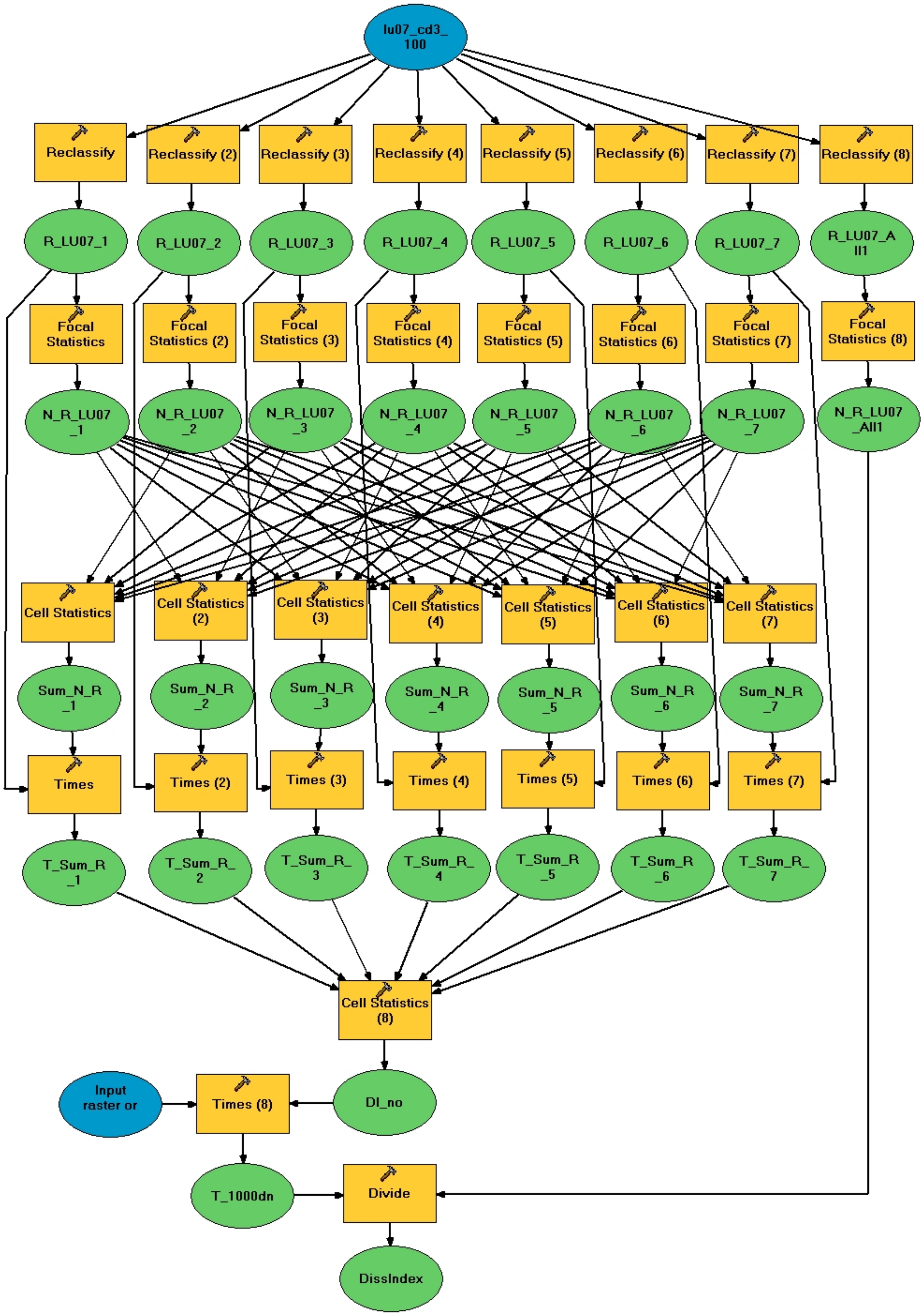
Zhang, M. (2006). Travel choice with no alternative: Can land use reduce automobile dependence? *Journal of Planning Education and Research*, *25*, 311-326.

Zietsman, J., & Rilett, L. (2002). *Sustainable transportation: Conceptualization and performance measures* (Report SWUTC/02/167403-1). College Station, TX: Texas Transportation Institute

Zietsman, J., Rilett, L., & Kim, S. (2003). *Sustainable transportation performance measures for developing communities* (Report SWUTC/03/167128-1). College Station, TX: Texas Transportation Institute.

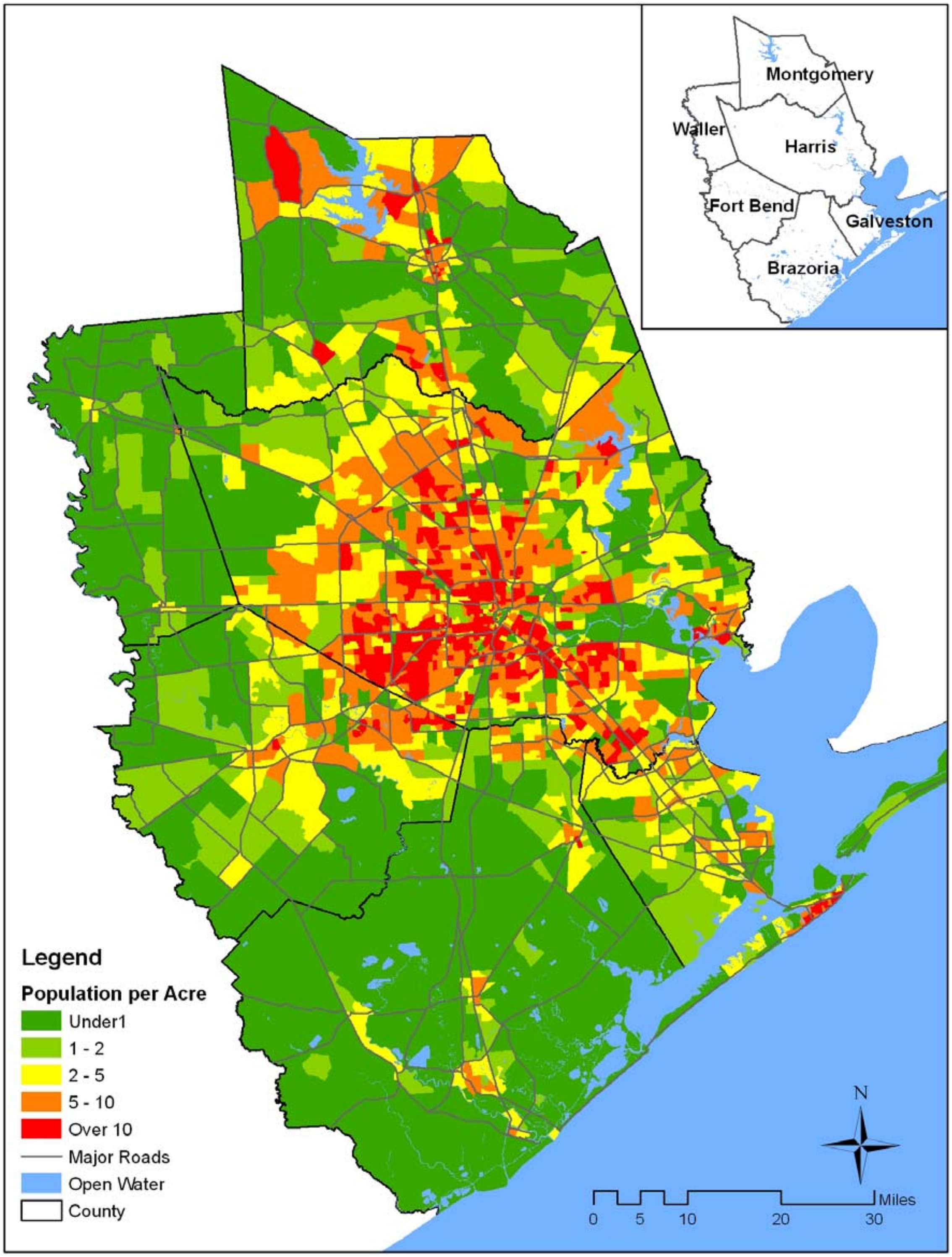
APPENDIX

A1. GIS Model for Computing Dissimilarity Index in the HGAC Region

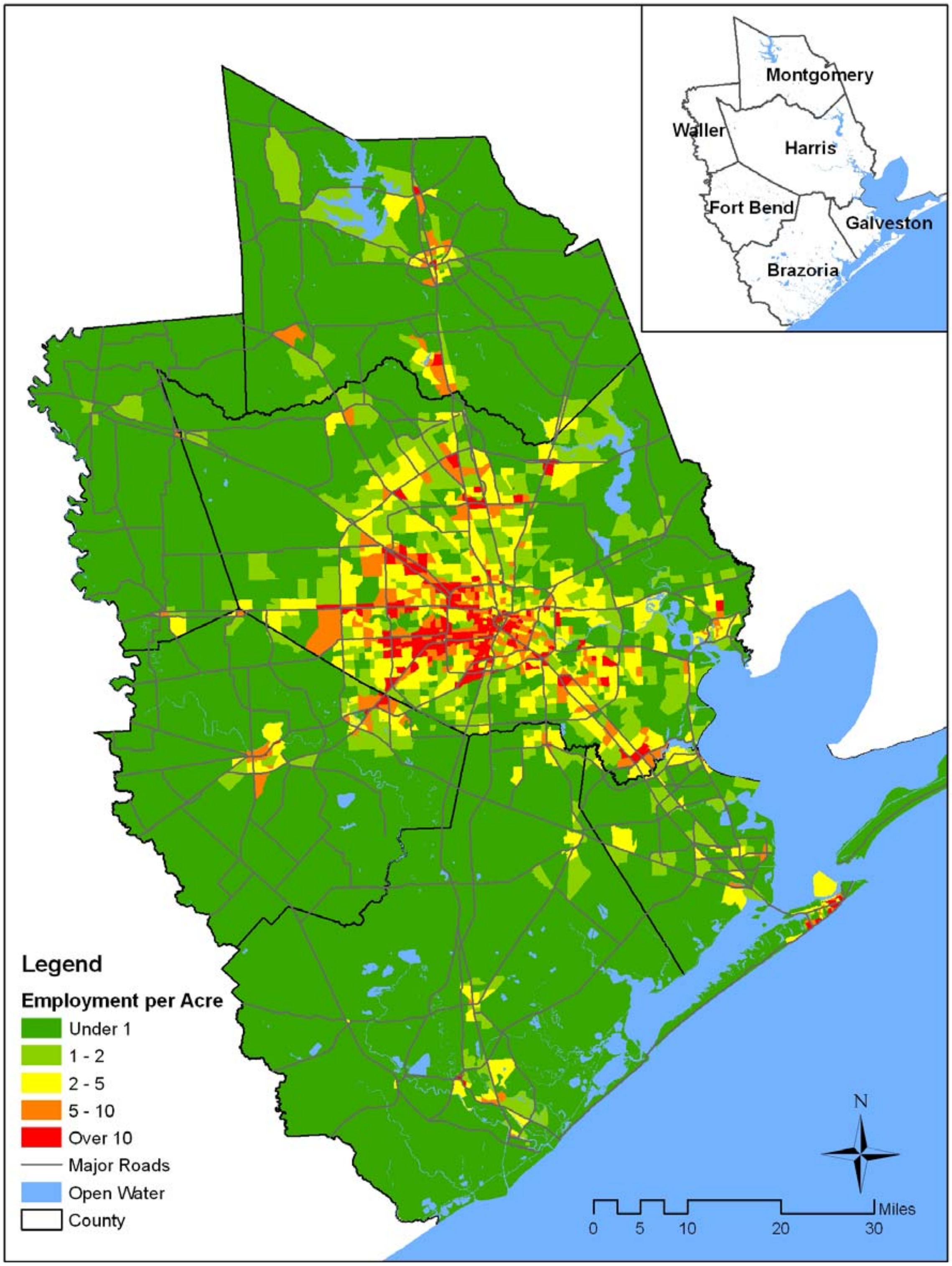


A2. Spatial Distribution of Land Use Measures in the HGAC Region

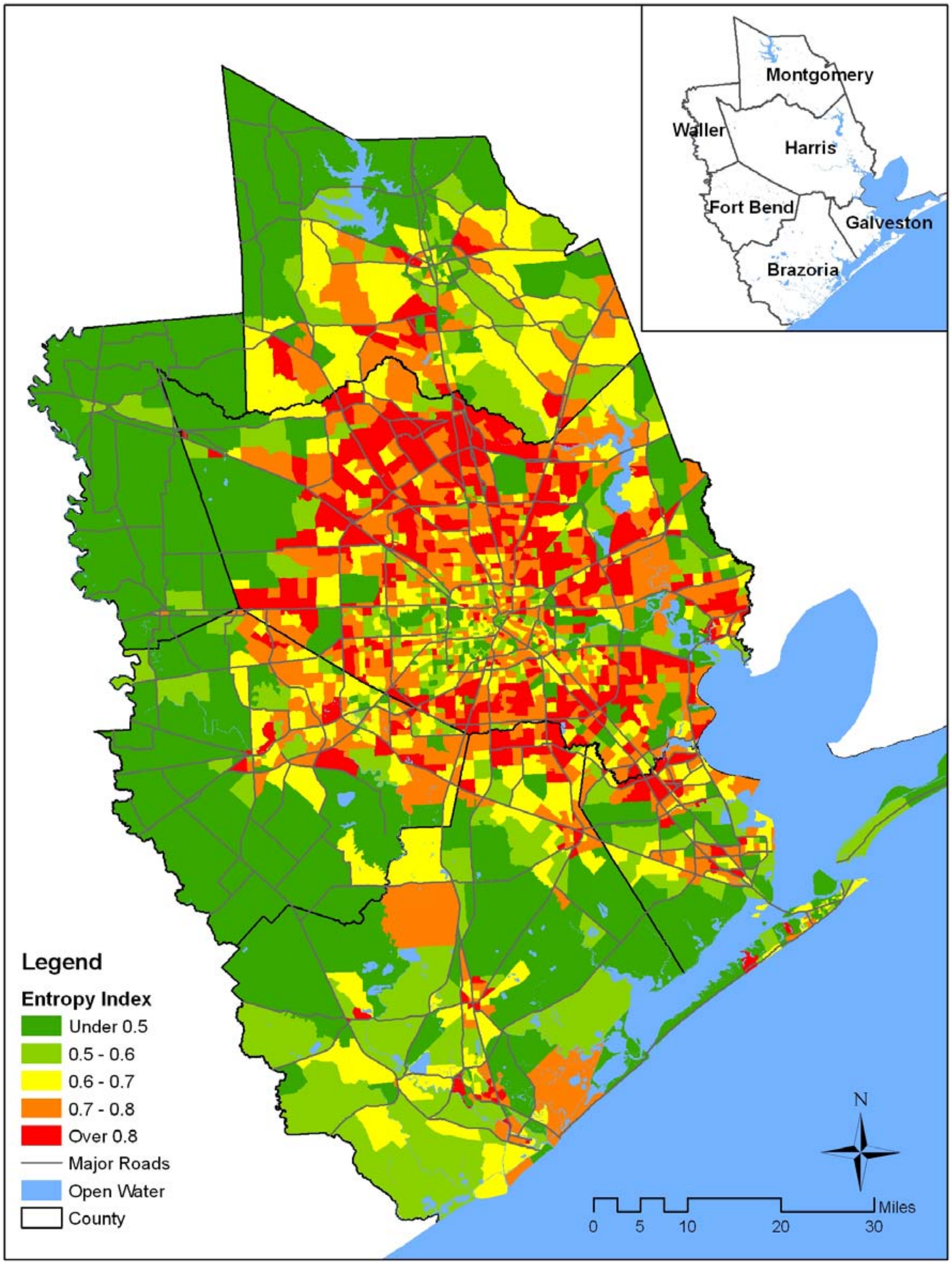
A2-1. Spatial Distribution of Population Density in the HGAC Region



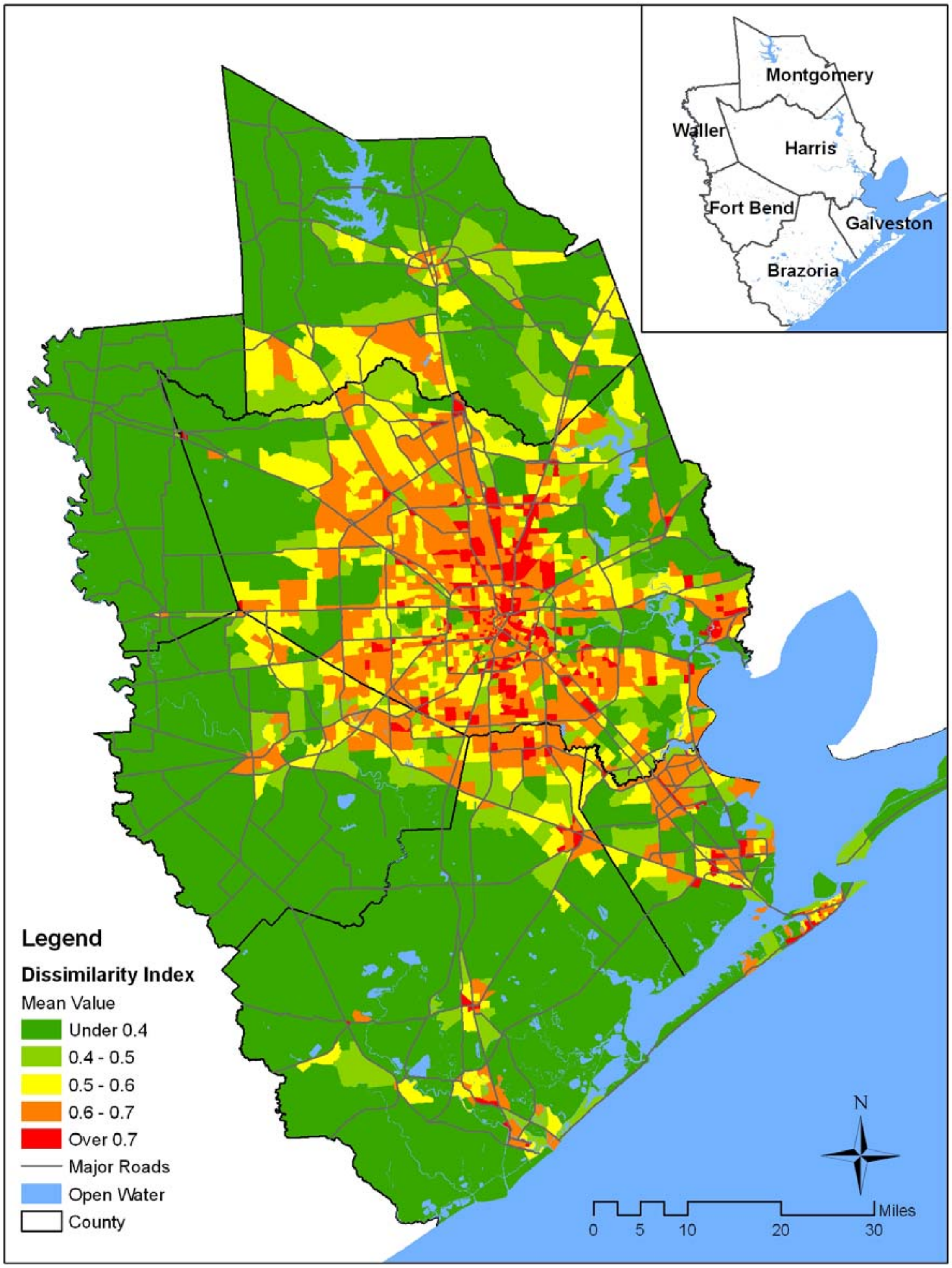
A2-2. Spatial Distribution of Employment Density in the HGAC Region



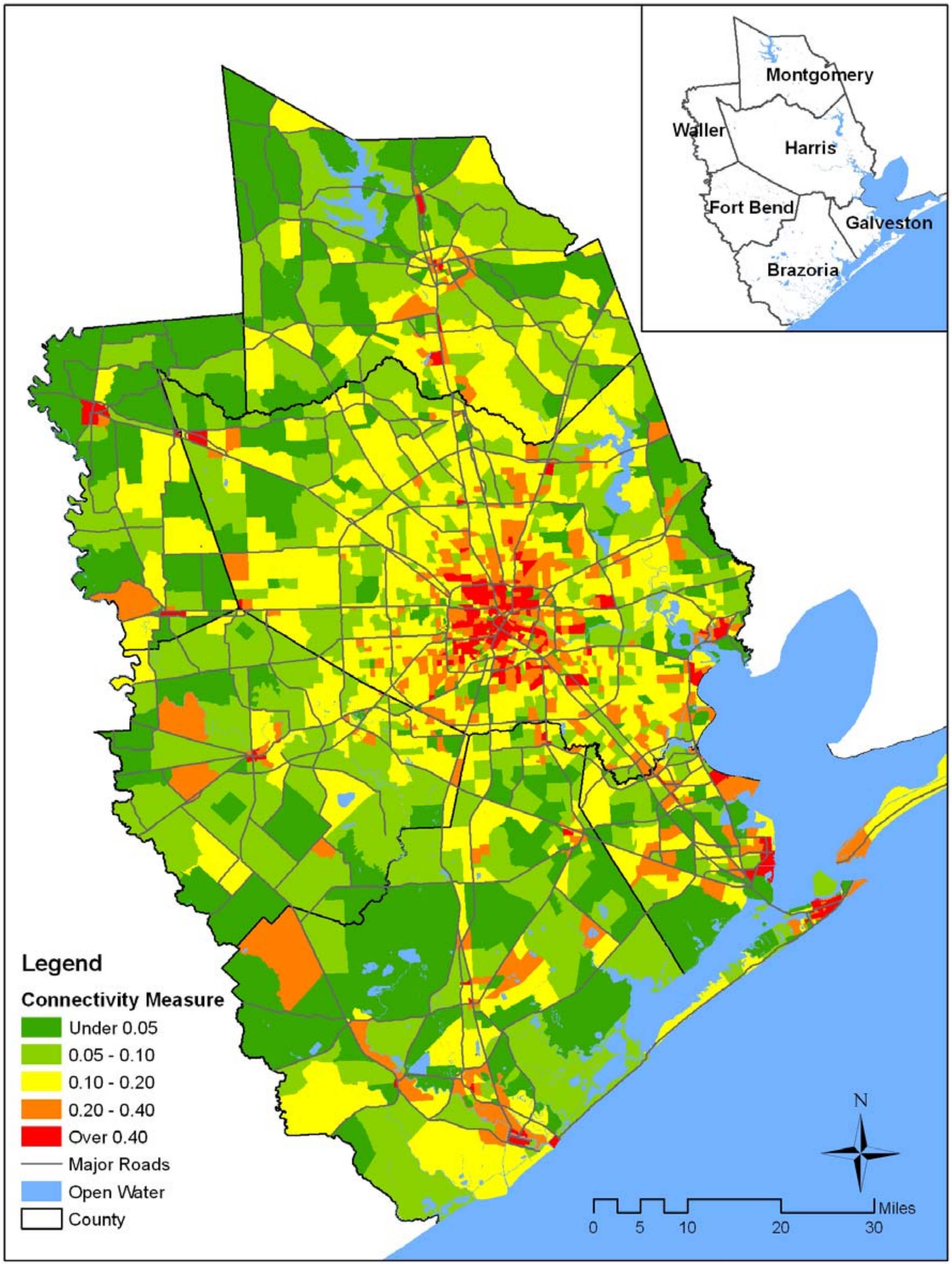
A2-3. Spatial Distribution of Entropy Index in the HGAC Region



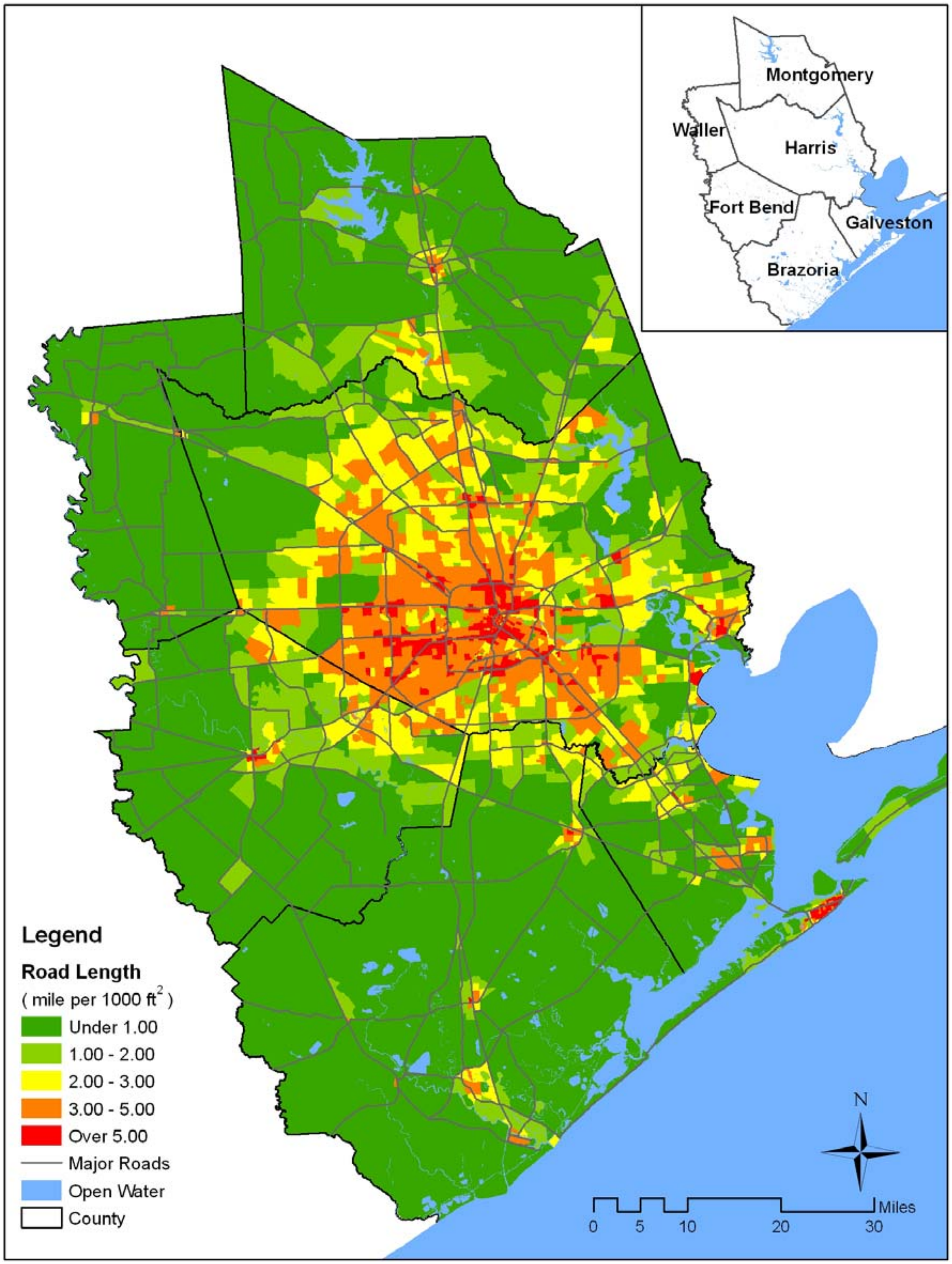
A2-4. Spatial Distribution of Dissimilarity Index in the HGAC Region



A2-5. Spatial Distribution of Connectivity Measure in the HGAC Region



A2-6. Spatial Distribution of Road Length Measure in the HGAC Region



A3. Correlation Matrices for Individual Mode Choice Models

A3-1. Correlation Matrix for HBW Trips (n = 6,239)

	auto	*ttdAU*	*bkus3*	*hhSi6*	*novh5*	*resS*	*inc15*	*opd2*	*dpd2*	*oed2*	*ded2*	*ozdis*	*dzdis*	*ordln*	*drdln*	*oconn*	*dconn*
auto	1.000																
ttdAU	0.094	1.000															
bkus3	-0.073	-0.004	1.000														
hhSi6	-0.004	0.187	0.030	1.000													
noVh5	0.070	0.130	0.014	0.476	1.000												
resS	0.092	0.049	0.017	0.144	0.244	1.000											
inc15	0.028	0.116	0.083	0.194	0.345	0.204	1.000										
opd2	-0.056	-0.152	-0.027	-0.115	-0.136	-0.159	-0.074	1.000									
dpd2	-0.024	-0.004	-0.002	-0.034	-0.056	-0.010	-0.044	0.224	1.000								
oed2	-0.149	-0.118	-0.025	-0.100	-0.118	-0.279	-0.043	0.314	0.071	1.000							
ded2	-0.031	0.092	-0.004	-0.015	-0.019	-0.019	0.091	0.067	-0.025	0.114	1.000						
ozdis	-0.035	0.027	-0.042	0.037	-0.061	-0.192	-0.132	-0.071	-0.006	0.227	-0.004	1.000					
dzdis	0.005	-0.080	0.022	-0.004	-0.019	0.005	-0.030	0.005	0.224	0.011	-0.041	0.027	1.000				
ordln	-0.040	-0.178	-0.016	-0.117	-0.123	-0.085	-0.058	0.571	0.164	0.316	0.076	-0.117	0.009	1.000			
drdln	0.004	0.119	-0.002	-0.041	-0.031	-0.001	0.025	0.201	0.374	0.107	0.259	-0.026	0.056	0.180	1.000		
oconn	-0.034	-0.157	-0.025	-0.087	-0.090	-0.029	-0.123	0.286	0.073	0.178	0.016	-0.003	0.002	0.439	0.073	1.000	
dconn	-0.014	0.037	-0.019	-0.028	-0.033	0.000	-0.033	0.116	0.166	0.065	0.253	-0.014	0.052	0.104	0.381	0.094	1.000

Note: For identifying variable names, refer to DAGs for appropriate model and travel purpose.

A3-2. Correlation Matrix for HBO Trips (n = 10,413)

	auto	*ttdAU*	*bkus3*	*hhSi6*	*novh5*	*resS*	*inc15*	*opd2*	*dpd2*	*oed2*	*ded2*	*ozdis*	*dzdis*	*ordln*	*drdln*	*oconn*	*dconn*
auto	1.000																
ttdAU	0.074	1.000															
bkus3	-0.115	-0.052	1.000														
hhSi6	-0.044	0.048	0.196	1.000													
noVh5	0.078	0.107	0.027	0.454	1.000												
resS	0.047	-0.008	0.023	0.129	0.185	1.000											
inc15	0.052	0.079	0.123	0.317	0.404	0.194	1.000										
opd2	-0.043	-0.117	-0.049	-0.154	-0.178	-0.142	-0.072	1.000									
dpd2	-0.009	-0.042	-0.039	-0.094	-0.112	-0.091	-0.037	0.393	1.000								
oed2	-0.016	-0.058	-0.021	-0.136	-0.134	-0.268	-0.047	0.325	0.209	1.000							
ded2	0.010	0.108	-0.021	-0.046	-0.041	-0.034	-0.006	0.082	0.034	0.090	1.000						
ozdis	-0.030	0.024	0.002	0.032	-0.060	-0.165	-0.115	-0.066	0.005	0.201	-0.011	1.000					
dzdis	0.042	-0.068	-0.010	-0.020	-0.038	0.004	-0.060	0.050	0.150	0.003	-0.072	0.073	1.000				
ordln	-0.032	-0.126	-0.041	-0.137	-0.166	-0.032	-0.042	0.585	0.311	0.282	0.051	-0.102	0.037	1.000			
drdln	0.022	0.021	-0.087	-0.097	-0.065	-0.042	-0.012	0.266	0.369	0.164	0.162	-0.031	-0.043	0.288	1.000		
oconn	-0.036	-0.084	-0.054	-0.103	-0.121	-0.011	-0.130	0.277	0.149	0.170	0.038	-0.002	0.068	0.420	0.114	1.000	
dconn	0.005	0.000	-0.033	-0.073	-0.052	-0.012	-0.080	0.098	0.164	0.075	0.137	0.032	0.004	0.088	0.305	0.161	1.000

Note: For identifying variable names, refer to DAGs for appropriate model and travel purpose.

A4. Correlation Matrices for Household Automobile Trip Generation Models

A4-1. Correlation Matrix for Total Trips (n = 3,976)

	trips	*costpt*	*hhSi6*	*novh5*	*resS*	*inc15*	*opd2*	*oed2*	*oent*	*ordln*	*oconn*
trips	1.000										
costpt	-0.199	1.000									
hhSi6	0.410	0.031	1.000								
novh5	0.371	0.134	0.474	1.000							
resS	0.128	0.024	0.147	0.211	1.000						
inc15	0.316	0.098	0.263	0.378	0.172	1.000					
opd2	-0.047	-0.164	-0.137	-0.164	-0.200	-0.074	1.000				
oed2	-0.043	-0.125	-0.120	-0.121	-0.287	-0.031	0.320	1.000			
oent	-0.037	-0.032	0.029	-0.057	-0.130	-0.116	-0.008	0.198	1.000		
ordln	-0.034	-0.195	-0.114	-0.149	-0.092	-0.055	0.578	0.298	-0.023	1.000	
oconn	-0.040	-0.148	-0.087	-0.094	-0.029	-0.120	0.279	0.164	0.033	0.431	1.000

Note: For identifying variable names, refer to DAGs for appropriate model and travel purpose.

A4-2. Correlation Matrix for Total Home-based Trips (n = 3,973)

	trips	*costpt*	*hhSi6*	*novh5*	*resS*	*inc15*	*opd2*	*oed2*	*oent*	*ordln*	*oconn*
trips	1.000										
costpt	-0.140	1.000									
hhSi6	0.485	0.016	1.000								
novh5	0.434	0.146	0.474	1.000							
resS	0.138	0.034	0.146	0.211	1.000						
inc15	0.315	0.115	0.263	0.377	0.172	1.000					
opd2	-0.054	-0.183	-0.136	-0.163	-0.199	-0.073	1.000				
oed2	-0.050	-0.125	-0.120	-0.120	-0.286	-0.031	0.320	1.000			
oent	-0.038	-0.035	0.030	-0.056	-0.129	-0.115	-0.009	0.197	1.000		
ordln	-0.024	-0.214	-0.115	-0.149	-0.092	-0.054	0.578	0.299	-0.023	1.000	
oconn	-0.031	-0.167	-0.087	-0.094	-0.028	-0.120	0.278	0.164	0.033	0.430	1.000

Note: For identifying variable names, refer to DAGs for appropriate model and travel purpose.

A4-3. Correlation Matrix for HBW Trips (n = 2,539)

	trips	*costpt*	*hhSi6*	*novh5*	*resS*	*inc15*	*opd2*	*oed2*	*oent*	*ordln*	*oconn*
trips	1.000										
costpt	-0.067	1.000									
hhSi6	0.193	0.107	1.000								
novh5	0.363	0.108	0.474	1.000							
resS	0.105	0.096	0.162	0.228	1.000						
inc15	0.136	0.111	0.222	0.321	0.186	1.000					
opd2	-0.010	-0.201	-0.132	-0.159	-0.192	-0.072	1.000				
oed2	-0.053	-0.147	-0.116	-0.119	-0.273	-0.030	0.321	1.000			
oent	-0.043	-0.008	0.035	-0.034	-0.131	-0.111	-0.012	0.204	1.000		
ordln	-0.001	-0.231	-0.137	-0.151	-0.103	-0.053	0.576	0.314	-0.049	1.000	
oconn	-0.009	-0.202	-0.095	-0.095	-0.038	-0.124	0.285	0.172	0.022	0.436	1.000

Note: For identifying variable names, refer to DAGs for appropriate model and travel purpose.

A4-4. Correlation Matrix for HBO Trips (n = 3,461)

	trips	*costpt*	*hhSi6*	*novh5*	*resS*	*inc15*	*opd2*	*oed2*	*oent*	*ordln*	*oconn*
trips	1.000										
costpt	-0.138	1.000									
hhSi6	0.391	-0.058	1.000								
novh5	0.251	0.089	0.466	1.000							
resS	0.096	0.018	0.138	0.205	1.000						
inc15	0.209	0.043	0.282	0.385	0.182	1.000					
opd2	-0.032	-0.175	-0.140	-0.161	-0.184	-0.066	1.000				
oed2	-0.033	-0.088	-0.122	-0.122	-0.293	-0.032	0.323	1.000			
oent	-0.006	-0.053	0.028	-0.057	-0.121	-0.108	-0.027	0.198	1.000		
ordln	-0.013	-0.201	-0.107	-0.148	-0.078	-0.055	0.578	0.285	-0.032	1.000	
oconn	-0.017	-0.135	-0.092	-0.096	-0.032	-0.128	0.279	0.160	0.022	0.429	1.000

Note: For identifying variable names, refer to DAGs for appropriate model and travel purpose.

A5. Correlation Matrices for Household Total VMT Models

A5-1. Correlation Matrix for Total Trips (n = 3,976)

	vmt	*costmi*	*trips*	*hhSi6*	*novh5*	*inc15*	*opd2*	*oed2*	*oent*	*ordln*	*oconn*
vmt	1.000										
costmi	-0.019	1.000									
trips	0.573	-0.004	1.000								
hhSi6	0.330	0.082	0.410	1.000							
novh5	0.417	0.060	0.371	0.474	1.000						
inc15	0.345	0.077	0.316	0.263	0.378	1.000					
opd2	-0.172	-0.081	-0.047	-0.137	-0.164	-0.074	1.000				
oed2	-0.136	-0.025	-0.043	-0.120	-0.121	-0.031	0.320	1.000			
oent	-0.033	-0.056	-0.037	0.029	-0.057	-0.116	-0.008	0.198	1.000		
ordln	-0.173	-0.061	-0.034	-0.114	-0.149	-0.055	0.578	0.298	-0.023	1.000	
oconn	-0.149	-0.038	-0.040	-0.087	-0.094	-0.120	0.279	0.164	0.033	0.431	1.000

Note: For identifying variable names, refer to DAGs for appropriate model and travel purpose.

A5-2. Correlation Matrix for Total Home-based Trips (n = 3,973)

	vmt	*costmi*	*trips*	*hhSi6*	*novh5*	*inc15*	*opd2*	*oed2*	*oent*	*ordln*	*oconn*
vmt	1.000										
costmi	-0.002	1.000									
trips	0.527	0.013	1.000								
hhSi6	0.326	0.078	0.485	1.000							
novh5	0.432	0.064	0.434	0.474	1.000						
inc15	0.324	0.075	0.315	0.263	0.377	1.000					
opd2	-0.189	-0.083	-0.054	-0.136	-0.163	-0.073	1.000				
oed2	-0.136	-0.030	-0.050	-0.120	-0.120	-0.031	0.320	1.000			
oent	-0.026	-0.056	-0.038	0.030	-0.056	-0.115	-0.009	0.197	1.000		
ordln	-0.182	-0.065	-0.024	-0.115	-0.149	-0.054	0.578	0.299	-0.023	1.000	
oconn	-0.150	-0.041	-0.031	-0.087	-0.094	-0.120	0.278	0.164	0.033	0.430	1.000

Note: For identifying variable names, refer to DAGs for appropriate model and travel purpose.

A5-3. Correlation Matrix for HBW Trips (n = 2,539)

	vmt	*costmi*	*trips*	*hhSi6*	*novh5*	*inc15*	*opd2*	*oed2*	*oent*	*ordln*	*oconn*
vmt	1.000										
costmi	-0.028	1.000									
trips	0.480	0.021	1.000								
hhSi6	0.195	0.004	0.193	1.000							
novh5	0.297	0.069	0.363	0.474	1.000						
inc15	0.165	0.008	0.136	0.222	0.321	1.000					
opd2	-0.129	-0.076	-0.010	-0.132	-0.159	-0.072	1.000				
oed2	-0.146	-0.034	-0.053	-0.116	-0.119	-0.030	0.321	1.000			
oent	0.011	-0.023	-0.043	0.035	-0.034	-0.111	-0.012	0.204	1.000		
ordln	-0.150	-0.060	-0.001	-0.137	-0.151	-0.053	0.576	0.314	-0.049	1.000	
oconn	-0.164	-0.016	-0.009	-0.095	-0.095	-0.125	0.285	0.172	0.022	0.436	1.000

Note: For identifying variable names, refer to DAGs for appropriate model and travel purpose.

A5-4. Correlation Matrix for HBO Trips (n = 3,461)

	vmt	*costmi*	*trips*	*hhSi6*	*novh5*	*inc15*	*opd2*	*oed2*	*oent*	*ordln*	*oconn*
vmt	1.000										
costmi	-0.095	1.000									
trips	0.520	-0.034	1.000								
hhSi6	0.176	0.098	0.391	1.000							
novh5	0.253	0.033	0.251	0.466	1.000						
inc15	0.157	0.076	0.209	0.282	0.385	1.000					
opd2	-0.182	-0.044	-0.032	-0.140	-0.161	-0.066	1.000				
oed2	-0.105	-0.012	-0.033	-0.122	-0.122	-0.032	0.323	1.000			
oent	-0.015	-0.048	-0.006	0.028	-0.057	-0.108	-0.027	0.198	1.000		
ordln	-0.188	-0.034	-0.013	-0.107	-0.148	-0.055	0.578	0.285	-0.032	1.000	
oconn	-0.118	-0.024	-0.017	-0.092	-0.096	-0.128	0.279	0.160	0.022	0.429	1.000

Note: For identifying variable names, refer to DAGs for appropriate model and travel purpose.

VITA

Jae Su Lee

Email: nowwater@gmail.com

ADDRESS

Raemian Bangbae ArtHill Apt. 109-603, Bangbae 3-dong, Seocho-gu, Seoul, Republic of Korea 137-936

EDUCATION

Ph.D. Urban and Regional Planning, Texas A&M University, 2009.

M.S., Urban and Regional Planning, Seoul National University, Korea, 2001.

B.S., Urban Engineering, University of Seoul, Korea, 1999.

RESEARCH EXPERIENCE

Graduate Research Assistant, Center for Air Quality Studies, Texas Transportation Institute, September 2006 – December 2009.

Graduate Student Worker, Environmental Management, Texas Transportation Institute, July 2006 – August 2006.

Graduate Research Assistant, Department of Landscape Architecture & Urban Planning, Texas A&M University, September 2005 – May 2006.

Research Associate, Department of Urban Planning & Design, Seoul Development Institute, Seoul, Korea, January 2001 – July 2004.

PUBLICATIONS

Lee, J., and Li, M. (2009). The impact of detention basin design on residential property value: Case studies using GIS in the hedonic price modeling. *Landscape and Urban Planning*, *89*(1-2), 7-16.

Lee, S., and Lee, J. (2009). A study of the area management activities through the public/private partnership for sustainable urban generation: Case studies of core areas within large cities in Japan. *Journal of Korea Planners Association*, *44*(1), 45-59.

Farzaneh, M., Lee, J., Ramani, T., Higgins, L., and Zietsman, J. (2009). *Toward a green campus: A transportation strategy for Texas A&M University* (Report SWUTC/09/167174-1). College Station, TX: Texas Transportation Institute.