THE CORRELATIONAL AND CAUSAL INVESTIGATION INTO THE LAND USE–TRANSPORTATION RELATIONSHIPS: EVIDENCE FROM THE DALLAS-FORT WORTH METROPOLITAN AREA

A Dissertation

by

SANGKUG 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

August 2006

Major Subject: Urban and Regional Science

THE CORRELATIONAL AND CAUSAL INVESTIGATION INTO

THE LAND USE–TRANSPORTATION RELATIONSHIPS:

EVIDENCE FROM THE DALLAS-FORT WORTH METROPOLITAN AREA

A Dissertation

by

SANGKUG 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 Zhang
	Chanam Lee
Committee Members,	David A. Bessler
	Daniel Z. Sui
	Mark W. Burris
Head of Department,	Forster Ndubisi

August 2006

Major Subject: Urban and Regional Science

ABSTRACT

The Correlational and Causal Investigation into the Land Use–Transportation Relationships: Evidence from the Dallas-Fort Worth Metropolitan Area. (August 2006)

Sangkug Lee,

B.S., Chungnam National University;

M.S., Purdue University

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

The role of land-use and related policies in reducing automobile dependence has been the subject of heated policy debate for over two decades. Previous research has shed light on the correlations between land-use and travel. Yet a crucial knowledge gap still exists in establishing causality between the two. Do changes in land-use characteristics *cause* behavioral changes in individuals' decisions on what transportation means to use for travel? How does land-use as a contextual factor shape the decision process and outcome of trip frequency and travel mode choice? These questions remain largely unanswered.

Attempting to fill the gap, this study applied the directed acyclic graphs method to identify the causal relationship between land-use and travel in the 9-county Dallas-Fort Worth (D-FW) metropolitan area. The logit captivity (LC) model, an extension to the conventional multinomial logit, was utilized to capture the contribution of land-use in affecting individuals' decisions on travel mode choice. All the data for this study were

obtained from the North Central Texas Council of Governments (NCTCOG).

Evidence from the D-FW region confirms to a certain extent the causal effects of land-use on travel. For work trips, increases in regional accessibility, job density and share of commercial land-use reduce the use of automobiles. Higher regional accessibility, however, causes households to generate automobile trips and thus leads to the increase in vehicle miles of travel (VMT). For non-work trips, population density, job density and regional accessibility are direct causes of the choice of automobile, while only regional accessibility is causally connected to reducing automobile trips and VMT. The logit captivity model results indicate that land-use contributes to captive-driving choices for home-based work trips. Lack of land-use mix at trip origins increases the probabilities of trip-makers being captive to the automobile from 0.06% to 5.62% for driving-alone and from 0.38% to 3.55% for shared-ride.

DEDICATION

To my family, Yoonhee and Eugene

and

my late father, Mr. Jaegeun Lee

ACKNOWLEDGMENTS

I have had countless benefits from the wonderful persons and environments to study transportation planning and economics for my doctoral program. I would like to thank Dr. Ming Zhang, one of my co-chairs at the University of Texas at Austin, for his thorough and cordial guidance. Without his invaluable help, this dissertation would never have been completed. Also I would like to thank Dr. Chanam Lee, my co-chair, who has helped me make progress on my dissertation and given me invaluable comments on issues. I would like to give sincere gratitude to Dr. David Bessler who has helped me through the entire period of graduate study and has helped me smooth out the coarse edges of my scholarship and knowledge in analytical methods. Many thanks go to Dr. Mark Burris who taught interesting topics in the transportation economics course and for giving valuable comments for improving the quality of the dissertation. Dr. Daniel Sui gave constructive comments and suggestions. Mr. Kenneth Cervenka of the North Central Texas Council of Governments kindly provided the data required for this research. Special thanks go to Dr. Sharada Vadali who offered research opportunity in the Economics and Policy Program, Texas Transportation Institute. My gratitude is extended to my brother (Sang-Il), parents-in-law, and sister-in-law. Thanks to all who are not listed here but who encouraged me during my graduate study.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES (Continued)

TABLE Page

CHAPTER I

INTRODUCTION

Automobile dependence has been growing in the United States. According to the 2001 National Household Travel Survey (NHTS), the increase in the number of vehicles (179%) over the past three decades (1969 through 2001) far exceeded the growth in population (41%), household (72%), and workers (92%). Approximately 86 percent of average annual person trips per household relied on private vehicles in 2001. Daily vehicle miles traveled (VMT) per household increased by 40 percent between 1990 and 2001, from 41 miles to 58 miles. The average daily time spent in driving also increased from 50 minutes in 1990 to 62 minutes in 2001. Among the well-documented negative consequences of automobile dependence are air and water pollution, energy consumption, fatalities and injuries from traffic accidents, costs of traffic congestion, land consumption and environmental degradation, and many public health problems related to pollution and sedentary lifestyles such as obesity, cancer, cardiovascular diseases, and respiratory diseases (BTS 2004a, 2004b; WHO 2000).

The pattern of urban growth in the decades since World War II is partly responsible for the increase in automobile ownership and use. Urban growth in this time period can be characterized mainly by low-density development and employment decentralization (Mills 1992; Glaser and Kahn 2004). Along with the extensive interstate

---

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

highway construction and investments on other roadways, this dispersed and segregated land-use has made driving a necessity, not an option, for people's daily living.

There have been attempts to find effective land-use policies to reduce automobile dependence. Early research has focused on advancing knowledge about the interactions between travel behavior and land-use. There has been remarkable progress in the refinements of land-use (or urban form) measures (Kockelman 1997; Cervero and Kockelman 1998). Several recent studies have explored new methods of modeling (Crane 1996a; Crane and Crepeau 1998; Boarnet and Sarmiento 1998; Boarnet and Crane 2001; Cervero 2002; Kockelman 2002; Zhang 2004; Bento, et al. 2005). There have also been significant improvements in theorizing with respect to the interactions between transportation and land-use. Advances in model specifications and estimations have improved our understanding of land-use influence on travel decisions, along with other factors such as price and traveler's socio-demographic characteristics. Many have thus prescribed densification, land-use mix, and infill development as land-use policies reducing automobile dependence.

Nevertheless, questions remain with regard to the effectiveness of land-use policies to reduce automobile dependence and to manage transportation demand. The efficacy of such policies is assessed *a priori* by the consistency of empirical results, but those are often mixed, complicated, and ambiguous. The inherent complexity behind these empirical studies is that there are numerous confounding factors affecting travel decisions. Teasing out the independent effect of land-use policies on travel is extremely difficult. Many shortcomings of existing studies in this area pertain to the limitations of data and methodologies.

This study builds on the previous research conducted to date on the relationship between land-use and transportation, and investigates how land-use affects travel demand in the Dallas-Fort Worth metropolitan area.

Research Objectives

The objectives of this study are:

1. To examine the causal relationships between transportation and land-use with application of the directed acyclic graphs (DAG) method for individual mode choice (automobile versus non-automobile mode), household trip frequency, and household total VMT;
2. To identify the captivity factors attributable to land-use with application of the multinomial logit captivity (LC) models for different trip purposes and then to estimate the intensity (probability) of captive driving as it relates to the land-use environment;
3. To explore how the impact of land-use on travel may differ between work trips and non-work trips; and
4. To draw implications of land-use policies for the purpose of reducing automobile dependence and its associated undesirable consequences.

Contributions of the Study

The major contributions of the study are twofold, both on methodological grounds. First, the study innovatively applies the directed acyclic graphs (DAG) method to address the causality issue that has intrigued researchers studying transportation–land-use connections. To the author's knowledge, this is the first attempt in this subject area to apply the DAG method to improve understanding of the role of land-use in influencing travel. Reliability of the DAG study results is cross-checked with conventional regression methods for estimating the models of travel mode choice, trip frequency, and household vehicle miles traveled (VMT).

Second, the study applies the multinomial logit captivity (LC) model to address captive driving behavior attributable to land-use. Existing studies have applied the LC method for a binary choice situation. This study has expanded existing research by estimating the multimodal LC models of travel model choice as it relates to land-use as well as other socioeconomic and demographic factors.

The empirical contribution of the study is also worth noting. The Dallas-Fort Worth (D-FW) region is one of the fastest growing regions in the nation. Driving demand is also growing rapidly in the region. There have been few studies, however, examining the relationship between land-use and travel by focusing on the D-FW region as a whole. Evidence of transportation–land-use connections identified in the region contributes to the literature in the field and helps better inform land-use and transportation policy making for the region and for other parts of Texas as well.

Organization of the Study

The remaining chapters are organized as follows: Chapter II reviews past studies on the issues of automobile dependence and the causal linkage between transportation and land-use. Chapter III sets out the conceptual frameworks and presents the research hypotheses drawn from travel demand models, causal graphical models and the logit captivity model. Chapter IV discusses research methodology including data sources, variables and measurements, and analytical approaches. In Chapter V, the empirical results are presented and the implications of findings are discussed. The final chapter highlights the key findings, and concludes with discussions of land-use and transportation policy implications of the research findings.

CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter reviews the literature that deals with automobile dependence, and studies examining the relationships between transportation and land-use. The first section provides an overview of literature discussing the definitions, measurements, and possible causes and consequences of automobile dependence. The second section reviews literature analyzing the link between travel behavior and land-use, covering the full array of variables in travel demand models. Past research investigating the issues of causality in land-use–transportation linkage is also reviewed in this section.

Automobile Dependence: Overview

Automobile dependence has been characterized and measured by a series of gross indicators such as annual gasoline consumption per capita, number of cars per person, vehicle miles traveled, per capita automobile travel frequency, automobile-oriented land-use patterns, fewer available transportation modes, and per capita multi-modal facilities (Newman and Kenworthy 1989a, 1989b, 1999; Kenworthy and Laube 1999; Handy 2002; Litman and Laube 2002). It has also been addressed within the context of its impact on the environment, society, the economy and public health. Environmental effects include land and habitat loss, resource depletion, climate change, and emissions (UNEP 1993; Freund and Martin 1993; Wackernagel and Rees 1996). Social effects include traffic fatalities and injuries, and equity (Altshuler 1979; Litman

1997; Mensah 1995; Litman 2002). Economic effects include aggregate costs associated with motor vehicles, internal (consumer) and external (social) costs, and economic development costs (Delucchi 1996; Litman 2002; Litman and Laube 2002). Public health effects include accidents, respiratory disease, and obesity (WHO 2000). These studies recognize automobile dependence as having social costs, such as roadway congestion, degradation of air quality, depletion of energy and natural resources, and urban sprawl.

A variety of policies and strategies have been initiated to lessen the social problems incurred by a great deal of transportation demand (Meyer 1999; VTPI 2005). Most metropolitan areas now have such policies with specific strategies including: congestion reduction – road pricing, transit improvements, rideshare programs, HOV priority, parking management and pricing, flextime, etc.; energy conservation and emission reduction – clean vehicles focusing on emission and fuel efficiency standards, travel demand management including distance-based emission fees, fuel tax, non-motorized transportation, ridesharing, speed reductions, etc.; and improvement of public health, equity, and safety – non-motorized transportation promotion, user-pays-drive, benefit programs for lower income or disadvantaged people, traffic speed reductions, etc.

Studies showing the current status of automobile dependence using gross indicators have contributed, to some extent, to awareness of automobile dependence. Many of the aforementioned studies center on descriptions of observed problems (i.e., increased use of automobile and motorized-oriented developments) associated with current transportation patterns. Now, automobile dependence is perceived to have much to do with land-use patterns in most urban areas, and empirical efforts to understand the nature of automobile dependence should carefully consider the roles of land-use patterns. Recently, the availability of disaggregate land-use data and technological progress in GIS have enabled researchers to examine transportation–land-use linkage in a more objective and precise manner, and to find the variables associated with urban spatial structure that influences travel behavior.

Definitions and Measurements of Automobile Dependence

A seminal work on automobile dependence by Newman and Kenworthy (1989a, 1989b) and a subsequent work (Kenworthy, et al. 1999) addressed a series of convergent conditions for land-use and transportation in cities where people were confronted with reduced mode options other than automobiles. In Newman and Kenworthy's research, the term, "automobile dependence" was formalized as a simple relationship (measured in correlation) between urban density and per capita gasoline consumption. According to their observations, most automobile dependent cities display low-density development, dispersed land-use, and a high priority for car use. However, these studies have serious drawbacks in the choice of data used, methodology, and application to transportation dimension. Some criticisms have been levied on the use of aggregate data on urban density, simple correlation analysis, and cluster analysis (Gordon and Richardson 1989; Gomez-Ibanez 1991; Steiner 1994; Mindali, et al. 2004). Their work, nonetheless, has been recognized as a landmark spurring additional research and discussions on the relationship between transportation and land-use patterns.

Litman (2002) and Litman and Laube (2002) defined automobile dependence as transportation and land-use patterns that result in a high level of automobile use and reduced travel options. Their concept of automobile dependence is similar to that of Newman and Kenworthy. For the purpose of comparative understanding, balanced transportation is presented as an opposite concept. Furthermore, Litman discusses specific costs (internal cost and external costs)[1] associated with increased automobile dependence by comparing the per household transportation cost of auto-dependent and multi-modal communities. Litman's definition is similar to Newman and Kenworthy's concept in that automobile dependence as travel behavior is closely related to land-use. Noteworthy, however, is Litman's effort made to measure automobile dependence in terms of economic costs.

Goodwin (1997) proposed a concept different from Newman and Kenworthy's perception that is based on classical gross indicators connecting transportation and land-use. In his work, automobile dependence is understood as a dynamic social and individual behavioral process that forms and develops over time. In reality, the intensities and factors leading to automobile dependence vary among individuals and over time. This feature necessarily leads to the exploration of individual travel behavior. When his concept is applied to mode choice in urban travel, automobile dependence is

[1] Internal costs are consumer costs incurred by ownership and use and include vehicle expenses, parking costs, accidents, travel time and stress, and reduced exercise and enjoyment, whereas external costs are imposed on someone other than the user, and include infrastructure costs, traffic congestion, air pollution, land-use impact, and aesthetic degradation.

likely to be seen as either individual choices resulting from the superiority of the over other modes, or the absence of other alternatives given the individual's specific attitude, environments for transportation, and land-use patterns. Despite his intellectual foresight in defining automobile dependence, Goodwin did not provide empirical examination.

In addition to this common concept of automobile dependence closely connected to land-use, there are a few other concepts that take distinctively different approaches, such as the positive effect of automobile systems (Dupuy 1999), and absolute and relative measures of automobile dependence (Stradling 2001). According to Dupuy, automobile dependence is due mainly to the superiority of positive effects (of accessibility concerning only automobile-related services) within automobile system exceeding the negative effects (congestion and pollution). Land use is not a factor to be taken into account within the automobile system. Thus, policy implementation reducing automobile dependence generates positive effects by the diversification of vehicles and ownership and the modification of road networks through reaching political consensus. In another approach based on transport psychology, Stradling (2001) looked at automobile dependence as the extent to which individuals are dependent on automobiles to meet individual transportation needs. Two measures were taken in his research: absolute versus relative. While absolute measures looked at the number of trips made by car at both travel time by car and distance traveled by car per unit of time, relative measures focused on car use in both mode mix and activity mix. The second measure stressed the individual's specific attitude based on his or her psychological attachment to automobiles.

Possible Causes of Automobile Dependence

The causes of automobile dependence are identified with several broad categories in the literature: advances in transportation technology, transportation capacity improvements (such as road construction), land-use (such as low-density land-use, and zoning), reduced mode alternatives, socioeconomic factors, and individual preferences (such as attitudes) (Newman and Kenworthy 1989a; 1999; Gomez-Ibanez 1991; Goodwin 1997; Raad 1998; Dupuy 1999; Litman 2002; Bagley and Mokhtarian 2002; Handy, et al. 2005). In reality, there is no single cause that incurs automobile dependence. Rather, there are cyclic contributing causes, many of which would be the consequences at one stage and the causes at the next stage. As noted, the causes of automobile dependence tend to give and take the cyclic feedback that gradually reinforces transportation-relation problems. The possible causes influencing the level of automobile use are reviewed below.

First, changes in transportation technology have been recognized as the primary trigger for widespread automobile ownership and use. The widespread availability of automobiles brought to more people the mobility necessary to travel long distances at relatively high speeds. In reality, the growth of vehicle ownership outpaces increases in road capacity. With automobiles available to the majority of the adult population, transportation capacity improvements through the construction of new roads sped up low-density suburban developments, particularly in the United States. Economic activities such as manufacturing and retail did not have to concentrate on specific locations or downtowns any more. Automobile roadways already became a predominant accessibility option to residences and industries (Mumford 1953; Illich 1974).

Second, road construction and improvement (i.e., road widening without sidewalks or bikeways, disconnected roads, and roads for cars only, etc.) tend to spur low-density development where the provision of transit services or non-motorized modes is neither feasible nor efficient, therefore limiting travel options to driving. Particularly, curvilinear roads in suburban areas make difficult to access by modes other than automobile. In turn, such roads tend to encourage low-density development patterns with a resulting increase in the expense of transportation. Also due to the segregation between trip origins and destinations, areas with single land-use or single zoning generate longer travel distances compared to areas with mixed land-uses. Therefore, transit service in low-density areas requires increased subsidies, and such high subsidies eventually often result in the reduction or elimination of existing transit services. According to recent empirical analyses, road transportation improvements cause greater demand for automobile trips (Goodwin 1996; Hansen 1995). That is, these improvements induce additional amounts of traffic in the short and long terms, rather than actually relieving traffic congestion as originally intended.

Third, land-use patterns, such as segregated, low-density developments, require necessary public services such as schools and hospitals to be located far away from the residential and activity centers that are often along the perimeters of urbanized areas where larger parcels of land are available at a lower cost, further encouraging urban sprawl and automobile use. In terms of costs, once consumer costs are imposed offsetting the benefits of individual automobile use, the external costs not borne directly by the driver are imposed on a whole society (Litman 2003). This issue is a core subject in this research and will be reviewed in greater detail in the next section.

Fourth, decreased viability of transit service or reduced transit service discourages people from using it due to the increased inconvenience and inefficiency. This consequently makes driving a much more attractive choice (Kain 1999). Transit-hostile or automobile-oriented land-use patterns in suburban areas further preclude other alternatives such as walking or biking. Where automobile traffic dominates, it is easy to observe resulting phenomena such as more automobile trips, increased VMT, increased car ownership, less walking and biking, and less transit use (Newman and Kenworthy 1989a).

Fifth, socio-economic forces may be important factors influencing automobile dependence. These factors mainly depend on economic, social, and psychological conditions (Goodwin 1997; Liu and Ingram 1999). For example, if a person in a household purchases her or his own car, she or he will more be likely to choose driving over other modes of transportation. Other members of the household will also likely be affected by that person's decision and behavior. These factors usually include gender, age, education, income, and other personal (or household) characteristics.

Lastly, attitudinal factors such as personal preferences and life style are associated with people's preference of certain transportation modes. Individuals who feel that driving allows them to get more done or gives more comfort and freedom are likely to stick to driving. Those who have strong attitudes favoring automobile use are likely to be royal to automobiles. Attitudes may be formed by personal characteristics as well as built environmental characteristics. When the decision to drive an automobile is based on attitudinal rather than externally observed factors, it is the most powerful cause of automobile dependence.

Consequences of Automobile Dependence

The effects of transportation observed with an increase in automobile use have been widely reported in the literature. These effects have fueled a growing public concern about current transportation conditions. The effects reviewed here include three dimensions: environmental, social, and economic. First, there are extensive environmental problems (such as air and water pollution, imbalanced ecological functions caused by the loss of non-urban land, landscape degradation, and energy depletion) associated with automobile dependence. Growing concern is placed on air quality problems in urban areas. Emissions and pollutants resulting from vehicle operation include sulfur dioxide (SO_2), carbon monoxide (CO), nitrogen oxides (NO_x), volatile organic compounds (VOC), and so on. Ozone (O_3) resulting from NO_x and VOC combined in sunlight causes major urban air pollution. These pollutants are direct causes of various respiratory diseases. On-road (or highway) vehicles in 2002 emitted 55 percent of the nation's CO, 35 percent of the nation's NO_x, and 27 percent of the nation's VOC (BTS 2004a). Transportation accounted for 22 percent of carbon dioxide (CO_2), one of the major greenhouse gases; on-road vehicles accounted for 79 percent, with passengers cars accounting for 45 percent of all transportation CO_2 emission in 2002 (USEPA 2004). According to the Victoria Transportation Policy Institute, the air pollution cost of average car in 1996 was an estimated 5.2 cents per mile for urban off-peak and 6.2 cents per mile for urban peak time (VTPI 2005).

Second, automobile dependence has social effects on health and equity. Transportation-related fatalities and injuries are mainly attributed to collisions involving motor vehicles. According to the OECD (1995), although many countries with high traffic volume have low vehicle-occupant fatalities and low pedestrian traffic fatalities, accident rates and fatality risks are strongly correlated with VMT (Litman 1997). Road deaths and injuries caused by traffic accidents impose significant costs on society. Vehicle crashes are among the primary causes of death among Americans especially those under 37 years of age (Richardson 1997). Motor vehicle crashes constitute one of the largest transportation costs and totaled an estimated $358 billion in 1988 in the U.S. Currently transportation equities or inequities are also a serious problem with the discussion centered around: wealth and ability to pay, income and social classes, and group with limited mobility or transportation disadvantages. The need for increased mobility made communities more automobile dependent, and, in turn, the financial burden related to transportation increases. This increased transportation cost may be affordable to higher income people, but it imposes disproportionately larger financial burdens on lower income people (TRB 2001). According to consumer Expenditure report (1999), lower income households spend much higher proportions of their income on transportation (up to 38%), compared to higher income households (as low as 12%).

Finally, automobile dependence has both positive and negative economic effects (Litman 2003; Delucchi 1996). While benefits are related to increased automobile mobility that affects local or regional productivity and efficiency, costs come from various inefficiencies such as delay (or congestion cost), driver stress, vehicle cost, and vehicle crash cost, etc. Both positive effects (benefits) and negative effects (costs) can be explained by a counterbalancing relationship. In literature, the economic effects are documented by evaluating the increased mobility, vehicle expenditure, parking costs, traffic congestion, accident damages, automobile-oriented land-use, and reduced travel choices. The Texas Transportation Institute (TTI) estimates, using an engineering approach, an annual economic congestion cost of $67.5 billion for 75 metropolitan areas in 2002 (TTI 2002). Litman (2002) estimates household transportation costs per mile in 1999 in an automobile dependent community and compares it with that of a community with more balanced transportation. Household transportation in an automobile dependent community costs 40% more.

Link between Transportation and Land-Use

Historical Overview

There is a considerable amount of research investigating the linkage between transportation and land-use. The historical overview of the literature falls into two distinct streams. The first stream, at present inactive compared to the second stream, focuses on evaluating the effects of transportation investment on the patterns of urban developments. The highly debated second stream concerns the effects of land-use on travel behavior in both theoretical and practical aspects (Pickerell 1999; Badoe and Miller 2000; Crane 2000). The first stream is not in line with the current research and hence is discussed only briefly here, with most of the review focused on the second stream.

Research pertaining to the first stream answers mainly a key question about how transportation investments and related travel behavior affect land-use, especially development patterns. Many past studies presented the empirical evidence from testing the classical hypotheses[2] based on the location theory for both residence (housing) and firm (business) locations (Mills 1972; Fujita 1989; Anas, et al. 1998). Recent studies centered on the investigation of the effects of transportation investments such as expansion of highway capacity and mass transit on employment decentralization. However, although the historical effects of transportation on land-use were presented clearly in some instances, recent evidence shows only a limited influence on housing and business location patterns (Hamilton 1982, 1989; Small and Song 1992; Mieszkowski and Mills 1993; Giuliano and Small 1993; Giuliano 2004).

The second stream, sparked in the 1990s in response to a growing interest in land-use policies to reduce automobile dependence, deals primarily with a question of how we design urban spatial structure and effectively shape urban areas to reduce automobile dependence. Many studies addressing such a question have shed light on the relationship between transportation and land-use based on the geographical unit of analysis. Past studies are considerable in their cumulative amounts, and a few articles by Crane (2000), Badoe and Miller (2000), and Ewing and Cervero (2001) give excellent

---

[2] A classical key hypothesis from traditional location theory is that the transportation cost determines the land-use in which each piece of land-use is associated with a unique location over geographical space (Alonso 1964; Muth 1969; Mill 1972; Henderson 1977; Fugita 1989).

reviews with diverse points of view. Studies in this stream brought a new perspective to policy debates among planning professionals. From the late 1980s to the mid 1990s, a growing interest occurred in the use of land-use policies to manage transportation demand and led to policy debates on whether or not land-use would matter. Policy debates claimed two different approaches for policies reducing automobile dependence: “modifying land-use” through physical planning and urban design (Newman and Kenworthy 1989; Cervero 1991; Cervero and Landis 1995; Newman, et al. 1995), and “taking economic measures” such as pricing mechanism in the transportation markets, levying taxation or lifting subsidies (Gorden and Richardson 1989; Gorden, et al. 1989; Gomez-Ibanez 1991; Giuliano and Small 1993; Giuliano 1995). These policy debates led researchers to recognize the importance of policy research on land-use as a way to cope with automobile dependence. It was essential for the advocates of land-use policy to pursue the enrichment of research on the relationship between transportation and land-use.

As a result, it is noteworthy that methodological progress has been made in several significant ways. For example, the identification of various land-use measures at various geographical units of analysis, theoretical and analytical foundations for the link between transportation and land-use, statistical methods, and acquisition and use of disaggregate land-use data or travel diary and survey data are all excellent methodological contributions. Insightful inquiry into the land-use measures (i.e., density, land-use mix, and accessibility) related to travel outcome variables[3] has expanded, to a

[3] In many studies, travel-outcome variables fall into vehicle miles traveled (VMT), trip frequency, mode

considerable extent, the capacity to efficiently and effectively gauge urban spatial structure. Further, the employment of these measures for land-use policy research has been suggested. Several earlier works documented the relationship between transportation and land-use using regression analysis by incorporating the various measures of land-use and by controlling non-land-use variables that affect travel behavior (Frank and Pivo 1994; Cervero and Gorham 1995; Cervero 1996; Handy 1996; Levinson and Kumar 1997; Handy, et al. 1998). Also a growing body of empirical studies dealing with transit and non-motorized transportation behaviors has identified a number of detailed and disaggregated land-use measures associated with transit use, walking, and biking. While the associational (mostly correlational) relationship found by these analyses was greatly informative for further research, efforts to describe travel decision process were limited. Further these studies had no ability to explain causality in their models.

Kockelman (1997) and Cervero and Kockelman (1997) showed prominent insights in their efforts to define and document land-use variables in three principal dimensions (density, diversity, and design), and analyzed in detail these variables based on the integrated model for transportation–land-use link. Findings indicated that compact, mixed-use, and pedestrian-friendly designs can reduce automobile dependence. Their results, however, were still correlational rather than causal. The causal model with behavioral links between travel outcome and land-use can correctly estimate and forecast the effects of land-use policy changes. Both studies detailed land-use variables in the

---

choice, trip length and duration, departure time, route choice, auto ownership, trip purpose, etc.

linkage, but have not applied them to a travel demand model yet. Crane (2000) classified these models (including earlier models) as *ad hoc* distinguished from demand models.

In recent years, some attempts to build a conceptual framework consistent with consumer behavior theory have been made in an effort to find additional evidence for land-use planning. Domencich and McFadden (1975) earlier mentioned that land-use could be included in the travel demand model with traditional demand variables (i.e., modal attributes and socioeconomic characteristics). Some recent works have refined the travel models to include the full array of explanatory variables such as demand variables, personal or household characteristics, and land-use variables. These models could capture the short-term effects through demand variables as well as the long-term effects by land-use variables associated with the long-term behavior (Crane and Crepeau 1998; Boarnet and Greenwald 2000; Boarnet and Crane 2001; Cervero 2002; Zhang 2004). Travel model improvement brings, to a certain extent, more and more attention to the issues of causality between travel behavior and land-use, but the empirical investigation into causality is still limited due to lack of data availability and methodological difficulties.

Land-Use in Travel Demand Models

Models linking land-use and travel behavior have been generally called travel demand models despite the theoretical distinctions. The linkage models from the consumer theory of microeconomic foundation should be distinguished from the other linkage models. As addressed above, the travel demand model must include a comprehensive set of explanatory variables (travel time and travel cost, socioeconomic characteristics, and land-use variables). In particular, price variables (travel time and travel cost) should not be omitted. It is evident that omitted variables lead to biased results from a statistical standpoint. The travel demand model must have a behavioral framework to describe the causal relationship between travel outcome and land-use. Such a model is useful in using policy changes to forecast actual travel demand. Studies reviewed below are selected based on meeting these qualifications as travel demand models.

The models with transportation and land-use linkage have been developed in tandem with the incorporation of the full array of explanatory variables. Existing literature review articles provide a good composite understanding of the land-use variables and their effects on transportation (Badoe and Miller 2000; Crane 2000; Ewing and Cervero 2001). A literature review suggests that a fair amount of information is known, but the relationship between transportation and land-use remains too complicated to be fully comprehended from the existing evidence. Some analyses using regression techniques overlooked the behavioral frameworks, and the theoretical and/or statistical considerations of causal relationships (Handy 1993; Frank and Pivo 1994; Cervero and Gorham 1995; Cervero 1996; Levinson and Kumar 1996; Kockelman 1997; Cervero and Kockelman 1997; Kitamura, et al. 1997; Handy, et al. 1998).

In recent works, several transportation planning scholars have been paying some attention to the causal relationship, establishing the travel demand models derived from the economic theory. Travel demand models incorporate mainly the land-use variables and the traditional demand variables into the analytical framework. In urban economics literature, monocentric models[4] (Muth 1969; Fujita 1989; Wheaton 1998) and some modifications to monocentric models (White 1988; Bento, et al. 2003) suggest that a household travel demand depends on the distribution of population and employment throughout a city, the size of city, its road and transit networks, the density of the road network, a marginal time cost, and a marginal price.

As noticed earlier in Domencich and McFadden (1975), travel demand model must be inherently causal through the behavioral link between travel behavior and decision variables (i.e., land-use variables and economic variables). Several recent studies stress the model specification in this context (Crane 1996a; Crane and Crepeau 1998; Boarnet and Sarmiento 1998; Boarnet and Greenwald 2000; Boarnet and Crane 2001; Kockelman 2001; Cervero 2002; Zhang 2004; Bento, et al. 2005). Table 2.1 summarizes the selected studies dealing with the transportation–land-use linkage in terms of the variable specification. A detailed review is conducted for the travel demand models including the land-use and economic variables. These variables are likely to clarify the role of decision variables in both short-term and long-term policy questions for urban transportation planning. The interest in land-use variables from the standpoint of transportation-policy analysis lies primarily in the question of whether planners can influence travel behavior by land-use policies, as the interest in economic variables appeals to economists with regard to the role of pricing policy.

---

[4] Monocentric model fundamentally depends on the rent gradient it faces and on the marginal cost of travel which varies with distance form CBD. Thus household travel model depends on the demand variables (marginal time cost and/ or marginal price) and on the urban form variables.

Outcome variables to measure travel are distinct in literature: trip frequency (rates), trip length (vehicle miles traveled or person miles traveled), mode choice, trip duration (vehicle hour traveled), departure time, route choice, vehicle ownership, and so on. Particularly, trip frequency, trip length and mode choice are most frequently used to examine travel behaviors directly associated with automobile dependence. Contrary to travel outcomes, land-use variables are featured by the various dimensions of land-use and urban form. Urban spatial structure has been identified and measured empirically by spatial activity outcomes, urban design characteristics, and transportation infrastructure. Spatial activity outcomes are often characterized by indicators such as density, land-use mix and accessibility. The design characteristics are measured both subjectively as perceived quality and objectively using GIS, and include safety related to transportation, aesthetic quality of the roadside environment, etc. Transportation infrastructure often includes the patterns of roads, street connectivity, and provisions for sidewalks or bike paths, availability of transit services, etc. More importantly, the spatial structure would somewhat depend on the geographical scales (i.e., neighborhood, census block and tract, zip code, community, city, and region), and the geographical scale may reinforce or attenuate the status of spatial structure (Handy 1993). At the neighborhood or community levels, the local street patterns and the residential location are likely to influence the urban spatial structure. At the regional level which may encompass multiple abutting large cities and metropolitan areas, the urban spatial structure may be influenced by main roadways, transit systems, major transportation terminals, and employment centers.

Table 2.1 Linkage Models of Transportation–Land-Use

Authors	Travel Measures	Explanatory Variables								
						Land-use dimensions				
		Modal Attribut-es (T/C*)	Travel supply/ Trip attr -ibutes	Socio-demo-graphics (P/H/I)*	Atti-tudes	Den-sity	Diver -sity	Access-ibility	Street features	Walk/ Bike prov-isions
Frank & Pivo (1994)	mode: SOV, transit, & walking					√				
Cervero & Gorham	mode: transit commuting					√			√	
(1995)				√						
Cervero (1996)	mode: Automobile commuting		√ √	√			√			
	Transit commuting		√	√			√			
	Walk/bike commuting		√ √	√			√			
	vehicle ownership (# of cars)		√ √	√			√			
	distance b/w home & job-place		√ √	√			√			
Levinson & Kumar	commute time			√		√				
(1997)	commute distance			√						
	commute speed			√						
Handy, et al. (1998)	strolling trips			√ √					√	√
	walks to the store			√					√	√
				√ √						
				√						
Kockelman (1997)	VMT for all trips per household			√ √			√	√		
	VMT HBNW trips per household			√ √			√	√		
	auto ownership			√ √		√	√	√		
	mode: personal vehicle		√	√ √		√	√	√		
	walk/bike		√	√ √			√	√		
Cervero &	person VMT for all trips		√	√			√	√	√	
Kockelman (1997)	person VMT for non-work trips		√	√			√		√	
	mode: Non-SOV for non-work		√	√ √			√		√	√
	mode: Non-PV for non-work		√	√ √					√	√
	for personal business		√	√ √						√
	for work trips		√ √	√ √					√	√

* T/C means trip time and trip cost, and P/H/I means personal, household characteristics, and income respectively.

Table 2.1 (Continued)

		Explanatory Variables												
										Land-use variables				
Authors	Travel Measures	Modal Attributes (T/C)*		Travel supply/ Trip attributes		Socio-demographics (P/H/I)*			Attitudes	Density	Diversity	Accessibility	Street features	Walk/ Bike provisions
Kitamura, et al. (1997)	total # of person trips					√	√	√					√	
	of transit trips					√	√	√				√		
	of non-motorized trips											√		√
Krizek (2003)	# of tour, # of trips per tour						√					√		
	VMT						√					√		
Greenwald (2003)	trip ratios					√		√		√	√		√	
Handy, et al (2005)	VMT					√			√		√	√	√	√
	Change in driving / walking					√			√		√	√	√	√
Boarnet & Sarmiento (1998)	non-work automobile trips		√			√	√	√		√	√		√	
Crane & Crepeau (1998)	car trip frequency	√				√	√	√		√	√		√	
	mode choice	√				√	√	√		√	√		√	
Boarnet & Greenwald (2000)	non-work auto trips per person	√				√	√	√		√				
Kockelman (2001)	# of trip per household	√	√					√				√		
Boarnet & Crane (2001)	trip frequency (Orange Co/ LA)	√				√	√	√		√			√	
	trip frequency (San Diego)	√				√	√	√		√	√		√	
Cervero (2002)	mode: driving-alone (DR)	√	√			√	√			√	√	√		√
	transit	√	√			√	√			√	√	√		√
	DR/group-ride/transit	√	√			√	√			√	√	√		√
Zhang (2004)	mode choice	√	√	√	√	√	√	√		√		√	√	
Bento, et al. (2005)	mode choice		√	√	√	√	√	√		√	√			
	VMT		√	√	√	√	√	√		√	√			

* T/C means trip time and trip cost, and P/H/I means personal, household characteristics, and income respectively.

Crane and Crepeau (1998) criticized earlier works which only included land-use variables such as urban density and four-way intersections. New specifications were made for improving the model and were applied to explain travel behavior as a function of behavioral variables [5] (prices and preferences) and land-use variables. In their empirical analysis, household travel diary and GIS data for San Diego were used for testing the effect of land-use (specifically, neighborhood street pattern and portion of land-use in census tract) on non-work car trip frequency at both the household and the personal levels, and on mode choice between car and walking. They concluded that higher street density reduces car trip frequency only on the household level, the higher commercial share of land-use increases the number of trips on the personal level, and there is no empirical evidence to show that street design patterns influence the likelihood of selecting driving. Although there is little support for the claims of new urbanism regarding the impact of land-use on travel behavior, this work improves the *ad hoc* models to travel demand model by including a theoretical causal structure.

Another study (Boarnet and Sarmiento 1998) specified the behavioral nature of the link between transportation and land-use derived from Crane's theoretical framework. In this research, the demand for non-work automobile trips was defined as a function of travel time cost, individual income, and socio-demographic variables. Here, travel time cost was assumed to be influenced by land-use variables (density, street grid, the mix of commercial and residential use). Empirical results, using travel diary data from southern

---

[5] Price (travel time and cost) variables are behavioral variables in travel demands because travelers can make decisions when confronted with alternative choices. These ideas are found in refined manner in Boarnet and Greenwald (2000), and Boarnet and Crane (2001a, 2001b).

California residents, showed that land-use variables measured at the census block/tract and the zip code levels do not support the new urbanist principles. However, although theoretical specification seemed to be more appealing than previous linkage models, the statistical methods used for the model are more or less questioned.

A similar study (Boarnet and Greenwald 2000) tests the hypotheses of empirical specification by including a comprehensive set of variables such as socio-demographic, trip cost, and land-use for the zip code and census tract from the 1994 Portland-Oregon travel diary. New to this research is the use of sophisticated estimation methods such as two-step procedures and instrumental variables. Findings from travel demand models report that the link between land-use and non-work trips seems weak but apparently exists, and that the consideration of appropriate geographical unit of analysis is critical.

Boarnet and Crane (2001) investigated the causal links between land-use and travel behavior in terms of the model specification and estimation issues. They found that many foregoing studies poorly incorporated the behavioral theory of travel demand and poorly addressed estimation issues. Their fundamental assumption that the pattern of land-use captures all price variations was substantiated by the theoretical framework and well incorporated into the estimation procedure. However, their price variables using trip length and speed for trip time were not likely to be captured completely by land-use characteristics. Data from travel diaries from Los Angeles and San Diego were used for empirical analysis. Some lessons from their findings inform us of how land-use variables are linked to price variables and then to travel behaviors, how important the geographical scale is, and how the causal flows are sometimes erroneously assumed by correlational associations in the literature.

Cervero (2002) stressed the adequate specification of the relationship between built environment factors and travel behaviors by paying special attention to the theory and methodology. Using 1994 Household Travel Survey from Montgomery County, Maryland, he estimated mode choice models with socio-economic characteristics (travel time and cost, and demographics) and built environment factors (density, diversity and design) at travel origins and destinations using the TAZ as the unit of analysis. Findings reveal that density and mixed land-use significantly influence the choice for driving-alone, group-ride, and transit, but the effect of urban design is trifling in ways not adequately appreciated in many policy discussions. Of special note is the model specification in his research that can be a prototype for conducting similar research in geographically diverse regions. In particular, he envisions a normative analytical framework based on consumer choice and travel demand theory allowing policy-makers to make informed decisions on land-use and transportation proposals.

An empirical study analyzes the influence of land-use on mode choice for both work and non-work trips using individual travel survey data in metropolitan Boston and Hong Kong (Zhang 2004). Analytical model was specified as travel demand model with mode attributes (travel time and travel cost), personal characteristics, and land-use variables. The study reports four major findings centering on the policy and methodological issues of travel demand models: there is considerable benefit in model improvement from the linkage to travel demand model, land-use matters when traditional demand variables are controlled, the influence of land-use on driving decision is potentially as strong as pricing depending on the combined land-use elasticities of driving probability, and the performance of land-use variables is weaker for non-work trips and some dimensions of land-use are no longer significant, while travel time and cost are still significant in explaining the mode choice for both types of trips.

Bento, et al. (2005) reported the effect of urban form on travel behavior by households in 114 urban areas with data from the 1990 Nationwide Personal Transportation Survey. In their model, measures of urban form include city shape measured in circularity, road density, population density, population centrality, job-housing balance, bus route miles, and railroad miles. The empirical models of commute mode choice, vehicle ownership and annual VMT depend on household characteristics, income, travel cost (city-specific gasoline price), and measures of urban form. In particular, they estimate a multinomial logit model of the number of vehicles owned and an equation for the annual VMT, conditional on owing vehicles. Empirical analysis finds that when the probability of driving to work is low, population centrality is high and road density is low. VMT is also influenced by population centrality, job-housing balance, city shape, and road density. The research concludes that in the U.S., heavily dependent on the automobile, urban form affects travel demand.

Causality in the Linkage

The direction of causal flows is usually defined *a priori* by a theory or a conceptual framework, but it is difficult to define the nature of relationships between travel behavior and land-use due to the complexity of land-use dimensions interwoven among themselves and to the potentially different roles of land-use variables when measured and analyzed at the different geographical units of analysis (Crane 1996a). Empirical relationships reported by previous research vary and remain to be correlational rather than causal. As pointed out by Boarnet and Crane (2001) and Cervero (2002), almost all past studies had theoretical and/or statistical misspecifications of models which were mainly attributed to the shortage of available data and the limitation of methodologies. Limited or little attention has been paid to the issue of causality in the link between travel and land-use in the previous literature. Recently, the issue of causality has started to attract much interest centering on causal mechanism linking land-use to travel behavior in the travel-demand model which is specified by the set of observed variables (Carne 1996a; Boarnet and Crane 2001a, 2001b), and checking the stability of causal relationships between travel and land-use by accounting for the possibility of self-selection based on unobserved preferences such as attitudes (Bagley and Mokhtarian 2002; Handy, et al. 2005).

Crane (1996a) developed a theoretically different framework from previous models linking land-use with travel behaviors, and offered the comparative-static effects of land-use change on the travel demand derived from the utility maximization theory. Trip demand (measured by the number of trips) for each mode (automobile, walking, and bus or other transit) is a function of price vector in which its change is caused by the change in land-use variables (grid, traffic calming, and mixed use). That is, the change in land-use shifts the number of trips by each mode through the change of price (generalized cost). In the previous models for trip generation, the demand for travel was mainly defined as a function of socio-demographics and land-use. Now, the price is assumed as completely captured by land-use characteristics in transportation market, and thus the vectors of price, land-use, and socio-demographics are included in the trip-generation model. However, land-use variables are unlikely to perfectly determine the price (trip length and speed), because trip time and trip cost can be directly observed in transportation markets rather than the implied prices by hedonic pricing (Crane and Crepeau 1998; Boarnet and Sarmiento 1998; Boarnet and Greenwald 2000; Boarnet and Crane 2001a, 2001b).

When the unobserved (or subjective) preferences such as intensions and attitudes are taken into account in the linkage model, the assumed causality for the linkage based on the revealed (or objective) preferences may be masked or reversed due to subjective preferences. Such preferences may be correlated with personal or household characteristics or the built environments, and in turn reinforce or suppress the observed preferences. The issue is oftentimes addressed by self-selection; for example, those who like to walk or use transit may choose to live in a neighborhood that has sidewalks and transit services that support their preference. Currently there are a few studies dealing with the issue of self-selection using cross-sectional data or panel data, but only one study tests self-selection directly. These studies are inclined to parsimoniously specify the models in the absence of potentially important variables (i.e., insufficient land-use dimensions or omission of demand variables), and such specification problems may outweigh the benefits of addressing causality in the model.

Another study deals with relationships among attitudes, residential location choice and travel behavior by using nine structural equations from data on 515 residents from five neighborhoods in the San Francisco Bay area (Bagley and Mokhtarian 2002). For the use of endogenous variables, the system of equations includes residential locations (traditional vs. suburban), attitudes (pro-high density, pro-driving, and pro-transit), travel demand (in log-transformed vehicle miles, transit miles, and walk/bike miles), and job location measured by commute distance. This study finds the multi-directions of causality to independently influence nine endogenous variables. According to the results, attitudinal variables affect transit miles and walk/bike miles, but residential location (suburban) shows no evidence of influencing travel behavior. That is, the inclusion of attitudes into the model changes the observed relationship between residential location and travel behavior, suggesting that residents with specific attitudes are self-selective in the specific type of neighborhoods in which they live. In particular, this approach tests the relationships between empirical data and the assumed causal structure, and suggests the possibility of multi-directions of causality to explain travel behavior.

A work by Krizek (2003) analyzed the relationship between urban form (neighborhood accessibility and regional accessibility) and travel behavior (vehicle miles traveled, person miles traveled, number of trips, number of trips per tour, and mode split[6]) using data from the Puget Sound Transportation Panel Survey. This study used a longitudinal design to test the impact of the change in urban form on travel behavior from a total sample of 6,144 focusing on 430 households that relocated between 1989

[6] In his paper, the results for mode split were not reported due to reason that the effects of the policy-relevant variables were not significant.

and 1997. The empirical results suggest that changes in urban form reduce both vehicle miles traveled (VMT) and person miles traveled (PMT), and change in neighborhood accessibility tends to increase trip generation. In the meantime, mode split (walking/biking, transit, and auto) is not influenced by changes in urban form. Krizek pinpoints that households prefer to remain fixed in terms of mode and are unwilling to change to alternatives, and suggests in a roundabout way that there is little influence of unobserved forces (self-selection) on travel behavior without measuring the household unobserved preferences. This study supports the assumed causality such that urban form still influences travel behavior although there exists the possibility of self-selection.

Handy, et al. (2005) used a sample of two groups of residents (relocated vs. non-relocated) in four neighborhoods (traditional vs. suburban) of Northern California to investigate a causal relationship between the built environment and travel behavior. Their cross-sectional analysis shows that attitudes toward transportation (pro-bike/walk, driving-safe, and car dependent) and socio-demographics contribute causally to explain travel behavior (vehicle miles driven), but the built environment does not. In quasi-longitudinal analysis, the change-in-driving is caused by land-use (change-in-accessibility factor) and attitudes (car dependent and pro-bike/walk), and the change in walking depends on change-in-accessibility factors and pro-bike/walk. These results show that there exists evidence supporting the presence of self-selection. An increase in change-in-accessibility leads to a decrease in the change-in-driving, while car-dependent people tend to drive more, pro-bike/walk people drive less, and walk more. Testing self-selection in the data shows that the decrease in driving caused by the built environment (an increase in accessibility) is somewhat suppressed by car-dependent attitudes but the increase in accessibility has the greatest effect on driving less. This study supports a causal relationship that the built environment influences travel behavior after accounting for attitudes. However, this study does not answer the nature of causality, including the magnitude of associations and the causal direction between attitudes and built environment, and the multi-directions of causality between all the explanatory variables.

In understanding the impact of land-use on travel behavior, it is important to identify and clarify the direction of causality based on the models with the full array of potential regressors. Currently there is a gap in the linkage models in terms of the causal notion for the explanatory variables. In the consumer demand-typed linkage models, prices have been viewed as important factors in influencing travel demands (such as mode split as well as, following Crane, trip generation and VMT) along with other shifters (socio-demographics and land-use), while in conventional linkage models, socio-demographics and land-use were important factors affecting travel demands, and attitudes seemed to play a role in determining (reinforcing or suppressing) the stability of assumed direction of causality between travel behavior and land-use. Relative or absolute belief in the efficacy of land-use policies might lead to different causal notions and to suggestions for different policy mix or practices to reduce automobile dependence.

Cross-sectional data has been pointed out as a primary constraint for research investigating the causal mechanism of the linkage models, but nonetheless, valuable efforts have been made to shed light on the issue of causality. These efforts are observed in research designed to search for the full array of explanatory variables developing behavioral theories (especially, travel-demand model linking land-use to travel behavior), the use of the instrumental variables modeling a system of structural equations, and further longitudinal analysis utilizing panel data. Nonetheless, the nature of the causal relationship between travel behavior and land-use is still poorly understood, although it is true that the limitations of the available empirical data and their relation to the assumed causality have been in part addressed by a small number of recent studies employing a longitudinal method. In addition to the need for more rigorous and extensive longitudinal studies, the strength of associations and the causal directions between attitudes, prices, socio-demographics, and land-use variables must be clarified to provide more conclusive evidence explaining the relationship between travel behavior and land-use. Much still remains to be understood regarding the nature of causality.

CHAPTER III

CONCEPTUAL FRAMEWORK AND HYPOTHESES

This chapter sets out the conceptual framework exploring the linkage between transportation and land-use, and considers the causal relationships between the two and among the explanatory variables. The conceptual framework serves as a basis for the research hypotheses of this dissertation.

Conceptual Framework

To examine the land-use-travel relationships, let us look at ways that travelers decide on whether to make trips, where and when to go, which mode to use, and which route to take. These travel decisions could be made based on the traveler's needs, transportation systems, trip time and cost, attitudes and intensions, socioeconomic characteristics, or the features of places where trips start and terminate, oftentimes called land-use characteristics (or built environment or urban form).[7] In this context, a specific behavioral travel model is needed to understand travel decisions made by each individual traveler or each household, to make it possible to test hypotheses derived from the assumptions of the model, and to be stable with the issue of causality. Such a travel model is useful for examining the effect of land-use on travel behavior and capturing the fundamental patterns of causality that the data contain. To deal with the

[7] "Land-use" is interchangeably used with either urban form or built environment in this research. Land-use is conceptually policy-oriented term at any spatial scale. "Urban form" in literature is a broader concept with morphological meaning than land-use, especially oriented to the transportation systems and urban design features, whereas "built environment" includes everything built in spatially dispersed areas as a result of human intervention through various activities in the natural physical world.

aforementioned issues, a framework for causality between land-use and transportation is conceptualized in a belief that land-use policies will be effective in reducing automobile dependence. Recently many metropolitan planning organizations in the U.S. initiated land-use strategies in an effort to curb low-density land development. They assume that low-density land development encourages people to choose automobiles because of its urban spatial structure favorable to automobiles. Furthermore, it is argued that such automobile-friendly conditions in low-density development patterns exert a positive influence on a trip-maker's captive choice. If true, this implies that efforts to discourage this type of development can be meaningful in reducing automobile dependence. In this context, a captive choice of automobile associated with land-use is also conceptually framed.

Issue of Causality

The issue of causality among the studies dealing with transportation–land-use linkage can be characterized by distinctions such as data, variable inclusion, causal notion, test for causality, causal structure among independent variables, etc. Shortcomings have been seen in many past studies but have not been highlighted seriously because of the relatively loose modeling traditions in planning. More seriously, travel demands theoretically derived from consumer choice theory have been under- or mis-specified in many empirical models (i.e., omitted variables). While the mode choice model has been rigorously specified in many travel demand models (Cervero 2002; Zhang 2004), other models such as trip frequency and VMT are often poorly handled.

The conceptual frameworks of the linkage models begin with an entire set of explanatory variables (prices, socio-demographics, and land-use). In particular, the inclusion of price variables (travel time and travel cost) is very important because the price variables are likely to interact with the conditions of land-use. A comparison between the existing models and the proposed models for the causal model is presented in Table 3.1.

Distinct features lie in the inclusion of variables, functional dependencies, and causal structure among independent variables. The conceptual framework does not include attitudinal variables which are not directly observable and not applicable to the travel forecasting model. When attitudinal variables are included in the linkage model, each individual is treated as a 'black box' because unobserved preferences, such as attitude, act as intermediaries between the environment and travel behavior. Consumer behavior literature identifies the sources of attitude formation as personal experience, friends and family, and media (Fishbein 1975; Schiffman and Kanuk 1996). In this context, an extension of the measurement model requires structural equation modeling (SEM) to be formalized to explain both attitude formation and travel behavior simultaneously. This can be shown as functional dependencies that the functional forms are written as structural equations. However, the most difficult task in modeling structural equations is to identify endogenous variables without a strong theoretical framework. This research is not extended to SEM.

Causal explanations in past studies mostly relied on an assumed causal link based on the theoretical foundation. In contrast to the assumed causality, theories often lack a clear explanation of the exact characteristics of the link. If theory does not explain the causal structure regarding the impact of land-use on travel, the causal structure should otherwise be supported by the data itself. The linkage model assumes the direct causal link between land-use and travel behavior, but in fact, land-use might influence travel behavior as well as prices (time and cost) in the travel market – the author does not agree fully with the assumption of Boarnet and Crane (2001a) that "land-use completely captures prices." Under this assumption, the effect of land-use on travel will be biased because land-use is a common cause of price and travel outcome. Where causal inferences are at stake, the influence of land-use on travel behavior remains elusive in the absence of a plausible explanation demonstrated by the data.

Table 3.1 Comparison between Existing Models and Proposed Models

	Existing Models	Proposed Models
Data	observational /unobservational	Observational
Variable inclusion (vector notation)	Partial /full array of variables: - sociodemographics (**S**) - land-use (**L**) - travel time & cost (**p**) - attitudes (**A**)	Full array of variables: - sociodemographics (**S**) - land-use (**L**) - travel time & cost (**p**)
Travel outcomes (**x**)	trip generation (frequency) mode choice vehicle miles traveled , etc	trip generation (frequency) mode choice vehicle miles traveled, etc
Functional dependencies	$\mathbf{x} = f(\mathbf{L}, \mathbf{S}, \mathbf{A})$ $\mathbf{p} = f(\mathbf{L})$, $\mathbf{A} = f(?)$	functional dependencies can follow causal structure
Causal notion (causality directions)	assumed direct causes / self-selection (unidirectional)	not assumed but determined by data
Causal structure among the variables	assumed independent (actually not identified)	identified as direct or indirect causes
Test for causality	based on equation	based on data
Travel forecasting	not applicable for attitude	applicable

The assumed causality that land-use influences travel behavior is plagued with self-selection bias and the interdependence of variables, with the exception of a few recent studies (Bagley and Mokhtarian 2002; Handy, et al. 2005). In the current study, a causal direction is considered to be established only if the causal relationships among the observed variables are established with statistical inference. A set of interactions among the explanatory variables is graphically represented to show the impact on travel behavior (Badoe and Miller 2000). The graphical representation is not obtained through the construction of a causal model, but is drawn from a review of empirical studies.

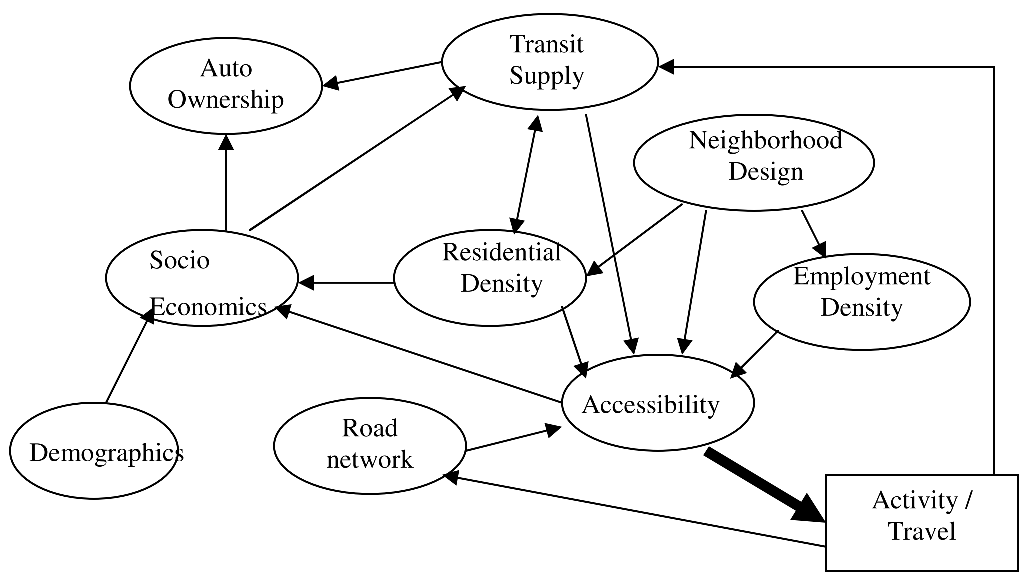


Figure 3.1. Urban Form Impacts on Travel. Source: Badoe and Miller (2000).

Badoe and Miller (2000) simplify the interactions among the variables but do not show any additional work making inferences from statistical data to causal structure.

Rather, they suggest the model specification of urban form is the endogenous component of the system. Figure 3.1 notably uncovers the causal relations among the variables in the linkage model, which is also helpful for initiating land-use policies. Pertaining to improving the analytical approaches to disentangle relationships among variables, a new methodology directly dealing with the causality issue will be explored in the next chapter.

Structural equation modeling (SEM) is concerned with how model variables are related to one another. It handles measurement problems by checking the entire structure of data assumptions and requires a well-developed theory among variables. However, SEM neither offers a method to test causal models nor provides the causal explanations of estimated parameters. Thus, the causal interpretation of SEM is generally questionable (Pearl 2000). As such, a newly developed method called directed acyclic graphs (DAG) is desired for understanding the causal relationships among the variables in the linkage model. Detailed discussion of this method is presented in the following chapter.

Captivity to Automobile

If the built environment (i.e., low-density, no transit service, no sidewalk, etc.) in a community is automobile-dependent, the residents will likely to be captive to automobiles, compared to those living in an environment with multiple transportation options available. In a sense, a captive choice to use automobiles can be conceptually distinguished from the free choice without captivity to automobiles. People may choose a specific mode because of some captive factors (constraints) associated with personal factors, place of residence, and trip characteristics. Such factors may reduce the number of choices from which an individual trip maker can choose, and may lead even to a single choice with no other alternative. This study conceptualizes the possibility of mode choice that may be incurred by (or tied to) land-use features.

If a choice of automobiles in low-density development patterns can be evidenced from a trip-maker's captive choice of automobiles, efforts aiming at discouraging such land development will need to focus on land-use strategies in reducing automobile dependence. A mode choice model is conceived as a behavioral framework to explore captivity from land-use-travel relationships, but it is not possible to apply directly a unique concept to the choice-modeling process.

There is no single commonly accepted method to measure captivity from the land-use perspective, and in general the features of captivity may be conceptually captured in different ways. One way is to segment individual trips into either a captive choice or a free choice based on identifiable constraints of the built environment (i.e., street connectivity, or grid-like patterns) (Beimborn, et al. 2003). Another way captures the choice set in probabilistic term by parameterizing (or estimating) captive variables in the scheme of a choice model. The second idea has been applied using various approaches in the literature, but mostly starts with the two-stage choice model of Manski (1977). A framework is conceptually understood to capture captivity solely attributable to land-use features as the probability of mode choice. Methodological procedure will be discussed in detail in the next chapter.

Hypotheses

This section presents the hypotheses for this dissertation which will test the influence of land-use on travel by employing empirical models. The maintained hypothesis is that urban form does affect travel patterns in cities that are heavily dependent on automobiles. These hypotheses are grouped into three sets based on the specified models; traditional travel demands, causal models for travel demands, and captivity choice. Specific variables used to measure various concepts mentioned in the hypotheses are presented in the Research Methodology chapter.

Hypotheses for Travel Demand Models

Hypotheses are stated based on the three parametric statistical models commonly used in the literature: individual's mode choice (walk/bike, bus, driving-alone, or shared-ride), household total number of automobile trips, and household total VMT.

$H_{mode.P1}$: *Job density of the TSZ at both trip ends (origin and destination) is associated with trip mode choice*. Specifically, an increase in job density at the origin TSZ induces the choice of walk/bike or transit (bus), but the increase of job density at destination TSZ lowers the probability of driving-alone or shared-ride. Workers in high job density areas are usually faced with making tough decisions associated with higher costs for housing, transportation, and other urban services near their work place. Therefore, workers take a utility-maximizing behavior given the income and cost constraints.

$H_{mode.P2}$: *Higher residential land-use share at origin is associated with increased* *non-motorized modes and transit use for home-based other (HBO) trips, and with reduced automobile use for HBO trips at destination*. If the residential development of land-use is favorable to non-work activities within a given proximity of neighborhood, higher residential land-use is associated with reduced automobile use.

$H_{mode.P3}$: *Higher commercial shares of land-use at both trip ends is associated with increased non-motorized transportation, but with decreased probability of choosing automobiles at destinations*. This variable measures land-use diversity to some extent, but does not explain the degree of the combinational share of land-use like a land-use balance (entropy index). In the literature, the entropy index is often believed to be a significant factor lowering the probability of choice for automobiles.

$H_{mode.P4}$: *Regional accessibility to jobs is correlated with a reduced likelihood of choosing automobiles for home-based work or non-work trips.* Since auto-based regional accessibility (calculated as gravity formula) is attractive for jobs and businesses, locations with good automobile accessibility may come with the high costs (such as congestion, toll, parking cost, etc.) incurred by the heavy use of private vehicles. Such costs may be burdensome to certain trip-makers, possibly encouraging them to shift mode choice. But despite the expense of automobile use resulting from the increased accessibility, others may still choose to drive with the expectation of higher earning opportunities.

$H_{freqency.P1}$: *The residential share of land at origin is positively correlated with the number of automobile trips*. Areas that are solely or predominantly residential are naturally isolated or separated from other types of land-uses. As the residential share increases, more automobile trips are likely to be made. In fact, many residential developments during the past decades are made up of low-density, detached single-family housing.

$H_{frequency.P2}$: *Regional accessibility is associated with decreased home-based non-work trips but increased home-based work trips*. As mentioned above, HBO trips may incur a higher relative cost than HBW trips, depending on income expectation from opportunities (jobs) induced by an increase in regional accessibility. That is to say, HBW trips are associated with productive activities generating income, but HBO trips are associated with consumption activities at an additional cost involved with increased accessibility

$H_{VMT.P1}$: *Population density at household location is negatively associated with household VMT for home-based work or non-work trips*. Household locations in high population density areas are usually vibrant with many urban activities (i.e., work, shopping, recreation and sports, public meetings, and cultural events, etc.). And the close proximity between homes and activities in this type of environment reduces trip lengths, often making non-motorized and transit use viable alternatives to driving.

$H_{VMT.P2}$: *Regional accessibility is associated with households' driving negatively for HBO trips but positively for HBW trips*. Regional accessibility is reliant upon the individual/household responsiveness to the relative costs involved in the trips. For an expected income, HBW trips linked to production activities have a lower relative cost than HBO trips made primarily for consumption.

***Hypotheses for Causal Model*s**

Hypotheses are derived from observed data for: dichotomous mode choice (non-automobile versus automobile), household total automobile trips, and household total VMT. While the main hypothesis is that some land-use factors directly cause a reduction in automobile dependence, also it is hypothesized that land-use also indirectly causes a reduction in automobile dependence through travel time (or generalized cost) which directly causes people to drive less.

$H_{mode.C1}$: *Job density at both trip ends directly causes a reduction in the choice of automobiles.* While places with high-density jobs may be highly attractive, particularly, for HBW trips, more cost may be involved in using automobiles due to congestion, toll, parking, and so on. And these areas often have convenient transit services and non-motorized transportation options available.

$H_{mode.C2}$: *The commercial share of land-uses at both trip ends directly causes discouragement in the choice of automobiles*. Commercial land-use, composed of office, retail, and hotel, will come with a dense and urban use of land, and in turn tend to enhance the expectation of earnings for people working there. People involved in jobs within the commercial land-use areas must pay more for their automobile trips around workplaces. This increased cost is likely to reduce the choice of automobiles for all purposes of trip.

$H_{mode.C3}$: *Regional accessibility is a direct cause in decreasing the choice of automobiles*. With more attractive opportunities (jobs) in a given driving travel time, a higher cost for such things as non-fuel cost (i.e., congestion, toll, parking cost, etc) will be required in order to reach destinations by automobile. In this context, people may be willing to substitute automobile travel with non-automobile travel.

$H_{freqency.C1}$: *The residential share of land at origin directly causes more automobile trips to be induced.* This is related to residential development patterns such as low density and segregated.

$H_{frequency.C2}$: *Regional accessibility causes an increase in HBW trips and a decrease in HBO trips.*

$H_{VMT.C1}$: *Population density at origin directly causes a reduction in the household VMT for HBW and HBO trips.* A variety of urban activities are available in places with a high density population, thus trips are likely to be shorter than those in places with a low-density population.

$H_{VMT.C2}$: *Regional accessibility directly causes less driving for HBO trips but more driving for HBW trips.*

Table 3.2 summarizes these hypothesis statements with the expected signs and the expected direct causes of land-use variables.

Hypotheses for Captivity Choice

The hypothesis is that the individual mode choice model that controls for low-density residential land-use provides a more accurate prediction of automobile captivity. Individual trip-makers with singleton choice sets (here, driving-alone and shared-ride) are captive to automobiles. Hence, the testable null hypothesis is: dominance of low-density residential land-use affects a captive choice of automobiles.

Table 3.2 The Hypothesized Signs and Direct Causes in Travel Demand Models and Causal Models.

Travel Outcomes	Land-Use Variables	Travel Demand Models: Hypothesize Correlation		Causal Models: Hypothesize Causality	
		HBW	HBO	HBW	HBO
Mode Choice	Population density at O / D	+ (−) / + (−)	+ (−) / + (−)	√ (− / −)	√ (− / −)
	Job density at O / D	+ (−) / +(−)	+ (−) / +(−)	√ (− / −)	√ (− / −)
	Residential use share at O / D	? / ?	+ (−) / + (−)	?	?
	Commercial use share at O / D	+ (−) / + (−)	+ (−) / + (−)	√ (− / −)	√ (− / −)
	Regional accessibility at D	(−)	(−)	√ (−)	√ (−)
	Entropy index at D	(−)	(−)		
Auto Trip Frequency	Population density at O	−	−	√ (−)	√ (−)
	Job density at O	−	−	?	?
	Residential use share at O	+	+	√ (+)	√ (+)
	Commercial use share at O	?	−	?	?
	Regional accessibility at O	+	−	√ (+)	√ (−)
	Entropy index at O	−	−	?	?
VMT	Population density at O	−	−	√ (−)	√ (−)
	Job density at O	−	−	?	?
	Residential use share at O	+	+	√ (+)	√ (+)
	Commercial use share at O	?	−	?	?
	Regional accessibility at O	+	−	√ (+)	√ (−)
	Entropy index at O	−	−	?	?

a. Expected signs in travel demand models: ( ) indicates the signs for drive-alone and shared-ride and ? indicates compounding variables.

b. Expected direct causes in causal models: √ indicates the expected direct causes, and ( ) presents the expected signs when assuming the direct causes. ? indicates compounding causes.

CHAPTER IV

RESEARCH METHODOLOGY

This chapter discusses the methodology for empirical analysis associated with the research questions addressed in this study. First, the data sources used for this study are presented together with the study area. Second, the travel outcomes and the explanatory variables used in the empirical analysis are discussed, and the operational processes for measuring the variables are explained. Third, empirical methods for both parametric and non-parametric analyses are discussed in detail in order to analyze the causal structure of the linkage between travel and land-use. In this section, parametric analysis and non-parametric analysis are addressed.

Data Sources

All the data for this study were obtained from the North Central Texas Council of Governments (NCTCOG), a metropolitan-wide association of local governments for regional planning and sound regional development of the Dallas-Fort Worth area. Travel data were obtained from the 1996 Dallas-Fort Worth Household Activity Survey (hereafter called the 1996 D-FW Travel Survey). This survey was the revealed preference survey using two sampling methods[8]: random digital dialing and intercept. Of 9,398 total recruited households, 3,996 households provided information on travel activities undertaken by all members of each household. Surveyed households were

[8] For details of survey design, see 1996 Dallas-Fort Worth Household Travel Survey.

stratified to assure reliability by the assigned day of the week at three levels: geographical location, household size, and vehicle ownership. Compared to data from the 1990 U.S. Census for Dallas-Fort Worth Consolidated Metropolitan Statistical Area (CMSA), the sample of stratified surveyed households falls mostly into the differences of 3% points or more (see Appendix A1). This survey recruited the large number of households, but a relatively small portion of trips made by transit was reported.

Trip records (37,065) did not originally contain the origin and destination for each trip. It was necessary to extract this information from the preceding and succeeding non-trip records. Records in which the origins and/or destinations were not identified were deleted along with records that were erroneous, leaving a file with 3,048 households and 22,316 trips. From these trip records, only trips made by adults (defined as over 17 years of age) were considered. Finally, a trip file with both 2,848 households and 15,138 trips were taken into account for the empirical analysis of this research.

The 1995 NCTCOG's land-use GIS and TransCAD data were used to capture local land-use characteristics at the level of traffic survey zone (TSZ). The data provide the spatial distribution of land-use in dimensions such as density, diversity, design, and accessibility. Various GIS techniques were employed to compute land-use measures (clipping land-use by categories from TSZs, intersecting the land-use and TSZs, spatial geo-coding, etc).

Other sources of data include ravel time and automobile operating cost for travels between each pair of the 4,874 traffic survey zones in the Dallas-Fort Worth metropolitan planning area. These data vary by travel mode. This study identifies four distinct mode splits (walk/bike, bus, driving-alone, shared-ride) based on the data analysis of the 1996 D-FW Travel Survey. Travel times for three travel modes (walk, driving-alone, and shared-ride), automobile operating costs for automobile (driving-alone and shared-ride), and travel distances between origins and destinations were obtained by skimming the NCTCOG model system's roadway network by time of day (AM period: 6:30~8:59 AM, PM period: 3:00~6:29 PM, and OP period: 9:00 AM to 2:59 PM and 6:30 PM to 6:29 AM), suggested by the NCTCOG's regional travel demand documentation.[9] The volume-delay function of the Dallas-Fort Worth Regional Travel Model is similar in form to the BPR-type functions used in other regional models.[10]

Table 4.1 summarizes data bases used in this study.

Table 4.1 Summary of Data Bases

Data Source	Description	Use
1996 D-FW Household Activity Survey	- Originally 37,065 trip records from 3,996 households surveyed. - Trip file (with 2,848 households and 15,147 trips) is used for the empirical analysis	- Trip characteristics: trip mode, trip duration, trip length, etc. - Personal and household characteristics: age, gender, income, household type, numbers of workers, etc.
1995 Land-use GIS data	- 20 land-use categories (codes) - GIS data format	- Local land-use characteristics: land-use mix, street features, entropy (concentration index), etc.
TransCAD data – TSZ	- Population and employment by sector in 1995, 1999, and 2025	- Local population and employment densities , job-housing balance, etc
TransCAD – Trip table	- O-D trip tables by time of day and by mode	- for skimming travel time and cost
TransCAD data – Roadway network	- Roadway network (1999)	- Roadway link capacity ratio total trips at TSZs, etc.

[9] Travel demand modeling documentation is available at www.nctcog.org.
[10] The general form of NCTCOG's volume-delay function is defined as travel time = free flow time $+ Min\{\alpha \cdot e^{\beta(v/c)}, \gamma\}$, where α, β and γ are delay function parameters. Each value for α, β and γ are different from freeway and non-freeway links.

Study Area

This study focuses on the NCTCOG's metropolitan planning area (MPA) where transportation planning efforts are currently concentrated. Figure 4.1 is the metropolitan area boundary (reddish area) which has 4,874 traffic survey zones (TSZ). Assuming all trips occur within this boundary, the coverage of trips includes internal trips between zones and within zones (Figure 4.2). The GIS data and TransCAD data also cover this boundary.

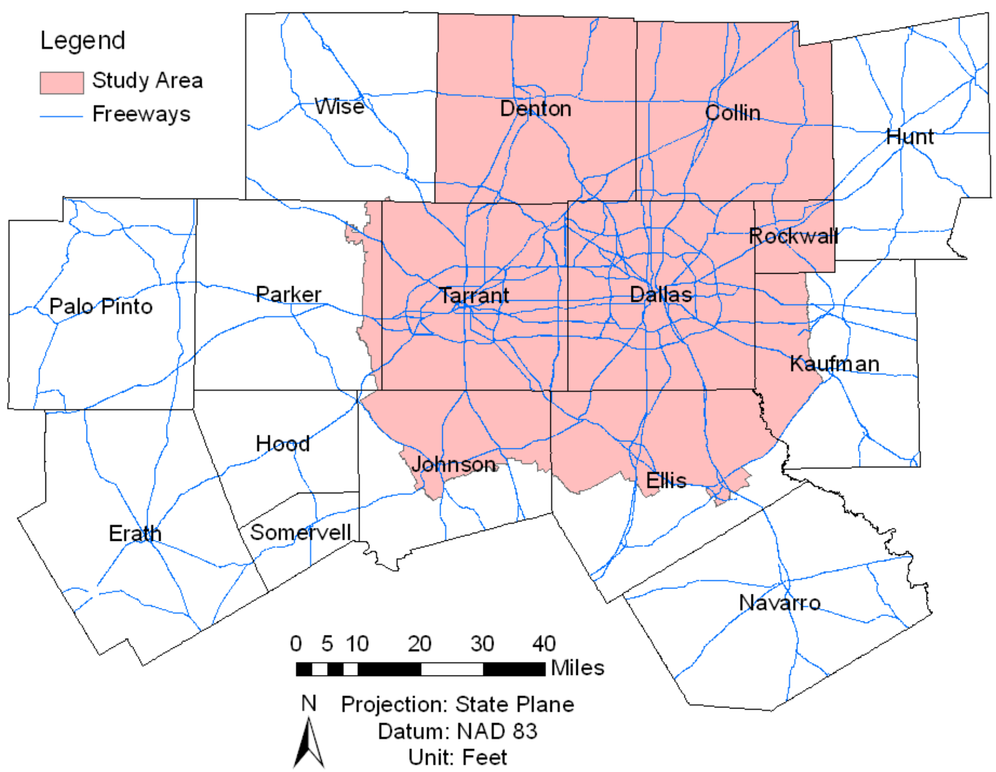


Figure 4.1 NCTCOG's Metropolitan Boundary

The MPA includes five full counties (Collin, Dallas, Denton, Tarrant, and Rockwall) and four partial counties (Ellis, Johnson, Kaufman, and Parker). These counties surround Dallas and Fort Worth as primary cities. The MPA boundary appears large enough for the analysis since it covers most locations where activities occur through trips in the Dallas-Fort Worth metropolitan area.

Variables and Measurements

This section discusses variables used to measure concepts that were identified in the hypotheses. Major components for the empirical analysis are travel outcomes and the set of potential factors, such as travel time (or generalized cost), socio-demographics, and land-use variables.

Dependent Variables

Dependent variables used in this study are travel outcomes which are observed as the results of individual or household travel decision-making (Table 4.2). These travel outcomes are mode choice, automobile trip frequency, and vehicle miles traveled (VMT). Mode choice is analyzed at the level of individual travel behavior, whereas trip frequency and VMT are taken at household level.

Table 4.2 Dependent Variables and Measurements Used for Travel Behavior Models

Variable	Type	Measurements
Travel mode	Discrete	If a trip maker drives alone, otherwise 0. (0 = walk/bike, 0 = bus, 1 = drive alone, 0 = shared-ride)
Trip frequency	Continuous / Count	Total number of auto trips for each household. Auto trips include driving-alone and shared-ride.
VMT	Continuous	Vehicle miles traveled by household members

Travel modes considered as alternative mode choices are walk/bike, bus[11], drive-alone, and shared-ride. These choices are reduced to a binary choice (automobile or non-automobile) to construct the causal models using directed acyclic graphs. Currently, NCTCOG operates three auto-and-transit-based mode choice models (HBW nested logit, HBN nested logit, and NHB multinomial logit).[12] Although mode choice has four alternatives, automobile use (driving-along and shared-ride) is overwhelming, as shown in Table 4.3. It indicates that individual trips in the Dallas-Fort Worth area are highly automobile-dependent. In particular, driving-alone is highly used in home-based work (HBW) trips as compared to home-based other (HBO) trips and non-home-based (NHB) trips. To enhance understanding of the spatial distribution of all trips, a digitized map with the trip end-points of 15,138 is presented in Figure 4.2. Many trips are observed in a cluster centering on both cities of Dallas and Fort Worth.

Table 4.3 The Distribution of Trips by Mode and by Trip Purpose

		Trip Purpose			Total
		HBW	HBO	NHB	
Mode	Walk/bike	56 (1.2)*	284 (4.0)	281 (8.4)	621 (4.1)
	Bus	276 (5.9)	66 (0.9)	48 (1.4)	390 (2.6)
	Drive-alone	3,917 (83.8)	3,484 (49.0)	1,941 (57.9)	9,342 (61.7)
	Shared-ride	423 (9.1)	3,278 (46.1)	1,084 (32.3)	4,785 (31.6)
Total		4,672	7,112	3,354	15,138

* Parenthesis indicates the percentage of trips.

[11] 1996 D-FW household activity survey recorded a large number of households, but the number of trips made by transit was relatively small. At that time, Dallas Area Rapid Transit (DART) was not opened yet. After opening DART, a set of data was added by including the records of transit-on-board surveys from DART in 1998 and Fort Worth Transit Authority (FWTA) in 1996.

[12] NCTCOG's mode choice is modeled from the dataset added by the surveys of DART in 1998 and FWTA in 1996. Currently, mode choice includes 1) auto-drive-alone, 2) auto-two occupants, 3) auto-three or more occupants, 4) transit-auto access, and 5) transit-walk access (NCTCOG 2005c).

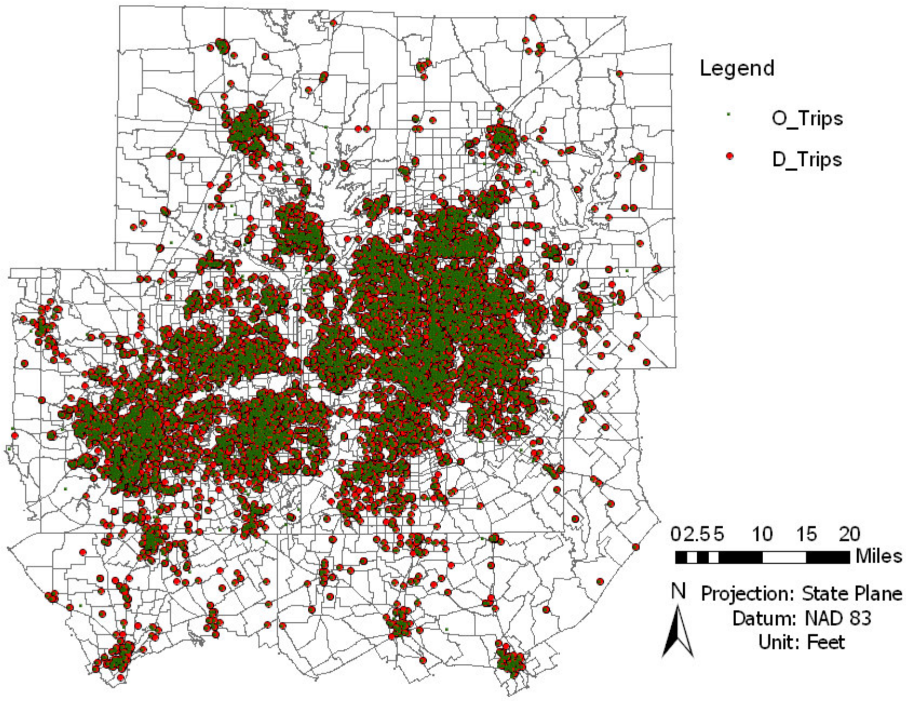


Figure 4.2 Trip Points at Origins and Destinations

Trip frequency is a travel outcome often used for household trip generation studies. Automobile trip frequency is defined as the number of trips generated by automobiles (i.e., drive alone per personal vehicle or carpool) for each household after eliminating duplicate trips. In this study, automobile trip frequency is used as a dependent variable for the trip frequency models. Household vehicle miles traveled (VMT) is a composite travel outcome computed as the sum of trip lengths over each origin-destination pair after eliminating duplicate trips and accounting for occupancy in a vehicle. Trip length with the shortest path between origin and destination was used for calculating household VMT. Sample sizes for automobile trip frequency and vehicle miles traveled (VMT) at household level are presented in Table 4.4.

Table 4.4 Automobile Trip Frequency and Total VMT at Household Level

	Trip Purpose		
	HBW + HBO + NHB	HBW	HBO
Number of households	2,749	1,955	2,072
Number of all auto trips	14,127 (5.1)*	4,340 (2.2)	6,762 (3.3)
Total VMT for all households (in miles)	148,579 (54.0)**	57,574 (29.5)	60,515 (29.2)

* and ** are average number of auto trips per household and average VMT per household respectively.

Independent Variables

Independent variables were cautiously explored from the varying databases (see Table 4.1) in order to account for the full array of factors thought to affect the travel outcomes discussed above. A full array of explanatory variables includes price variables (travel time or generalized cost), socioeconomic characteristics (personal or household characteristics), and land-use variables with the dimensions of density, diversity, design, and accessibility.

Travel Time and Generalized Cost

Travel time by mode (walking, driving-alone, and shared-ride) was skimmed through the shortest network paths between each origin-destination pair. In particular, travel times by driving-alone and shared-ride were obtained by skimming the shortest paths by the time-of-day (AM peak, PM peak, and off-peak). Walking travel time was obtained by converting trip length to minutes at the speed of 3 miles per hour (NCTCOG

2005c). Estimates of roadway travel times in the Dallas-Fort Worth Regional Travel Model (DFWRTM) include a combination of “free” speed travel time, delay time, and intrazonal travel time. First, two travel times are calculated through the traffic assignment volume-delay function from interzonal trips, whereas intrazonal travel times are obtained separately from interzonal trips (NCTCOG 2000, 2005a; Vadali and Lee 2005). Travel time for taking transit (bus) was not able to be skimmed due to the non-existence of transit routes and the complexity of multimodal use for trip points. Hence, travel time by transit was calculated by applying both transit operational performance (4.35 minutes per mile) and NCTCOG’s maximum walk access time (20 minutes). [13]

Some drawbacks of using travel times through the shortest network paths should be noted. The NCTCOG roadway network data do not provide travel times for public transit having its own fixed routes. Therefore, the calculated travel times by transit may have limitation used for mode choice model. Network travel times skimmed for this study are likely shorter than the actual travel times made by trip makers because they do not take into account access time and stops made on the way to destination. Also, some people may make trips through certain routes relatively familiar to them rather than the shortest paths. The drawbacks pinpointed here may exist in a similar fashion for travel cost (automobile operating cost) as well as travel length (for the use of calculating VMT) skimmed using the roadway network data.

---

[13] According to national transit databases (FHWA FTA 2006), the average speed for transit (rail + non-rail) passengers was 20.3 miles per hour (mph) in 1997. Rail speed and non-rail speed were 26.1 mph and 13.8 mph respectively. DART rail was under construction at survey time, and hence 13.8 mph for non-rail average speed (or 4. 35 minutes per mile) was applied for calculating transit time. NCTCOG mode choice model assumes a value of 1 mile as the maximum walk access time to transit at an assumed 3.0 mph walking speed.

In the roadway traffic assignment module of the Dallas-Fort Worth Regional Travel Model (DFWRTM), each link's generalized cost is composed of travel time and automobile operating cost. Herein, the automobile operating cost primarily includes the fuel cost only as influence[14] on travel cost, assuming 7.3 cents per mile in 1999 constant dollar. Link tolls are adjusted to the constant 1999 dollar but have little influence on travel cost. Thus, the automobile operating cost with the fixed cents per mile shows a high correlation with travel time or trip length (see Appendix A3). In this respect, the operating cost is inappropriate for use as travel cost. The total cost (generalized cost)[15] of traveling through a roadway link is calculated and is used as a proxy of travel time in this study.

Socioeconomic Characteristics

Socioeconomic characteristics are important factors in trip generation and mode choice. This study initially identified several variables in this category; age, sex, household size, household income, number of workers in a household, number of vehicles owned by the household, and household dwelling-type. Among those variables, age, gender, and household dwelling-type (multifamily or single-family housing) are typically confounding variables and have mostly minor statistical significance in empirical analysis. Therefore, household size, household income, number of workers in a household, and number of vehicles owned by the household are incorporated into the

[14] Fuel cost is primarily composed of a short-term cost, while long-term cost includes insurance, car-buying cost, repairs, etc.
[15] Generalized cost used in the user equilibrium module of DFWRM is defined as sum of *auto operating cost* + (*VOT*)*(*travel time*) where the VOT is $10.00 / hour ($0.167 / minute) for auto-based vehicle classes (driving-alone, shared-ride with HOV, and shared-ride without HOV) (NCTCOG 2002, 2005c).

empirical models (see Table 4.5).

Land-Use Measures

Various land-use measures have been generated in 4,874 traffic survey zones (TSZ) in the D-FW area. These measurements typically rely on the use of geographical information system and the availability of spatial data at the regional or local level. Land-use measures are usually classified into dimensions of density, diversity, design, and accessibility in literature. Overall, this study adopts similar measures employed in the existing literature

The density of population or employment is the most popular measure of land-use which concentrates on the intensity of development in a developed area. Population density (per acre) and employment density (per acre) are adopted at the trip points of origin and destination. Employment is summed over retail, service, and basic (industrial) sectors. The density is measured on a per acre basis within each TSZ, not within a quarter- or a half-mile radius at a typical local level. The use of a unit-mile radius (i.e., 'around each trip point') provides a good measurement of a neighborhood but may be biased by assuming a uniform density for the distributions from different TSZs. Of course, per acre density within TSZ exhibits the possibility of attenuating or hiding a 'right' density in the unit of neighborhood. Despite such a drawback, this study uses per acre density based on TSZ in order to keep the consistency of geographical unit of analysis as well as to reduce the modifiable areal unit problem (MAUP) or ecological fallacy associated with a geographical scale and aggregation.

Per developed acre intensities of land-use categories are measures that indicate land-use composition as share (or percentage) to total developed area of land-use classified as residential, commercial, industrial, government/education, and infrastructure, etc. Of several components, the residential share of land (single-family, multi-family, and mobile home) in TSZ, and the commercial share of land (office, retail, and hotel and model) in TSZ were computed at both the trip origin and destination. High percentages of classified land-use result in less diversity of land-use. In addition, two indicator variables are derived from the residential or commercial share of land-uses in order to represent the residential land-use dominance. If land is dominantly used for residential land with no commercial use at TSZ origins or destinations, this indicates a dominant residential use of the total developed area. This variable is used for mode captivity in order to test for the captive effect of dominant residential land-use on choosing automobiles. The entropy index[16] measures land-use balance to show how the land of a certain area is used in accordance with various land-uses. However, this measure fails to explain the compositional difference in land-use. For example, it is difficult to figure out the composition of land-use from the index of range (0.4~0.6). The entropy index was computed with four categories (residential, commercial, industrial, and public use). Another old popularized measure, job-housing balance, was excluded

---

[16] The formula of entropy index is as follows. $E_K = -\left(\sum_{i=1}^{K}(p_i \cdot \ln p_i)\right) / \ln K$ where p_i is the proportion of land-use category i of the considered developed area, and K is the number of land-use categories (K=4). The index value varies from 0 to 1 (from single-use to equal land-use). This formula is not defined when calculating no proportion of land-use, i.e., *ln*0. An alternative measure can be applied to measuring land-use balance. It is defined as $CR_k = {}^1\!/_k \cdot \sum_{i=1}^{K}(p^2{}_i)$ where p_i is the proportion of land-use category i, and K is the number of land-use categories (K=4). The value of CR_k ranges from 1 to k.

because of no or bad performance in forecasting in this study.

Density is addressed as a proxy for accessibility in literature (Kockelman 1997; Cervero and Kockelman 1998), but the correlation between job (or population) density and gravity-type regional accessibility in the Dallas-Fort Worth area provides no empirical support. Regional accessibility (RI) refers to the number of opportunities of a place for other places. It is defined as the terms of automobile access to a TSZ from all other TSZs in this research, and is normalized to a range of 0 to 1 divided by a scale factor to be used as 'relative' index rather than 'absolute'. Regional accessibility to jobs is computed as

$$RI_i = k \cdot \sum_{j=1}^{J-1} \left(\frac{jobs_j}{e^{\beta \cdot t_{ij}}} \right) \tag{4.1}$$

where i is the TSZ in question and j (1,2,…, J-1) is other TSZs with access to i. And t_{ij} is travel time between i and j, and β is a parameter of Bessel function in the NCTCOG gravity model. The parameter varies along trip purpose (HBW = 0.00156, HBO = 0.0042, and NHB = 0.001515). k is a scale factor, divided by a maximum regional accessibility. However, this measure is a composite function of both travel time and the number of jobs, and is likely to have a different impact on travel by type of trips.

This study does not include design factors such as street patterns and pedestrian amenities due to data constraints. Neither are provisions for public transit included. An effort was made to measure design or transportation provisions, intersection control index and roadway link capacity to per hour trips at TSZ were created as proxy variables but these variables were declined during pre-analysis due to poor performance in forecasting. These variables are not considered in the main study.

Table 4.5 summarizes independent variables and their measurements used throughout this study. Descriptive statistics presented in Tables 4.6 and 4.7 provide a good grasp of explanatory variables to explain trip data at both individual level and household level.

Table 4.5 Independent Variables and Measurements

Variable	Measurements
Travel times by mode	Transit minutes is calculated by applying operation performance and maximum walk access time, and walking minutes and vehicle minutes for driving-alone and shared-ride are skimmed.
Generalized cost	Auto operating cost + (VOT)*(travel time) for automobile in US$.
Age	Age in year for trip maker
Sex	Gender of trip maker, male = 1, female = 0
Dwelling type	Multifamily housing = 1, otherwise = 0
Household size	Number of household members
Household income	Household income as estimated from one of twelve income brackets or the sum of household members' income
Number of workers	Number of workers in the household members
Number of vehicles	Number of vehicles owned by the household
Population density at origins	Population density per acre in TSZ at origins
Population density at destinations	Population density per acre in TSZ at destinations.
Jobs density at origins	Employment density per acre in TSZ at origins. (employment = jobs in service + jobs in retail + jobs in basic sector)
Jobs density at destinations	Employment density per acre in TSZ at destinations
% residential use at origins	Percentage residential use to total developed acre in TSZ at origins. Total developed area includes residential, commercial, industrial, government/education, and infrastructure use
% residential use at destinations	Residential use share to total developed acre in TSZ at destinations
% commercial use at origins	Commercial use share to total developed acre in TSZ at origins
% commercial use at destinations	Commercial use share to total developed acre in TSZ at destinations
Entropy index of land-use mix	Formula measure of land-use balance
Regional accessibility	Measure of opportunities of a place to other places in gravity model form
Residential land-use dominance at origins	If land at TSZ origin is dominantly used for residential land with no commercial use, then 1, otherwise 0.
Residential land-use dominance at destinations	If land at TSZ destination is dominantly used for residential land with no commercial use, then 1, otherwise 0.

Table 4.6 Descriptive Statistics for Individual Trips Data

Variables	HBW		HBO		NHB	
	Mean	Std. D	Mean	Std. D	Mean	Std. D
Age of trip maker	41.7	12.02	43.9	15.47	42.0	12.66
Sex of trip maker (male =1, female=0)	0.55	0.5	0.42	0.49	0.43	0.5
Household income (US$)	60097	34.19	60134	35.32	63654	36.98
Household size	2.8	1.37	3	1.39	2.8	1.32
Number of workers in HH	1.9	0.84	1.5	0.98	1.7	0.83
Number of vehicle in HH	2.1	1.05	2.1	1.03	2	0.98
Multi-family housing (yes =1, no=0)	0.18	0.39	0.14	0.34	0.18	0.39
Travel time by walking (A) (min)	240.48	209.88	182.47	247.15	197.48	253.99
Travel time by driving (B) (min)	20.2	13.51	14.5	15.29	15.3	15.54
Travel time by transit (C) (min)	63.43	58.29	77.96	49.50	67.07	59.70
Travel time difference (A-B) (min)	220.3	197.02	167.9	232.19	182.2	238.78
Population density (per acre) at O	5.4	5.69	6.2	5.04	4.8	4.94
Population density (per acre) at D	5.2	5.73	6.1	4.88	4.8	5.19
Job density (per acre) at O	27.2	159.45	4.8	40.37	41.1	185.2
Job density (per acre) at D	31.2	165.41	7.2	58.44	37.1	166.92
% residential Use at O	58.0	35.64	69.6	28.54	47.1	36.25
% residential Use at D	54.4	36.88	69.4	28.9	47.4	36.21
% commercial Use at O	15.8	23.64	11.9	17.25	23.2	27.97
% commercial Use at D	17.4	25.31	11.9	17.92	23.4	27.95
Regional accessibility to jobs	0.9698	0.01	0.8943	0.03	0.9839	0.01
Land-use balance (entropy) at D	0.4168	0.24	0.4196	0.24	0.4215	0.24
Residential use dominance at O	0.1579	0.3647	0.1483	0.3554	0.1032	0.3042
Residential use dominance at D	0.1597	0.3663	0.1513	0.3583	0.1011	0.3014
Sample size	4,672		7,112		3,354	

Table 4.7 Descriptive Statistics of Household Level Data

	HBW		HBO	
Variables	Mean	Std. D	Mean	Std. D
Household income (US$)	60453	34.48	57732	35.08
Household size	2.7	1.33	2.7	1.37
Number of workers in HH	1.7	0.77	1.5	0.91
Number of vehicle in HH	2.1	0.97	2.0	1.01
Travel time per mile (minutes)	1.73	0.52	4.03	10.76
Travel time per trip (minutes)	20.61	11.87	21.11	22.38
Generalized cost per mile (US$)	0.06	0.01	0.17	0.49
Generalized cost per trip (US$)	4.31	2.60	4.49	4.89
Population density (per acre) at O	6.7	5.80	6.7	5.28
Job density (per acre) at O	4.5	46.36	3.7	39.72
% residential Use at O	74.3	25.29	75.4	24.23
% commercial Use at O	8.9	13.45	8.6	12.67
Regional accessibility to jobs	0.9643	0.03	0.9367	0.04
Land-use balance (entropy) at O	0.3836	0.22	0.3795	0.22
Sample size	1,955		2,072	

Research Design

All the empirical models are analyzed using both regression method and causal graphical analysis, and include a full array of factors such as price variables, socioeconomic characteristics, and land-use variables. From the standpoint of variable inclusion, the author assumes that overall travel behavior is completely influenced by an entire set of variables, particularly focusing on the impact of land-use on travel. Price variable (travel time or generalized cost) is assumed as a factor connected to socioeconomic characteristics and/or land-use variables. It is intended to show the role of price variable in the travel demand model. Land-use variables are captured at from-where-to (origin to destination) measures for mode choice and from-where measures (origin or household location) for household trip frequency and VMT.

Conventional regression models are estimated for three travel outcomes: individual mode choice for four alternatives (walk/bike, bus, drive-alone, and shared-ride), household trip frequency, and household total vehicle miles traveled (VMT). These outcomes are explored by trip purpose such as home-based work (HBW) trips, home-based other (HBO) trips, and non-home-based (NHB) trips. All the models include travel time (or generalized cost) and socioeconomic characteristics in the base model. Land-use variables are added to the extended model to test for the improvement of model. The scheme of mode choice is extended to logit captivity to explore contribution to travel behavior attributable to land-use. Logit captivity is discussed later in this chapter.

Directed acyclic graphs (DAG) method is also implemented to examine the causal structure based on the same models as in conventional regression methods:

individual mode choice, and household trip frequency, and household total VMT. DAG is applied for a binary choice (automobile versus non-automobile) rather than multinomial choice because, to date, choice model has not been defined in the directed graph. Similar to the conventional approach, travel outcomes are modeled with a full set of explanatory variables by trip purpose. This study develops the DAG to discover the causal information flows imbedded in observational data. This method has never been employed in studies concerned with transportation–land-use linkage, and thus is unique in this study. The increasing use of DAG in applied sciences supports an external validity.

The research strategy used in this study goes beyond the traditional analysis of cross-sectional and observational data used to infer associative results under the assumed causality. A newly developed method, directed acyclic graphs (DAG) is employed to shed light on the connection between causality and data. Such a connection does not necessarily require experimental randomization in application. Longitudinal data have been suggested to shed light on causal relationships by analyzing changes in households using a household relocation data. However, despite the benefits obtained from the use of longitudinal data, the sufficiency of sample size, the inclusion of variables, and the analytical methods for the purpose of relocation have been criticized (Krizek 2003; Handy, et al. 2005). DAG originally was developed to discover the causal relationships within the cross-sectional and observational data using a series of algorithms derived from research in Artificial Intelligence. This method permits research ideal for analyzing causality in individual as well as household travel using either cross-sectional or time-series data.

Directed Acyclic Graphs

Over the last two decades, a group of philosophers and computer scientists have developed a graph-based analysis of causal structure and have shed light on the relationship between particular causal orders and relationships of conditional independence embedded in the statistical function. Pearl (2000) and Spirtes, et al. (2000) advocate the dominant position accounting for the graphical causal models based on non-experimental (or observational) data, and have as the basis of their work a nonparametric analysis, but not structural equation modeling.[17] The central feature of the model is structured by a directed acyclic graph (DAG). The main idea of this approach is to deal with the independence relations of variables entailed by the application of a causal graph under the causal Markov assumption: *a variable X is independent of every other variable (except X's effects) conditional on all of its direct causes*.

It is helpful to define some terms used in graph theory. A graph is formally composed of an ordered triple $\langle \boldsymbol{V}, \boldsymbol{M}, \boldsymbol{E} \rangle$ where $\boldsymbol{V}$ is a non-empty set of variables (or vertices or nodes), $\boldsymbol{M}$ is a non-empty set of marks (or symbols) at the endpoints of undirected edges, and $\boldsymbol{E}$ is a set of edges (or links) with the ordered pairs of variables and marks. Causal connections between variables are indicated by edges - any two variables connected by an edge are adjacent - that may or may not have the symbols of

[17] Structural equation modeling (SEM) is a modeling framework to deal with unobservable (or latent) variables and endogeneity among variables associated with measurement problems and pre-specifies direct, indirect, and associative relationships between variables that corresponds with theory and expectation (Washington et al. 2003). However, when we do not know the 'true' system, the SEM is paralyzed in handling the causal information embedded in data and the causal interpretation is elusive (Pearl 2000).

arrowheads indicating the direction of causation. If we have a *causally sufficient*[18] set of variables $\{V_1, V_2, V_3, V_4\}$, the graphs contain undirected edges (V_1 — V_2), directed edges ($V_2 \rightarrow V_3$), and bi-directed edges ($V_3 \leftrightarrow V_4$). However, when not assuming causal sufficiency, a partially oriented inducing path graphs contain directed edges ($\rightarrow$), partially directed edges (o$\rightarrow$), non-directed edges (o—o), and bi-directed edges ($\leftrightarrow$). Here, directed acyclic graphs (DAG) contain no directed cycles (or no self-loops).

Directed acyclic graphs defined by the usual graphic theory are now united with a probability theory with a focus on conditional independence, and with philosophy involved in causation among variables. The DAG specifies a class of probability distributions in a way given by the Markov condition, and the resulting probability is decomposed as a recursive product

$$P(v_1, \ldots, v_n) = \prod_{i=1}^{n} P(v_i \mid pa_i) \tag{4.2}$$

where P is the joint probability of variables $v_1, \ldots, v_n$ and pa_i represents the possible realizations of any subset of just immediate parent (or direct cause) variables V_i in order $V_1, V_2, \ldots, V_n$. The above representation of conditional independence shown in equation 4.2 is characterized by d-separation proposed by Pearl (1988, 1995, 2000), which is equivalent to a more general graphical relation. D-separation (directional separation) is a relation between three disjoint sets of variables V_1, V_2, V_3 in a DAG, and its basic idea is to check whether a set of variables V_2 blocks all connections of a certain type between V_1

[18] When we draw a causal graph, we assume that the set of variables in the graph is causally sufficient unless there are measurement errors in the variables. If a set of variables $\boldsymbol{V}$ includes all the common (direct or indirect) causes of pairs of variables in $\boldsymbol{V}$, then we say $\boldsymbol{V}$ is causally sufficient. For the example of X $\leftarrow$Y $\rightarrow$ W $\rightarrow$ Z, the set {X, Y, W, Z} is causally sufficient, while the set {X, W, Z} is not.

and V_3 in the graph G. Then, it is said that V_1 and V_3 are d-separated by V_2 in G. A notation for independence introduced by Dawid (1979) is formally used as $V_1 \perp\!\!\!\perp V_3 | V_2$ which means: V_1 and V_3 are independent conditional on V_2.

Hausman (1984) and Papineau (1985) realized that it is possible to capture the asymmetry of causation by adding a third variable to the systems in which it was not captured for the systems of two variables. In the context of the notion of d-separation, when causal meaning is attributed to the arrows in the graph, three different types of DAGs showing the causal directions with triples of variables V_1, V_2 and V_3 help figure out the intuition behind d-separation. First, if a variable V_2 takes each piece of information stemming from adjacencies (V_1 and V_3) but is no longer open for other variable(s), then the variable V_2 is a collider, and the causal graph will represent two causes having a common effect:

$$V_1 \longrightarrow V_2 \longleftarrow V_3$$

Information flows from this graph indicate that all forces caused by the variables V_1 and V_3 come together (or collide) on V_2 without going through it. Here, the variables V_1 and V_3 are *d-separated* by themselves in the DAG and, thus, the unconditional association (or correlation) is zero. However, if we condition on V_2 for the purpose of opening up a path to another variable, say, V_4 (this variable should be a child of V_2), then V_1 and V_3 are d-connected as a conditional association (or correlation). An example of such a causal relation in the linkage of travel and land-use can be addressed as follows: 'drive less' (V_2) cannot cause an increase in population density (V_1) and mixed land-use (V_3), but the increase in population density (V_1) and the increase in mixed land-use (V_3) cause one to drive less (V_2); then one opens a path to a health indicator for obesity (V_4) resulting from 'drive less'.

Second, consider that all information flows originate from a common cause. It is easy to see that V_1 and V_3 are not independent because both variables depend on V_2 (i.e., V_1 and V_3 are said to be d-connected as unconditional correlation). Here, it is intuitive that V_1 is independent of V_3 conditional on their common cause (V_2).

$$V_1 \longleftarrow V_2 \longrightarrow V_3$$

In other words, conditioning on the common cause V_2 nullifies the association between V_1 and V_3 (conditional association will be zero) and eventually V_1 and V_3 are *d-separated*. For example, an increase in population density (V_2) causes one to drive less (V_1) as well as to walk / bike more (V_3).

Finally, a further example of a causal chain illustrates how to apply the idea of *d-separation*. In a causal chain, V_1 and V_3 are dependent (d-connected as unconditional association), but independent conditional on V_2 (*d-separated* as conditional association).

$$V_1 \longrightarrow V_2 \longrightarrow V_3$$

We offer a simple example for the above causal chain: transit-oriented development (V_2) depending on high population density (V_1) in an urban area causes one to drive less (V_3) over areas where transit is easily accessible. As shown in the above three causal graphs, when a third variable is added to the system of two variables, causal structure can induce conditional dependence (for a case of collider) as well as eliminate unconditional dependence (for cases of both a common cause and a causal chain).

PC Algorithm

A procedure computing d-separation in any graph has been incorporated into an algorithm that computes all the directed acyclic graphs by Spirtes, et al. (2000). So far, several algorithms [19] have been developed, one of which is called PC algorithm computed by the TETRAD II or III (Scheines, et al. 1994; Spites, et al. 1996). The PC algorithm conducts the ordered commands sequentially: first, start with an undirected graph connected by every variable, second, remove edges (or lines) between each pair of variables through testing for conditional independence based on the partial correlation of order k (i.e., 0, 1, 2,…, k), and finally, orient the remaining edges based on separation set (or sepset)[20] and the away-from-a collider test.[21] A related algorithm using unshielded colliders instead of sepset is an inductive causation (IC) algorithm by Pearl (2000).

More specifically, the PC algorithm starts with the empirical distribution of a set of variables represented by the variance-covariance matrix or the unconditional correlation matrix and implements the test for probabilistic independence using partial (or conditional) correlations following order conditioning. The statistical significance of the conditional correlation is tested using Fisher's z-statistic

$$z = \frac{\sqrt{(n-3-k)}}{2} \times \ln\left(\frac{1+\rho_{ij|k}}{1-\rho_{ijk}}\right) \qquad (4.3)$$

---

[19] Several algorithms are described in detail by Spirtes et al. (2000): PC algorithm (p. 84), Modified PC algorithm (p. 125), Causal inference algorithm (p. 139), and Fast casual inference algorithm (p. 144).

[20] Sepset is the conditioning subset that renders variables X and Y independent given an unshielded undirected graph, X — W — Y. The undirected graph can be directed simply by determining whether W is a member of sepset (X,Y). If Z is not a member of the sepset to be d-separated, W is a collider and directed as X → W← Y.

[21] After every potential unshielded collider has been fully directed, the away-from-a collider test is applied immediately: if X →W, W and Z are adjacent, X and Z are not adjacent, and there is no arrowhead at W, then direct W — Z as W → Z (Scheines et al. 1994).

where $\rho_{ij|k}$ is population partial correlation of variable *i* and *j* conditional on *k*, and *n* is the number of observations used to estimate the correlations. *k* is the number of conditioning variables. If all variables (*i, j* and *k*) used for calculating the partial correlations are normally distributed, then the partial correlations will also follow a standard normal distribution.

Figure 4.3 (i)-(v) shows how PC algorithm works. The true causal graph with an unshielded collider is depicted in Figure 4.3 (i). The true structure determines which correlations will be found in the data, and which can be eliminated or oriented in each step of the algorithm. Starting from the unconditional correlation matrix calculated by four variables (V_1, V_2, V_3, and V_4), the algorithm begins with a graph (i) in which every variable is linked with each other with no direction. It then eliminates an edge (or link) between V_1 and V_2 by an unconditional (or zero-order partial) correlation test, shown in (ii). Next, it tests for the 1st order partial correlation of each pair of variables conditional on one variable (V_3), and leaves the edges as shown in (iii). In principle, it would continuously test for the k^{th} order partial correlation of each pair of variables conditional on *k* variables. For each triple of variables (V_1, V_2, V_3) in Figure 4.3 (iii), a pair V_1,V_3 and another pair V_2,V_3 are each adjacent to V_3, but V_1, and V_2 are not adjacent to V_3. If conditioning on V_3 renders V_1 and V_2 correlated, then the edges are oriented as arrows pointing into V_3 in (iv). Since V_3 is identified as a collider on V_1 — V_3 — V_2, the unshielded collider is fully directed. Next, we know from (iv) that V_3 screens off the correlation between V_1 and V_4 or between V_2 and V_4. Usually this means $V_1 \rightarrow V_3 \rightarrow V_4$ or $V_2 \rightarrow V_3 \rightarrow V_4$. This is consistent with the away-from-a collider test applied after every potential unshielded collider has been fully directed. The last process leads to the directed acyclic graph in Figure 4.3 (v).

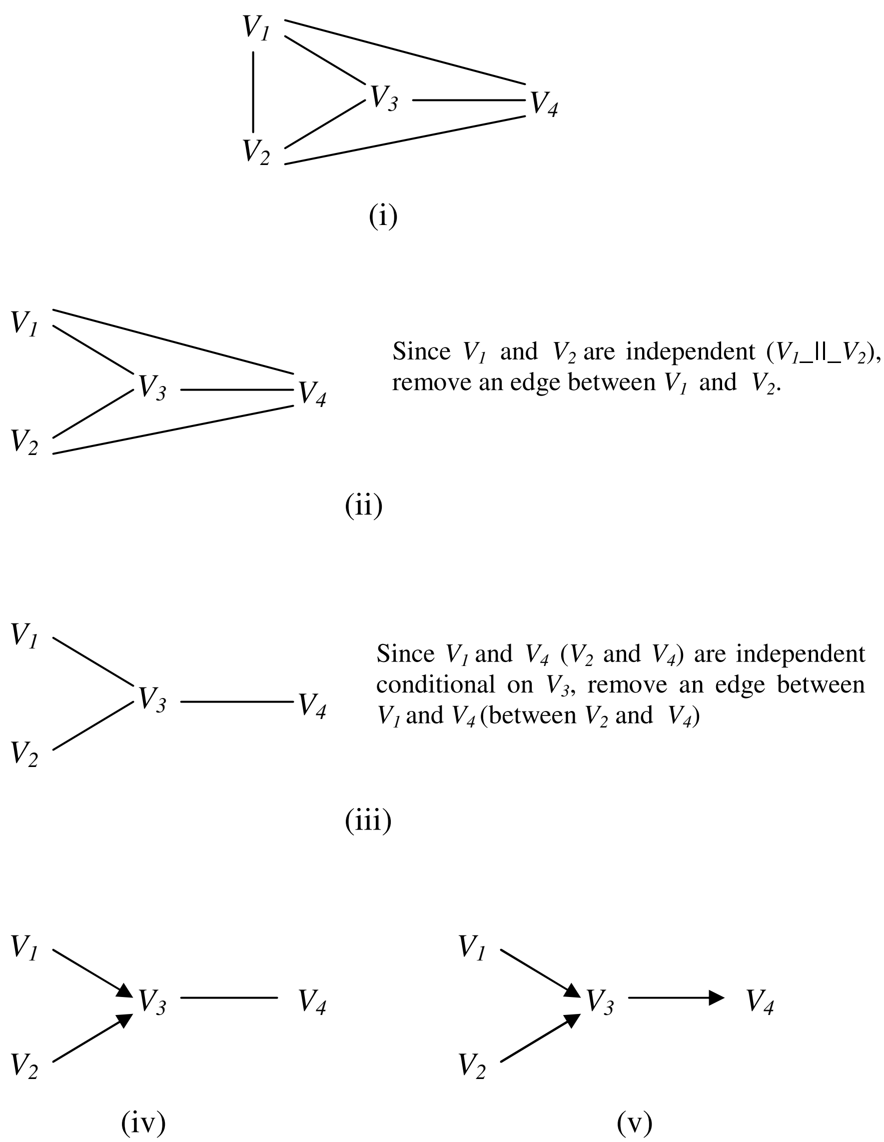


Figure 4.3 How Does the PC Algorithm Work?

Caveat

The approach on directed acyclic graphs is currently used in applied sciences[22] for classification, forecasting, and predicting the effects of interventions, but much of this approach is quite new and is generally unknown to planning professionals. Existing research using this method assumes that the causally ordered data are cross-sectional while time-series data are not directly applied to this method (Demiralp and Hoover 2003). As a statistical method, the method of directed acyclic graphs can test and potentially discover cause-effect relationships between variables in circumstances in which it is not possible to conduct controlled experiments. In this vein, the PC algorithm is completely successful in identifying the correct causal structures with reasonable reliability. However, the theory and scientific practice of directed acyclic graphs in terms of causal structure depends on the following assumptions: causal sufficiency, causal Markov condition, and faithfulness condition.

These assumptions may be violated when the observational data are employed in empirical research. One should have a *causally sufficient* set of variables which includes all the common causes of the measured variables. In other words, there should be no omitted variable that causes two or more included variables. Failure to do so may lead to spurious causal flow between two or more included variables.[23] The next requirement is the *causal Markov assumption* which states that all the relevant probabilistic information (or distribution) about a variable must be fully captured from its just parents or its direct

[22] Some applications have been taken by Druzdzel and Glymour (1999), Roh, Bessler, and Gilbert (1999), Shipley (1999, 2000), and Bessler and Loper (2001).

[23] Also see footnote 18. If X is a common cause of Y and Z but is omitted from the current analysis, then a causal flow, if any, between Y and Z may be spurious due to the fact that X causes Y and Z.

causes. The last assumption, *faithfulness*, focuses on the relationship between d-separation and probabilistic independence. This condition means that special combinations of causal strengths can result in unfaithful probability distribution information when captured by a graph. Although this is a very limiting case, the quantitative causal effect of two variables along different graphs exactly cancels each other out. The PC algorithm is applied with these three assumptions. When observational data are used for analysis, any result obtained from this application should be interpreted with caution in situations in which any assumption may be violated.

Choice Model Structure

Multinomial Logit (*MNL*)

The multinomial logit (MNL) model structure is based on the utility maximization theory. Each of the available modes (walk/bike, bus, driving-alone, and share-ride) has an associated utility that is a function of individual or household characteristics, mode attributes, and land-use characteristics. Binary logit is a simple case applied to two alternatives, specified with non-automobile (walk/bike and bus) and automobile (driving-alone and shared-ride) for the purpose of examining the causal structure of the choice model in the directed acyclic graph. Utility theory states that an individual trip maker chooses a mode that maximizes her or his utility. For a given observation, the utility of mode i of the trip-maker is given as

$$U_i = V_i + \varepsilon_i \qquad (4.4)$$

where U_i is the utility of mode i to the trip maker, V_i is the deterministic (observed) component of utility, and ε_i is the error (unobserved) component of utility (Train 2003).

The MNL structure depends on two basic assumptions. First, the error components of the utility function are an extreme value type I distribution referred to as the Gumbel distribution. This is the most commonly used distribution leading to a closed-form model for the choice probabilities. The second assumption requires equal variance for modes and for all individuals that there is no correlation between the error terms of modes and between the error terms of individuals. It is assumed that the error components are identically and independently distributed (IID) across observations as well as across modes (Horowitz 1986; Train 2003). The MNL structure is well-known for simple formulation and easy application. It provides the probability that the individual will choose a given mode based on the observable portion of the utility of the mode. Using MNL, the probability that a given individual chooses mode i from j modes is

$$P(i) \quad = \quad \frac{e^{\mathrm{V_j}}}{\sum_{J=1}^{J} e^{\mathrm{V_j}}} \quad \text{where} \quad V_j = \boldsymbol{\beta}'\mathbf{x} + \boldsymbol{\varepsilon} \qquad (4.5)$$

The benefit of MNL relies on how the maintained hypotheses of the study would be tested with the generic and/or specific specifications in the choice model. However, MNL in itself is plagued with the independence of the irrelevant alternative (IIA) property, which implies that for any given individual, the ratio of the choice probabilities of two alternatives is independent of all other alternatives. The MNL model, because of this property, overestimates the probability of taking either of the similar modes (for example, red bus and blue bus) compared to intuitive judgment, but underestimates the distinct mode. Thus, there is a need for each alternative to be identified distinctly.

As a goodness-of-fit, the likelihood ratio index (ρ^2 or $\bar{\rho}^2$) is often used with logit models to measure how well the models fit the data. The likelihood ratio index ranges from 0 to 1, but it should be noted that it is interpreted differently from R^2 used in regression. R^2 indicates the percentage of the variation in the dependent variable that is explained by the estimated model, while ρ^2 is the percentage increase in the log-likelihood function above the value taken at zero parameters. Another goodness-of-fit is the percent-correctly-predicted (%CP) which is calculated by identifying for each individual traveler the alternative with the highest probability, based on the estimated model, to the actual choice which the individual traveler made. The likelihood ratio test can be used to test for the model improvement by the land-use variables.

Ordered Logit

The number of trips taken by a household is discrete and thus can be ordered. The discrete and ordinal nature of the dependent variable as an outcome can be fairly captured by the ordered probit and logit models rather than with ordinary regression, MNL or probit. Also, ordinary regression analysis would err. It should be noted that probability in the ordered logit model incorporates a binary logit formula, but the ordered logit model has only one utility function with multiple choices to represent the level of utility unlike a binary logit with two utility functions. Household trip frequencies in the samples of work and non-work trips are categorized by order (1 to 3 for work trips, and 1 to 4 for non-work trips) depending on the identification of empirical distribution of households by automobile trip frequency. The decision can be represented as; "subsistence" (household trips ≤ 2), "moderate" (2 < household trips ≤ 4), "high" (household work trips > 4 or 4 < household trips for non-work ≤ 7), and "very high" (household non-work trips > 7).

Logit Captivity

Methodology chosen for creating a feasible choice set is important because the assumptions made in the analytical process can affect the results and validity of the model (Thill 1992). The most straightforward approach in the research of mode choice was to assume that every individual has the same choice set. However, this is not realistic: every trip maker would have a different choice set based on the various constraints associated with individual preferences as well as surrounding environments.

When each individual is confronted with a choice situation, his or her choice is based on a non-empty subset of universal choice space (M) which includes all of the possible modes used by people involved in trips. If one determines the dimensions of the universal choice space, the number of the non-empty subsets of the universal choice space grows with 2^M-1. For example, when the universal choice space is composed of three different types of mode (walk, auto, and transit), the non-empty choice subsets are numerated as {walk}, {auto}, {transit}, {walk, auto}, {walk, transit}, {auto, transit}, and {walk, auto, transit}. A choice set composed of {walk, auto, transit} is a universal choice set, free to choose a mode given three available modes. A trip maker with two-space choice sets, {walk, auto}, {walk, transit} and {auto, transit}, is still free to choose a mode given two available modes, although the freedom to choose a mode is slightly reduced. Here, single choice sets such as {walk}, {auto}, and {transit}, are equivalent to mode captivity. These singleton choice sets arise not only from mode captivity but also from the availability constraints faced by the trip maker.

In earlier times, Manski (1977) proposed the two-stage choice model to consider all the non-empty subsets of the universal choice set. The general form to capture the probability to choose a mode is expressed as

$$p_n(i) = \sum_{C \in G} P_n(i \mid C) \cdot Q_n(C) \tag{4.6}$$

where $P_n(i)$ is the probability of an individual (n) choosing mode i and $P_n(i|C)$ represents the probability of an individual choosing mode i among the modes contained in the choice subset C. $Q_n(C)$ is the probability that the individual considers choice subset C inside the all possible choice subsets G. Let us simply apply this to a case of two modes for choice of travel: driving (d) or non-driving (t). We assume that an individual has been (historically) captive to driving by automobile or free to choose either mode between driving and non-driving. Now he or she is newly confronted with a choice of mode. His or her probability of choosing driving (d) is calculated as

$$P(d) = P(d \mid C_d) \times Q(C_d) + P(d \mid C_t) \times Q(C_t) + P(d \mid C_{d,t}) \times Q(C_{d,t}) \tag{4.7}$$

where $P(d)$ is the probability that the individual chooses to drive in all choice situations, and $P(d|C_d)$, $P(d|C_t)$ and $P(d|C_{d,t})$ are the conditional probabilities of the choice subsets {driving}, {non-driving}, {driving, non-driving}, respectively. $Q(C_d)$, $Q(C_t)$ and $Q(C_{d,t})$

are consideration probabilities for each choice subset containing mode(s), respectively. Consideration probability for singleton mode indicates the intensity of mode captivity.

The specification of a two-stage choice model may lead to estimation results different from the specification of a choice model from usual practice, assuming that all individuals have the same choice dimension equal to the universal choice space. An example is presented in Table 4.8 which shows the joint probability density function of two choices (driving, non-driving) and three choice subsets ({driving}, {non-driving}, {driving, non-driving}) with each consideration probability $Q_i(C)$. The choice probabilities of driving (*d*) and non-driving (*t*) from usual practice are equivalent to the marginal probability for driving (0.6) and for non-driving (0.4), respectively, while the choice probabilities from the practice of consideration subsets are 0.658 and 0.342, respectively. The main interest of the study of automobile dependence is the estimation of the captivity coefficients attributable to land-use variables and the exploration of the probabilities of captivity represented by as the bold-faced numbers in Table 4.8.

Table 4.8 Hypothetical Probabilities of Choice and Consideration Set

		Consideration Set			Marginal Probability	Choice Probability
		$\{d\}$	$\{t\}$	$\{d, t\}$		
Choice	Driving (*d*)	0.3	0.1	0.2	0.6	0.658
	Non-driving (*t*)	0.1	0.2	0.1	0.4	0.342
Marginal Probability		0.4	0.3	0.3	-	-
$Q_i(C)$		**0.3**	**0.1**	0.6	-	-

The probability of having the full set of available modes is given 60%, whereas the probability of having the singleton set of a mode accounts for 30% for driving, and 10% for non-driving, respectively. The potential importance of mode captivity from the hypothetical probabilities of choice set should be noted (Table 4.8). The hypothetical result implies that trip makers have a 0.4 probability of being captive to a single mode. This may lead to a significant reduction of the impact of land-use relative to the MNL model, and there may be also biases in the estimated parameters and elasticities resulting from the MNL estimation.

The logit captivity model was first theoretically derived by McFadden (1976) and Ben-Akiva (1977), was developed by the 'dogit' model of Gaudry and his colleagues (Gaudry and Dagenais 1979; Gaudry and Wills 1979), and was generalized to a 'parameterized logit captivity' (PLC) model dealing with random constraints to choice-set formation by Swait and Ben-Akiva (1987a, 1987b). Research focusing on the decision-making process is still underway, utilizing either a probabilistic choice set model (Ben-Akiva and Boccara, 1995) or a choice set generation model (Swait, 2001; Basar and Bhat, 2004). The functional form of the model is written as the multinomial logit (MNL) form with the same structure of likelihood function. The search process, however, is very complex because the log-likelihood of the model is not globally concave. The functional form to be estimated is given as

$$
\begin{aligned}
P(i) = \frac{e^{\boldsymbol{\beta}'\mathbf{x}_i} + e^{\boldsymbol{\gamma}'\mathbf{z}_i}\sum_j e^{\boldsymbol{\beta}'\mathbf{x}_j}}{\sum_j e^{\boldsymbol{\beta}'\mathbf{x}_j}\left(1+\sum_{j\in C} e^{\boldsymbol{\gamma}'\mathbf{z}_j}\right)} \quad &= \quad \frac{e^{\boldsymbol{\gamma}'\mathbf{z}_i}}{\left(1+\sum_{j\in C} e^{\boldsymbol{\gamma}'\mathbf{z}_j}\right)} + \frac{1}{\left(1+\sum_{j\in C} e^{\boldsymbol{\gamma}'\mathbf{z}_j}\right)} \cdot \frac{e^{\boldsymbol{\beta}'\mathbf{x}_i}}{\sum_j e^{\boldsymbol{\beta}'\mathbf{x}_j}} \quad i, j = 1, \dots, n \\
&= \quad \frac{\theta}{(1+\theta)} \quad + \quad \frac{1}{(1+\theta)} \cdot \frac{e^{\boldsymbol{\beta}'\mathbf{x}_i}}{\sum_j e^{\boldsymbol{\beta}'\mathbf{x}_j}} \qquad (4.8) \\
&= \quad Q_{C_i} \quad + \quad Q_C \cdot P_{j|C}
\end{aligned}
$$

where vector **z** represents the variables to explain captivity to mode *i*, while vector **x** include a set of variables impacting the choice of mode *i*. Both vectors may partially or totally overlap. β and γ are the estimated coefficient vectors, particularly, the later is the vector of estimated captivity coefficients. Q_{C_i} is the probability that an individual is captive to mode *i*, or is referred to as captivity odds, and measures the intensity of automobile dependence as a probability for driving-alone or shared-ride. Q_C is the probability that the individual becomes a free choice user, and $P_{j|C}$ is the probability of choosing mode *i* given that the individual is a free choice user. If theoretically there is no captivity to mode ($\gamma' Z_i = 0$) or statistically there is no significant coefficient, the model is equivalent to the standard multinomial logit (MNL). As noted, this model may provide realistic estimates for choice captivity.

CHAPTER V

EMPIRICAL RESULTS

This chapter presents and interprets empirical results which estimate the specified models of individual mode choice, a household trip frequency, and household VMT. Then, logit captivity results follow. Prior to treating with the empirical results, land-use as the status quo is examined with the focus on land-use balance in the D-FW metropolitan area.

Status Quo of Land-Use

According to the North Central Texas Council of Governments (NCTCOG), 58 percent of the total developed land in the sixteen counties in question is residential. Residential land-use accounts for nearly three times the land area used for commercial, industrial, and institutional uses combined. Similar trends are seen from land-use in the D-FW metropolitan boundary. As shown in Table 5.1, single-family residential land-use appears to dominate the urban development patterns that have formed the D-FW area. In 1995, residential land-use in the D-FW area accounted for 65 percent of total land-use considered (59 percent for single-family residential). Currently, the D-FW metropolitan boundary includes more than 150 municipal cities, many of which have their own zoning regulations to limit the density of new residential development.

Residential development is likely to be affected by zoning regulations in major metropolitan areas which highly rely on smaller municipalities for land-use planning. In fact, zoning regulations enforced for a low-density development pattern may lead people to drive farther in order to meet their activity needs. Cities are highly motivated to build too much parking and multi-lane arterials for cars rather than developing pedestrian-friendly (walkable, bikable or active-living) communities (Levine 2005). No empirical study, however, supports the connection that the urban design template written into land-use (zoning) regulations leads to more driving. This is a testable hypothesis requiring future study.

Table 5.1 Land-Use in the D-FW Metropolitan Area

Land-Use Classification		1990	1995		2000	
Residential	Single family	362,593	421,472	58.8	435,583	57.8
	Multi-family	26,932	27,202	3.8	27,326	3.6
	Mobile home	18,149	16,550	2.3	22,555	3.0
Commercial	Office	10,588	10,707	1.5	21,819	2.9
	Retail	39,891	45,458	6.3	45,105	6.0
	Hotel / Motel	911	1,048	0.1	1,080	0.1
Industrial		67,988	67,291	9.4	64,947	8.6
Institutional		na	35,644	5.0	45,791	6.1
Infrastructure		na	81,774	11.4	81,895	10.9
Under construction		na	10,211	1.4	7,071	0.9
Total acres		527,053	717,359	100.0%	753,171	100.0%

There are many ways to measure land-use features which fall into categories such as density, diversity, accessibility, and design. These measures are directly or indirectly related to humans or their activities, but diversity measures, such as land-use balance (entropy), mixed-use indicator (ratio), and specific land-use share, represent the physical sizes or combinational portions of land-use per unit of area. As noted, the status quo analysis of land-use balance (entropy index) will provide a good diagnosis for how land in study area is used (or developed) in concert with a variety of human needs and activities.

The distribution of entropy index at TSZ is presented over each focused region in Table 5.2. The zonal distributions for four land-use types (residential, commercial, industrial, and public use) were considered for the computation of the entropy index. To compare the regional distribution to the county level, Dallas and Tarrant counties were also examined. The regional average entropy index is 0.3832, and 784 TSZs fall into a range of 0.4~5.0. The same trends are observed from two counties. TSZs with zero entropy are used for either no land-use or only of the four types of uses. Single-family-residential use accounts for 199 out of 521 TSZs in D-FW. Out of 1,870 TSZs with a 'low' (≤ 0.3) entropy index, 59 percent (1,094 TSZs) is used for single-family oriented residential, exceeding more than ten times the land area used for commercial, industrial, and institutional uses combined in terms of land area.

Table 5.2 Distribution of Entropy Indices for Four Land-Use Types

Entropy Index	D-FW-TSZs		Dallas Co.-TSZs		Tarrant Co.-TSZs	
0.0000	521*	10.7	179*	8.2	115*	8.2
0.0001~0.1000	344	7.1	156	7.1	97	6.9
0.1001~0.2000	461	9.5	217	9.9	115	8.2
0.2001~0.3000	544	11.2	261	11.9	148	10.6
0.3001~0.4000	613	12.6	304	13.9	173	12.4
0.4001~0.5000	785	16.1	364	16.6	240	17.2
0.5001~0.6000	544	11.2	243	11.1	164	11.7
0.6001~0.7000	468	9.6	204	9.3	153	11.0
0.7001~0.8000	386	7.9	176	8.0	120	8.6
0.8001~0.9000	131	2.7	54	2.5	72	5.2
0.9001~1.0000	77	1.6	34	1.6	0	0.0
Total	4,874	100.0%	2,192	100.0%	1,397	100.0
Land-use types	Average	EI Range	Average	EI Range	Average	EI Range
Four	0.5999	0.1010~0.9930	0.5996	0.0968~0.9930	0.5908	0.0270~0.9900
Three	0.4356	0.0098~0.7923	0.4356	0.0098~0.7923	0.7919	0.0209~0.7919
Two	0.2504	0.0022~0.5000	0.2464	0.0022~0.5000	0.2562	0.0033~0.5000
One or Zero	0.0000	0.0000~0.0000	0.0000	0.0000~0.0000	0.0000	0.0000~0.0000
Total	0.3832		0.3883		0.3883	

* 199 TSZs are used for single-family residential land in the D-FW area, 60 TSZs in Dallas county, and 28 TSZs in Tarrant county respectively.

Figure 5.1 portrays land-use balance (entropy index) at TSZ level. TSZs with relatively 'high' land-use balance (> 0.6000, pink- or brown-colored TSZs) are spatially heterogeneous among the four land-use types, while TSZs with white, light-blue, and light green colors (less than 0.4 in entropy index) have reduced land-use types (mostly singleton or two type uses). In the latter case, land-use is likely to be somewhat separated or isolated from other land-uses. As shown below, high-entropy TSZs are observed mainly along major arterials in the D-FW metropolitan area. It is thought that land along major roadways or within proximity of these roads has been developed for a variety of land-uses because of the convenient automobile access they provide.

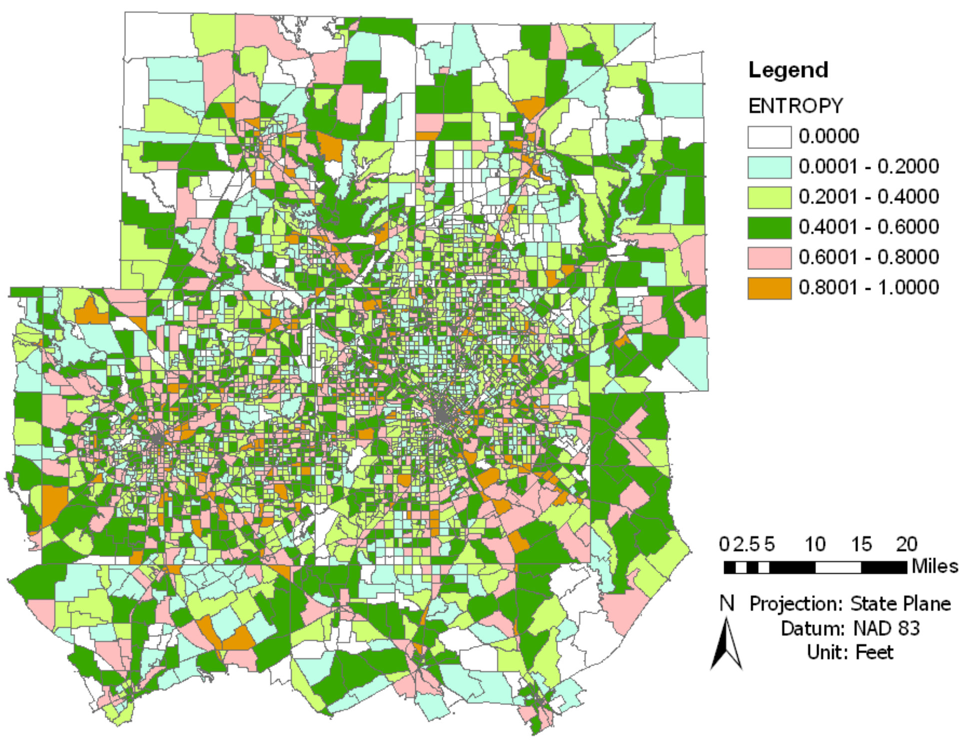


Figure 5.1 Land-Use Balance (Entropy Index) at TSZ Level in the D-FW Area

Mode Choice Results

MNL Results and Choice Elasticities

For each trip purpose, such as home-based work (HBW) trips, home-based other (HBO) trips, and non-home based (NHB) trips, a sufficient sample size was available to estimate the multinomial logit (MNL) choice models across the four alternatives of walk/bike, transit, drive-alone, and shared-ride. The reference mode is transit, and, hence each coefficient for constant term on the utility function should be interpreted with reference to this category. Normally, the MNL models are sensitive to generic or alternative-specific specifications. While generic coefficients were assumed to have the same influence on mode options, alternative-specific coefficients were estimated under an assumption that individual travelers are influenced differently by the factors of different mode options. Results for the MNL models for HBW, HBO, and NHB trips appear in Table 5.3 and Table 5.4. Mode choice for each purpose was estimated in two model schemes (base model vs. extended model). Both tables display the estimated coefficients, t-values, goodness-of-fit, model improvement test, and choice prediction.

The estimation results under the extended models maintain the same patterns as the base models for socio-economic variables, but there are variations for mode attributes (travel times). Walk time for HBW, HBO, and NHB trips in the extended model was significant at 5% level. Transit time and auto time were significant at 5% and 10% for HBW trips, but not for both HBO and NHB trips. As suggested by theory, an increase in travel time decreases each probability to choose each mode, but people are less likely to walk to work than they are to take transit, drive or carpool. Since most work trips occur during a peak period, people are faced with higher time costs when choosing a mode with less mobility. Travel time is an important factor in influencing a decision to commute to work but not a significant factor in decreasing the choice of transit or automobile in non-work trips. From the results of mode choice estimation for

Table 5.3 D-FW Multinomial Logit Models of Mode Choice for Home-Based Trips

	Work Trips				Non-Work Trips			
	Base Model		Extended Model		Base Model		Extended Model	
Variables	Coef.	*t*	Coef.	*t*	Coef.	*t*	Coef.	*t*
Constant (W/B)	2.1086	**7.24**	2.9054	**9.30**	4.8686	**20.14**	4.6780	**16.62**
Constant (D)	1.0957	**3.32**	28.5118	**4.32**	2.1947	**6.67**	11.4340	**4.13**
Constant (S)	-1.8331	**-5.78**	25.5791	**3.88**	1.2085	**3.72**	10.4452	**3.78**
Walk/bike time (W/B)	-0.0532	**-6.63**	-0.0566	**-6.88**	-0.0622	**-9.95**	-0.0610	**-9.02**
Transit time (T)	-0.0070	***-1.82***	-0.0159	**-2.35**	-0.0168	**-2.21**	-0.0159	-1.25
Auto time (D,S)	-0.0335	-1.54	-0.0455	***-1.76***	-0.0837	**-2.07**	-0.0490	-1.05
Age (W/B, T)	0.0239	**4.54**	0.0225	**3.83**	-0.0227	**-5.05**	-0.0245	**-5.35**
Sex (D, S)	0.3006	**2.26**	0.1388	0.93	-0.1784	-1.33	-0.1932	-1.41
HH Income (D,S)	0.0237	**6.83**	0.0308	**7.52**	0.0179	**5.56**	0.0184	**5.62**
HH Size (T,S)	0.2384	**7.29**	0.2286	**6.80**	0.3873	**18.44**	0.3891	**18.49**
#. of Workers (D)	0.0081	0.15	-0.0094	-0.17	0.1493	**5.28**	0.1508	**5.33**
Vehicles in HH (D,S)	1.3388	**13.02**	1.2200	**10.64**	0.8609	**8.77**	0.8030	**8.10**
MF Housing (W/B,T)	0.5389	**3.61**	0.3906	**2.21**	0.9642	**6.55**	1.0197	**6.34**
Pop density at O (W/B,T)			-0.0268	***-1.68***			-0.0168	-1.31
Pop density at D (D, S)			0.0327	***1.87***			0.0223	1.59
Job density at O (W/B, T)			0.0020	**6.25**			0.0024	1.06
Job density at D (D, S)			-0.0020	**-6.06**			-0.0056	**-6.56**
Resid. share at O (W/B,T)			0.0000	0.00			0.0003	0.09
Resid. share at D (D,S)			0.0013	0.42			-0.0030	-0.84
Com. share at O(W/B, T)			0.0181	**5.17**			-0.0091	-1.57
Com. share at D (D,S)			-0.0090	**-2.70**			-0.0008	-0.16
Accessibility at D (D,S)			-0.2748	**-4.07**			-0.1054	**-3.45**
Entropy index at D (D,S)			-0.2324	-0.77			0.2636	0.95
Sample size	4,672		4,672		7,112		7,112	
LLF($\mathscr{L}$) at converge	-2,275.3		-2109.1		-5480.5		-5,438.8	
Goodness-of-fit : ρ^2, $\bar{\rho}^2$	0.6487, 0.6467		0.6744, 0.6709		0.4441, 0.4428		0.4483, 0.4460	
Model improvement test: -2[$\mathscr{L}$ (B)- $\mathscr{L}$ (E)] (in extended models)	χ^2 = 332.40, df = 10, Prob. < 0.001				χ^2 = 83.40, df = 10, Prob. < 0.001			
	W/B T (Bus) D S Total				W/B T(Bus) D S Total			
Actual choice share	1.2% 5.9% 83.8% 9.1% 100.0%				4.0% 0.9% 49.0% 46.1%, 100%			
% correctly predicted	0.2% 2.7% 83.4% 0.0% 86.3%				1.7% 0.1% 33.2% 23.4%, 58.4%			

a. W/B = walk/bike, T = bus, D = driving-alone, and S = shared-ride

b. Parenthesis in variable column indicates the mode(s) to which the variable is specified.

c. *t*-values in bold-face are significant at 95% level and in italic bold-faces at 90% respectively.

travel times, the coefficient estimate for transit time appears to be underestimated, compared to the coefficient estimate for auto mode in HBW, HBO, and NHB trips. It is thought that transit operational performance of 13.8 mph for non-rail is relatively low in the Dallas-Fort Worth area where many freeways have been developed to maintain high speed. Also unitary application of maximum walk access time (20 minutes) for different trip lengths is thought to contribute to underestimation.

The effects of socioeconomic characteristics on mode choice are in accord with the literature. Personal characteristics were typical confounding variables: sex was not a significant factor in the extended models, while age is positively associated with non-automobile choice in work trips and negatively associated with automobile choice in non-work trips. Household income, household size, the number of workers in household, and the number of vehicles owned by household have statistically significant effects on the probability that a trip maker walks for non-work trips, takes a bus, drives alone, or shares a ride. In three samples, higher household income and the number of vehicles owned by a household are more likely to depend on automobiles (i.e., drive alone or share a ride) than they are for walking or taking a bus. Those who dwell in multi-family housing have a higher probability of walking in work or non-work trips. From this finding, multi-family housing location appears to have much to do with density development pattern which may encourage residents to walk more rather than drive.

The most robust effect of land-use as measured at destination by job density and regional accessibility is to decrease the probability of choosing automobiles in any type of trip. Conversely job density and regional accessibility at destination increase the chances that a trip-maker walks or takes a bus in any trip. The effect of land-use at trip origins stands out by job density and the commercial share of land for only the work-trip sample, increasing the probability of walking in work or taking a bus to work. However, the entropy index, a typical measure for a variety of land-uses, does not explain mode choice in the D-FW metropolitan area.

Table 5.4 D-FW Multinomial Logit Model of Mode Choice for Non-Home-Based Trips

	Non-Home Based Trips			
	Base Model		Extended Model	
Variables	Coef.	*t*	Coef.	*t*
Constant (W/B)	5.7015	**19.54**	4.5529	**13.72**
Constant (D)	2.3725	**5.66**	28.2657	**2.87**
Constant (S)	1.6491	**4.01**	27.5586	**2.80**
Walk/bike time (W/B)	-0.1201	**-13.69**	-0.0913	**-9.65**
Transit time (T)	-0.0283	**-3.69**	-0.0233	-1.64
Auto time (D,S)	-0.1540	**-3.86**	-0.0816	-1.58
Age (W/B, T)	-0.0035	-0.58	-0.0090	-1.29
Sex (D, S)	-0.3240	**-2.13**	-0.2971	***-1.73***
HH Income (D,S)	0.0128	**4.27**	0.0160	**4.60**
HH Size (T,S)	0.1552	**4.94**	0.1530	**4.85**
#. of Workers (D)	0.1744	**3.49**	0.1810	**3.59**
Vehicles in HH (D,S)	0.7059	**6.58**	0.7741	**6.35**
MF Housing (W/B,T)	-0.1487	-0.78	-0.0771	-0.36
Population density at origin (W/B,T)			-0.0217	-1.02
Population density at destination (D, S)			0.0291	1.38
Job density at origin (W/B, T)			0.0031	**6.95**
Job density at destination (D, S)			-0.0033	**-6.16**
Residential share of land (W/B,T)			-0.0013	-0.35
Residential share of land (D,S)			0.0025	0.64
Commercial share of land (W/B, T)			-0.0011	-0.29
Commercial share of land (D,S)			-0.0008	-0.21
Regional accessibility at destination (D,S)			-0.2734	**-2.75**
Entropy index at destination (D,S)			0.4891	1.39
Number of Observations	3,354		3,354	
LLF(L) at converge	-2,637.8		-2,522.9	
Goodness-of-fit : ρ^2 , $\bar{\rho}^2$	0.4327, 0.4299		0.4574, 0.4524	
Model improvement test: -2[L (*B*)- L (*E*)] (in extended models)	χ^2 = 229.80, *df* = 10, Prob. < 0.001			

	W/B	T (Bus)	D	S	Total
Actual choice share	8.4%	1.4%	57.9%	32.3%	100.0%
% correctly predicted	6.0%	0.2%	56.3%	0.5%	63.0%

a. W/B = walk/bike, T = bus, D = driving-alone, and S = shared-ride

b. Parenthesis in variable column indicates the mode(s) to which the variable is specified.

c. *t*-values in bold-face are significant at 95% level and in italic bold-faces at 90% respectively.

A few resulting statistics such as like the goodness-of-fit (ρ^2), the χ^2 model improvement test statistic, and the percentage-correctly-predicted (%CP), support that expended models outperformed base models. As expected, the inclusion of land-use variables improved overall predictability in the extended models. Hence, land-use variables matter somewhat in explaining the travel decisions of mode options available to D-FW area residents.

Elasticity is used for measuring a choice probability in response to a change in some decision variable and, by definition, is the percentage change in one variable that is associated with 1% change in another variable. Table 5.5 shows the elasticities of mode choices that are associated with changes in selected land-use variables and travel time. Similar to the available literature, the travel time elasticities of mode options are higher

Table 5.5 Mode Choice Elasticities

Trips	Variables and Model Specifications*	Elasticities			
		W/B	T (Bus)	D	S
HBW	Travel time (min) [W/B], [T], [D, S]*	-13.4405	-0.9463	-0.1485	-0.8357
	Job density at origin [W/B, T]	0.0529	0.0032	-0.0449	-0.0048
	Job density at destination [D, S]	0.0619	0.0037	-0.0525	-0.0057
	Commercial share at origin [W/B, T]	0.2832	0.0169	-0.2403	-0.0259
	Commercial share at destination [D,S]	0.1551	0.0093	-0.1316	-0.0142
	Regional accessibility [D,S]	0.2633	0.0158	-0.2235	-0.0241
HBO	Travel time (minutes) [W/B], [D, S]	-10.6893	-1.2270	-0.3634	-0.3841
	Job density at destination [D, S]	0.0386	0.0004	-0.0197	-0.0185
	Regional accessibility [D, S]	0.0905	0.0009	-0.0462	-0.0435
NHB	Travel time (minutes) [W/B], [D, S]	-16.5216	-1.5400	-0.5259	-0.8448
	Job density at origin [W/B, T]	0.1150	0.0018	-0.0727	-0.0406
	Job density at destination [D, S]	0.1134	0.0018	-0.0716	-0.0400
	Regional accessibility [D, S]	0.2464	0.0038	-0.1557	-0.0869

* Brackets, [ ], indicate alternative-specific specifications, and elasticities are calculated at a weighted-average choice probability. For alternative-specific counterparts, the signs of estimated coefficients are reversed, and then the resulting elasticities are calculated.

than those of land-use variables. The effect of land-use on the decision to drive less, though statistically significant, is generally smaller in absolute magnitude (|elasticity| < 0.3) than travel time. For example, a 10% increase in the regional accessibility of automobile use reduces driving-alone by 2.2% in HBW trips, by 0.4% in HBO trips, and by 1.6% in NHB trips, respectively. A 10% increase in job density at destination increases the probabilities of walk/bike by 0.6%, 0.4%, and 1.1% in HBW, HBO, and NHB trips, respectively.

Results on Directed Graphs

The estimation of choice models depends on the multivariate distribution of variables for each utility function for all choices, while directed graphs rely on a multivariate distribution of variables from data. It is not known whether directed graphs can be constructed for a multinomial choice, but a binary choice with alternative-specific specification is likely to construct the directed graphs. For the directed graphs four choices (walk/bike, transit, drive-alone, and shared-ride) in the MNL models are grouped into two choices: *non-automobile* and *automobile*. The estimated results of binary logit are presented in Table 5.6 to compare with the directed graphs.

The directed graphs of mode choice are presented for three types of trips (HBW, HBO, and NHB). The analysis proceeds from the lower triangular form of a correlation matrix between each of fifteen variables: automobile choice (*AUTO*), travel time differential (*T_TIME*), household income (*INC*), household size (*HHSZ*), the number of workers in household (*WRKRS*), the number of vehicles owned by household

(*VEHNUM*); for trip ends of origins and destinations, population densities (*O_POPDEN*, *D_POPDEN*), job densities (*O_JOBDEN*, *D_JOBDEN*), residential shares of land (*O_RESID*, *D_RESID*), commercial shares of land (*O_COMM*, *D_COMM*), and regional accessibility at destination (*D_ACCESS*). The unconditional correlations between each of the variables are summarized in Appendix A4. Such a correlation matrix provides the starting point for the analysis of causation using the directed graphs. As discussed in the previous chapter, the TETRAD II algorithm removes edges by taking into account the unconditional and conditional correlations between variables. Analysis begins by imposing two constraints on the orderings. First, socioeconomic characteristics precede

Table 5.6 The Estimation Results of Binary Logit (Auto vs. Non-auto) Models

Variables	HBW		HBO		NHB	
	Coef.	*T*	Coef.	*t*	Coef.	*t*
Constant	28.7718	**4.35**	9.6416	**3.62**	33.3258	**3.54**
Travel time diff. (walk – driving)	-0.0006	-1.63	-0.0102	**-8.14**	-0.0107	**-8.75**
Household income	0.0321	**7.99**	0.0243	**7.73**	0.0185	**5.77**
Household size	-0.2825	**-5.34**	-0.1116	**-2.48**	-0.1168	***-1.85***
Number of workers in household	0.0250	0.21	-0.2310	**-2.78**	-0.4112	**-3.28**
Number of vehicles in household	1.3368	**11.47**	1.0781	**10.96**	1.0612	**8.23**
Pop density at origin	0.0107	0.74	-0.0176	**-1.98**	0.0268	1.35
Pop density at destination	0.0191	1.18	-0.0095	-0.80	0.0319	1.52
Job density at origin	-0.0018	**-5.91**	-0.0027	***-1.68***	-0.0029	**-6.78**
Job density at destination	-0.0019	**-5.82**	-0.0047	**-4.91**	-0.0041	**-7.81**
Residential share at origin	0.0008	0.26	0.0012	0.35	0.0008	0.25
Residential share at destination	0.0014	0.49	-0.0022	-0.67	0.0019	0.57
Commercial share at origin	-0.0185	**-5.45**	0.0040	0.79	-0.0013	-0.42
Commercial share at destination	-0.0094	**-2.90**	-0.0022	-0.48	-0.0045	-1.42
Reg. accessibility at destination	-0.2901	**-4.28**	-0.1035	**-3.53**	-0.3402	**-3.57**
Number of Observations	4,672		7,112		3,354	
Goodness-of-fit : $\rho^2, \bar{\rho}^2$	0.7828	0.7781	0.8028	0.7997	0.7385	0.7316
	Nonauto Auto Total		Nonauto Auto Total		Nonauto Auto Total	
Actual choice share	332 4,340 4,672		350 6,762 7,112		329 3,025 3,354	
% correctly predicted	3.0% 92.4% 95.4%		1.1% 94.8% 95.9%		4.5% 89.7% 94.2%	

a. *Non-auto* is reference mode in the binary logit, and the estimated coefficients should be interpreted with reference to non-auto mode. Only travel time differential is specified as generic specific

b. *t*-values in bold-face are significant at 5% significance level, and in italic bold-face at 10% level.

land-use variables and travel time. Second, the land-use variables at origin do not affect land-use variables at destination, and vice versa. A 1 % significance level is used for removing edge at sample size (e.g., 0.1 at sample sizes between 100 and 300) as suggested by Spirtes, et al. (2000).

As shown from a directed graph in Figure 5.2, the choice of automobile is explained for HBW trips by household income (*INC*: +), the number of vehicles in a household (*VEHNUM*: +), job densities at both origins (*O_JOBDEN*: −) and destinations (*D_JOBDEN*: −), the commercial shares of land at both origins (*O_COMM*: −) and destinations (*D_COMM*: −), and regional accessibility at destination (*D_ACCESS*: −). The directed arrows with these immediate causal factors found at 1% significance are causally connected to automobile choice for home-based work trips. The directed graph contains colliders (*T_TIME*, *O_RESID*, and *D_RESID*) in which information flows running from other variables collide. There are also the precedent (parent) variables (*O_POPDEN*, *D_POPDEN*, and *WRKRS*) of colliders of which the causal flow is blocked by colliders. Focusing on travel time and land-use variables shows that travel time and residential shares at both trip ends absorb causal flow running from their parent variables but never transmit causal flow to child variables. Population densities at both ends are parent variables that open up the path to a collider ‘travel time’. Thus, travel time, residential shares, and population densities are not causally connected to the decision of choosing an automobile. The direct causes found in Figure 5.2 are consistent with the estimated coefficients of significance in the binary logit model for HBW trips.

Household size (*HHSZ*) is causally connected to automobile choice running through the regional accessibility. This causal flow, however, is quite difficult to interpret in terms of how household size works its way through accessibility. The bidirected edges between population density at destination and each land-use variable (regional accessibility, and residential and commercial shares at destination), and between residential share at destination and regional accessibility are difficult to understand but suggest the existence of a latent variable between two variables. Potential possibilities could be considered as zoning practices and regulations to qualitatively measure land-use. One interesting result in HBW trips is that travel time is not a direct cause of automobile choice. This result may be understood by the nature of work trips in auto-dependent cities. Work trips are especially associated with earning income and potential income sources. A rational trip-maker who owns a car will drive with the expectation of higher income, and hence travel time by driving is less likely to matter for his or her work trip. Another finding is that the number of vehicles owned by a household is causally connected to several land-use variables (population densities at both ends, job densities at both ends, and regional accessibility). A causal connection between the number of vehicles and the land-use variables is typically observed in low-density and auto-dependent cities.

Figures 5.3 and 5.4 present the directed graphs at the 1% significance level for HBO and NHB trips, respectively. Data on HBO trips are causally explained by travel time (-), household income (+), the number of vehicles (+), job density at origin (-), job density at destination (-), population density at origin (-), and regional accessibility (-).

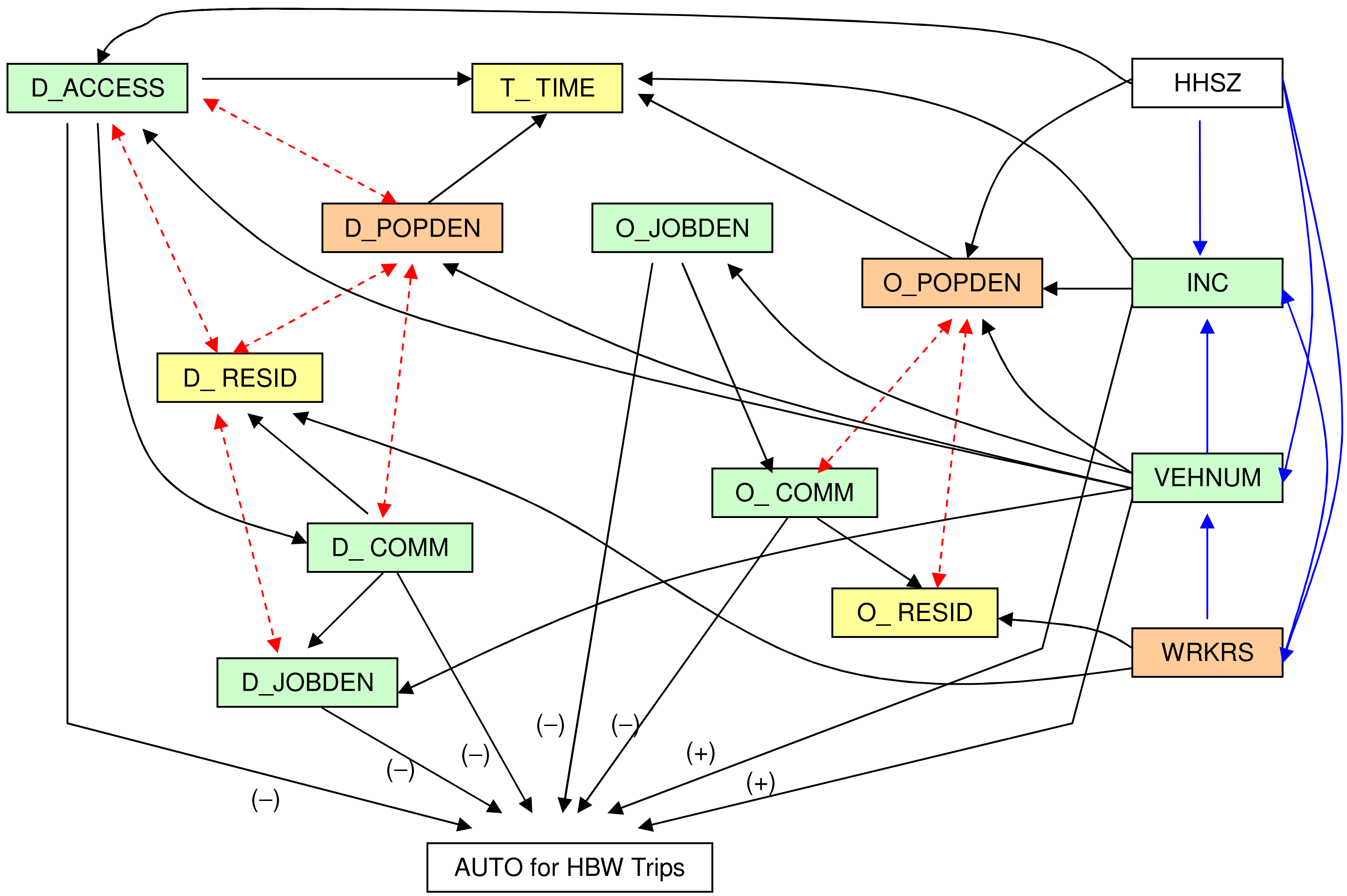


Figure 5.2 Directed Graphs from Data on Binary Choice (Auto vs. Non-auto) for HBW Trips at 1% Significance Level (Dotted Edges with Arrows Indicate a Need for a Common Cause between Two Variables).

Regarding travel time, the direction of causal flow runs from household size, population densities at trip ends, and regional accessibility to travel time, which opens up a path to automobile choice. Travel time is a primary causative factor of automobile choice as well as being influenced by some land-use variables (*O_POPDEN*, *D_POPDEN*, *D_ACCESS*). The directed graph for NHB trips given in Figure 5.4 may well be reflective of similar causations of travel time (*T_TIME*). Travel time is a direct cause of automobile choice for NHB trips. Regarding travel time, the direction of causal flow runs from regional accessibility (*D_ACCESS*), population density at origin (*O_POPDEN*), and commercial share of land at origin (*O_COMM*) to travel time.

Four bidirected edges between each pair of land-use variables are identified in the directed graph for HBO trips in Figure 5.3, while a directed graph for NHB trips in Figure 5.4 presents a bidirected edge between commercial share at origin (*O_COMM*) and residential share at origin (*D_RESID*). As suggested previously, zoning practices and regulations, and pedestrian-friendly environments can be candidates for the appropriate latent variables between these bidirected edges.

There is no collider to sink any causal paths in the directed graph given in Figure 5.3, but non-directed edges are present between job density at origin (*O_POPDEN*) and residential share at origin (*O_RESID*). Let us presume that non-directed edges do not have much to do with causal connections between the variables (albeit it is more or less ambiguous). Thus, the residential share of land-use at origin (*O_RESID*) is virtually a sink where information flow stops. Once backing to its parent, causal flow running from the commercial share at origin (*O_COMM*) is blocked by the sink child, residential share at origin (*O_RESID*). Both variables are not causally connected to the choice of an automobile for non-work trips. This intuition may be applied to model specification incorporating the appropriate variables into an empirical model to hold the fundamental pattern of causality. In contrast, the directed graph for NHB trips given in Figure 5.4 displays a collider (*O_RESID*).

Look at the colliders and collider-blocked parent variables identified in the directed graphs. These variables may be candidates for irrelevant variables in empirical models. The identification of irrelevant variables can address somewhat an inconsistency between the direct causes of directed graph and the estimated significant coefficients of a regression model. These variables can be excluded from alternative regression models. For example, three colliders (*T_TIME*, *O_RESID, D_RESID*) and three colliders' parents (*O_POPDEN*, *D_POPDEN*, *WRKRS*) may be dropped off for a parsimonious model for HBW trips. Two assumed colliders (*O_RESID, D_COMM*) in HBO trips and a collider (*O_RESID*) in NHB trips can be removed respectively for more parsimonious models.

The entire set of variables used for the directed graphs numbered fifteen because of the limitation of the number on variables handled in TETRAD II. Some variables of socio-demographics (i.e., age, sex, and multi-family housing) were not initially included in the directed graphs, while household characteristics (household size, number of vehicles, household income, and number of workers) included in the directed graphs resulted in undirected edges. The undirected edges among household characteristics did not characterize the connection between probability and causality. These undirected edges in three samples were directed on the basis of author's judgment (blue-colored solid edges): household size as source is a direct common cause to increase the number of workers, vehicle ownership, and household income, the number of workers in household is a direct common cause of number of vehicles owned by household, and more household income, and the vehicle ownership is a cause of more household income.[24] Land-use balance (entropy index) was not included in all the directed graphs models based on both the statistical test results of binary logit and no edge found between the entropy index and the other variables of the system at 1% and 5% significance.

---

[24] According to Cervero et al. (2002), car ownership significantly increased the odds that someone switched welfare-to-work. It is inferred further that car ownership results in income increase through job accessibility

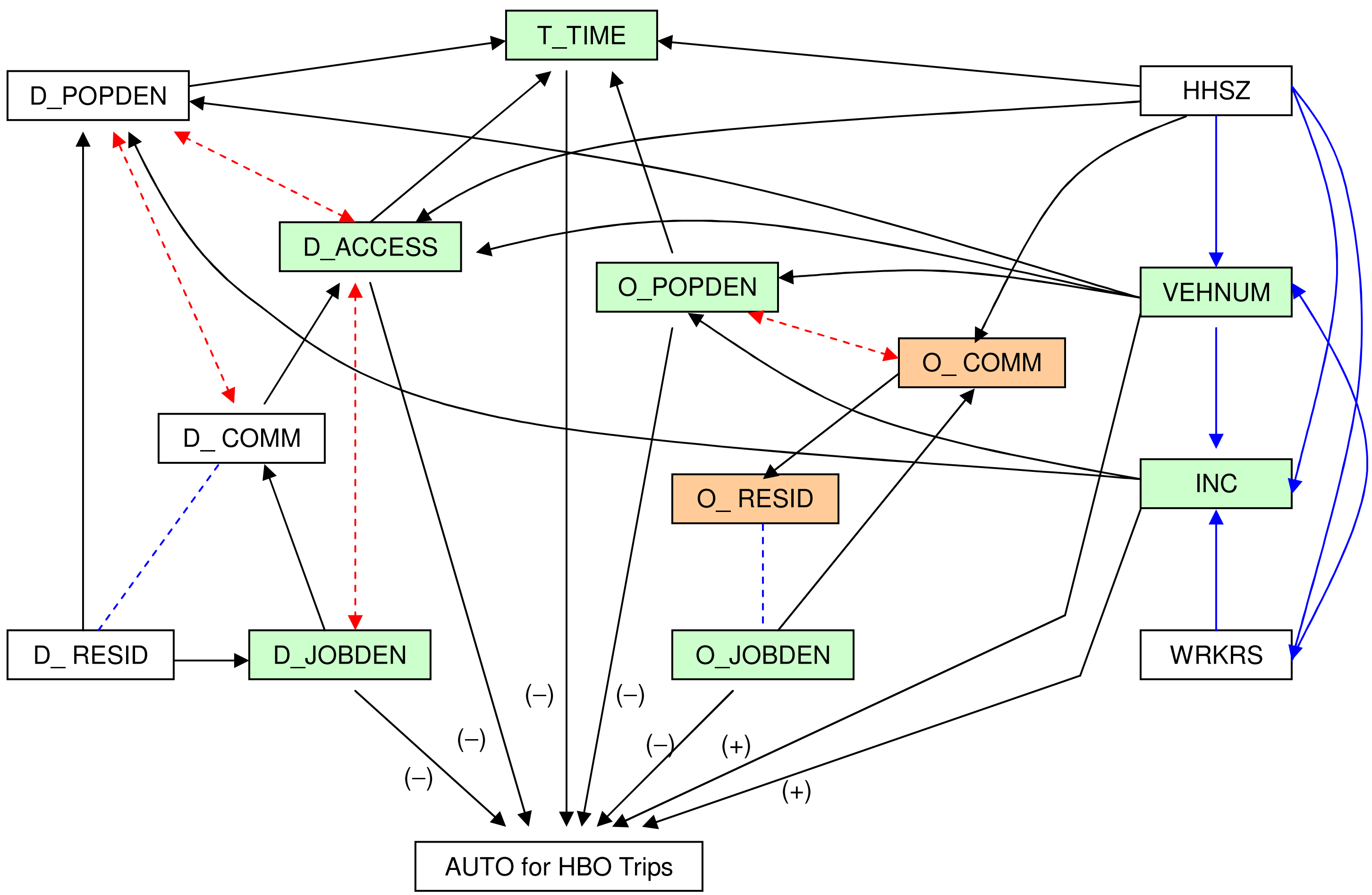


Figure 5.3 Directed Graphs from Data on Binary Choice (Auto vs. Non-auto) for HBO Trips at 1% Significance Level (Dotted Edges with Arrows Indicate a Need for a Common Cause between Two Variables).

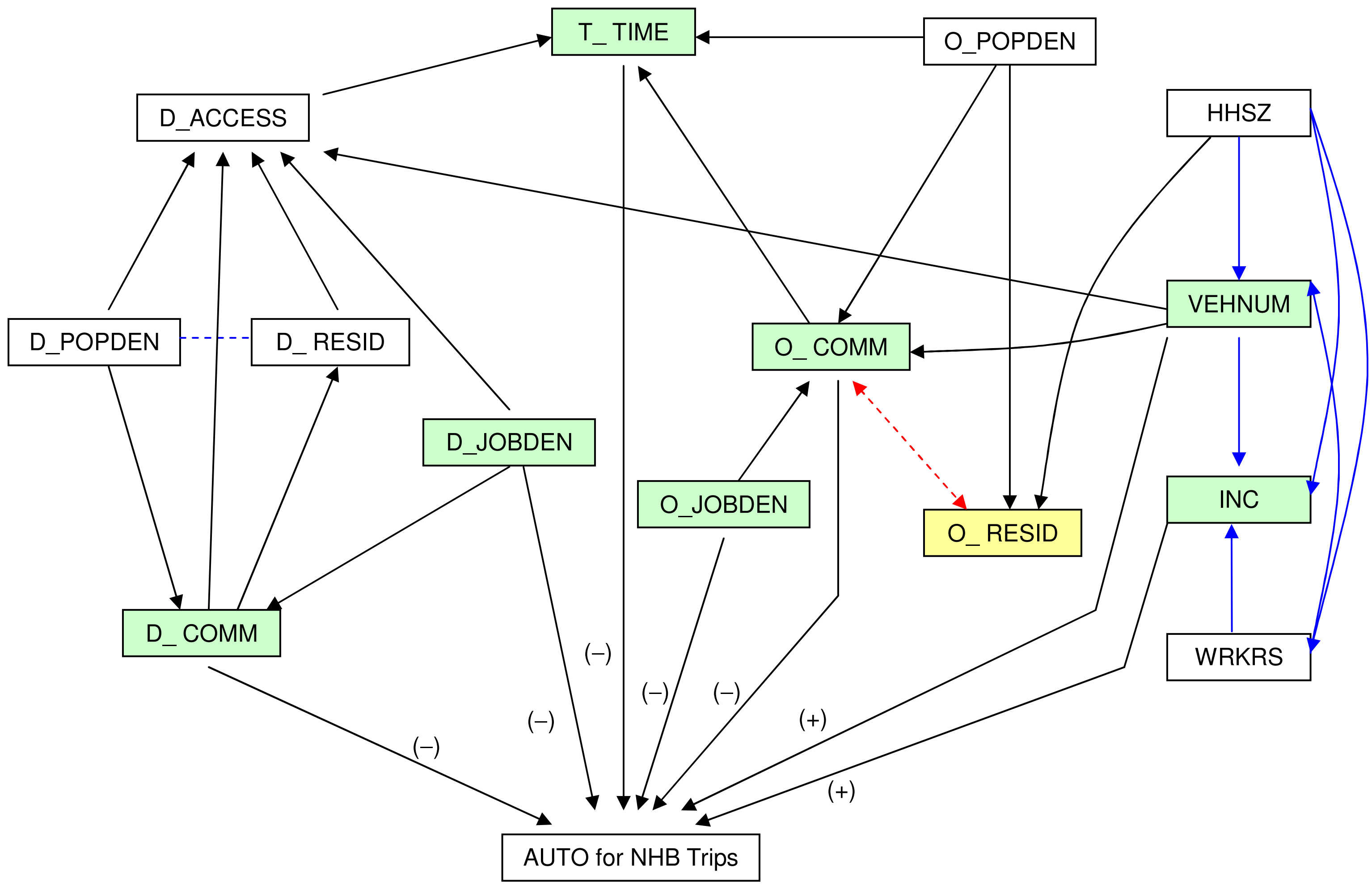


Figure 5.4 Directed Graphs from Data on Binary Choice (Auto vs. Non-auto) for NHB Trips at 1% Significance Level (Dotted Edges with an Arrow Indicate a Need for a Common Cause between Two Variables).

Trip Frequency Results

The ordered logit results of household trip frequency in work and non-work automobile trips are presented Table 5.7. Explanatory variables included in the ordered logit models are the generalized cost per trip, household characteristics (household income, household size, number of workers, number of vehicles), and land-use variables at trip origins (population density, job density, residential share of land, commercial share of land, and regional accessibility, and entropy index). The results of extended models were compared to the base models by testing the influence of a set of land-use variables on household trip generation. That is to say, a likelihood ratio (LR) test was performed to assess the incremental contribution of the inclusion of land-use variables. The LR test results support the fact that models slightly increase the ability of prediction (goodness-of-fit was increased by 10.3% in work trips and by 8.2% in non-work trips) for the household trip rates when land-use variables are included.

The generalized cost computed per trip is significant at the 1% level in both samples. As expected, higher generalized costs reduce household trip generation. The inclusion of price variables (i.e., generalized cost) improve the model, fitting better than the base model otherwise does, although the LR test result is not presented here. Household income is significant at the 5% level for non-work trips and at the 10% level for work trips, contributing to the generation of both. Household size appears to matter in non-work trips, but does not in work trips based on the 1% statistical significance level. As the number of workers increases in a household, the work trips are significantly positive at the 1% level, but not significant for non-work trips. Both household size and the number of workers in a household, have the opposite effect on the trip rates in the samples of work and non-work trips. The number of vehicles in a household is positively associated with both types of trip.

Table 5.7 Household Auto Trip Frequency Models for Home-Based Trips

Variables	Work Trips				Non-Work Trips			
	Base Model		Extended Model		Base Model		Extended Model	
	Coef.	z	Coef	z	Coef.	z	Coef.	Z
Gen. cost ($ / trip)	-0.1346	**-5.64**	-0.1406	**-5.64**	-0.1636	**-11.55**	-0.1638	**-11.41**
HH income in $1K	0.0034	**2.08**	0.0031	***1.83***	0.0067	**5.07**	0.0074	**5.43**
Household size	-0.0745	-1.52	-0.0541	-1.10	0.3076	**8.35**	0.3016	**8.10**
Number of workers	1.1420	**12.58**	1.1621	**12.63**	-0.0731	-1.28	0.0122	0.66
Number of vehicles	0.1366	**2.18**	0.1706	**2.67**	0.2059	**4.14**	0.1917	**3.80**
Pop. density at O			-0.0074	-0.66			0.0050	0.53
Job density at O			-0.0022	-0.71			-0.0013	-0.63
% Resid. use at O			0.0095	**2.46**			0.0018	0.61
% Comm. Use at O			0.0105	***1.92***			0.0036	0.79
Reg. accessibility			10.7086	**4.62**			-5.6185	**-5.50**
Entropy index			0.3655	1.02			-0.3152	-1.10
Cutoff 1	-2.8541	**-21.73**	-16.6779	**-7.21**	-3.8699	**-23.14**	1.1620	1.14
Cutoff 2	-5.2911	**-14.33**	-14.2009	**-6.18**	-2.2870	**-15.93**	2.7642	**2.76**
Cutoff 3					-0.9199	**-6.84**	4.1493	**4.13**
Number of Obs.	1,955		1,955		2,072		2,072	
LLF($\mathscr{L}$) at converge	-1,192.17		-1,175.26		-2,189.71		-2,171.66	
Goodness-of-fit: ρ^2	0.1213		0.1338		0.0840		0.0915	
Model improvement test: -2[$\mathscr{L}(B)$- $\mathscr{L}(E)$]	χ^2 = 33.82, df = 6, Prob. < 0.001				χ^2 = 36.10, df = 6, Prob. < 0.001			

a. z-values in bold-face are significant at 95% level, and in italic bold-face are significant at 90 %.

Among the land-use variables, the residential share of land at origin (or household location) and regional accessibility are significantly positive with work trips. The positive impact of residential share on trip rates can be construed from the status quo of land-use in the D-FW area where low-density and single-family residential development has been dominant due to the high increase in the rate of population, and major employment centers have been contiguous to major highways and arterials.

Automobile-based regional accessibility, computed in gravity formula, evaluates the extent of opportunities (normally referred as to jobs) to a specific place from other locations given travel time. Such accessibility can be seen from the possibility of expected income at the expense of travel time to get to jobs. In this context, while work trips linked to higher expected earnings are positively associated with regional accessibility, non-work trips mainly oriented to consumption or non-work activities are decreased by an increase in such accessibility. That is to say, regional accessibility induces people to drive more to workplaces, while reducing trips for people driving to non-work activities as shown in Table 5.7.

Directed Graphs on Trip Frequency

To search the statistical causal models of household trip frequency in TETRAD II, the lower triangular correlation matrix was made up of twelve variables (see Appendix A5): household trip frequency (*FREQ*), generalized cost per trip (*G_COST*), household income (*INC*), household size (*HHSZ*), the number of workers in a household (*WRKRS*), the number of vehicles owned by a household (*VEHNUM*); for trip origins, population density (*POPDEN*), job density (*JOBDEN*), residential share of land (*%RESID*), commercial share of land (*%COMM*), regional accessibility (*ACCESS*), and entropy index (ENTROPY). Figure 5.5 and 5.6 show the directed graphs of the models for HBW and HBO trips, respectively, at the 1% significance level. A restriction imposed on the construction of directed graphs is that socioeconomic variables are the only causes of the generalized cost and land-use, but the opposite case never happens. The directed edges (i.e., the connection between probability and causality) were not found between household characteristics. Author's judgment was made to direct the undirected edges (as blue-colored solid edges) for both work and non-work trips like mode choice models.

Household trip frequency for HBW trips in Figure 5.5 is explained by the immediate causal factors such as the generalized cost (*G_COST*: −), the number of workers in a household (*WRKRS*: +), residential share *(%RESID*: +), and regional accessibility *(ACCESS*: +). This directed graph results in a collider (%COMM) that absorbs causal information running from parents but never opens up to its child. Also this model includes the bidirected edges between population density and residential share and between commercial share and residential share, suggested by the existence of latent variables like zoning practices and regulations. One interesting finding is that population density and household income are causally connected to the trip frequency running through the generalized cost as the same causal paths were found in the directed graph on the mode choice for HBW trips. Another interesting finding is that land-use balance and job density are causally connected to trip frequency running through the residential share of land. From this causal path and the impact sign for trip frequency, it is inferred that the residential land-use oriented to low-density and single-family housing development in the D-FW area induces people to drive more. However, such a development pattern has the possibility of being affected by other causal factors such as land-use balance and job density.

The directed graph of household trip frequency for HBO trips indicates that the generalized cost (*G_COST*: −), the number of vehicles in a household (*VEHNUM*: +), household income (*INC*: +), household size (*HHSZ*: +), and regional accessibility (*ACCESS*: −) are the immediate causal factors at the 1% significance level as shown in Figure 5.6. Two colliders (% *RESID* and %*COMM*) are identified due to absorbing causal information running from other variables. In addition, two variables, *JOBDEN* and *ENTROPY,* are colliders' parents which are not causally connected to trip frequency because both colliders block causal flow. The finding is useful for constructing a statistical model with the fundamental pattern of causality. A regression model from the directed graph of household trip frequency for HBO trips is suggested to include the remaining variables after excluding two colliders and their colliders' parents. Also this model includes the bidirected edges in which the existence of latent variables is implied.

In Figure 5.6, the direction of causal flow to the generalized cost runs from population density, regional accessibility, and household size, and then the generalized cost opens up its path to trip frequency. These causal paths empirically support a hypothesis addressed in Chapter VI that land-use is a cause of price (travel time or generalized cost) which directly causes people to drive less. That is to say, travel time or travel cost affects travel patterns through land-use variables. It is noteworthy that population density and regional accessibility in HBO trip sample have causal flows from only socioeconomic variables such as household income and vehicle number to *POPDEN*, and household size and household income to *ACCESS*.

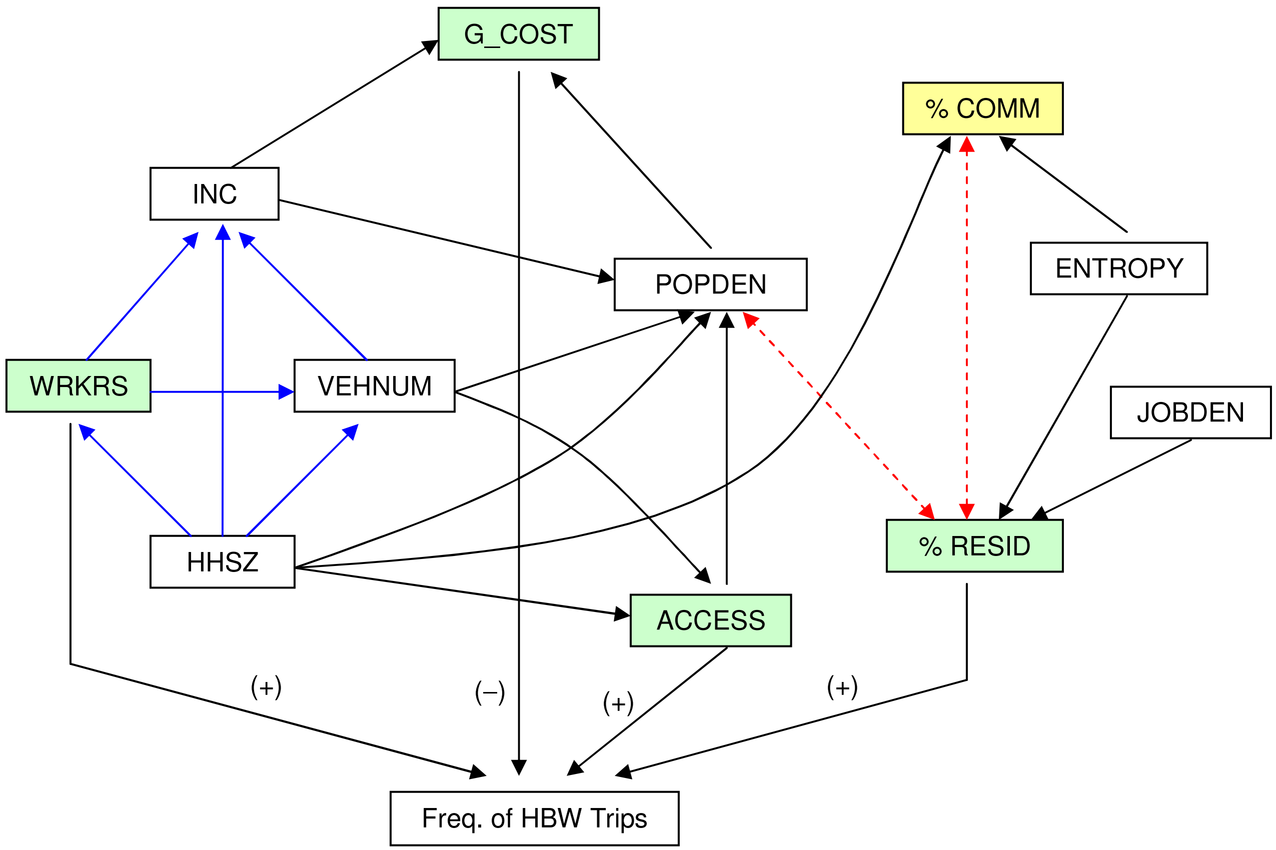


Figure 5.5 Directed Graphs on Household Trip Frequency Model for HBW Trips at 1% Significance Level (Dotted Edges with Arrows Indicate a Need for a Common Cause between Two Variables).

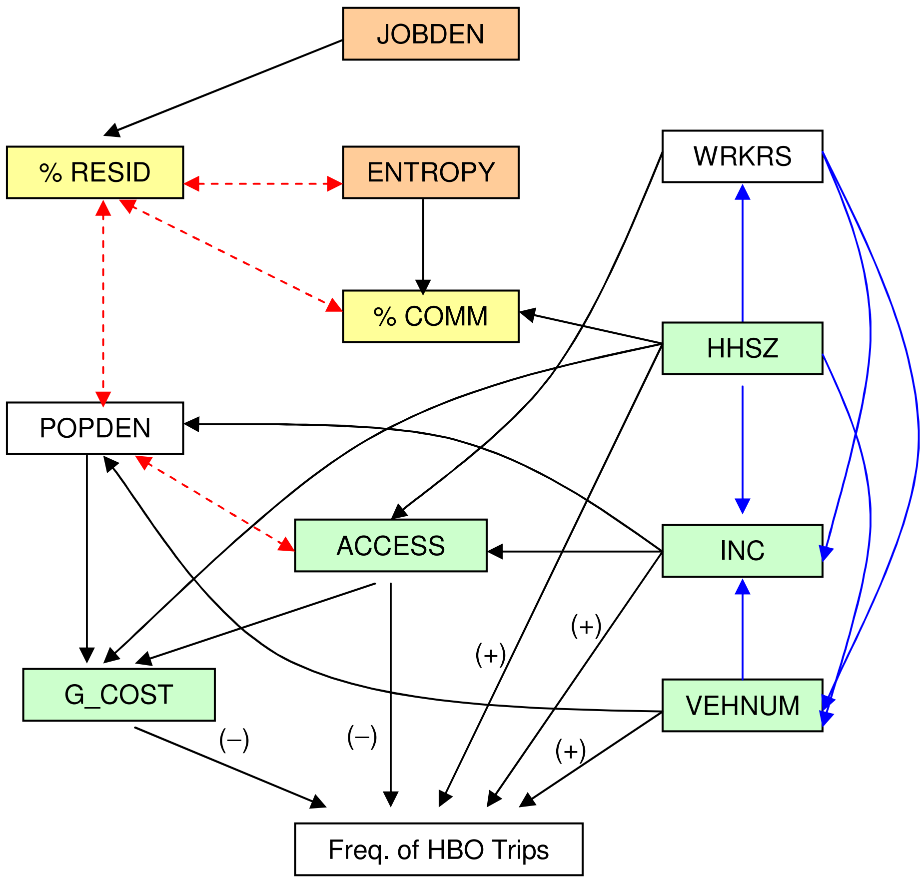


Figure 5.6 Directed Graphs on Household Trip Frequency Model for HBO Trips at 1% Significance Level (Dotted Edges with Arrows Indicates a Need for a Common Cause between Two Variables).

Household VMT Models

The VMT models were estimated by regressing the generalized cost, socioeconomic characteristics, and land-use variables on the vehicle miles traveled (VMT) in both trip samples of home-based work (HBW) and home-based non-work (HBO). The dependent variable was transformed to the logarithm of VMT to better fit the model and an ordinary least squares method was used for regression. Table 5.8 presents the results of household VMT models estimated for HBW trips and HBO trips. The results of extended models were compared to the base models by testing the influence of set of land-use variables. When generalized cost and socioeconomic variables (household income, household size, number of workers, and number of vehicles) are controlled, 54.9% of the variation in VMT is explained for work trips and 50.6% for non-work trips, respectively. The explained variations of the models appear high, compared to most of the previous studies in which mode attributes (travel time and travel cost) have never been considered. A great deal of variation is explained by the addition of generalized cost, even if the results of the controlled model were not presented. The control variables are consistent with the expected signs.

For the extended model of non-work trips, population density, and regional accessibility are significant predictors of the household VMT. For work trips, the residential share of land adds a marginal contribution to the prediction of VMT at the 1% significance level. That is to say, household VMT is positively influenced by an increase in residential land-use share which is mainly characterized by the residential-dominant development oriented to low-density-single-family housing pattern in the D-FW metropolitan area, appearing as isolated or segregated land-use. It is inferred from this finding that the low-density-single-family residential development pattern, easily observed in the D-FW area, is apparently not conductive to reducing automobile driving. In fact, dense population areas normally come with mixed land-use (i.e., vertically mixed with residences, retail stores, and offices) and are built for less automobile driving. It is also noted that regional accessibility is positively associated with work trips but negatively associated with non-work trips in explaining household VMT at the 1% significance level. From these findings, the land-use condition for less driving in the D-FW area is associated with an increase in population density, the reduction of residential land-use share (or an increase in land-use diversity), and an increase in accessibility.

Table 5.8 Household total VMT Models for Home-Based Trips

	Work Trips				Non-Work Trips			
	Base Model		Extended Model		Base Model		Extended Model	
Variables	Coef.	t	Coef.	t	Coef.	T	Coef.	T
Constant	4.7726	**68.32**	1.5489	**2.85**	2.9961	**53.89**	4.4165	**9.76**
Gen. cost (US$ /mile)	-1.3399	**-45.92**	-1.3362	**-45.90**	-7.0872	**-44.94**	-7.0403	**-44.61**
HH Income in 1000$	0.0029	**6.24**	0.0025	**5.29**	0.0008	1.45	0.0009	1.55
Household size	-0.0284	**-2.17**	-0.0257	**-1.98**	-0.0356	**-2.15**	-0.0432	**-2.60**
Number of workers	0.1684	**6.98**	0.1658	**6.69**	-0.0244	-0.93	0.0030	0.11
Number of vehicles	0.0677	**3.66**	0.0708	**3.84**	0.1148	**5.03**	0.1031	**4.47**
Pop. density at origin			-0.0072	**-2.51**			-0.0089	**-2.17**
Job density at origin			-0.0002	-0.45			0.0002	0.47
% Resid. use at origin			0.0044	**4.82**			-0.0005	-0.41
% Comm. use at origin			0.0001	0.07			-0.0022	-1.15
Regional accessibility			3.0315	**5.58**			-1.4408	**-3.12**
Entropy index			0.0835	0.91			0.1123	0.89
Number of Obs.	1,955		1,955		2,072		2,072	
Sum of squared error	871.8		841.9		1,625.5		1,608.1	
R^2, $\bar{R}^2$	0.5491, 0.5480		0.5646, 0.5621		0.5064, 0.5052		0.5115, 0.5190	
Model improvement test	F = 11.10, m = 6, df_E = 1,943 Prob. < 0.001				F = 3.67, m = 6, df_E = 2,060 Prob. < 0.01			

a. t-values in bold-face are significant at 95% level.

b. $F = \frac{(SSE_B - SSE_E)/m}{SSE_E / df_E}$, where m $= (df_E - df_B)$ and follows the F distribution with m and df_E.

The magnitude impact of land-use variables on household VMT is trivially small by the land-use elasticities of household VMT computed from the estimation results: a 10% increase in population density reduces household VMT by only 0.016% for work trips and 0.02% for non-work trips, and a 10% decrease in residential share (or increase in land-use diversity) resulting from the increase in land-use mix (commercial, industrial, and public purpose) leads to a 1% decrease in household VMT for work trips. A 10% increase in regional accessibility, on the other hand, induces one to drive more by 1% for work trips, but reduces household VMT by 0.5% for non-work trips.

Directed Graphs on Household VMT

Twelve variables were used for creating the lower triangular correlation matrices (Appendix A6). The only difference from the household trip frequency is the use of a log-transformed VMT variable instead of trip frequency. Other variables are the same: household vehicle miles traveled (*VMT)*, generalized cost per mile in a household (*G_COST*), household income (*INC*), household size (*HHSZ*), the number of workers in a household (*WRKRS*), the number of vehicles owned by a household (*VEHNUM*), for trip origins, population density (*POPDEN*), job density (*JOBDEN*), residential share of land (*%RESID*), commercial share of land (*%COMM*), regional accessibility (*ACCESS*), and the entropy index (ENTROPY). The directed graphs in Figure 5.7 and 5.8 were constructed for work trips and non-work trips respectively at the 1% significance level.

The same restriction imposed on trip frequency models was applied to the construction of directed acyclic graphs on household VMT: socioeconomic variables cause land-use and travel cost, but the opposite never occurs. Also author's judgment was made to direct the undirected edges (as blue-colored solid edges) among household characteristics which did not show the connection between probability and causality: household size as source is a direct common cause to increase the number of workers, vehicle ownership, and household income, the number of workers in household is a direct common cause of number of vehicles owned by household, and more household income, and the vehicle ownership is a cause of more household income.

According to Figure 5.7, household VMT in work trips is affected by direct causes such as the generalized cost (*G_COST*: −), the number of vehicles in a household (*VEHNUM*: +), household income (*INC*: +), household size (*HHSZ*: −), and regional accessibility (*ACCESS*: +). However, careful interpretation should be made for the number of workers in household (*WRKRS*). The variable in Table 5.8 is a significant coefficient when household characteristics are included in the VMT model of home-based work (HBW) trips, while the number of workers (*WRKRS*) is not the direct cause of household VMT. From such a gap, it is thought that a latent variable may exist between the numbers of workers and household VMT in work trips. Workers may be affected differently by the cost of automobile commuting trip or type of job (part- or full time). The identification of latent variable suggests the existence of another path. The directed graph identifies two colliders (*%RESID* and *%COMM*) which do not open up each path to child in question. Both colliders prevent the transmission of causal effects along each path from their parents (*JOBDEN*, *ENTROPY*) and then do not get to be causally connected to the VMT. That is to say, changes in either job density or entropy index may provoke changes in either residential share or commercial share, but will not provoke changes in household VMT. Thus, only population density among land-use variables is found as indirectly connected to household VMT in work trips.

Similar Findings were evidenced by the results of the directed graph for non-work trips (Figure 5.8). Compared to work trips, population density (*POPDEN*: −) was added to the direct causal factors of the VMT model, while household income (*INC*: +) and household size (*HHSZ*: −) were left out of the VMT model for non-work trips. Three colliders (*%RESID*, *% COMM*, *ENTROPY*) and a collider-parent (*JOBDEN*) were found in relation to directed path to VMT. These four variables are causally independent of household VMT in non-work trips. This means that changes in job density may provoke changes in residential share, but will not provoke any change in VMT because of the prevention of information transmission on three colliders. This insightful explanation supports the statistical insignificance of land-use variables resulting from the regression analysis of the VMT model for non-work trips, as presented in Table 5.8. Now, the entire set of land-use variables can be grouped into a causally-independent set (*%RESID*, *% COMM*, *ENTROPY*, *JOBDEN*) and directed-path group (POPDEN, ACCESS).

Bidirected edges (dashed line with arrows) between land-use variables suggest the existence of an unmeasured common cause (or latent variable). However, it is difficult to justify latent variable(s), although zoning practices or regulations were earlier suggested among land-use variables. The generalized cost affects household VMT and is also affected by accessibility, supporting the hypothesis addressed earlier. One thing that should be carefully addressed from this causal model is the number of vehicles owned by a household (*VEHNUM*), as automobile ownership is a direct cause to VMT as well as an indirect cause to population density, accessibility, and the generalized cost.

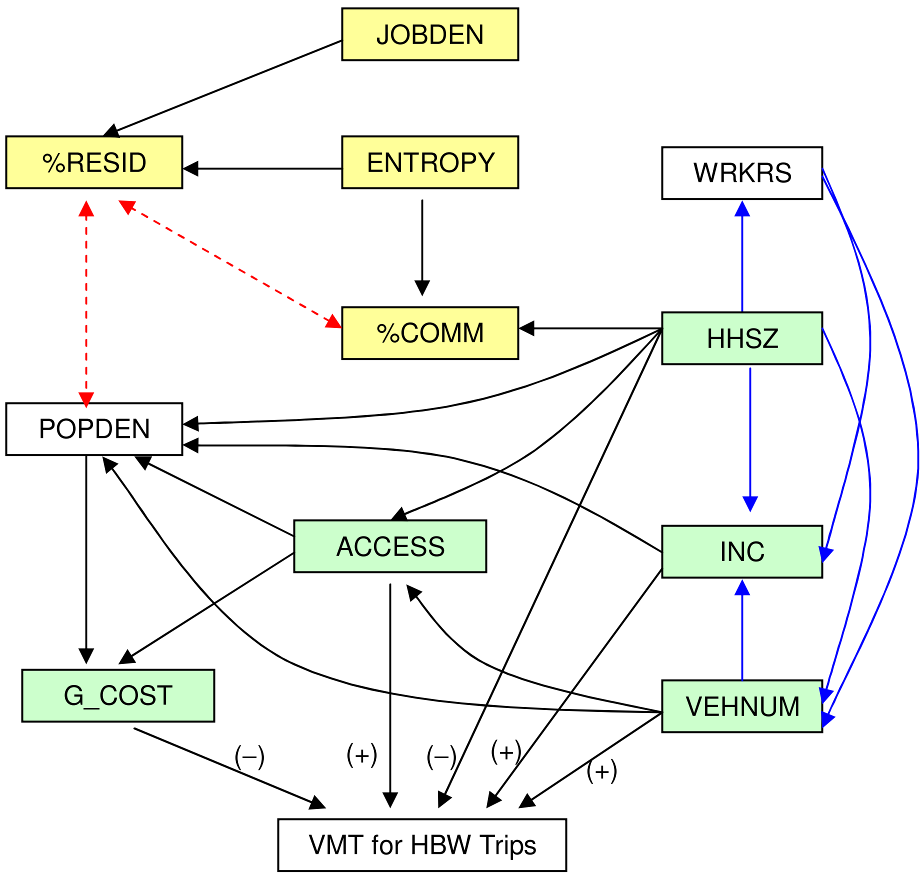


Figure 5.7 Directed Graphs on Household Total VMT Model for Work Trips at 1% Significance Level (Dotted Edges with Arrows Indicate a Need for a Common Cause between Two Variables).

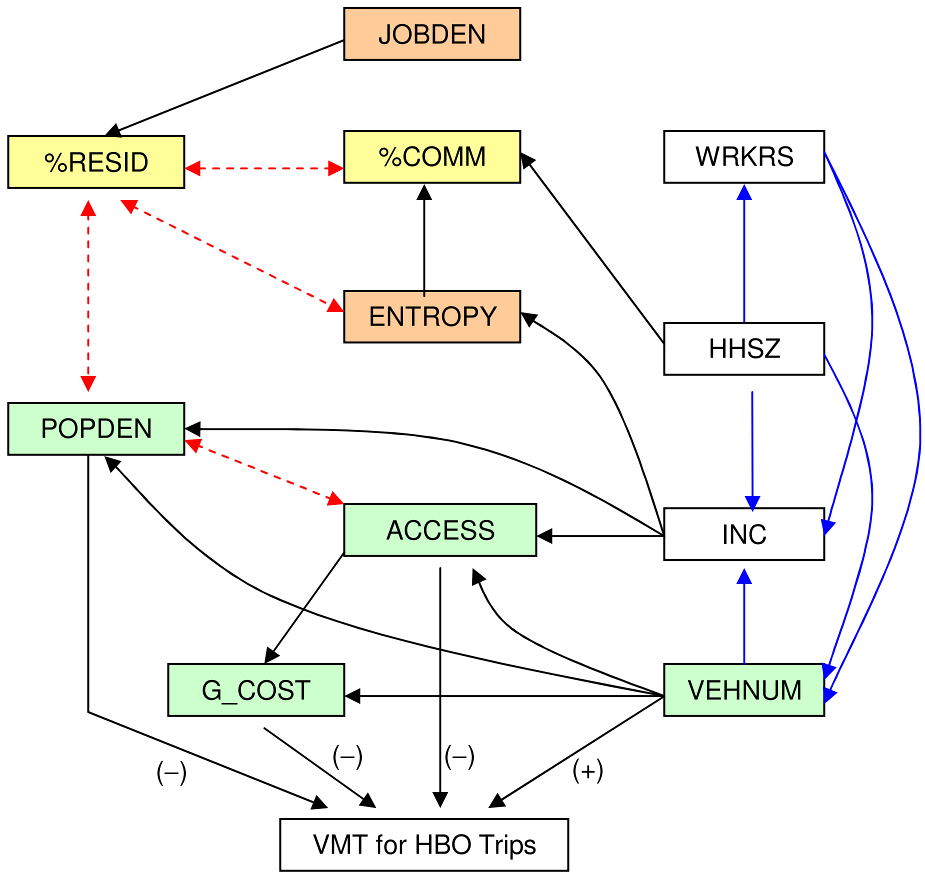


Figure 5.8 Directed Graphs on Household Total VMT Model for Non-Work Trips at 1% Significance Level (Dotted Edges with Arrow Indicate a Need for a Common Cause between Two Variables).

Logit Captivity Results

A main emphasis in investigating the potential captive factors of the logit captivity model for only home-based work (HBW) trips was placed on land-use variables, although some socio-demographic variables may be better candidates for captive factors. Initially, the same model specifications as the models in Table 5.3 and Table 5.4 were made for a comparison of estimation results. In addition, the models without transit time were specified in consideration of the underestimated coefficient for transit time. However, trip samples with other purposes (HBO and NHB) did not give convergent estimates under the given number of iterations for estimation. HBW trip sample resulted in convergent estimates for the model in which transit time was left out. Therefore, only the result of HBW trips was reported from the estimation of logit captivity model.

As shown in Table 5.9, two models of mode captivity were estimated and compared in an attempt to identify land-use factors which played an important role in choice set determination. *Captivity model I* has a notion that some trip-maker's choices are based on singleton choice sets in which alternatives are not considered, while *captivity model II* is further parameterized by including variables that could explain singleton choice sets. That is to say, the first model captures captivity behavior distinguished from choice behavior as each constant term for each singleton choice mode in transportation markets, while the second model assumes that the probability of a singleton choice set for driving-alone or shared-ride depends on the land-use variables which indicate residential land-use dominance.

Table 5.9 The Estimation Results of Logit Captivity Models for HBW Trips

	Captivity Model I		*Captivity Model II*	
Variables	Coefficients	*t-values*	Coefficients	*t-values*
Constant (W/B)	1.3623	**5.97**	2.2710	**4.45**
Constant (D)	30.4460	**2.97**	31.9974	**2.16**
Constant (S)	27.6244	**2.67**	29.1550	**1.96**
Age (W/B, T)	0.0198	**5.37**	-0.0048	-1.54
Sex (D, S)	0.0672	1.44	-0.1275	**-2.05**
HH Income (D,S)	0.0322	**2.59**	0.0220	***1.85***
HH Size (T,S)	0.1348	0.60	-0.0215	-0.41
#. of Workers (D)	-0.0746	**-5.21**	-0.1135	**-4.75**
Vehicles in HH (D,S)	1.1300	**4.57**	0.4328	**5.08**
MF Housing (W/B,T)	0.4741	0.11	-0.3620	-0.11
Walk time (W/B)	-0.0347	**-7.10**	-0.0497	**-7.13**
Auto time (D, S)	-0.0140	-1.48	0.0377	0.99
Population density at O (W/B,T)	-0.0304	-1.20	-0.0050	-1.31
Population density at D (D, S)	0.0213	1.31	0.0210	1.42
Job density at O (W/B, T)	0.0019	**4.20**	0.0027	**3.70**
Job density at D (D, S)	-0.0028	**-4.17**	-0.0049	**-3.45**
Resid. share of land at O (W/B,T)	0.0010	0.03	0.0028	0.04
Resid. share of land at D (D,S)	0.0030	0.18	-0.0023	-0.25
Comm. share of land at O (W/B, T)	0.0202	**3.49**	0.0057	**2.51**
Comm. share of land at D (D,S)	-0.0066	**-2.40**	0.0057	**2.04**
Regional accessibility at D (D,S)	-0.2965	**-2.81**	-0.3200	**-2.13**
Entropy index at D (D,S)	-0.2417	-0.48	0.1575	0.56
Captivity Variables				
Walk – constant	-23.7187	**-7.27**	-6.8558	**-4.81**
Transit – constant	-19.6046	**-3.66**	-23.5230	-1.61
Driving alone – constant	-7.4525	**-3.22**	-2.8399	**-2.12**
Residential-use dominance at O (D)			0.3988	**2.08**
Residential-use dominance at D (D)			-0.0280	-1.00
Shared ride – constant	-5.5674	***-1.82***	-3.2897	-1.49
Residential-use dominance at O (S)			0.3050	***1.65***
Residential-use dominance at D (S)			0.0085	1.19
Number of observations	3,354		3,354	
LLF($\mathcal{L}$) at zero	-6,476.7		-6,476.7	
LLF($\mathcal{L}$) at converge	-2,088.6		2,084.4	
Goodness-of-fit : ρ^2 , $\bar{\rho}^2$	0.6775, 0.6736		0.6782, 0.6735	
	W/B T D S Total		W/B T D S Total	
Actual choice share	1.2% 5.9% 83.8% 9.1% 100%		1.2% 5.9% 83.8% 9.1% 100%	
% correctly predicted	0.2% 2.7% 83.4% 0.0% 86.3%		0.2% 2.7% 83.4% 0.0% 86.3%	

a. W/B = walk/bike, T = bus, D = driving-alone, and S = shared-ride

b. Parenthesis in variable column indicates the mode(s) to which the variable is specified.

c. *t*-values in bold-face are significant at 95% level, and *italic* bold face at 90 % level respectively.

A comparison of the estimated coefficients in Table 5.9 reveals a general decline in the effects of socioeconomic variables on mode choice in contrast to the MNL model in Table 5.3. Household income and the number of vehicles owned by a household have positive effects on the choice of automobiles. The effect of land-use variables on mode choice is similar to that of MNL, although the magnitude of the effect decreases in the logit captivity models. The constant terms in both models are mostly significant, and the constants of *captivity model I* have larger magnitudes than those of the second model. A meaningful result was obtained in *captivity model II* when explanatory variables and residential land-use dominance indicators at both trip ends were specific to both modes (driving-alone and shared-ride). Residential land-use dominance at trip origins was found to increase the captivity probabilities of driving-alone and shared-ride at the 10% significance level. The alternative hypothesis that the observed captive behavior is generated by this model is accepted.

The choice set probabilities (or captivity odds) were calculated for each mode separately and for the full set of modes by taking the exponentials of the predicted values at the mean values of captive variables, given in Table 5.10. These probabilities lead to important implications regarding mode captivity and trip-maker mode choice sets. First, relative to a standard MNL model for HBW trips (see Table 5.3), the captivity model results are consistent with a notion that some trip-makers' choices of automobile, although the probabilities are extremely small, are based on singleton choice sets. There is a 99.56% probability that trip-makers are free to choose a mode from all considered modes and therefore a 0.44% probability that trip-makers face a singleton choice set. The finding of singleton choice sets (driving-alone and shared-ride) has an impact on trip-maker sensitivity to changes in the explanatory (residential land-use dominance) variables.

Second, the captivity probability that trip-makers face when presented with a non-singleton set of choices was decreased when residential land-use dominance variables were included. For example, the non-singleton choice probability decreased from 99.56% to 90.73%, and the probabilities of trip-makers captive to automobiles increased from 0.06% to 5.62% for driving-alone and from 0.38% to 3.55% for a shared-ride. The change in captivity probabilities depends on the model parameterized by including variables that could explain singleton choice sets. In this study, residential land-use dominance was identified as a relevant land-use factor and was found to increase the probability of using automobiles. The inclusion of captive variables in a mode choice model could better explain captivity behavior as well as predict choice behavior.

Table 5.10 Choice Set Probabilities

Choice Set		Probabilities for *Captivity Model I*	Probabilities for *Captivity Model II*
Singleton choice set	Walk/Bike	0.0000	0.0010
	Bus	0.0000	0.0000
	Driving-alone	0.0006	0.0562
	Shared-ride	0.0038	0.0355
	Sub-total	0.0044	0.0927
Non-singleton choice set		0.9956	0.9073

Although the captive odds to automobile (driving-alone and shared-ride), targeted for a particular land-use variable, are small and viewed as tentative and arbitrary, it is worthy to note that some trip-makers are forced to make a captive choice of automobile. Five percent of automobile-trip makers are in a segmented travel market with no alternative (4.7% for driving-alone and 0.3% for shared-ride).[25] This finding points to the need for a better understanding of how land-use pattern constrains mode choice in transportation market. This may be useful for the planning of residential development reducing automobile dependence. However, despite the scope advantage for the purpose of this analysis, the logit captivity model generally includes high cost in terms of estimation and difficulty in identifying relevant factors to be parameterized into the model.

[25] Actual choice shares in work-trips are 83.8% for driving-alone and 9.1% for shared-ride, and captive odds attributable to residential land-use dominance are 5.62% for driving-alone and 3.55% for shared-ride. Thus, the captive factor segments mode choice market into 4.7% for driving-alone and .0.3% for shared-ride.

CHAPTER VI

CONCLUSION AND DISCUSSION

This chapter concludes by summarizing the important findings from this study and by discussing the land-use policy implications of the major findings. Then, the limitations and the future extensions of research are addressed.

Conclusion

There has been a vital and pressing debate over the role of land-use and its relevant policies among academic scholars as well as planning professionals in an effort to reduce the problems associated with automobile dependence over recent decades. Much research has shed light on the relationships between transportation and land-use in parallel with the advances in using diary data, measuring land-use thanks to GIS techniques, specifying empirical models and applying analytical methods. Regardless of these advances in research exploring the predictability of effects of land-use on travel behavior, most empirical models linking land-use to transportation still maintain theoretical weakness and ignore causality issue. Moreover, the mixed and complex results in previous studies appear somewhat associative rather than causal in interpretation. To fill such a gap, this study investigates the causal effects of land-use on travel patterns using the datasets of the 1996 D-FW household survey, level-of-service, and land-use.

Land-use variables assessed at the TSZ level may have the limitation of fully reflecting the spatial distribution of current land-use associated with a variety of human needs and activities over the D-FW area. Nonetheless, land-use variables measured at the TSZ level were used for the exploration of their causal effects on travel behavior. The land-use status quo analysis of the D-FW area preceded the analysis of empirical models. Then, conventional regression methods were used for the estimation of travel demand models, and a causal graphical analysis was performed to study the causal relationship among the variables. In addition, logit captivity model was utilized for exploring captive contribution to mode choice attributable to land-use.

Empirical Findings

The status quo analysis of D-FW area land-use indicates that single-family residential use has dominated urban development patterns, accounting for nearly three times the land for commercial, industrial, and institutional uses combined. Such development patterns appear to be low-density-oriented and affected by zoning regulations of municipalities for land-use planning. Average land-use balance (entropy index) also appears relatively low with 0.38 and single-family-residential land-use accounts for 22 percent of total TSZs. Heterogeneous land-use (with entropy index > 0.6) is mainly observed along major arterials. These observations by status quo analysis address the urban form of the D-FW area as low-density, single-family-residential-use oriented, isolated from other land-uses, and heavily reliant upon automobile access to out-of-home activities.

The urban form addressed in the status quo analysis is examined by the empirical models of transportation–land-use linkage. The impact of land-use on travel behavior (individual mode choice, household trip frequency, and household VMT) is summarized for work trips and non-work trips.

What land-use measures influence mode choice in the D-FW area in both work and non-work travel? The significant land-use factors in the conventional regression models are compared to the causal factors found in causal graphical models. The major findings of mode choice from multinomial logit and directed acyclic graphs are summarized below.

1) *Multinomial choice results suggest that some measures of land-use have a small but statistically significant effect on travel demand.* Job densities and regional accessibility are significant factors in reducing the choice of automobiles. The commercial shares of land at both trip ends are also significant factors in reducing automobile choice in HBW trips. However, their effects on travel demand are small, compared to that of travel time. For example, a 10% increase in regional accessibility lowers the chance that a trip maker drives to work, to non-work places, and from non-home to other places by 2.3%, 0.4%, and 1.5%, respectively.
2) *Direct causes on directed graphs from HBW, HBO, and NHB trips are very consistent with the estimated coefficients of the significance level in binary logit models.* The indirect causes are eventually connected to automobile choice through a complex connection with other variables on the way, however,

colliders are not connected to automobile choice. Such a causal structure contributes to articulating the relationships among variables from the data.

3) *Automobile choice for home-based work (HBW) trips is explained by the direct causes of the directed acyclic graph: household income, the number of vehicles owned by a household, job densities at both trip ends, the commercial shares of land-use at both trips ends, and regional accessibility*. Travel time and residential shares at trip ends absorb causal flow running from their parent variables but never transmit to the child variables, and population densities at both ends are parent variables to open the path to travel time only. Thus, travel time, residential shares, and population densities are not causally connected to the decision of choosing an automobile.

4) *The directed acyclic graph of mode choice for HBO trips relates direct causation to socioeconomic characteristics (income, and number of vehicles owned by a household), travel time, and some land-use variables (population density at destination, job densities at both trip ends, regional accessibility).* Population density, job density, and residential share, which measure the characteristics of trip destination, are indirectly connected to mode choice, while residential and commercial shares at origins are not causally connected to mode choice. In addition, travel time is explained by a function of land-use and household size.

5) *Job densities and commercial shares of land-use at both trip ends, travel time, household income, and number of vehicles in a household are direct causes of*

automobile choice in NHB trips. The logit results for the commercial shares of land-use are not significant, while the causal graphical analysis results in the direct causes of automobile choice, as expected theoretically. Population densities at both trip ends, residential share at destinations, and regional accessibility are indirect causes of automobile choice in NHB trips.

What land-use factors reduce household automobile-trip generation in work trips and non-work trips? How are the generalized cost, socioeconomic characteristics, and land-use variables are causally connected to household trip rates? The following summarizes major findings from both the ordered logit model and directed acyclic graphs of household behavior of trip frequency.

1) *Regional accessibility influences household trip frequency for both work and non-work automobile trips, and an increase in residential share or commercial share of land-use induces people to make more HBW trips by automobiles*. While work trips are likely to increase when automobile-based accessibility to jobs increases, non-work trips decrease due to the relative decrease in accessibility. As the residential or commercial share of total land-use at TSZ level gets higher, work trips made by automobile increase. The commercial share of land-use shows a different result from what is expected in HBW trips.

2) *Household trip frequency for HBW trips is explained by causal factors like the generalized cost, the number of workers in a household, the residential share, and regional accessibility.* Commercial share is not a direct cause leading to auto-dependence in work trips by a causal graphic model. D-separation predicts

that there is no dependence between commercial share and trip frequency, when the collider set is removed due to a lack of a causal connection to trip frequency. Population density, job density, and land-use balance are indirect causal factors.

3) *Regional accessibility is added to the direct causes of HBO trip frequency and negatively contributes to trip frequency, but other land-use factors (residential share, commercial share, job density, and land-use balance) are not causally connected to trip frequency.* The direction of causal flow runs from household size, regional accessibility, and population density to the generalized cost. Residential share and commercial share serve as colliders to absorb information running from their parents. Job density and entropy are running to each collider.

How do land-use variables contribute to reducing household vehicle miles traveled (VMT)? How causally are the variables connected to less driving in work trips and non-work? Major findings are summarized from the results of regression and directed acyclic graphs.

1) *According to the regression results, population density negatively affects household vehicle miles traveled (VMT), but the impact of regional accessibility on VMT is positive with work trips and negative with non-work trips.* Where the spatial distribution of population is more compact, households are likely to drive less. As regional accessibility is improved, households are more likely to drive further to work but less likely to drive shorter distances for non-work trips. Surprisingly, the residential share of land is positively associated with VMT in HBW trips, and likely to contribute to urban sprawl in the D-FW area. However,

the quantitative effect of these variables on household VMT is very small.

2) *Regional accessibility is causally connected to household VMT for work and non-work trips, and population density causes a decrease in VMT for non-work trips. Other land-use factors (residential share, commercial share, job density, and land-use balance) are not causally connected to VMT.* The direction of causality for the generalized cost runs from only land-use variables (population density and accessibility) to household VMT for work trips, while it runs from both regional accessibility and the number of vehicles owned by household to household VMT for non-work trips.

If some trip-makers respond to mode choice in a captive manner, how such a choice behavior attributable to land-use can be captured to better predict choice behavior?

1) *Land-use contributes to captive-driving choices for home-based work trips*. Residential land-use dominance (or single-use for residence) at trip origins explains somewhat captivity behavior in the choice of automobile choice as well as better predicts choice behavior.

2) *Lack of land-use mix at trip origins increases the probabilities of trip-makers being captive to the automobile from 0.06% to 5.62% for driving-alone and from 0.38% to 3.55% for shared-ride.*

The empirical results of regression models suggest that some land-use measures have a small but statistically significant effect on travel demand in the Dallas-Fort Worth metropolitan area that is heavily dependent on automobiles. The direct causes derived from the causal graphical models are mostly consistent with the significant results of regression models, but there are a few discrepancies. *For work trips*, increases in regional accessibility, job density, and share of commercial land-use reduce the use of automobiles. Higher regional accessibility, however, causes households to generate automobile trips and thus leads to the increase in vehicle miles of travel (VMT). *For non-work trips*, population density, job density, and regional accessibility are direct causes of the choice of automobile, while only regional accessibility is causally connected to degenerating automobile trips and VMT. In brief, density measures like population density, job density, and regional accessibility are the causal factors to reduce automobile dependence particularly in non-work trips, but the compositional (residential or commercial) share of land-use are least likely to be factors causally connected to automobile travel patterns. Land-use balance (entropy index) is not causally connected to travel behavior either. Logit captivity model results indicate that land-use contributes to captive-driving choices for home-based work trips.

Policy Implications

Communities across the most metropolitan areas of the United States are now initiating various land-use strategies to reduce the negative impacts of automobile dependence and to attain “smart” urban growth. Currently, enhancing travel mode options and preserving air quality are among the top priority in those initiatives, and accordingly land-use policies are getting greatly emphasized to evaluate transportation alternatives. Then, what implications can be drawn from this study for planning initiatives in the Dallas-Fort Worth metropolitan area? The discussion will focus on using land-use approach to urban transportation problems from the standpoint of land-use densification.

One needs first to think about non-automobile travel choices (i.e., public transit, and pedestrian facilities). According to the trip distribution of data used for this study, the actual mode choice of automobile accounts for 93% in HBW trips and for 95% in HBO trips, while non-automobile choice just accounts for 7% (HBW) and for 5% (HBO). Walking and biking are perceived as the most sustainable modes in terms of resource consumption, but currently there is little indication that low-density-and-single-family residential land-use in the D-FW area has much to do with relatively higher share of walking and biking. The level of bus transit accessibility is increased in the many urban centers of D-FW area, but there is no indication that bus ridership is increasing either (although the DART ridership is slightly increased over years). These facts are likely to make an appeal for the increase of density enough for public transit through changes in urban development pattern over long term. This approach will widen travel choices to eventually reduce the negative impacts of automobile dependence as well as to improve opportunities for transportation minorities who oftentimes have limited options for travel. As suggested by logit captivity results, an automobile choice with no alternative attributable to land-use may impose more burdens on the transportation expenditure and even limit the opportunities of social activities and participation to transportation minorities

Finding that diversity measures such as compositional shares of land-use and entropy index are not causes of reducing automobile dependence dissents from with the current thought of mixed-use impact found in many empirical studies. Such a discrepancy appears to have much to do with the status quo of D-FW area land-use oriented to low-density-and-single-use for residential development. Residential land-use has extremely high share oriented to single-use development over overall area and the mixed-use development are not zoned encouragingly by many local governments. Moreover, lots of single-use areas are connected to automobile access only in the region.

Density measures may result in recommendations for efforts to reduce automobile dependence. Strategies for increasing density can include the infill mixed-use development in single-use area and infill non-automobile options like rail transport or bike routes. Eventually these strategies should focus on compositional development through infill development, mixed-use development, and change of residential development's paradigm. However, efforts to increase density probably ought to be conducted with the cooperation of local governments. A large gap can exist between empirical results and policies to modify land-use at the municipality level. Local governments mostly act to limit areas of increased density or mixed-use to protect established neighborhoods through zoning ordinances or regulations. For example, zoning regulations such as lot coverage, floor-area ratio, number of unrelated persons living together, minimum parking standards, engineering or architectural building requirements, and so on, can be used to limit density. In the role of zoning to shape metropolitan form, Levine (2006) suggests a rationale for land-use policy reform, indicating that municipal regulations may lead to low-density and automobile-oriented development.

This study suggests that land-use intensification is effectual in reducing the overwhelming use of automobile in the D-FW. Practical planning and development strategies for land intensification include increasing employment and housing density, increasing activity-mix, clustering jobs, commercial (retail) activities in closer proximity to residence, although these strategies may differ from regional scales: local and regional. As discussed so far, these strategies should be a part of the solutions for urban transportation problems in the regional area.

Limitations and Future Extensions

This study is methodologically new in drawing causal relationships using cross-sectional data, and is reliable in incorporating the entire set of independent variables into the empirical models. However, there is a data deficiency in using the land-use variables with a variety of dimensions. In particular, lack of design measures (i.e., street connectivity, and street types) at TSZ level may limit the evaluation of land-use impact on travel behavior.

Another potential problem may exist in the geographical unit of analysis, the traffic study zone (TSZ). Recently, land-use variables at the TSZ level have been criticized for a deficiency of reliability in representing land-use features at trip points. Most trip points are located along or on streets, while land-use variables normally represent the features of areas encompassing street boundary as an average or total values. In reality, land-use does not explain the land-use features centering on trip points, and may lead to a bias in spatial distribution. Seemingly looking better, spatial features are captured by the grid-cutting or spatially encompassed area within 1/4 mile.

The primary purpose of this study is to enhance the understanding of the causal influence of land-use (or urban form) variables on travel behavior. Thanks to causal graphical analysis, variable interactions are causally interpreted. Difficulty remains, however, in finding latent variables (represented as the bidirected edges) between land-use variables, and including them in the empirical models. Much remains to be made in the latent variables.

The logit captivity model as choice set generation is more likely to explain the mode choice better for individual trip-makers who are faced with land-use constraints. Within empirical results, it is worth noting that the captivity of users of automobiles is attributable to land-use constraints for a rationale of specific land-use policy. However, despite the benefit of having better results, this model has high cost in terms of estimation (converging problem) and difficulty in estimating the captive variables into the model.

REFERENCES

Altshuler, A. (1979). *The urban transportation system: politics and policy innovation*. Cambridge, MA: MIT Press.

Alonso, W. (1964). *Location and land-use*. Cambridge, MA: Harvard University Press.

Anas, A., Arnott, R., & Small, K. (1998). Urban spatial structure. *Journal of Economic Literature, 36*, 1426-1464.

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

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

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

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

Bento, A., Cropper, M., Mobarak, A., & Vinha, K. (2005). The effects of urban structure on travel demand in the United States. *The Review of Economics and Statistics*, *87*, 466-478

Bessler, D., & Loper, N. (2001). Economic development: evidence from directed acyclic graphs. *The Manchester School, 68*, 457-476.

Bhat, C. (1997). An endogenous segmentation mode choice model with an application to intercity travel. *Transportation Science, 31*, 34-48.

Bhat, C., Govindarajan, A., & Pulugurta, V. (1998). Disaggregate attraction-end choice modeling. *Transportation Research Record, 1645*, 60-68.

Boarnet, M., & Crane, R. (2001a). The influence of land-use of travel behavior: specification and estimation strategies. *Transportation Research A*, *35*, 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 behavior? A study of the link between non-work travel and land-use characteristics. *Urban Studies*, *35*, 1155-1169.

Bureau of Labor Statistics (BLS). (2001). *Consumer expenditure in 1999*. Washington, DC: U.S. Department of Labor.

Bureau of Transportation Statistics (BTS). (2002). *National transportation statistics 2002*. Washington, DC: U.S. Department of Transportation.

Bureau of Transportation Statistics (BTS). (2003). *NPTS: Highlights of the 2001 National Household Travel Survey*. Washington, DC: U.S. Department of Transportation.

Bureau of Transportation Statistics (BTS). (2004a). *2004 national transportation statistics*. Washington, DC: U.S. Department of Transportation.

Bureau of Transportation Statistics (BTS). (2004b). *Transportation statistics annual report*. Washington, DC: U.S. Department of Transportation

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 environment and mode choice: toward a normative framework. *Transportation Research D*, *7*, 265-284.

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*, 199-219.

Cervero, R., & Landis, J. (1997a). The transportation–land-use connection still matters. *Access*, *7*, 2-10.

Cervero, R., & Landis, J. (1997b). Twenty years of Bay Area rapid transit system: land use and development impacts. *Transportation Research A*, *31*, 309-333.

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

Cervero, R., Sandoval, C., and Landis, J. (2002). Transportation as a stimulus of welfare-to work: private versus and public mobility. *Journal of Planning Education and Research*, *22*, 50-63.

Carne, R. (1996a). On form versus function: will the new urbanism reduce traffic, or increase it? *Journal of Planning Education and Research*, *15*, 117-126.

Crane, R. (1996b). Cars and drivers in the new suburbs: linking access to travel in neotraditional planning. *Journal of the American Planning Association*, *62*, 51-65.

Crane, R. (2000). The impact of urban form on travel: an interpretative review. *Journal of Planning Literature*, *15*, 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.

Cullianane, S., & Cullianane, K. (2003). Car dependence in a public transport dominated city: evidence from Hong Kong. *Transportation Research* D, *8*, 129-138.

Dawid, A. (1979). Conditional independence in statistical theory. *Journal of the Royal Statistical Society B*, *41*, 1-31.

Delucchi, M. (1996). Total cost of motor-vehicle use. *Access*, *8*, 7-13.

Demiralp, S., & Hoover, K. (2003). Searching for the causal structure of a vector autoregression. *Oxford Bulletin of Economics and Statistics*, *65* Supplement, 745-767.

Domencich, T., & McFadden, D. (1975). *Urban travel demand: A behavioral analysis*. Amsterdam, Netherlands: North Holland Publishing Company.

Downs, A. (1992). *Still stuck in traffic: coping with peak-hour traffic congestion*. Washington, DC: Brookings Institution.

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. 521-540). Cambridge, MA: MIT Press.

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

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

Federal Highway Administration (FHWA). (1995). *Highway statistics summary to 1995*. Washington, DC: U.S. Department of Transportation.

Federal Highway Administration (FHWA). (2001). *Highway statistics 2001*. Washington, DC: U.S. Department of Transportation.

Federal Highway Administration (FHWA). (2004). *Summary of travel trends: 2001 National Household Travel Survey*. Washington, DC: U.S. Department of Transportation.

Federal Highway Administration (FHWA), Federal Transit Administration (FTA) (2006). *Status of the nation's highway, bridges, and transit: 2004 conditions and performance*. Report to Congress. Washington, DC: U.S. Department of Transportation.

Fishbein, M. (1975). Attitude, attitude change and behavior: a theoretical overview. In P. Levine (Eds.), *Attitude research bridges the Atlantic* (pp. 127-155). Chicago, IL: American Marketing Association.

Frank, L., & Pivo, G. (1994). Impacts of mixed use and density on utilization of three modes of travel: single-occupant vehicle, transit, and walking. *Transportation Research Record*, *1466*, 44-52.

Freund, P., & Martin, G. (1993). *The ecology of the automobile*. Montreal, Canada: Black Rose Books Ltd.

Fujita, M. (1989). *Urban economic theory: land-use and city size*. Cambridge, UK: Cambridge University Press.

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

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

Giuliano, G.. (2004). Land use impacts of transportation investments: highway and transit. In S. Hanson & G. Giuliano (Eds.), *The geography of urban transportation* (pp. 237-273). New York: Guilford Press.

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

Glaeser E., & Kahn, M. (2004). Sprawl and growth. In V. Henderson & J. Thisse (Eds.), *Handbook of urban and regional economics: cities and geography* (pp. 2481-2527). Amsterdam, Netherlands: Elsevier.

Gomez-Ibanez, J. (1991). A global view of automobile dependence. *Journal of the American Planning Association*, *57*, 376-9.

Goodwin, P. (1996). Empirical evidence on induced traffic. *Transportation, 23*, 35-54

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

Gorden, P., Kumar, A., & Richardson, H. (1989). The influence of metropolitan spatial structure on commuting times. *Journal of Urban Economics*, *26*, 138-149

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

Greenwald, M. (2003). The road less traveled: new urbanist inducements to travel model substitution for work trips. *Journal of Planning Education and Research*, *23*, 39-57.

Hamilton, B. (1982). Wasting commuting. *Journal of Political Economy*, *90*, 1035-1051.

Hamilton, B. (1989). Wasting commuting again. *Journal of Political Economy*, *97*, 1497-1504.

Handy, S. (1993). Regional versus local accessibility: implication for nonwork travel. *Transportation Research Record*, *1993*, 58-66.

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

Handy, S (1996b). Understanding the link between urban form and nonwork travel behavior. *Journal of Planning Education and Research*, *15*, 183-198.

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

Handy, S. (2004). *Critical assessment of the literature on the relationship among transportation, land-use, and physical activity*. Paper prepared for the 2004 Transportation Research Board and the Institute of Medicine Committee on Physical Activity, Health, Transportation, and Land Use. Department of Environmental Science and Policy, University of California at Davis.

Handy, S., Cao, X., Buehler, T., & Mokhatarian, P. (2005). *The link between the built environment and travel behavior: correlation and causality*? Paper presented at the 84th Annual Meeting of the Transportation Research Board, Washington DC.

Handy, S., & Clifton, K. (2001). Local shopping as a strategy for reducing automobile travel. *Transportation*, *28*, 317-346.

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*. Research report SWUTC/98/465650-1, The University of Texas at Austin.

Hansen, Mark. (1995). Do new highways generate traffic? *Access* 7, 16-22.

Hausman, D. (1984). Causal priority. *Noûs*, 18, 261-279.

Hausman, D., & Woodward, J. (1999). Independence, invariance and the causal Markov condition. *British Journal of the Philosophy of Science*, *50*, 521-583.

Henderson, J. (1977). *Economic theory and the cities*. New York: Academic Press.

Holland, P. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*, 945-960.

Horwitz, J., Koppelman, F., and Lerman, S. (1986). *A self-instructing course in disaggregate mode choice modeling-FTA*. Federal Transit Administration. Retrieved March 31, 2004, from http://ntl.bts.gov/DOCS/381SIC.html

Illich, I. (1974). *Energy and equity*. New York: Harper and Row.

Ingram, G., and Liu, Z. (1999). Determination of motorization and road provision. In J. A. Gomez-Ibanez, W. B. Tye, & C. Winston (Eds.), *Transportation economics: essays in honor of Professor John R. Meyer* (pp. 325-356). Washington, DC:

Brookings Institution.

Kain, J. (1999). The urban transportation problem: a reexamination and update. In J. A. Gomez-Ibanez, W. B. Tye, & C. Winston (Eds.), *Transportation economics: essays in honor of Professor John R. Meyer* (pp. 359-401). Washington, DC: Brookings Institution.

Kenworthy, J., & Laube, F. (1999). Patterns of automobile dependence in cities: an international overview of key physical and economic dimensions with some implications for urban policy. *Transportation Research A*, 33, 691-723.

Kenworthy, J., Laube, F., Newman, P., Barter, P., Raad, T., Poboon, C., & Guia, B. (1999). *An international sourcebooks of automobile dependence in cities, 1960 -1990*, Boulder: University Press of Colorado.

Kitamura, R., Mokhtarian, P., & Laidet, L. (1997). A micro-analysis of land-use and travel in five neighborhoods in the San Francisco Bay Area. *Transportation*, *24*, 125-158.

Kockelman, K. (1997). Travel behavior as a function of accessibility, land-use mixing, land-use balance: evidence from the San Francisco bay area. *Transportation Research Record*, *1607*, 79-87.

Kockelman, K. (2001a). Modeling traffic's flow-density relation: Accommodation of multiple flow regimes and traveler types. *Transportation*, *28*, 363-374.

Kockelman, K. (2001b). A model for time- and budget-constrained activity demand analysis. *Transportation Research B*, *35*, 255-269.

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

Levine, J. (2005). *Zoned out: regulation, markets, and choices in transportation and metropolitan land-use*. Baltimore, MD: Johns Hopkins University Press.

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

Litman, T. (1997). Distance-based vehicle insurance as a TDM strategy. *Transportation Quarterly, 51*, 119-137.

Litman, T. (2002). *The cost of automobile dependency and the benefits of balanced transportation.* Victoria Transport Policy Institute. Retrieved May 30, 2004, from http://www.vtpi.org

Litman, T. (2003). *Transportation cost and benefit analysis: techniques, estimates and implication.* Victoria Transport Policy Institute. Retrieved May 30, 2004, from htt://www.vtpi.org

Litman, T., & Laube, F. (2002). *Automobile Dependency and Economic Development.* Victoria Transport Policy Institute. Retrieved May 30, 2004, from http://www.vtpi.org

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

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

McNally, M., & Ryan, S. (1997). Comparative assessment of travel characteristics for neotraditional designs. *Transportation Research Record*, *1607*, 105-115.

Mensah, J. (1995). Journey to work and job search characteristics of the urban poor: a gender analysis of survey data from Edmonton. *Transportation*, *22*, 1-19.

Mieszkowski, P., & Mills, E. (1991). Analyzing urban decentralization: The case of Houston. *Regional Science and Urban Economics*, *21*, 183-199.

Mieszkowski, P., & Mills, E. (1993). The causes of metropolitan suburbanization. *Journal of Economic Perspective*, *7*, 135-147

Miller, T. (1991). *The cost of highway crashes*. Washington, DC: U.S. Department of Transportation, Federal Highway Administration.

Mills, E. (1972). *Urban economics*. Glenview, IL: Scott, Foresman and Company.

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

Mumford, L. (1953). *The highway and the city*. New York: Harcourt Brace Jovanovich.

Muth, R. (1969). *Cities and housing: the spatial pattern of urban residential land-use*. Chicago, IL: The University of Chicago Press.

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.

North Central Texas Council of Governments. (1996). *1996 Dallas-Fort Worth Household Travel Survey: Report on survey methods*. Los Angeles, CA: Applied Management & Planning Group.

North Central Texas Council of Governments, (2000). *Dallas-Fort Worth regional travel model* (*DFWRTM*): *description of the multimodal forecasting process*. Retrieved May 23, 2003, from http://www.nctcog.org/trans/modeling/documentation

North Central Texas Council of Governments, (2005a). *DFW regional travel model documentation – roadway traffic assignment*. Retrieved April 20, 2005, from http://www.nctcog.org/trans/modeling/documentation

North Central Texas Council of Governments, (2005b). *DFW regional travel model documentation – transCAD base model overview*. Retrieved April 20, 2005, from http://www.nctcog.org/trans/modeling/documentation

North Central Texas Council of Governments, (2005c). *DFW regional travel model documentation – NCTCOG mode choice model estimation*. Retrieved April 20, 2005, from http://www.nctcog.org/trans/modeling/documentation

Organization for Economic Cooperation and Development (OECD). (1995). *Urban travel and sustainable development*. Paris, France: OECD.

Papineau, D. (1985). Causal asymmetry, *British Journal of Philosophy of Science*, *36*, 279-289.

Pearl, J (1988). *Probabilistic reasoning in intelligent systems*. San Mateo, CA: Morgan and Kaufman.

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

Pearl, J. (2000). *Causality*. Cambridge, UK: Cambridge University Press.

Pickrell, D. (1999). Transportation and land-use. In J. A. Gomez-Ibanez, W. B. Tye, &

C. Winston (Eds.), *Transportation economics: essays in honor of Professor John R. Meyer* (pp. 403-435). Washington, DC: Brookings Institution.

Pucher, J. (1988). Urban travel behavior as the outcome of public policy. *Journal of the American Planning Association*, *54*, 509-520.

Pushkarev, B., & Zupan, J. (1977). *Public transportation and land-use policy*. Bloomington, IN: Indiana 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 MS thesis. The University of British Columbia, Vancouver, Canada.

Richardson, H. (1997). *Motor vehicle crashes as a leading cause of death in the U.S. 1992*. Washington, DC: U.S. Department of Transportation, National Highway Traffic Safety Administration.

Roh, J., & Bessler, D. (1999). A study of traffic fatalities using directed 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.

Schiffman, L., & Kanuk, L. (1996). *Consumer behaviour*, 6th ed., Englewood Cliffs, NJ: Prentice-Hall International.

Schimek, P. (1996). Household motor vehicle ownership and use: how much does residential density matter. *Transportation Research Record*, *1552*, 120-125.

Scheines, R., Spirtes, P., Glymour, C., & Meek, C. (1994). *TETRAD II: Tools for causal modeling - User's manual and software*. Manwah, NJ: Lawrence Erlbaum Associates.

Shen, Q. (2000). Spatial and Social Dimensions of Commuting. *Journal of the American Planning Association*, *66*, 68-82.

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

Shipley, B. (1999). Testing causal explanations in organismal biology: causation, correlation and structural equation modeling. *Oikos*, *86*, 374-382.

Shipley, B. (2000a). A new inferential test for path models based on directed acyclic graphs. *Structural Modeling*, 7, 206-218.

Shipley, B. (2000b). *Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference*. Cambridge, UK: Cambridge University Press.

Siegan, B. (1990). Land use regulations should preserve only vital and pressing governmental interests. *Cato Journal*, *10*, 127-1158.

Siegan, B. (1997). *Property and freedom: the constitution, the courts, and land-use regulation*. New Brunswick, NJ: Transaction Publishers.

Small, K. & Song, S. (1992). Wasteful commuting: a solution. *Journal of Political Economy*, *100*, 888-898.

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

Steiner, R. (1994). Residential density and travel patterns: review of the literature. *Transportation Research Record*, *1466*, 37-43

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

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. (1986). Constraint on individual travel behavior in a Brazilian city. *Transportation Research Record*, *1085*, 75-85.

Swait, J., & Ben-Akiva, M. (1987a). Increasing random constraints in discrete models of choice set generation. *Transportation Research B*, *21*, 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*, 103-115.

Tayal, T., Anantuni, K. & Burns, E. (2001). *Measuring auto dependence in metro Phoenix using GIS*. Paper presented at the 21annual ESRI international user conference, San Diego, CA.

Thill, J. (1992). Choice set formation for destination choice modeling. *Progress in Human Geography*, *16*, 361-382.

Train, K. (1986). *Qualitative choice analysis*. Cambridge, MA: MIT Press.

Train, K. (2003). *Discrete choice methods with simulation*. Cambridge, UK: Cambridge University Press.

Transportation Research Board (2001). Critical issues in transportation 2002. *TR News 217* (Nov./Dec. 2001), 4-11.

U.S. Environmental Protection Agency (USEPA). (2004). *Inventory of U.S. greenhouse gas emissions and sinks: 1990-2002*. Washington, DC.

United Nations Environment Programme (UNEP). (1993). *Environmental Data Report, 1993-94*. Oxford: Blackwell Publishers.

Vadali, S., & Lee, S. (2005). *Travel time savings: LBJ-IH635 project in Dallas, TX*. Technical memorandum prepared for TXDOT Dallas District, College Station, TX: Texas Transportation Institute

Victoria Transport Policy Institute (VTPI). (2004). *Transportation cost and benefit analysis – air pollution cost*. Retrieved July 6, 2004, from http://www.vtpi.org

Victoria Transport Policy Institute (VTPI). (2005). *Online TDM encyclopedia*. Retrieved November 4, 2005, from http://www.vtpi.org/tdm/index/php

Wackernagel, M. & Rees, W. (1996) *Our ecological footprint: reducing human impact on the earth*. Gabriola Island, BC: New Society Publishers.

Washington, S., Karlaftis, M., & Mannering, F. (2003). *Statistical and econometric methods for transportation data analysis*. Boca Raton, FL: Chapman & Hall/CRC.

Wheaton, W. (1998). Land use and density in cities with congestion. *Journal of Urban Economics*, *43*(2), 258-272

White, M. (1988). Location and choice and commuting behavior in cities with decentralized employment. *Journal of Urban Economics*, *24*(2), 129-152

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

Zhang, M. (2004). The mobility role of land-use: evidence from travel choice in two world cities. *Journal of the American Planning Association*, *70*, 344-360.

APPENDIX

A1. Stratified Sampling in 1996 Dallas-Fort Worth Household Activity Survey.

Table A1-1. Stratified Sampling in 1996 D-FW Household Activity Survey, Source: 1996 Dallas-Fort Worth Household Travel Survey: Report on Survey Methods.

Level	Response	Survey Households		D-FW CMSA	Difference (% point)
		N	%		
County	Collin	372	9.3%	6.8%	2.5%
	Dallas	1,633	40.9%	49.8%	**-8.9%**
	Denton	440	11.0%	7.2%	**3.8%**
	Ellis	64	1.6%	1.8%	-0.2%
	Johnson	117	2.9%	2.1%	0.8%
	Kaufman	13	0.3%	0.3%	0.0%
	Parker	20	0.5%	0.2%	0.3%
	Rockwall	47	1.2%	0.6%	0.6%
	Tarrant	1,264	31.6%	31.1%	0.5%
	(Refused)	26	0.7%	-	
Household size	1	1,028	25.7%	25.1%	0.6%
	2	1,424	35.6%	30.3%	**5.3%**
	3	651	16.3%	17.6%	-1.3%
	4	585	14.6%	15.6%	-1.0%
	5	209	5.2%	7.0%	-1.8%
	6	67	1.7%	2.4%	-0.7%
	7+	32	0.8%	1.7%	-0.9%
Vehicles available	0	207	5.2%	6.4%	-1.2%
	1	1,316	32.9%	35.0%	-2.1%
	2	1,731	43.3%	41.6%	1.7%
	3	521	13.0%	12.8%	0.2%
	4	148	3.7%	3.3%	0.4%
	5+	73	1.8%	1.0%	0.8%

Source: 1996 Dallas-Fort Worth Household Travel Survey: Report on Survey Methods, NCTCOG

The distributions of surveyed households are compared to data from the 1990 U.S. Census for Dallas-Fort Worth Consolidated Metropolitan Statistical Area (CMSA). Differences in the survey dataset of 3% points or more are identified with boldface letters in Table A1-1.

A2. Travel Time Components in the Dallas-Fort Worth Regional Travel Model

Estimates of roadway travel times used in calculating the value of time (VOT) savings include a combination of uncongested "free" speed travel time, delay time, and intrazonal travel time. First two travel times are calculated through the traffic assignment volume-delay function from interzonal trips, whereas intrazonal travel times are obtained separately from interzonal trips.

For interzonal trips, travel times between zones are calculated and travel time within a zone is assumed as zero. However, this assumption does not reflect reality because any trips within the zone are necessarily accompanied by travel times. The time matrices used in this analysis include interzonal as well as intrazonal trips.

Volume-Delay Function (Interzonal Travel Times)

Volume-delay function in the Dallas-Fort Worth Regional Travel Model (DFWRTM) similar in form to the BPR-type functions, in that link speed decreases as the (volume/capacity) ratio increases. The general form of NCTCOG's volume-delay function is as follows.

$$Travel\ Times = Free\ Flow\ Time + Min[\alpha \cdot e^{\beta(v/c)}, \gamma] \tag{1}$$

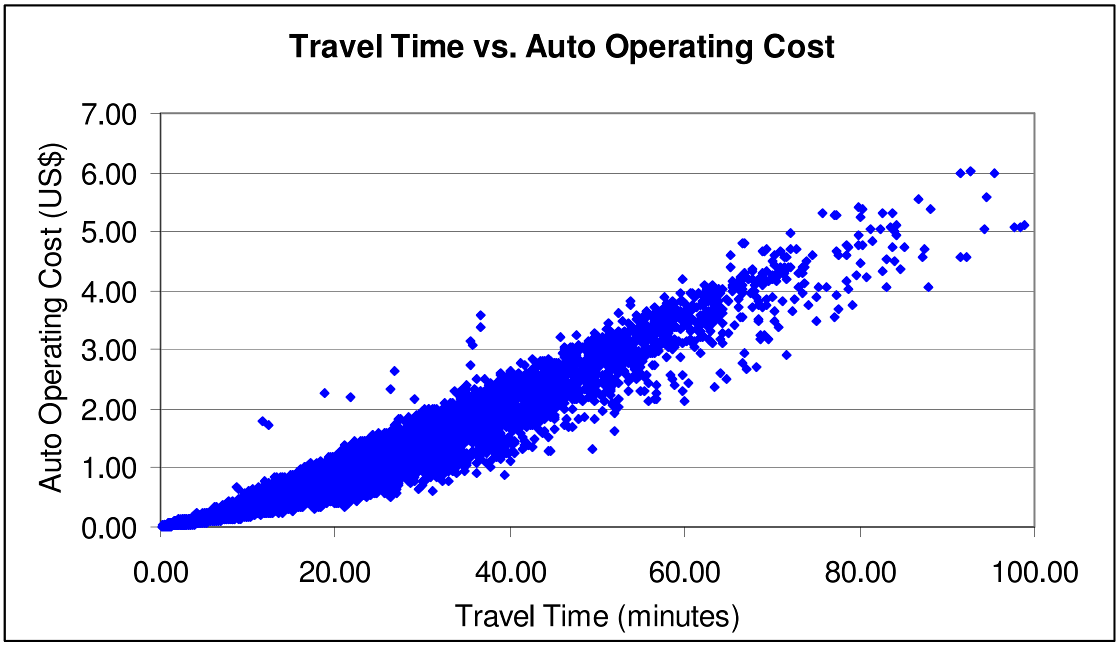
where α, β, and γ are delay function parameters (see NCTCOG 2005). According to NCTCOG's volume-delay function, when volume/capacity ratio for roadway links is no more than 0.3, vehicles are driving in speed taking free flow time for either freeway or arterial. However, traffic volume tends to depend on the total cost of traveling through a roadway link. Such a total cost of traveling can be calculated as each link's generalized cost consisting of operating cost, toll cost, and travel time. After taking initial steps to create network file and to identify centroids in transCAD, the key steps to run trip assignment by time period of day are taken in the setting of four modal classes (DA, SRHOV, SRNoHOV, and Truck). Then total traffic volume for each time period is calculated and new link travel times are estimated as shown in equation (1).

Intrazonal Travel Times

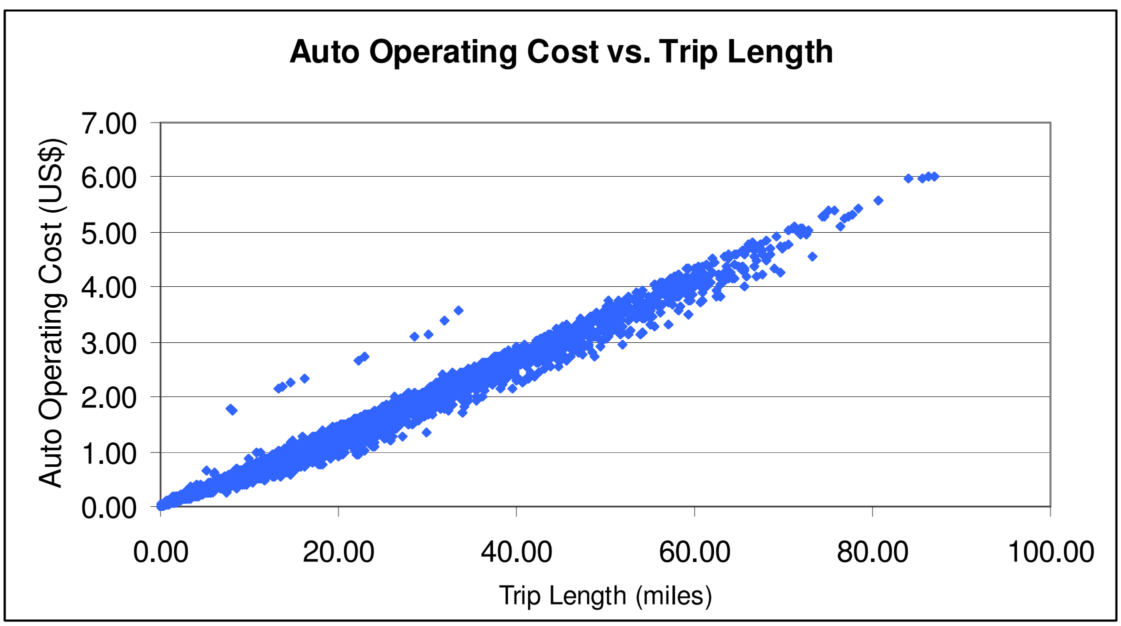
Intrazonal travel time estimates are obtained from the path-building process to calculate a more precise value for each zone. In NCTCOG's travel model, a zone is divided into 13 concentric squares, and the average distance from the center of the zone to the perimeter of each square is determined for zones at distances greater than walking distance. For each distance, a cost-per-mile value is applied to convert the distance to travel time. The cost-per-mile values vary by time-of-day (AM, PM, OP) and 5 area types.

A3. Relationships among Travel Time, Auto Operating Cost, and Trip Length

A3-1. Correlation($\rho =$ 0.9714) between Travel Time and Auto Operating Cost



A3-2. Correlation($\rho =$ 0.9947) between Travel Time and Auto Operating Cost



A4. Correlation Matrices for Mode Choice Models

A4-1. Correlation Matrix for HBW Trips (n = 4,672)

Corr(HBW) =

auto	*tt*	*inc*	*hhsz*	*wrkrs*	*vehnum*	*opopd*	*dpopd*	*ojobd*	*djobd*	*oresid*	*dresid*	*ocomm*	*dcomm*	*access*
1.00														
0.02	1.00													
0.24	0.06	1.00												
0.04	-0.01	0.18	1.00											
0.14	-0.01	0.29	0.48	1.00										
0.29	0.05	0.40	0.36	0.50	1.00									
0.00	-0.15	-0.08	-0.09	-0.02	-0.11	1.00								
-0.01	-0.10	-0.07	-0.05	-0.00	-0.08	-0.08	1.00							
-0.26	-0.01	-0.05	-0.01	-0.03	-0.07	-0.13	0.06	1.00						
-0.24	-0.01	-0.04	-0.01	-0.03	-0.06	0.05	-0.14	0.00	1.00					
0.11	-0.06	0.06	0.02	0.07	0.07	0.49	-0.20	-0.25	0.04	1.00				
0.09	-0.01	0.03	0.06	0.07	0.05	-0.21	-0.51	0.05	-0.25	-0.30	1.00			
-0.21	-0.01	-0.03	-0.05	-0.05	-0.09	-0.20	0.13	0.42	-0.02	-0.56	0.17	1.00		
-0.16	-0.02	-0.01	-0.06	-0.04	-0.07	0.14	-0.20	-0.03	0.40	0.15	-0.55	-0.06	1.00	
-0.16	-0.13	-0.05	-0.11	-0.09	-0.17	0.17	0.14	0.04	0.14	0.02	-0.25	0.05	0.30	1.00

A4-2. Correlation Matrix for HBO Trips (n = 7,112)

Corr(HBO)=

auto	*tt*	*inc*	*hhsz*	*wrkrs*	*vehnum*	*opopd*	*dpopd*	*ojobd*	*djobd*	*oresid*	*dresid*	*ocomm*	*dcomm*	*access*
1.00														
0.12	1.00													
0.19	0.00	1.00												
0.03	-0.09	0.20	1.00											
0.07	-0.01	0.25	0.45	1.00										
0.22	0.04	0.38	0.32	0.42	1.00									
-0.12	-0.15	-0.14	-0.08	-0.04	-0.14	1.00								
-0.09	-0.16	-0.12	-0.05	-0.03	-0.12	0.24	1.00							
-0.04	0.01	-0.01	-0.01	-0.01	0.00	-0.05	0.28	1.00						
-0.16	0.00	-0.04	-0.02	-0.02	-0.05	0.04	-0.09	0.04	1.00					
-0.01	-0.06	0.05	0.08	0.01	0.04	0.35	-0.03	-0.18	-0.01	1.00				
0.02	-0.03	0.06	0.07	0.02	0.03	-0.02	0.38	0.00	-0.22	-0.08	1.00			
-0.02	-0.01	-0.06	-0.10	-0.03	-0.06	-0.08	0.09	0.20	0.03	-0.63	0.07	1.00		
-0.05	-0.03	-0.06	-0.08	-0.03	-0.07	0.08	-0.11	0.04	0.26	0.06	-0.62	0.00	1.00	
-0.11	-0.19	-0.02	-0.10	-0.07	-0.13	0.22	0.31	0.02	0.09	0.02	-0.09	0.04	0.17	1.00

A4-3. Correlation Matrix for NHB Trips (n = 3,354)

Corr(NHB)=

auto	*tt*	*inc*	*hhsz*	*wrkrs*	*vehnum*	*opopd*	*dpopd*	*ojobd*	*djobd*	*oresid*	*dresid*	*ocomm*	*dcomm*	*access*
1.00														
0.20	1.00													
0.18	0.02	1.00												
0.07	-0.02	0.23	1.00											
0.07	0.00	0.24	0.49	1.00										
0.21	0.06	0.41	0.39	0.50	1.00									
0.10	-0.09	-0.04	0.03	-0.04	-0.03	1.00								
0.10	-0.08	-0.04	0.01	-0.02	-0.03	0.14	1.00							
-0.39	-0.08	-0.06	-0.06	-0.01	-0.06	-0.17	-0.09	1.00						
-0.42	-0.08	-0.04	-0.03	-0.01	-0.05	-0.12	-0.17	0.43	1.00					
0.19	0.06	0.00	0.08	-0.01	0.04	0.56	0.09	-0.27	-0.18	1.00				
0.20	0.08	-0.02	0.05	0.00	0.05	0.10	0.53	-0.16	-0.27	0.20	1.00			
-0.26	-0.12	0.00	-0.06	-0.03	-0.08	-0.22	-0.06	0.37	0.25	-0.57	-0.17	1.00		
-0.25	-0.12	0.01	-0.04	-0.01	-0.05	-0.07	-0.21	0.24	0.37	-0.18	-0.58	0.29	1.00	
-0.17	-0.30	0.02	-0.07	-0.06	-0.14	0.05	0.10	0.09	0.15	-0.11	-0.23	0.18	0.27	1.00

A5. Correlation Matrices for Household Trip Frequency Models

A5-1. Correlation Matrix for HBW Trips (n = 1,955)

$$
Corr(\mathrm{HBW}) = \begin{array}{c}
\begin{array}{cccccccccccc} \mathit{freq} & \mathit{g\_cost} & \mathit{inc} & \mathit{hhsz} & \mathit{wrkrs} & \mathit{vehnum} & \mathit{o\_popd} & \mathit{o\_jobd} & \mathit{o\_resid} & \mathit{o\_comm} & \mathit{access} & \mathit{entropy} \end{array} \\
\begin{bmatrix}
1.00 & & & & & & & & & & & \\
-0.12 & 1.00 & & & & & & & & & & \\
0.10 & 0.13 & 1.00 & & & & & & & & & \\
0.19 & -0.01 & 0.18 & 1.00 & & & & & & & & \\
0.38 & 0.00 & 0.25 & 0.47 & 1.00 & & & & & & & \\
0.22 & 0.09 & 0.34 & 0.36 & 0.47 & 1.00 & & & & & & \\
0.03 & -0.19 & -0.14 & -0.14 & -0.07 & -0.18 & 1.00 & & & & & \\
-0.04 & -0.02 & -0.03 & -0.03 & 0.01 & -0.01 & -0.04 & 1.00 & & & & \\
0.11 & 0.00 & 0.12 & 0.04 & 0.07 & 0.07 & 0.30 & -0.18 & 1.00 & & & \\
-0.04 & -0.07 & -0.12 & -0.12 & -0.10 & -0.12 & 0.00 & 0.13 & -0.57 & 1.00 & & \\
0.08 & 0.00 & -0.03 & -0.12 & -0.08 & -0.13 & 0.15 & 0.01 & -0.01 & 0.07 & 1.00 & \\
-0.04 & -0.02 & -0.13 & -0.05 & -0.09 & -0.10 & -0.17 & -0.05 & -0.65 & 0.46 & 0.02 & 1.00
\end{bmatrix}
\end{array}
$$

A5-2. Correlation Matrix for HBO Trips (n = 2,072)

$$
Corr(\text{HBW}) = \begin{array}{cccccccccccc}
\textit{freq} & \textit{g\_cost} & \textit{inc} & \textit{hhsz} & \textit{wrkrs} & \textit{vehnum} & \textit{o\_popd} & \textit{o\_jobd} & \textit{o\_resid} & \textit{o\_comm} & \textit{access} & \textit{entropy}
\end{array}
$$

$$
\begin{bmatrix}
1.00 & & & & & & & & & & & \\
-0.28 & 1.00 & & & & & & & & & & \\
0.17 & 0.02 & 1.00 & & & & & & & & & \\
0.28 & -0.14 & 0.23 & 1.00 & & & & & & & & \\
0.10 & 0.04 & 0.30 & 0.48 & 1.00 & & & & & & & \\
0.19 & 0.01 & 0.36 & 0.36 & 0.45 & 1.00 & & & & & & \\
-0.02 & -0.06 & -0.17 & -0.13 & -0.06 & -0.20 & 1.00 & & & & & \\
-0.02 & 0.00 & -0.03 & -0.02 & 0.01 & -0.01 & -0.03 & 1.00 & & & & \\
0.10 & -0.07 & 0.09 & 0.05 & 0.02 & 0.06 & 0.27 & -0.16 & 1.00 & & & \\
-0.07 & 0.04 & -0.11 & -0.12 & -0.07 & -0.11 & 0.03 & 0.06 & -0.56 & 1.00 & & \\
-0.11 & 0.07 & 0.16 & 0.07 & 0.28 & 0.08 & 0.10 & 0.02 & -0.02 & 0.01 & 1.00 & \\
-0.08 & 0.05 & -0.11 & -0.07 & -0.05 & -0.09 & -0.15 & -0.02 & -0.69 & 0.49 & -0.03 & 1.00
\end{bmatrix}
$$

A6. Correlation Matrices for Household VMT Models

A6-1. Correlation Matrix for HBW Trips (n = 1,955)

Corr(HBW) =

ln(VMT)	*g_cost*	*inc*	*hhsz*	*wrkrs*	*vehnum*	*o_popd*	*o_jobd*	*o_resid*	*o_comm*	*access*	*entropy*
1.00											
-0.71	1.00										
0.15	-0.01	1.00									
0.09	-0.03	0.18	1.00								
0.21	-0.06	0.25	0.47	1.00							
0.18	-0.05	0.34	0.36	0.47	1.00						
0.03	0.13	-0.14	-0.14	-0.07	-0.18	1.00					
-0.03	0.00	-0.03	-0.03	0.01	-0.01	-0.04	1.00				
0.05	0.09	0.12	0.04	0.07	0.07	0.30	-0.18	1.00			
-0.08	0.00	-0.12	-0.12	-0.10	-0.12	0.00	0.13	-0.57	1.00		
0.10	-0.06	-0.03	-0.12	-0.08	-0.13	0.15	0.01	-0.01	0.07	1.00	
-0.03	-0.06	-0.13	-0.05	-0.09	-0.10	-0.17	-0.05	-0.65	0.46	0.02	1.00

A6-2. Correlation Matrix for HBO Trips (n = 2,072)

$$
Corr(\text{HBO}) = \begin{array}{cccccccccccc}
ln(VMT) & g\_cost & inc & hhsz & wrkrs & vehnum & o\_popd & o\_jobd & o\_resid & o\_comm & access & entropy
\end{array}
$$

$$
\begin{bmatrix}
1.00 & & & & & & & & & & & \\
-0.71 & 1.00 & & & & & & & & & & \\
0.09 & -0.07 & 1.00 & & & & & & & & & \\
0.04 & -0.06 & 0.23 & 1.00 & & & & & & & & \\
0.03 & -0.03 & 0.30 & 0.48 & 1.00 & & & & & & & \\
0.14 & -0.09 & 0.36 & 0.36 & 0.45 & 1.00 & & & & & & \\
-0.10 & 0.05 & -0.17 & -0.13 & -0.06 & -0.20 & 1.00 & & & & & \\
0.01 & -0.01 & -0.03 & -0.02 & 0.01 & -0.01 & -0.03 & 1.00 & & & & \\
0.00 & -0.03 & 0.09 & 0.05 & 0.02 & 0.06 & 0.27 & -0.16 & 1.00 & & & \\
-0.05 & 0.05 & -0.11 & -0.12 & -0.07 & -0.11 & 0.03 & 0.06 & -0.56 & 1.00 & & \\
-0.10 & 0.07 & 0.16 & 0.07 & 0.28 & 0.08 & 0.10 & 0.02 & -0.02 & 0.01 & 1.00 & \\
-0.02 & 0.05 & -0.11 & -0.07 & -0.05 & -0.09 & -0.15 & -0.02 & -0.69 & 0.49 & -0.03 & 1.00
\end{bmatrix}
$$

A7. GAUSS Statistical Code for Multinomial Logit Captivity Model

```
/**************************************************************************
 *  GAUSS Code for Parameterized Captivity Logit:                          *
 *                                                                         *
 *     Variables from a vector, x                                          *
 *   0 Dependent: 1=walk/bike, 2=bus, 3=driving-alone, 4=shride-ride       *
 *   1 ONE      : constant                                                 *
 *   2 AGE      : age in year                                              *
 *   3 SEX      : 1 = male,  0 = female                                    *
 *   4 INC      : Household income in $US                                  *
 *   5 HHSZ     : Number of people in household                            *
 *   6 WRKRS    : Number of workers in household                           *
 *   7 VEHNUM   : Number of vehicles in household                          *
 *   8 DW_MFH   : Dwelling - multi-family housing ( yes = 1, no =0)        *
 *   9 WKTT     : Walking time  bike (minutes)                             *
 *  10 DRTT     : Travel time by driving-alone (min)                       *
 *  11 SRTT     : Travel time by shared ride (min)                         *
 *  12 o_popden : Population density per acre at Os in TSZ                 *
 *  13 d_popden : Population density per acre at Ds in TSZ                 *
 *  14 o_jobden : Total job density per acre at Os in TSZ                  *
 *  15 d_jobden : Total job density per acre at Ds in TSZ                  *
 *  16 o_resid  : % Residential LU of total developed area at Os in TSZ    *
 *  17 d_resid  : % Residential LU of total developed area at Ds in TSZ    *
 *  18 o_comm   : % Commercial LU of total developed area at Os in TSZ     *
 *  19 d_comm   : % Commercial LU of total developed area at Ds in TSZ     *
 *  20 d_ai     : Accessibility index at destination (gravity measure)     *
 *  21 d_entpy  : Land use balance at destination                          *
 *  22 o_ld_res : Low-density Residential use at origin                    *
 *  23 d_ld_res : Low-density Residential use at destination               *
 **************************************************************************/

new;
cls;

output file = c:\gauss\captive\DFW4672.out  reset;
load data[4672,23]=c:\gauss\captive\DFW4672.txt;
format /m1 /rd 8,3;

n = rows(data);
c = 4;                                  @ number of alternatives  @
y0 = data[.,1];

y= zeros(n,c);
i=1;
do until i > n;
       j=1;
              do until  j > c;
                     if y0[i,1] == j; y[i,j]=1;
                                   else;
                            endif;
                     j=j+1;
              endo;
       i=i+1;
endo;

one = ones(n,1);
zero = zeros(n,1);

x=one~data[.,2:23 ];
```

```
mean_x=meanc(x);

b0={ 4.7726, 11.6597, 10.5052,
    -0.0267,  0.0218, 0.0161, 0.3461, -0.0018, 1.3578, 0.9159, -0.0611, 0.0164, -0.0165,
     0.0103,  0.0034, -0.0080,  0.0024,  -0.0019, -0.0059,  0.0033,  -0.1127,  0.3112,
    -5.6939, -5.5955, -7.1521, -3.6793,
    -0.0465, -0.8806,  0.1787, -0.3094     };         @ initial values @

/************  maxlik.src  MALIK: Maximum Likelihood Estimation   ****************/

library maxlik;
#include maxlik.ext;
maxset;
max_CovPar=    1;        @ Computation method for Var-covariance matrix  @
max_GradTol=0.001;
max_MaxIters=500000;
max_Algorithm=5;         @ Use BHHH algorithm   @

{ beta, f, g, h, ret } = maxprt(maxlik( x, 0, &llf, b0 ));

se=sqrt(diag(h));
t= (beta./se);
at=abs(t);
k1=rows(b0);
df=n-k1;
p=cdftc(at,df)*2;

/*************************  Log-likelihood Function   ************************/

proc llf(b0,x);
local  ev1, ev2, ev3, ev4, ez1, ez2, ez3, ez4, sumev, sumez, p1, p2, p3, lnL;
       ev1 = exp( x[.,1 2 8 9 12 14 16 18 21]     *b0[1  4 10 11  13  15 17 19  22] );
       ev2 = exp( x[.,2 5 8     12 14 16 18 21]   *b0[   4  7 10  13  15 17 19  22] );
       ev3 = exp( x[.,1 3 4 6 7 10 13 15 17 19 20]*b0[ 2 5 6 8 9 12 14 16 18 20 21] );
       ev4 = exp( x[.,1 3 4 5 7 11 13 15 17 19 20]*b0[ 3 5 6 7 9 12 14 16 18 20 21] );

       ez1 = exp( x[.,1     ] *b0[23  ] );
       ez2 = exp( x[.,1     ] *b0[24  ] );
       ez3 = exp( x[.,1  22 23] *b0[25 27 28 ] );
       ez4 = exp( x[.,1  22 23] *b0[26 29 30 ] );

       sumev = ev1 + ev2 + ev3 + ev4;
       sumez = ez1 + ez2 + ez3 + ez4;

       p1 = (1./(1 + sumez)) .*(ez1 + (ev1./sumev));
       p2 = (1./(1 + sumez)) .*(ez2 + (ev2./sumev));
       p3 = (1./(1 + sumez)) .*(ez3 + (ev3./sumev));

lnL= y[.,1].*ln(p1) + y[.,2].*ln(p2) + y[.,3].*ln(p3) + y[.,4].*ln(1-p1-p2-p3) ;

retp(lnL);
endp;

/******************************  Run Time  ***********************************/


/*---------------------------- Logit Captivity Results  -----------------------*/
?""; ?"";
print "Variables Estimates  Std.err   t-value    p-value  ";
print "===================================================================";
```

```
let names = c1 c3 c4
            age sex inc hhsz wrkrs vehnum dw_mfh
            watt_tt auto_tt
            o_pden d_pden o_jobden d_jobden o_resid o_resid o_comm d_comm access entropy
            cp1 cp2 cp3 cp4
            o_ld_r3  d_ld_r3 o_ld_r4  d_ld_r4;


mask = 0~1~1~1~1;      /* To be printed as string(0) or numeric(1); */
fmt={"-*.*s" 8 8, "*.*lf" 10 4, "*.*lf" 10 4, "*.*lf" 10 4, "*.*lf" 10 4   };
out = names~beta~se~t~p;
call printfm(out,mask,fmt);         /* printfm :GAUSS Reference p.3-622  *//

?"";

ilf = 0.47549;                     @ log-likelihood function value at 1st iteration @
format /m1/rd 9,4;
print "Log-likelihood at 'converge'  =  "  n*f;
print "Log-likelihood at ' 0 '       =  "  -n*ilf;
print "LR test                       =  "  -2*(-n*ilf -n*f);
print "Goodness-of-fit index         =  "  1-(n*f/(-n*ilf));
print "Crtd goodness-of-fit index    =  "  1-(n*f-k1)/(-n*ilf);

/**********************************End of Gauss Code ******************************/
```

VITA

Name: Sangkug Lee

Address: 2500 Central Park Ln, #603, College Station, TX 77840

E-mail Address: leesk86@gmail.com

Education: Ph.D. in Urban and Regional Science, Texas A&M University, 2006
Specialty: Transportation Planning, and Travel Demand Modeling
Dissertation: The Correlational and Causal Investigation into the Land Use–Transportation Relationships: Evidence from the Dallas-Fort Worth Metropolitan Area
Co-chairs: Ming Zhang and Chanam Lee
M.S. in Economics, Purdue University, West Lafayette, Indiana, 1996
B.S. in Ag-Economics, Chungnam National University, Korea, 1990