Graduate Theses and Dissertations   Graduate School

11-8-2010

The Interaction Between Urban Form and Transit Travel

Sisinnio Concas
University of South Florida

Follow this and additional works at: http://scholarcommons.usf.edu/etd

Part of the American Studies Commons

The Interaction Between Urban Form and Transit Travel

by

Sisinnio Concas

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Economics
College of Arts and Sciences
University of South Florida

Major Professor: Joseph S. DeSalvo, Ph.D.
Brad Kamp, Ph.D.
Ram M. Pendyala, Ph.D.
Gabriel A. Picone, Ph.D.

Date of Approval:
November 8, 2010

Keywords: urban economics, residential location, travel behavior, activity-based modeling

Dedication

I dedicate this dissertation to my wife Martha and to our daughters Isabella and Alessia. Thank you, Amore, for being part of this journey, for your understanding, and for always being there for me. I dedicate this work to you, Isabella and Alessia, hoping you will grow fond of learning, and to wish you to achieve this and much more.

Acknowledgements

I would like to extend my gratitude to my major professor, Dr. Joseph S. DeSalvo, for his supervision, advice, and guidance. Thank you, Dr. DeSalvo, for all our conversations held throughout this effort, which at times spanned far beyond the scope of this work to embrace other aspects of research and practical life. Thank you Dr. Brad Kamp, Dr. Ramp Pendyala, and Dr. Gabriel A. Picone for agreeing to serve on my committee and for your assistance and insightful comments. Special thanks go to Dr. Josefina Ramoni and Prof. Gianpaolo Orlandoni for their review and feedback on the empirical models. Finally, I would like to thank Stephen L. Reich, my program director at the Center for Urban Transportation Research. Thank you, Steve, for your support and for allowing me flexibility in my schedule to accommodate for both work and study.

Table of Contents

vi

List of Tables

List of Figures

The interaction Between Urban Form and Transit Travel

Sisinnio Concas

Abstract

This study presents an analytical model of the interaction between urban form and the demand for transit travel, in which residential location, transit demand, and the spatial dispersion of non-work activities are endogenously determined. In this model, travel demand is considered a derived demand brought about by the necessity to engage in out-of-home activities whose geographical extent is affected by urban form. In a departure from the urban monocentric model, residential location is defined as a job-residence pair in an urban area in which jobs, residences, and non-work activities are dispersed. Transit demand is then determined by residential location, work trips, non-work trip chains, and goods consumption.

Theoretically derived hypotheses are empirically tested using a dataset that integrates travel and land-use data. There is evidence of a significant influence of land-use patterns on transit patronage. In turn, transit demand affects consumption and non-work travel. Although much reliance has been placed on population density as a determinant of transit demand, it is found here that population density does not have a large impact on transit demand and, moreover, that the effect decreases when residential location is endogenous. To increase transit use, urban planners have advocated a mix of residential and commercial uses in proximity to transit stations. In this study, it is found that the impor-

tance of transit-station proximity is weakened by idiosyncratic preferences for residential location. In addition, when population density and residential location are jointly endogenous, the elasticity of transit demand with respect to walking distance to a transit station decreases by about 33 percent over the case in which these variables are treated an exogenous.

The research reported here is the first empirical work that explicitly relates residential location to trip chaining in a context in which individuals jointly decide residential location and the trip chain. If is found that households living farther from work use less transit and that trip-chaining behavior explains this finding. Households living far from work engage in complex trip chains and have, on average, a more dispersed activity space, which requires reliance on more flexible modes of transportation. Therefore, reducing the spatial allocation of non-work activities and improving transit accessibility at and around subcenters would increase transit demand. Similar effects can be obtained by increasing the presence of retail locations in proximity to transit-oriented households. Although focused on transit demand, the framework can be easily generalized to study other forms of travel.

Chapter 1: Introduction

Motivation

Among the challenges posed by evolving trends in transport and land-use is pro-
viding a better explanation of the role of non-work travel in residential location decisions.
Greater mobility and a shift from monocentric to polycentric urban forms have substan-
tially increased non-work travel, further weakening the relevance of the classical com-
muting-based theory of residential location (Anas, Arnott, and Small 1998).

Although the transportation literature on non-work travel has grown in recent
years, it has largely done so without providing a generally accepted behavioral frame-
work. Recent attempts to unify the economic theory of urban residential location and
transportation highlight the relevance of non-work travel to residential location (Anas
2007; Ben-Akiva and Bowman 1998). Central to this endeavor is the notion that in
choosing a residential location, the household considers the pattern of non-work trips that
its members are likely to make from that residential location. Accessibility to non-work
opportunities is likely to be important and, for many households, perhaps more important
than accessibility to jobs. In addition, the extent to which households self-select into
communities that support their preferences for transportation and other amenities compli-
cates the effort to uncover causality between urban form and travel behavior.

In recent years, urban policies intended to reduce presumed negative externalities
associated with suburbanization have focused on reducing auto travel by manipulating

urban form to reduce trip frequencies and travel distances. Specifically, locating residences in proximity to businesses should, it is hypothesized, reduce travel distances because nearby destinations will be preferred to more distant ones. In addition, it is assumed that shorter distances provide added opportunities to link more destinations in a single trip chain (Noland and Thomas 2007).

The empirical work on the efficacy of such policies provides mixed evidence. This is so because the research is based on ad-hoc empirical specifications, lacking a formal behavioral framework that considers travel the result of activities planned and executed through space and time. It is the purpose of this dissertation to provide such a behavioral framework and test its implications empirically.

Critique of Previous Work on Transit Travel Behavior and Urban Form

The policies discussed above form what is currently defined as *transit-oriented development* (TOD). The underlying assumption of TOD is that increased public transportation will reduce auto travel. The effectiveness of TOD, however, depends importantly on individual self-selection to residential location. Thus, ignoring such idiosyncratic preferences toward residential location may lead to over-reliance on TOD by urban planners as well as to overestimation of its impact on travel behavior by empirical researchers.

Despite a significant amount of academic and practitioner-oriented research, the practice of choosing the right transit service to support desired development and the right development to support transit ridership relies on findings that no longer apply to the current urban landscape. Early studies estimated the housing and job densities necessary to support different transit modes (Pushkarev and Zupan 1977). Such studies did not con-

sider changes in urban structure, such as transit-oriented development, that have recently emerged. At the same time, the urban landscape has evolved from monocentricity, where the CBD is the predominant employment center, to polycentricity, where multiple employment centers characterize an urban area and where households can locate anywhere in an increasingly suburban environment. Employment decentralization, coupled with the increased relevance of non-work travel, has had a profound impact on the way transit responds to urban form, making the earlier studies obsolete.

Debate has shifted from the need to determine minimum density thresholds that support transit to the need to provide reliable information to guide decision makers about what mix of land-use policies would better promote transit use. In most previous work, density is treated as exogenous and is assumed not to be impacted by transportation system changes. It is now recognized that this approach is inadequate and that what is needed is an empirically estimable behavioral model conducive to generalization and applicability.

The bulk of previous research is empirically oriented. It uses multivariate techniques to estimate the effect of measures of travel behavior (commute length, vehicle-miles of travel, mode choice) on measures of residential and employment density, while controlling for travelers' demographic characteristics. These studies report the statistical significance, sign, and magnitude of the estimated coefficients. A statistically significant negative coefficient leads one to conclude that an inverse relationship exists between travel and density, that is, higher density leads to shorter commutes, fewer vehicle-miles of travel (VMT), or a shift from auto transportation to alternative modes, such as transit. The abundance of such studies has led to the conclusion that policy interventions to in-

crease density would reduce automobile use. These studies have undergone systematic criticism, however, mainly of their ad hoc specifications and failure to recognize that the relationship between urban form and travel might entail simultaneity and endogeneity.

Research Objectives

The objectives of this dissertation are to (1) define a theoretical model of the interaction between urban form and the demand for transit, in which residential location, transit demand, and the spatial dispersion of non-work activities are endogenously determined, and (2) to test the hypotheses of that model.

The research:

1. Controls for idiosyncratic preferences toward residential location to test the hypothesis that land-use characteristics affect non-work travel behavior.

2. Shifts the focus from monocentric measures of urban form to polycentric ones.

3. Utilizes a framework that better accounts for the influence of space on travel patterns, by shifting the focus from a single-purpose trip-generation analysis to one that accounts for scheduling and trip-chaining effects.

4. Accounts for the trade-off between commute time and non-work activities.

Outline of Remaining Chapters

Chapter 2 reviews the literature on the relationship between transit travel behavior and urban form. An analytical framework in which residential location and travel behavior are simultaneously determined is presented in Chapter 3. First, Chapter 4 describes the development of the dataset used in the empirical research. Then, the dataset is used to

4

test the relationships hypothesized in Chapter 3. This chapter also discusses the validity of the empirical work and identifies issues that might potentially affect a generalization of the findings. Chapter 5 concludes and provides direction for further research.

Chapter 2: Urban Form and Transit Travel Behavior, A Review of the Literature

Introduction

Non-work travel is the result of engaging in activities, other than commuting, through time and space. Consisting of travel for shopping, social and recreational activities, and family and personal errands, non-work travel accounts for almost 85 percent of all daily trips undertaken at the household level (BTS 2001). The latest statistics from the U.S. Department of Transportation National Household Travel Survey (NHTS) report that non-work travel now constitutes 56 percent of trips during the AM peak and 69 percent of trips during the PM peak, with a ten-year growth of 100 and 35 percent, respectively (NHTS 2007).

These trends in non-work travel follow closely the pattern of urban growth in the United States, consisting of residential and employment decentralization. Policy responses to the potential negative externalities associated with decentralization and its effects on land-use and travel behavior now include attempts to limit urban growth or to change its form. In particular, proponents of *neo-traditional* or *transit-oriented* development (TOD) advocate the idea that land use can be manipulated to serve congestion management, air quality, or other related transportation objectives (Cervero et al. 2004). The policies most often associated with reduced automobile dependence are mixed-land-use, high-density environments that reduce the distance between residence and non-work travel activities.

In the last decade, more than 50 empirical studies have examined the linkages between urban form and travel behavior. Crane (2000), Badoe and Miller (2000), and Ewing and Cervero (2001)summarize the most relevant empirical work published in the literature of transportation research. The bulk of this research is empirically oriented and based on the application of multivariate techniques that regress various measures of travel behavior (commute length, vehicle miles of travel, mode choice) on measures of residential and employment density, while controlling for the demographic characteristics of travelers. The abundance of these types of studies has led to the conclusion that policy interventions to increase density are capable of reducing automobile use (Burchell et al. 1998; Cao, Mokhtarian, and Handy 2006; Ewing 1997). These studies have undergone systematic criticism, however, due to their ad-hoc specifications and omitted variable bias problems, the latter due to the possibility that the relationship between urban form and travel might entail simultaneity and endogeneity (Badoe and Miller 2000; Crane 2000).

In this chapter, we review studies that look at the influence of transit on urban form, the influence of urban form on transit patronage, and the simultaneous relationship between transit and urban form. The intent is to provide a critical assessment of the various methodologies employed in these studies, their control for relevant factors associated with transit patronage, and the general validity of their findings.

We uncovered the following issues that to-date have been addressed but not completely resolved. In particular, it is widely recognized that we lack a behavioral framework that can be applied to empirical work and that is conducive to generalization of findings and applicability (Badoe and Miller 2000; Crane 2000; Ewing and Cervero 2001). Studies that related population and employment density to travel behavior are

monocentric-based and fail to account for the employment and residential decentraliza-

tion now characterizing the urban landscape.

We found that there has been a shift from the study of density threshold levels

that make transit cost-feasible to an analysis of the role of urban design and land-use mix

on transit usage, after controlling for density levels. The issue is no longer at what densi-

ty threshold it makes sense to supply transit. Instead, the issue is the set of policies af-

fecting urban design and land-use mix that best influences the spatial arrangements of

activity locations so that individuals are more likely to utilize transit. This shift is reflect-

ed in a growing number of studies that are dedicated to studying the relevance of transit

oriented development (TOD) policies in a context where households or individuals tend

to prefer certain urban settings to others. Not accounting for these inherent idiosyncratic

preferences prevents the unraveling of the true impact of TOD.

There is a lack of empirical work that examines the relationship between urban

form and travel behavior within an activity-based framework and that takes into account

the complexity of travel (i.e., that accounts for trip-chaining). Those studies that have

employed activity-based modeling have failed properly to account for endogeneity and

have disregarded spatial mismatch effects (Dong et al. 2006). Activity-based modeling, a

relatively new and growing field of research, is characterized by the recognition that tra-

vel is a derived demand, a focus on constrained patterns or sequences of behavior instead

of discrete trips and the interdependence of decisions usually made within a household

context (Jones, Koppelman, and Orfeuil 1990). This framework is better suited than

those previously used to analyze the impact of land-use on travel patterns, as it fully ac-

knowledges the presence of trip-chaining behavior. In this context, a *trip chain* is defined

as a sequence of trips that starts from home and/or ends at home. Sometimes called stop-making behavior, trip-chaining behavior in activity-based modeling describes the importance of multi-purpose trip-making rather than single-purpose trip-making. Numerous studies have examined trip-chaining or stop-making models using the frequency of stops on the way home and/or on the way to work as dependent variables (Bhat 1999; Chu 2003; Concas and Winters 2007; Shiftman 1998).

To date, no empirical work has been done that explicitly relates location to trip-chaining behavior in a context where individuals jointly decide location, the optimal trip chain, and the area of non-work activities, based on the optimal trade-off between commute time and non-work travel activities.

Boarnet and Crane (2001) recognize that there is no best way of organizing a literature survey of this subject area. We organize our survey around the alternative ways of viewing the relationship between transit and urban form: (1) studies that have examined the influence of urban form on travel behavior, (2) studies that have examined the influence of transit use on urban form, and (3) studies that have examined the simultaneous nature of the relationship between transit and urban form.

This literature review is not comprehensive because it ignores empirical work involving only anecdotal accounts or descriptive analyses without an analytical framework of any sort. Descriptive studies have the benefit of assessing actual behavior without the need to establish causal links. Such studies are limited, however, in providing any useful perspective or guidance in the development of a theoretical or analytical model. As such, these studies are not deemed relevant to the objectives of this research effort. An assessment and review of recent anecdotal studies has been informally conducted by Taylor and

Miller (2003). For the same reason, we do not summarize Transit Cooperative Research Program (TCRP) reports that discuss the impact of the built environment on physical activity (Committee on Physical Activity 2005) or report successful case studies of TOD projects (Evans et al. 2007).

The Demand for Travel and Urban Form

The demand for travel is a derived demand generated by the necessity to engage in activities that are located outside one's place of residence (Domencich and McFadden 1975; McFadden 1973). This recognition requires studying the determinants of the demand for out-of-home activities as well as the characteristics of the environment affecting the choice of one mode of transport over another. In this context, urban form affects the demand for travel in two ways. First, the location of employment affects the probability that an individual will choose a given mode, given its supply. Second, the spatial extent of goods, services, and activities affects mode use for non-work travel purposes.

The influence of urban form on travel behavior is complicated by the evolution of the built environment itself. Since the development of the monocentric urban model (Alonso 1964; Mills 1972; Muth 1969), the urban landscape has evolved into one where multiple employment centers characterize an urban area and where households can locate anywhere in an increasingly suburban environment. The empirical fact of polycentricity complicates the development of a theoretical model of the relationship of transit and urban form. Nevertheless, as we shall see, transit patronage is still assumed to be largely dependent on the presence of a major employment center although the literature is evolving in the direction of how best to supply transit services in a polycentric urban landscape (Casello 2007; Modarres 2003).

Throughout this review, the term *urban form* refers to various measures of land-use density and urban design. Land-use density encompasses both residential and employment densities, while the term *urban design* refers to both the characteristics and arrangements of land-uses that affect accessibility to both transit services and activity locations.

Existing Critical Literature Reviews

Crane (2000), Boarnet and Crane (2001), Badoe and Miller (2000), and Ewing and Cervero (2001)summarize empirical work on the relationship between transportation and urban form.

Crane (2000) describes research methods, data, and results by dividing empirical work into two main categories: ad-hoc studies and theoretically-oriented studies. His review focuses only on studies that use statistical techniques to uncover the relationship between travel behavior and urban form. Most studies he reviews find that higher density patterns are correlated with less car travel. Crane concludes, however, that these ad-hoc studies are typically difficult to generalize and lack sufficient robustness to be used as a basis for policy. Crane uses these findings to justify the development of a behavioral framework consistent with the microeconomic theory of demand for travel, as discussed in detail in the next section of this chapter.

Badoe and Miller (2000) review the empirical literature up to 2000 with the objective of pinpointing the shortcomings that lead to what are considered questionable and contradictory results. The analysis deals with studies of the relationship between land use and travel behavior at the macro (density) and micro (design) levels. In conformity with Crane's findings, Badoe and Miller uncover weaknesses either in the data used or the me-

thodology employed. For example, some studies work with variables aggregated into zones that are not homogeneous with respect to neighborhood design, land use, and socioeconomic characteristics; this increases data measurement errors. Other studies ignore relevant variables, such as measures of transit supply, thereby contributing to omitted variable bias.

Ewing and Cervero (2001) summarize more than 50 empirical studies up to the year 2000 that examine the linkage between urban form and travel behavior. They focus on presenting findings that, at a minimum, "make some effort to control for other influences on travel behavior (p. 870)." Their review does not cover papers that explicitly treat trip-chaining behavior because of a lack of empirical work relating trip chaining to land-use and design variables. They find that while trip frequencies are primarily a function of socioeconomic characteristics rather than a function of urban form, trip lengths are primarily a function of the built environment.

Studies Analyzing the Influence of Urban Form on Transit Patronage

The most relevant early work under this heading is Pushkarev and Zupan (1977). This publication presented "land-use thresholds" at which different types of transit become feasible. The methodology used single-equation ordinary least square (OLS) regression analysis. The choice of this method was dictated by the paucity of data available at the time as well as the desire to present results as nomograms. A nomogram is a graph with which one can find the value of a dependent variable given the values of two or more independent variables, with only the use of a straightedge. The nomograms were designed to facilitate a planner's choice of a feasible transit alternative, given values of current or expected density levels and other relevant variables.

The determinants of transit demand used by Pushkarev and Zupan were the size of the central business district (CBD), measured in non-residential floor space; the distance of a site from the CBD; and residential density. The study also accounted for socio-demographic characteristics affecting transit patronage, such as vehicle ownership levels, household size and income. In an update of their 1977 study, Pushkarev and Zupan (1982) examine the feasibility of fixed guide-way transit under the assumption that all work travel was to the CBD. This assumption would be quite restrictive today, given the multi-centered character of many metropolitan regions.

In a report for the Transit Cooperative Research Program (TCRP), Zupan et al. (1996) provide guidance on the land-use characteristics that could cost-efficiently support new fixed-guideway transit services. The authors find that, in a transport corridor, ridership rises exponentially with both CBD employment and employment density. They present separate models for light rail and commuter rail. For both models, the dependent variable is a natural log transformation of total daily transit boardings for 261 stations across 19 rail lines located in 11 cities.

Multicollinearity impairs the reliability of these estimates, as recognized by the authors. Determination of causality is also a problem, for the estimated elasticities merely support a direct relationship between transit patronage and population density. This causality problem, which affects most findings in this research field, is discussed in a later section of this study. Finally, the authors do not employ a model that accounts for inherent, unobserved region-specific characteristics that might affect the reliability of estimates.

Following Zupan et al. (1996), Kuby et al. (2004) examine the determinants of light rail transit ridership with a multiple regression model using weekday boardings for 268 stations in nine cities. For each city, five categories of independent variables accounting for land use and other factors are used. The authors assume that employment within walking distance of each station is the most important factor for work trips. The model also controls for the relevance of nearby airports and for city-specific unobserved effects that might affect weekly boarding, such as the presence of an international airport. The study finds that an increase of 100 persons employed within walking distance of a station increases boardings by 2.3 passengers per day while an increase of 100 persons residing within walking distance of a station increases boardings by 9.2 passengers per day. The study also finds higher residential population to be associated with higher weekly boardings and that the CBD variable is not statistically significant, indicating that centrality is no longer relevant in determining light rail ridership. This result could, however, be due to faulty test statistics produced by the high correlation between the model's measures of centrality and the CBD dummy.

Kuby et al. (2004) make some important improvements to the methodology of Zupan et al. (1996). First, instead of examining ridership at non-CBD stations, they capture the effect of the CBD on boardings by introducing a dummy variable for CBD location. Second, Kuby et al. include employment near non-CBD stations instead of just employment within the CBD. Third, they include accessibility to non-CBD stations. Zupan et al. (1996) compute distances from the stations to the CBD but ignore stations' accessibility to other stations. Finally, Kuby et al. use residential population in the CBD as an independent variable.

While the studies so far discussed use aggregate data, the increasing availability of disaggregate (or micro) data after about 1995 provides the opportunity to study travel behavior at the individual or household level. The availability of disaggregate data brought about a paradigm shift in travel behavior analysis.

Reilly and Landis (2002) provide an early use of micro data to study the relationship between urban form and travel. In a study of the 1996 San Francisco Bay Area Travel Survey (BATS96), they test the relationship between measures of urban form and mode choice. Using geographic information system (GIS) methods, they obtain small scale measurements of land-use diversity, intersection density, and average lot size. To obtain these measurements, they generate a map of the study area, subdivided into a set of 10,000 grid-cells of one square meter each, called rasters. Then they proceed to obtain land-use measurements, such as the number of transit stops within a grid-cell. The authors fit gross population density and the amount of residential land area at the census block level into the grid-unit level to compute density values. The results of a multinomial logit mode-choice model show that an increase in average density of 10 persons per hectare (about four persons per acre) within one mile of an individual's residence is associated with a 7 percent increase in the probability of walking or taking transit (p. 24). As in most of the studies reviewed, this study does not determine causality between urban form and travel behavior.

Studies Analyzing the Influence of Transit on Urban Form

Other research examines the influence that transit has on urban form. In this context, the vast majority study impacts on urban form in terms of changes in land values at the station-area level (Baum-Snow and Kahn 2000; Bollinger 1997; Bowes and Ihlanfeldt

2001; Cervero and Landis 1997; McDonald and Osuji 1995; Nelson et al. 2007; Zheng and Kahn 2008; Kahn 2007). Most of these studies also examine the economic benefits of rail systems at the regional and local level. Economic benefits may accrue because transit improves productivity, which increases regional product and income because of accessibility improvement.

TCRP Report 16 (TCRP 1995) finds that transit raises residential property values near stations. Furthermore, there is support that both CBD's and subcenters benefit from transit development at the station-area level. In the case of CBD's, transit development helps centers retain their dominance. In the case of subcenters, regional rail systems contribute to the decentralization of both population and employment. This evidence is provided by way of anecdotal case studies, not empirical investigations based on quantitative analysis.

In an in-depth analysis of the Bay Area Rapid Transit (BART) system, Cervero and Landis (1997) find that transit investment had localized impacts on land use that were limited to downtown San Francisco, Oakland, and a few subcenters. Some studies looked at gentrification effects associated with transit systems. For example, Kahn (2007), shows that access to transit in the form of "Walk and Ride" positively impacts the gentrification trend. Gentrification is a phenomenon where old, deteriorated neighborhoods go through a process of renovation leading to land-value appreciation. This is brought about by population cohorts sorting themselves out in residential clusters.

Bollinger and Ihlanfeldt (1997) provide a more sophisticated analysis of the impact of rail transit on economic development. They present a simultaneous equation model that accounts for the simultaneity between population and employment density in

16

proximity to rail stations in the area covered by the Metropolitan Atlanta Rapid Transit Authority (MARTA). Results indicate that MARTA has had no discernible impact on total employment and population around stations.

Studies Analyzing the Contemporaneous Relationship between Transit and Urban Form

Apart from the instances outlined above, the vast majority of empirical work on the relationship between transit and urban form considers this relationship as one in which the urban landscape influences, in a unidirectional fashion, transit supply levels. Existing critical literature reviews identify the shortcoming of this assumption, the results of which fail to account for any underlying unobserved endogeneity between urban form and travel.

There are a few empirical efforts that provide an explicit analytical framework based on clearly defined behavioral assumptions (Badoe and Miller 2000; Boarnet and Crane 2001; Boarnet and Crane 2001, 2001; Boarnet and Sarmiento 1996; Boarnet and Sarmiento 1998; Crane and Crepeau 1998b, 1998a; Moshe and Bowman 1998; Schimek 1996; Voith 1991, 1997). These analyses use either multiple regression analysis or discrete choice models, techniques that better account for the interrelationship between the built environment and travel behavior than the approaches of previous research. Next, we summarize those studies most relevant to our research.

Using the 1990 Nationwide Personal Travel Survey (NPTS), Schimek (1996) applies a multiple regression model that accounts for simultaneity between a household's pick of neighborhood density and the amount of travel. The model is specified as

$$V = f(\beta X) \tag{2.1}$$

$$D = f(V, \beta \boldsymbol{X})\tag{2.2}$$

where V is the number of vehicles per household, D represents vehicle use (measured by VMT or trips) per household, $\boldsymbol{X}$ is a vector of demographic and geographic characteristics, and β is a column vector of parameters to be estimated.

Schimek substitutes (2.1) into (2.2) to obtain a reduced form equation of (2.2) that he estimates by linear regression. Endogeneity arises between urban form variables and vehicle usage variables because these variables affect vehicle ownership levels and, in turn, vehicle ownership affects residential location. Endogeneity is controlled by using an instrumental variable (IV) regression with the following instruments: race (white and Hispanic), location of household within the New York City standard metropolitan statistical area (SMSA), a dichotomous variable indicating if a household is located within an SMSA of three million or more, and a dichotomous variable indicating if a household is located within an SMSA of one million or more. Schimek justifies race as an IV by arguing that race and urban form variables in $\boldsymbol{X}$ are linked by spatial and housing market discrimination. He acknowledges that these variables might violate the basic IV assumption of no correlation with the exogenous variables of the reduced form equation. He does not, however, perform any tests for exogeneity or over-identification restrictions. The results are indeed impaired by the choice of weak instruments, as these are correlated with the other exogenous variables. The model's results show that a 10-percent increase in density leads only to a 0.7 percent reduction in household automobile travel. By comparison, a 10-percent increase in household income leads to a 3-percent increase in automobile travel. The results are similar when vehicle trips are used as the dependent variable.

Boarnet and Sarmiento (1998) provide a more robust analytical framework, which has been adapted by Boarnet and Crane (2001) and Crane and Crepeau (1998b). Boarnet and Sarmiento address some of the shortcomings of previous work, namely the a priori specification of a behavioral model from which a series of hypotheses is tested. They specify a non work-trip demand function in reduced form, where trip frequencies and VMT are a function of a set of socio-economic characteristics, land-use factors, and costs of travel

$$N = f(p, y; S) \tag{2.3}$$

$$p = f(L) \tag{2.4}$$

where N is the number of non-work auto trips, p is trip cost, y is income, S is a vector of socio-demographic variables, and L is a vector of land-use characteristics.

The authors argue that the cost of travel is itself affected by land-use and that land-use is endogenous because individuals tend to cluster in residential areas based on idiosyncratic preferences for residential location. They formalize this assertion by adding two equations that relate land use, L, to residential location

$$L = f(ResLoc_i) \tag{2.5}$$

$$ResLoc_i = f(C_i, A_i) \tag{2.6}$$

where $ResLoc_i$ indicates individual residential location; C_i represents individual socio-demographic characteristics (essentially the same as S); and A_i represents characteristics of residential locations, such as amenities. In particular, A_i is a vector of IV's for L in a two-stage least squares (2SLS) regression of Equation (2.3).

Boarnet and Sarmiento expect the qualitative effect of each of the independent variables to be indeterminate. A set of neighborhood amenity variables is used as IV's: the

proportion of block-group or census-tract population that is black, the proportion Hispanic, the proportion of housing stock built before 1940, and the proportion built between 1940 and 1960. The authors justify their choice of these IV's based on evidence that neighborhood demographic composition and age of housing are determinants of residential location choice. They argue that these IV's are good instruments since they are correlated with land use but not with transport (VMT or trips) and are, thus, exogenous to the error term. Boarnet and Sarmiento conclude that there is limited evidence of an effect of land use on transportation behavior; the most important result is that land use is endogenous to transportation behavior.

Several issues, related to the IV's being employed, impair the validity of these results. In this kind of analysis, good IV's must be correlated with land use, but they must not be correlated with transportation. It is easy to show that race or minority status affects both location and transportation demand (Arnott 1998). Race is correlated with location because minorities' choice set is more constrained than that of whites. The same argument applies to minorities' transportation choice set as, for example, when race and income are determinants of auto ownership and, therefore, impact both trips and VMT (the dependent variable employed by the authors). We conclude that the IV's chosen by the authors are poor, even if they pass a test for exogeneity based on over-identification, as outlined in Wooldridge (2002).

Crane and Crepeau (1998b) introduce a set of trip-demand functions (they report the demand for auto trips) as a function of travel time and income derived from a Cobb-Douglas specification

$$a\left(p_a, y, \tau\right) = \frac{\alpha y \tau}{p_a} \tag{2.7}$$

20

where α represents a taste parameter; y represents income; τ represents land-use features, which serve as proxies for the cost of travel (time and distance); and p_a indicates the price of a trip. Travel time is equal to the ratio of trip length to travel speed (which are themselves choice variables).

The authors conduct the analysis at a disaggregated level with respect to travel choice and land use. Land-use data from the Census Bureau are merged with travel-diary data using Geographic Information System (GIS) techniques to match residential location with land-use data at the tract level. GIS visual inspection of the network within 0.5 miles of the household allows measuring the characteristics of street grids and the presence of cul-de-sacs (measures of design). Land-use characteristics enter the demand function as shift parameters. The empirical analysis examines the impact of cross-sectional changes of τ.

The problem with this approach is that travel distance and speed are both affected by land use and urban design, but the functions specified by the authors dismiss endogeneity between land use and travel demand. For example, the vector of time prices is a function of speed and trip length, but trip length is also a function of location and street design. The authors acknowledge the problem and run a 2SLS regression using instruments for the price variables, although without "satisfaction with the variables available in the data (p. 233)."

Greenwald and Boarnet (2001) use the preceding model with minor variations to assess the impact of land use on non-work walking trips, and Zegras (2004) applies it to the relationship between land-use and travel behavior in Santiago, Chile.

Voith (1991) analyzes transit-ridership response to fare levels. He models transit demand and transit supply in a context where changes in transit service affect residential location. In this model, the author assumes that changes in service affect location decisions around transit stations, which, in turn, affect transit demand and, recursively, transit supply. In an update of his earlier work, however, Voith (1997) concludes that, after controlling for prices and service attributes, demographic effects on transit are minimal.

Methodological faults affect other research that attempts to model transit demand and transit supply simultaneously. For example, although Taylor and Miller (2003) recognize the need to model demand and supply jointly to avoid misspecification issues, they provide a poorly specified model.

Inherent Complexity: Accessibility, Urban Design, and Self-Selection

In recent years, urban policies to reduce externalities associated with employment and residential decentralization have relied on influencing the choice and amount of auto travel by manipulating urban form. The rationale behind these policies is that car-travel reductions can be achieved by reducing trip frequencies and travel distances. Mixing residential and employment locations expands the choice set by clustering amenities, which reduces average travel distances because nearby destinations are preferred to more distant ones. Furthermore, offering increased public transportation choices further reduces auto travel.

Such policies drive the so-called *transit oriented development* (TOD) (Cervero et al. 2004) approach to land-use planning. An issue at the heart of TOD effectiveness, which has attracted the attention of transportation researchers, is individual self-selection to residential location. In other words, individual preferences for location if not explicitly

accounted for during empirical research, might lead to overestimation of the impact of TOD policies on travel behavior. Researchers have tested the effectiveness of TOD by examining aspects of the built-environment, such as the relationship between mixed land-uses (where residential and commercial land-uses are in close proximity) and accessibility measures to residential locations.

There is a recent vast and fast growing literature addressing whether or not urban form affects travel behavior and, if it does, then what is the structural formation of the linkage. Within this field of research, a topic that has been increasingly studied and debated is that of residential sorting or self-selection. This refers to the phenomenon that leads individuals or households to prefer a certain residential location due to idiosyncratic preferences for travel. In applied work, if residential self-sorting is not accounted for, findings tend to overstate the relevance of policies to impact travel behavior by changing the built environment.

Mokhtarian and Cao (2008) provide a comprehensive review of empirical work on residential self-selection. While this growing body of literature recognizes that unobserved idiosyncratic preferences for travel affect residential location decisions, there is still disagreement on how best to treat the most common consequence of not controlling for this problem, namely, the resulting omitted variable bias. The empirical treatment of omitted variable bias in the context of self-selection ranges from nested logit models Cervero,(2007) to sophisticated error-correlation models (Bhat and Guo 2004; Pinjari et al. 2007).

Cervero (2007) estimates the degree to which residential self-selection affects transit mode choice by using conditional probability estimates that control for idiosyn-

cratic preferences for location. He specifies a decision nest requiring the parameterization of two indirect utility functions, one function expressing residential location choice (specifically, residence within a mile of a rail stop) and one function expressing transit mode choice. Workplace proximity to a rail station, job-accessibility, and household and personal attributes are among the factors affecting location choice. He specifies the mode-choice indirect utility function to include a travel-time ratio (transit vs. auto), vehicle stock, personal attributes, and neighborhood density. The results of his analysis show significantly higher transit ridership shares associated with transit-oriented living, but that residential self-selection might account for about 40 percent of such shares.

Two issues related to the modeling technique and the choice of the observational unit cast doubt on the possibility of generalizing these findings. First the residential location utility function, although controlling for accessibility and socio-demographics, does not include any controls for neighborhood and housing characteristics. Second, it is not clear if the observational unit of analysis is the household or the individual (the subscript n in equation 1 on page 2,077 refers to the individual, but page 2,078 refers to a household). The implications of modeling household versus individual residential choice are non-trivial. For example, in a two-member household, even after controlling for household characteristics, the first person might have a transit stop near his or her work location, while the second person might not. This results in a different travel-time ratio (a control in the lower level mode-choice utility function). When estimating the nested logit regression, the predicted probabilities of residential location might differ, assigning the first person to the predicted choice of "near transit station" and the second person to the

predicted choice of "far." This results in having two individuals within the same household living at different locations.

Following the latest applications of discrete mode-choice modeling developed by Bhat and Guo (2004), Pinjari et al. (2007) propose a model of joint determination of residential location and mode choice where both choices influence each other by accounting for observed and unobserved individual taste heterogeneity. Findings suggest that, after accounting for self-selection, the built environment has an impact on commute mode-choice behavior.

The authors present two indirect utility functions, one describing mode choice and one defining residential location. The two functions are related by way of an error-term specification. They capture self-selection endogeneity by controlling for both observed and unobserved factors impacting residential location and commuting-mode choice. First the mode-choice indirect utility function (*indirect* here means that the function depicts a realized choice that reflects the primitive objective function; it is not the indirect utility function of economic theory) includes a term indicating observed socio-demographic factors influencing the mode-choice decision. Then, an unobserved term is added to capture taste heterogeneity linked to the location decision but affecting mode-choice. This takes the form of an error term that is correlated to the second indirect utility function related to location choice. A final independent and identically distributed error term is added to the equation. The second indirect utility equation works the same way, with an error term correlated with the mode-choice utility function.

The main issue with this methodology is related to the claim of simultaneously determining mode-choice and location. This approach prompts the question "is the

mode-choice decision really simultaneously determined with the location decision?" The authors seem at first to state this hypothesis, then, later, to refute it (p. 564) by admitting that, "The model structure assumes a causal influence of the residential location choice (and hence the built environment) on commute mode choice." This apparent contradiction is probably justified by the specific econometric approach that they take. Specifically, they assume that individuals simultaneously maximize two different, although interdependent, utility functions, subject to somewhat different constraints. As in the case of Cervero (2007), this problem is the result of ad-hoc specifications of indirect utility functions without knowledge of the primitive objective functions, as discussed by Jara-Díaz and Martinez (1999).

Another problem in the study of self-selection arises when residential choice is modeled as a discrete variable. The treatment of the location decision as a dichotomous variable inherently presents a problem that is at the very heart of residential self-selection research. When using discrete choice modeling, one must assume that all individuals can choose among all possible locations within an urban area. The treatment of mode choice and residential location in more sophisticated frameworks does not eliminate the need to determine ad hoc the residential choice set. For example, both Pinjari et al. (2007) and Bhat and Guo (2004), who adopt the more sophisticated multinomial logit-ordered structure that explicitly considers the correlation of unobserved factors simultaneously affecting both choices, must determine a priori the location choice set (in that case, any individual is assumed to be able to choose among 223 different locations). This assumption does not explicitly acknowledge that, subject to income and vehicle availability, some individuals have more constrained mode choices and residential location sets, with the

undesirable effects described by spatial mismatch theory (Kain 1968). The result is not being able to fully discern the influence of idiosyncratic preferences for location on residential choice from issues related to spatial mismatch.

An alternative to treating residential location as a discrete choice is instrumental variable regression that uses a set of properly tailored instruments, with leading examples discussed earlier (Boarnet and Sarmiento 1998; Crane 2000). Other researchers advocate the use of simultaneous equation modeling (SEM), where additional equations are added to account for simultaneity between urban form, attitudes toward travel, and other factors. Researchers justify preference for the latter approach on the basis of its capability to uncover causality between travel and urban form.

In many instances, research efforts that claim to uncover causality between urban design, travel behavior, and individual self-selection do not make appropriate use of the econometric techniques therein employed. Data constraints also affect the usefulness of this statistical technique. For example, while Bagley and Mokhtarian (2002), Handy et al. (2005), and Cao et al. (2007) discuss the advantages of SEM, assuming the availability of longitudinal data, they all use the same cross sectional dataset that employs a mix of secondary data and primary data from a travel attitude survey (the authors define this dataset as *quasi-longitudinal*). Furthermore, in the context of simultaneous equation modeling or instrumental variable regression, the validity of results hinges on the determination of the exclusion restrictions. That is, the researcher must determine a priori what explanatory variables are to be included and excluded from a given equation. The determination of the exclusion restrictions determines a model that is correctly specified in the sense that the matrix of the reduced form parameters to be estimated is unique in its re-

presentation of the more primitive structural matrix. Exclusion restrictions need to be drawn outside of the variables a researcher has available from a given dataset, i.e., they should be based on sound behavioral theory (Wooldridge 2002).

In all studies of residential self-selection employing SEM techniques previously reviewed, including the work of Bagley and Mokhtarian (2002), Handy et al. (2005), and Cao et al. (2007, 2006), there is no explicit treatment of the exclusion restrictions that can be traced back to a formalized theoretical framework.

An alternative approach is presented by Vance and Hedel (2007), who employ a two-part model consisting of probit and OLS estimation, using the German Mobility Panel survey (MOP 2006). In the first part of the model, a probit model that controls for socio-demographic factors (income, age, driving license) and urban form (commercial density, street density, commercial diversity) estimates the probability of owning a vehicle. The second stage, a regular OLS model, conditional on the first-stage predicted vehicle ownership, regresses vehicle use (distance traveled) on a vector of socio-demographic and urban form variables. The model is further enhanced by instrumenting the urban form variable using the set of instruments suggested by Boarnet and Sarmiento (1998). Although instrument validity is checked against exogeneity by applying selected diagnostic tests, the choice of instruments is limited to a set of controls for housing characteristics without the inclusion of neighborhood characteristics controls to capture a broader set of factors affecting residential location choice.

Measuring Accessibility

Accessibility measures are widely used in transportation planning to relate the pattern of land use and the nature of the transportation system. Various measures have

been employed when analyzing the efficacy of mixed land use or transit-oriented poli-

cies. A problem related to the use of accessibility is that its measurability is inherent in

its definition and quantification. For example, one definition is "the ease and conveni-

ence of access to spatially distributed opportunities with a choice of travel" (DOE 1996).

Obviously, the main difficulty is to quantify the ease of accessibility. We now turn to a

discussion of the most widely used measures of accessibility.

As recently summarized by Dong et al. (2006), there are essentially three meas-

ures of accessibility that have been employed to date: isochrones, gravity-based meas-

ures, and utility-based measures. The most widely employed are the gravity-based meas-

ures, which have the following generic form

$$Acc_i = \sum_j a_j f(c_{ij}) \tag{2.8}$$

where Acc_i means accessibility to zone i; j indexes the available destination zones that

can be reached from zone i; a_j measures the activity opportunities in zone j; and $f(c_{ij})$

represents an impedance, or decay, function of traveling from zone i to zone j. This trip-

based measure has been used in the recent work of Maat and Timmermans (2006), one of

the few studies examining the influence of land use on activity-based travel.

As pointed out by Dong et al. (2006), this measure is limited in that it neglects he-

terogeneity of preferences across individuals, which can lead to absurd conclusions, e.g.,

"a gravity measure of this type says that a retired grandfather and his college student

grandson who live together have identical values of accessibility (p.165)." Furthermore,

this measure is highly sensitive to the specification of the decay function.

All of the models showing a relationship between increased transit usage and im-

provement in accessibility rely on one of the above measures. We think that analyzing

the complexity of accessibility and travel behavior requires the use of accessibility measures that are strictly linked to the way activities are organized. These measures should be selected based on the relationship with the observed activity pattern.

Some attempts are now appearing in the literature, although not directly related to the field of transportation research, that take into account individual heterogeneity and preferences. For example, utility-based measures of accessibility, which are based on the random utility theory as originally exposited by Domencich and McFadden (1975), provide ways of relating accessibility measures to the characteristics of the alternative and the characteristics of the individual. The activity-based accessibility measure introduced by Dong et al.(2006), for example, is a utility-based measure. This measure is capable of capturing taste heterogeneity across individuals, combining different types of trips into a unified measure of accessibility, and of quantifying differing accessibility impacts on diverse segments of the population.

Urban Form Measures and Polycentric Cities

As discussed in the introduction to this chapter, another problem of empirical analyses of the relationship between travel and land use is the adoption of measures of urban form that are monocentric. Monocentric models only consider measures of the strength of the relationship between central business district (CBD) employment (and other activities located at the CBD) and travel behavior. For example, in their seminal work, Pushkarev and Zupan consider the relationship between transit service and density in a context where the CBD is the main determinant of transit trips. More recently, Bento et al. (2005) examine the effects of population centrality, jobs-housing balance, city shape, road density, and public transit supply on the commute-mode choices and annual vehicle-

miles of travel of households living in 114 urban areas in 1990. They found that the probability of driving to work is lower the higher the population centrality and rail miles supplied and the lower the road density. Road density, in this model, is defined as miles of road multiplied by road width (for different categories of road) and divided by the size of the urbanized area.

In recent decades the process of decentralization has taken a more polycentric form, with a number of clustered employment centers affecting both employment and population distributions. The majority of these centers is subsidiary to an older CBD. Such centers are usually called subcenters or sub-regional centers. McMillen (2001) suggests a more formal definition by defining a subcenter as a "site with (1) significantly larger employment density than nearby locations that has (2) a significant effect on the overall employment density function (pp.448–449)."

The transportation literature has seldom examined the influence of subcenters on travel behavior. An exception is Cervero and Wu (1998), who study the influence of subcenters on commute distances in the San Francisco Bay area. They conclude that employment decentralization has led to increased vehicle travel. These studies generally consider subcenters as exogenously determined either by assumption or by an empirical determination that makes use of specific density thresholds.

More recent developments in travel demand behavior and geographical science provide some insight on how better to capture the relationship between urban form and travel in a highly decentralized context. For example, Modarres (2003) proposes the use of GIS to determine subcenters using spatial clustering techniques to cluster patterns of major employers. He then considers the relevance of transit accessibility within the iden-

tified subcenters (accessibility is defined as the level of service provided by existing routes in each census tract) and concludes that spatial accessibility is high within these subcenters. Casello (2007) discusses the potential to increase and the impacts of increasing transit modal split in the polycentric metropolitan area of Philadelphia. By identifying "activity centers," i.e., areas where transit use is likely, he models transit competitiveness and system performance. Kuby et al. (2004) update and improve previous research and find that the same factors affecting CBD boardings also influence non-CBD (subcenter based) transit ridership.

The decreasing relevance of the CBD with respect to transit patronage is illustrated by its statistical insignificance in determining transit usage in the recent work of Brown and Neog (2007) and Thompson and Brown (2006). In particular, Brown and Neog examine aggregate transit ridership in 82 U.S. metropolitan statistical areas (MSA) using data from the National Transit Database as provided by the Florida Department of Transportation Transit Information System (FTIS). The authors use employment in the CBD and total metropolitan employment as proxies for urban form explanatory variables in a series of multivariate regression models. They find that transit ridership is not affected by the strength of a CBD, suggesting that improvement in ridership can be obtained by better serving decentralized urban areas.

These findings are supported by Brown and Thompson (2008), who employ a time series analysis of aggregate ridership data of the Metropolitan Atlanta Rapid Transit Authority (MARTA) in Atlanta, Georgia. The authors define two employment decentralization measures: number of employees within the MARTA service area located outside the Atlanta CBD (variable EMPMARTA) and the ratio of employment outside the

32

MARTA service area to employment inside the MARTA service area (variable RA-TIO_EMP). They specify a first difference autoregressive model with annual linked passenger trips per capita as the dependent variable as a function of transit supply measures and the above-mentioned decentralization variables. Results show that model performance is affected by inclusion of a time trend variable, as reflected by the standard error estimates of the variable RATIO-EMP across the two models' specifications. Notwithstanding these econometric issues, the authors conclude that there exists a positive association between decentralized employment growth (served by transit) and transit patronage.

Although these conclusions favor policies geared at servicing employment rather than population concentrations, a generalization of these findings to other spatial context is not warranted. The lack of relevance of the Atlanta CBD is due to the peculiar spatial characteristics that make it unique with respect to the rest of the U.S., and the world, as argued by Bertaud (2003). By comparing Atlanta's spatial arrangement of population and employment to other U.S. and world cities, Bertaud shows that the uniqueness of Atlanta (being highly polycentric) makes a supply-side policy cost-infeasible. In particular, Bertaud shows that with only 2 percent of the total jobs located at the CBD and only 8 percent within 5 kilometers of the city center, Atlanta's dispersion of employment would require the addition of about 3,400 kilometers of metro tracks and about 2,800 new metro stations to provide the same transit accessibility to a comparable, although monocentric-based, city, requiring only 99 kilometers of tracks and 136 stations. Bertaud uses these findings to justify congestion tolling and the provision of small, niche-type transit services to control the negative externalities usually associated with sprawl.

In summary, the literature provides contrasting results on the relevance of the CBD to the demand and supply of transit services. The strength of the CBD is conditional on the spatial characteristics of the neighboring suburban areas.

From Trip Generation to Activity-Travel Behavior

In examining the relationship between travel behavior and urban form, the literature reviewed above rarely accounts for the fact that the demand for travel is an indirect demand, which arises from the necessity to engage in activities requiring travel. The recognition that travel patterns are complex and that trips are the result of a decision-making process in which activities are organized and prioritized through space and time has led to what is generally considered a paradigm shift in the study of urban travel behavior (Pass 1985). This paradigm shift has paved the way for a new field of research, defined as *activity-based modeling*. Activity-based modeling is characterized by the recognition that travel is a derived demand, a recognition that shifts the research focus from single trips to trip chains and from individual decision making to household members' interdependent decision making (Jones, Koppelman, and Orfeuil 1990). Activity-based approaches are currently used to describe the activities individuals pursue, at what locations, at what times, and how these activities are scheduled within a transportation network characterized by opportunity and constraints (Bhat and Koppelman 1999).

Essentially, a trip chain may be defined as a sequence of trips that starts from home and/or ends at home. Different taxonomies defining trip-chaining complexity are possible depending on the purpose or mode of the trip for different classes of travelers. Sometimes called stop-making behavior, trip-chaining behavior in activity-based model-

ing describes the importance of multi-purpose trip-making rather than single-purpose trip making.

Numerous studies have examined trip-chaining or stop-making models using the frequency of stops on the way home and/or on the way to work as dependent variables (Bhat 1999; Chu 2003; Concas and Winters 2007; Shiftman 1998). In these studies, stop-making behavior describes stopping behavior made by a traveler, in particular a commuter, on the way to home or work. Under the assumption that a commuter follows a regular route, then stopping at a location other than home or work during the commute is treated as a deviation from the commute trip. In prior research, stop-making models were usually applied to trips linking non-work activities with work activities, including the morning commute, midday trips, evening commute, and trips before or after the commute.

The analysis of travel behavior within this context allows the recognition that trips are interrelated as opposed to the current transportation planning modeling assumptions of separation and independence of trips by purpose. Models based on microeconomic theory that explicitly treat the trade-offs involved in the choice of multiple-stop chains (i.e., the linking of several out-of-home activities and related trips into one tour) first appeared in the 1970's (Adler and Ben-Akiva 1979). In addition to work trips, non-work trips have also been investigated, where non-workers' trip-chaining is a series of out-of-home activity episodes (or stops) of different types interspersed with periods of in-home stays (Misra and Bhat 2002; Misra, Bhat, and Srinivasan 2003).

Although travel-demand forecasting models are now starting to incorporate trip-chaining behavior, only a limited number of studies exists that link the different aspects

of trip-chaining behavior (trip-tour frequency, complexity, duration) and urban form. There are some studies that relate trip chaining to regional accessibility or that compare trip-chaining behavior across regional subareas, for example, city versus suburbs, as summarized by Ewing and Cervero(2001). Maat and Timmermans (2006) represent a recent effort to examine the influence of land-use on trip-chaining behavior (by way of analyzing tour complexity). There is some research attempting to integrate activities and residential location by using discrete choice models of household residential location and travel schedules (Ben-Akiva and Bowman 1998).

We find to date no empirical work explicitly relating location to trip-chaining be-havior in a context in which individuals jointly decide location, the optimal trip chain, and the area of non-work activities, based on the optimal trade-off between commute time and non-work travel activities. We think that better insight on the relationship be-tween urban form and travel behavior would be gained by testing the hypothesis that an individual's residential location is based on utility maximizing behavior.

Summary and Implications for Integrated Models of Transportation and Land Use

The bulk of research reviewed in this chapter is empirically oriented and based on the application of multivariate techniques that regress various measures of travel behavior (commute length, vehicle-miles of travel, mode choice) on measures of residential and employment density, while controlling for the demographic characteristics of travelers. These studies examine the statistical significance, sign, and magnitude of the estimated coefficient on residential population density or employment density. A statistically sig-nificant negative coefficient leads one to conclude that a negative relationship exists be-tween travel and density. For example, higher density leads to shorter commutes, fewer

36

vehicle-miles of travel (VMT), or a shift from auto transportation to alternative modes, such as transit. The abundance of these types of studies has led to the conclusion that policy interventions directed to influence density are capable of reducing automobile use.

The literature review uncovered the following issues that, to date, have been addressed but not completely resolved. In particular, it is widely recognized that there is a lack of a behavioral framework that can be applied to empirical work and is conducive to generalization of findings and applicability. Studies that relate density (population and employment) measures to travel behavior are monocentric and, therefore, fail to account for the employment and residential decentralization now characterizing the urban landscape. In most of this work, density is treated as exogenous and is not assumed to be impacted by transportation system changes. These studies have undergone systematic criticism due to their ad-hoc specifications and because of omitted variable bias problems due to the possibility that the relationship between urban form and travel might entail simultaneity and endogeneity. In addition, most of the work that jointly estimate transit demand, transit supply, and factors affecting both supply and demand are affected by methodological faults, ranging from misuse of simultaneous equation modeling methods to improper functional specifications.

More recent developments in travel demand behavior and geographical science provide some insight on how better to capture the relationship between urban form and travel in a highly decentralized context. The significance of the CBD in determining transit ridership levels has been revisited and more relevance is now attributed to decentralized employment by examining the influence of subcenters in an increasingly poly-centric urban landscape.

While early work sought to provide a generalized analytical framework that made use of aggregate data, the more recent literature consists of papers that model the simultaneous decision of location and travel (as an application of improved discrete-choice modeling techniques) in a context where individuals choose locations based on specific travel preferences (for example, a preference about a specific mode) at the disaggregate level. Location decisions based on idiosyncratic preferences for travel define the term "residential self-selection behavior" to indicate how individuals with similar tastes and preferences tend to cluster together in given locations.

Finally, there is a lack of empirical work that studies the relationship between urban form and travel behavior within an activity-based framework, which takes into account the complexity of travel (i.e., that accounts for trip chaining). Those studies that have employed activity-based modeling have failed to properly account for endogeneity and have disregarded spatial mismatch effects. In examining the relationship between urban form and travel, it is crucial to distinguish the effects of land use from the effects of systematic socio-demographic differences of individuals.

It is the purpose of this dissertation to provide an estimable model for these failings of previous research.

Chapter 3: Methodology

Introduction

The objective of this chapter is to develop an empirically testable model of the relationship between transit travel behavior and urban form. Following the methodological issues highlighted by the literature review, the proposed framework seeks to address unresolved issues as follows:

- It controls for individual idiosyncratic preferences for residential location

- It shifts the focus from monocentric-based measures of urban form to polycentric ones

- It utilizes a framework that better accounts for the spatial influence on travel patterns, by shifting the focus from a single-purpose trip-generation analysis to one that accounts for trip chaining

- It accounts for the trade-off between commute time and non-work activities

In this model, travel demand is considered a derived demand brought about by the necessity to engage in out-of-home activities whose geographical extent is affected by urban form. Furthermore, budget-constrained utility-maximizing behavior leads to an optimization of the spatiotemporal allocation of these activities and an optimal number of chained trips. Socio-demographic factors directly influence residential location, consumption, and travel behavior. To date, no empirical work has been done that explicitly

relates location to trip chaining behavior in a context where individuals jointly decide location, the optimal trip chain, and the area of non-work activities, based on the optimal trade-off between commute time, leisure, and non-work travel activities and accounts for the other methodological problems noted above.
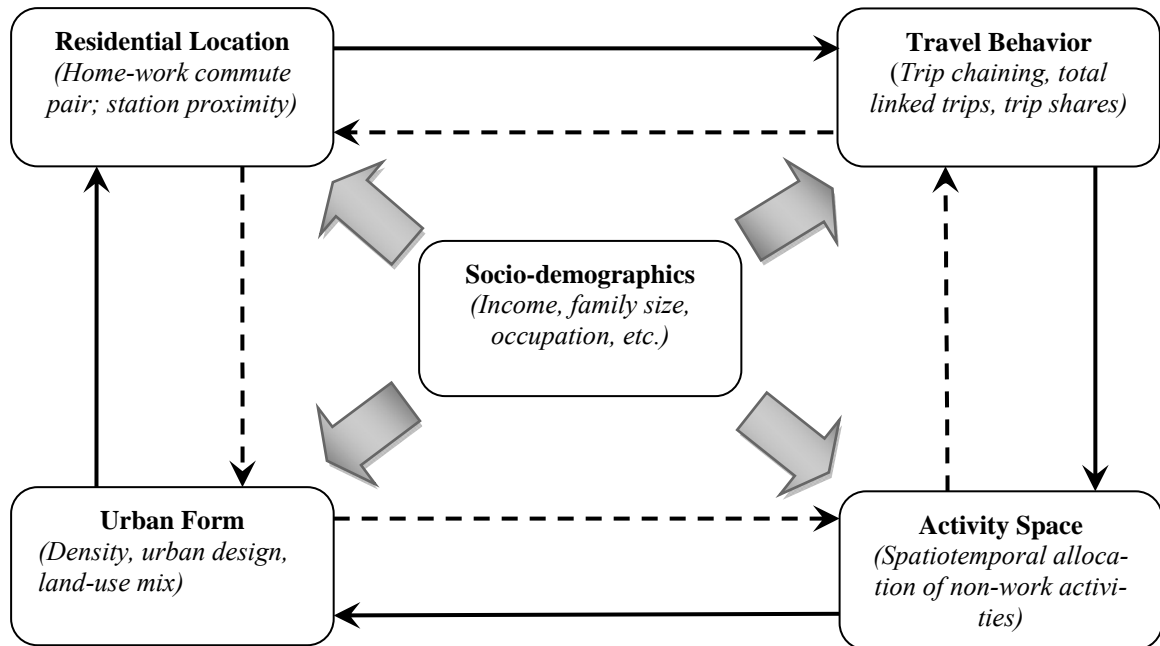
FIGURE 3.1 Conceptual Model of Urban Form and Travel Behavior

In this model, residential location, travel behavior, the activity space, and urban form are all endogenously determined. Following urban residential location theory, the location decision is assumed to be the result of a trade-off between housing expenditures and transportation costs, given income and the mode-choice set. In a departure from the monocentric model, the definition of residential location is taken from the polycentric model of Anas and his associates (Anas and Kim 1996; Anas and Xu 1999). In this work, residential location is defined as the optimal *job-residence pair* in an urban area in

which jobs and residences are dispersed. Following Anas (Anas 2007), the location decision is also based on idiosyncratic preferences for location and travel. In addition to determining optimal residential location, this approach also determines the optimal sequence of non-work trip chains, goods consumption, and transit patronage. It is within this framework that questions related to the interrelation between urban form, residential location, and transit travel demand are addressed. How do location decisions affect travel behavior? How does urban form relate to travel behavior? Do residential location and urban form affect travel behavior? What is the impact of higher density on travel behavior? To address these questions, we first introduce a basic travel demand model treating residential location and density as exogenous (Model I). We then consider subsequent extensions (Model II and Model III) that relax these assumptions to discuss what expected behavioral conclusions can be reached. This chapter presents the most relevant results of the comparative static analysis, while the complete derivation of the comparative statics and the necessary assumptions to carry them out are detailed in Appendix A.

Model I: Exogenous Residential Location and Density

In this specification, residential location, transit station proximity, and density are exogenous. Given these variables, the model jointly defines the activity space and the optimal trip chain. The joint determination of activity space and trip chain determines a travel demand function, given consumption and location decisions. The household (rather than the individual) is the unit of analysis because these decisions take place at the household level. Empirical studies on the relevance of transit station proximity to transit patronage show a strong relationship between transit use and station proximity (Cervero

2007; Cervero and Wu 1998). Therefore, this model includes this possibility. To include these considerations, Model I takes the following specific form

$$TC = TC(AS, RL, WD, X_{TC}) \qquad (3.1)$$

$$AS = AS(TC, D, X_{AS}) \qquad (3.2)$$

$$TD = TD(TC, AS, RL, WD, X_{TD}) \qquad (3.3)$$

where

TC = the number of non-work trip stops per commute-chain or chain length

AS = the activity space (measured as the geographic area surrounding the residence in which non-work trips are made)

TD = the demand for transit trips (measured as the number of transit trips)

RL = residential location (measured as the *job-residence pair* distance)

D = a vector of residential and employment density controls

WD= transit station proximity (measured as walking distance to the nearest transit station)

X_{TC} = a vector of controls specific to the TC function;

X_{AS} = a vector controls specific to the AS function

X_{TD} = a vector of controls specific to the TD function

This model permits testing the hypothesis that individuals living farther from the workplace engage in more complex tours characterized by a higher number of non-work trips linked to the commute tour. As in Kondo and Kitamura (1987), the number of non-work trip stops, TC, determines the length of the trip chain. In addition, TC, as it relates to transit patronage, is directly affected by transit station proximity and by other factors summarized by the vector of controls, X_{TC}. This vector, as explained in more detail in

Chapter 4, includes vehicle availability and the presence of young children among other factors likely to affect trip-chaining formation.

Trip-chaining behavior defines an activity space, *AS*, which is assumed to represent the optimized spatiotemporal allocation of non-work activities as affected by the built environment, summarized by the exogenous vector, *D*. For example, more densely populated urban areas have more densely clustered activity locations, which shrink the size of the activity space relative to less densely populated areas. A smaller activity space reduces trip chaining, *TC*, ultimately affecting the demand for travel, *TD*. As we shall see, *AS* captures the characteristics of activity locations as well as the spati-otemporal constraints linked to trip-chaining behavior.

This model is suited to either describe a situation where residential location is considered as predetermined, such as a short run time frame or can be used to cross compare decision making among households at any point in time. The model may be used to test the effect of urban design policies directly affecting travel distances and the land-use mix. Specifically, it may be used to test if higher density environments entail shorter travel distances, which in turn should affect the composition and complexity of trip chains and the overall amount of travel.

Residential Location, *RL*, and Transit Station Proximity, *WD*

The definitions of residential location and transit station proximity used here differ from those used in the current literature. For example, in studies of residential self-selection, the location decision is often presented as a dichotomous choice, i.e., whether to live near or far away from a transit station. Proximity is measured by a circular buffer around a station, often with a half-mile radius. The extent of this buffer is usually justi-

fied on empirical grounds. Cervero (2007), for example, used a half-mile radius in estimating a nested logit model of the joint determination of mode and location. This measure of transit proximity fails to account for barriers that prevent access to a station that lies within the half-mile radius. Some researchers have considered residential location as a choice to reside within a geographical unit, such as a traffic assignment zone (Bhat and Guo 2004; Pinjari et al. 2007).

The use of transit proximity as a proxy for residential location, while dictated by the need to sort out the influence of the built environment from self-selection, is not based on any other theoretical underpinnings about the decision-making process that is at the heart of urban residential location theory. That is, it does not take into consideration the trade-off between housing and transportation costs that, at the margin, determine where an individual decides to locate. For example, the standard theory of location shows that individuals choose an optimal distance between work and home given housing and transportation costs. In a monocentric model that only looks at travel between home and the CBD, individuals locate at a distance where the marginal cost of transportation is equal to the marginal housing cost savings obtained by a move farther from the CBD (Alonso 1964; Muth 1969). Recent departures from this view consider that individuals can locate anywhere in an urban area, choosing an optimal home-work distance that optimizes also the amount of non-work travel and non-work activities (Anas and Kim 1996; Anas and Xu 1999). Further explorations also consider the role of trip chaining behavior (Anas 2007).

Activity Space: Spatial Dispersion of Non-Work Activities

The concept of activity space, although not new to behavioral sciences, is novel in terms of its application to travel behavior. The relationship between urban form and geographical patterns of activities has been studied only recently, due to the availability of specialized travel diary data and increasingly sophisticated geospatial tools. A growing field of research that looks at the relationship between urban form and the spatiotemporal allocation of activities and travel provides additional insight on the impact of the built environment. Recent research describing travel behavior and the influence of urban morphology and entire patterns of daily household activities and travel demonstrates how households residing in decentralized, lower density, urban areas tend to have a more dispersed activity-travel pattern then their counterpart residing in centralized, high density urban areas (Buliung and Kanaroglou 2006; Maoh and Kanaroglou 2007).

This study explicitly accounts for the influence of the built environment in affecting the spatial dispersion of activities and how spatial dispersion affects the demand for travel and location decisions. This effect is accounted for by introducing the variable *activity space*, *AS,* into the model. The extent of the activity space is assumed to be affected by the built environment. Densely populated urban areas tend to cluster activity locations together thus shrinking the size of the activity space. This affects the spatial allocation of activities, thus affecting the demand for travel. As seen in the next chapter, there exist several ways empirically to measure the spatial dispersion of activities.

Trip Chaining, *TC*

According to activity-based modeling practice, trip chaining describes how travelers link trips between locations around an activity pattern. In this context, a trip from

45

home to work with an intermediate stop to drop children off at day care is an example of a trip chain. In the literature there is not a formal definition of *trip chain,* and different terms and expectations exist as to what kind of trips should be considered as part of a chain (McGuckin and Murakami 1999). Sometimes, the term *trip chain* is used interchangeably with the term *tour* to indicate a series of trips that start and end at home.

In this study, we hypothesize that trip chaining occurring on the home-job commuting pair saves time. These time savings in turn can be either allocated to additional non-work travel, thus increasing the overall demand for travel (e.g., total number of trips), or be used to determine a longer commute (i.e., a home-job commuting pair farther apart). The hypothesis of increased discretionary travel due to trip-chaining has recently been theoretically demonstrated (Anas 2007). The hypothesis of a positive relationship between more complex trip chains and the home-work commute is confirmed by empirical work. For example, in an analysis of trip chaining involving home-to-work and work-to-home trips using data from the 1995 nationwide personal transportation survey (NPTS), McGucking and Murakami (1999) found that people are more likely to stop on their way home from work, rather than on their way to work. About 33 percent of women linked trips on their way to work compared with 19 percent of men, while 61 percent of women and 46 percent of men linked trips on their way home from work. Using the 1991 Boston Household Travel Survey, Bhat (1997) found that about 38 percent of individuals made stops during the commute trip. Davidson (1991) found similar results from her analysis of commute behavior in a suburban setting, showing that travelers rely heavily on trip chaining in an urban context characterized by higher spatial dispersion of non-work activities. Other studies also provide empirical evidence of increased stop-making

46

during the commute periods (Bhat 2001) or how the ability to link trips is enhanced by the flexibility inherent in automobile use (Strathman 1995).

Travel Demand, *TD*

Travel demand is herein treated as a derived demand brought about by the need to purchase goods and services. Travel demand, *TD*, measures the number of work and non-work transit trips at the household level. The decision process behind the choice of the number of trips, as formalized by this framework, considers trip generation as a function of trip chaining and exogenous residential location and socio-demographic factors. The constrained maximization problem of the joint determination of activity space and trip-chaining defines an optimal vector of non-work trips, given residential location and urban form characteristics (e.g., residential and employment density levels, land-use mix). This treatment of travel demand as derived from the desire to engage in out-of-home activities departs in terms of behavioral sophistication from the treatment of trip generation as de-veloped by Boarnet and Crane (2001) in their analysis of travel demand and urban de-sign. In Boarnet and Crane (2001) trip demand functions are either directly affected by land use or indirectly (by influencing the cost of travel).

In contrast, in this model land use (i.e., urban form) directly affects the spatial al-location of activities. As shown by Anas, (2007), it is the budget-constrained utility-maximization behavior that defines optimal travel patterns. The complexity of this me-chanism is better shown in the ensuing comparative static analysis, which allows ascer-taining the effect that urban form exerts on the demand for travel.

Comparative Static Analysis

The basic theoretical implications of Model I can be explored by employing comparative static analysis. This section considers the impact of changes in exogenous density, D and exogenous residential location, RL, on travel demand, TD. Basically, starting from an equilibrium state, the impacts of an increase in density and residential location on the initial equilibrium are determined. The objective is to see what happens to transit demand as density levels change (for additional details on assumptions and derivation of the comparative statics, see Appendix A).

Effects of an Increase in Density, D

The effect of an increase in density on travel demand is obtained as

$$\frac{dTD}{dD} = \frac{\overbrace{\underset{(-)}{TD_{AS}}\underset{(-)}{AS_D}}^{\alpha} + \overbrace{\underset{(-)}{TD_{TC}}\underset{(+)}{TC_{AS}}\underset{(-)}{AS_D}}^{\beta}}{\underbrace{1 - \underset{(-)}{AS_{TC}}\underset{(+)}{TC_{AS}}}_{(+)}} > 0 \tag{3.4}$$

where subscripts denote a partial differentiation of the subscripted variable with respect to the variable abbreviated by the subscript. The product $\alpha = TD_{AS}AS_{TD}$ gives the increase in transit demand caused by a contraction in the activity space as a result of increased density. The product $\beta = TD_{TC}TC_{AS}AS_D$ gives the increase in transit demand caused by decreasing trip chaining as a result of increased density

Based on an assumed relationship between spatial dispersion of activities and trip chaining, the result of this analysis shows that changes in density levels exert two contrasting effects on the demand for transit trips.

48

This explanation is inherent in the determinants of trip chaining behavior. In higher density environments, as the spatial extent of non-work activities reduces, trip chaining needs decrease, but individual trips increase and individuals prefer to make non-chained trips. First, increased density reduces the activity space, which directly increases the demand for non-chained trips. Second, increased density reduces the activity space, which reduces the need to chain trips (as time-saving opportunities decrease) and thus the demand for transit trips.

Change in Residential Location, *RL*

Next, we derive the comparative statics of an increase in residential location, *RL*. Note that *RL* is considered as predetermined in Model I. The question to be answered is: "What happens to transit demand as the job-residence pair changes?" Using cross sectional data, this question can be translated as: "How does transit demand differ for those households facing long commutes from those making short commutes?"

The comparative static result describing the impact of a change in residential location on the demand for transit trips is

$$\frac{dTD}{dRL} = \frac{\overset{(\pm)}{\overbrace{TD_{RL}}} + \overset{(+)}{\overbrace{TC_{RL}}} \overset{(-)}{\overbrace{TD_{TC}}} + \overset{(-)}{\overbrace{AS_{TC}}}\left(\overset{(+)}{\overbrace{TC_{RL}}}\overset{(-)}{\overbrace{TD_{AS}}} - \overset{(+)}{\overbrace{TC_{AS}}}\overset{(\pm)}{\overbrace{TD_{RL}}}\right)}{\underset{(+)}{\underbrace{1 - \underset{(-)}{\underbrace{AS_{TC}}}\underset{(+)}{\underbrace{TC_{AS}}}}}} \gtrless 0 \qquad (3.5)$$

As previously discussed, an increase in residential location increases trip chaining ($TC_{RL} > 0$), which in turn positively affects both the size of the activity space, *AS*, and the demand for transit services. The overall effect on transit demand hinges on the sign of TD_{RL}. To the extent that an urban area is well served by transit, then the relationship between transit demand and residential location is positive. A positive relationship is ob-

served in older, more monocentric cities, where existing transit services support commuting. On the other hand, if supply constraints exist, transit demand declines as the job-residence distance increases. Therefore, the overall effect on transit demand due to a change in location depends on both the sign and magnitude of TD_{RL}.

Change in Walking Distance to Nearest Station, *WD*

A change in transit station proximity causes a change in transit demand equivalent to

$$\frac{dTD}{dWD} = \frac{\overset{(-)}{\overbrace{TD_{WD}}} + \overset{(-)}{\overbrace{TD_{TC}}}\overset{(\pm)}{\overbrace{TC_{WD}}} + \overset{(-)}{\overbrace{AS_{TC}}}\left(\overset{(\pm)}{\overbrace{TC_{WD}}}\overset{(-)}{\overbrace{TD_{AS}}} - \overset{(+)}{\overbrace{TC_{AS}}}\overset{(-)}{\overbrace{TD_{WD}}} \right)}{\underbrace{1 - \underset{(-)}{AS_{TC}}\underset{(+)}{TC_{AS}}}_{(+)}} \lesseqgtr 0 \qquad (3.6)$$

The overall effect of an increase in walking distance is ambiguous. An increase in distance to the nearest station directly reduces transit demand ($TD_{WD} < 0$). At the same time, reduced accessibility impacts and the ability to engage in trip chaining using transit, producing an ambiguous effect on transit demand. The sign hinges on the relationship between trip chaining and distance to the nearest transit station, ($TC_{WD} \gtreqless 0$), which is undetermined. On the other hand, the empirical literature provides unequivocal evidence of a negative relationship between distance to transit stops and the demand for transit services (Cervero 2007; Cervero and Kockelman 1997). The debate is mostly centered on the magnitude of this relationship, as high-lighted by the growing body of literature on residential self-selection.

Model II: Endogenous Residential Location, Exogenous Density

In this model, we relax the assumption of exogenous residential location. Treated as a choice variable, residential location is the outcome of a trade-off between transporta-

tion and housing costs. Taking into account idiosyncratic preferences for location, households choose an optimal home-work commute pair, while at the same time optimizing goods consumption and the ensuing non-work travel behavior (optimal non-work trip chaining and activity space). This model is specified as

$$TC = TC(AS, RL, WD, X_{TC}) \tag{3.8}$$

$$AS = AS(TC, D, X_{AS}) \tag{3.9}$$

$$TD = TD(TC, AS, RL, WD, X_{TD}) \tag{3.10}$$

$$RL = RL(TC, TD, X_{RL}) \tag{3.11}$$

where X_{RL} is a vector of controls specific to the RL equation and all other variables are as defined earlier.

Comparative Static Analysis

The complete comparative statics are presented in Appendix B. A discussion of the findings is presented below. Note that the inclusion of the endogenous residential location equation, RL, complicates the computation of the total partial derivatives.

Effects of an Increase in Density, *D*

The effect of an increase in density on travel demand is obtained as

$$\frac{dTD}{dD} = \frac{\overset{(-)}{\widehat{AS_D}}\left[\left(1-\overset{(+)}{\widehat{RL_{TC}}}\overset{(+)}{\widehat{TC_{RL}}}\right)\overset{(-)}{\widehat{TD_{AS}}}+\overset{(+)}{\widehat{TC_{AS}}}\left(\overset{(+)}{\widehat{RL_{TC}}}\overset{(\pm)}{\widehat{TD_{RL}}}+\overset{(-)}{\widehat{TD_{TC}}}\right)\right]}{\underbrace{1-\underset{(+)}{RL_{TC}}\underset{(+)}{TC_{RL}}-\underset{(\pm)}{RL_{TD}}\underset{(\pm)}{TD_{RL}}+\underset{(+)}{AS_{TC}}\left[-\underset{(\pm)}{RL_{TD}}\underset{(+)}{TC_{RL}}\underset{(-)}{TD_{AS}}+\underset{(+)}{TC_{AS}}\left(-1+\underset{(\pm)}{RL_{TD}}\underset{(\pm)}{TD_{RL}}\right)\right]-\underset{(\pm)}{RL_{TD}}\underset{(+)}{TC_{RL}}\underset{(-)}{TD_{TC}}}_{|J|}} \lessgtr 0 \tag{3.12}$$

In the long run, the activity space, transit demand, trip chaining and residential location are all jointly determined. Exogenous changes in density levels therefore affect all these variables. An increase in density directly contracts the activity space, whereas it indirectly reduces trip chaining and ambiguously affects transit demand through its effect on the

activity space. The effect on residential location operates through the effect on transit demand, but that effect is ambiguous. This renders the effect of density on transit demand ambiguous as well. Comparing equation (3.12) to equation (3.4), we see that the complexity of the relationship between transit demand and density increases substantially.

Change in Walking Distance to Nearest Station, *WD*

The comparative static effect of a change in transit station proximity on transit demand is

$$\frac{dTD}{dWD} = \frac{\overset{(-)}{TD_{WD}} + \overset{(-)}{TD_{TC}}\overset{\pm}{\overset{(\pm)}{TC_{WD}}} + \overset{(+)}{RL_{TC}}\left(\overset{(\pm)}{\overset{(\pm)}{TC_{WD}}}\overset{(\pm)}{TD_{RL}} - \overset{(-)}{\overset{(+)}{TC_{RL}}}\overset{(-)}{TD_{WD}}\right) + \overset{(+)}{AS_{TC}}\left(\overset{(\pm)}{\overset{(\pm)}{TC_{WD}}}\overset{(-)}{TD_{AS}} - \overset{(-)}{\overset{(+)}{TC_{AS}}}\overset{(-)}{TD_{WD}}\right)}{|J|} \lesseqgtr 0 \ (3.12)$$

With endogenous residential location, the sign, as well as the magnitude of dTD/dWD depends on both the sign and magnitude of TD_{RL} and TC_{WD}, all of which are unknown. As in Model I, the effect of *WD* is ambiguous.

Model III: Endogenous Residential Location, Endogenous Density

In this last extension to Model I, the assumption of exogenous density is relaxed. This model translates the conceptual framework of Figure 3.1 into the following analytical model

$$TC = TC(AS, RL, \ WD, X_{TC}) \tag{3.13}$$

$$AS = AS(TC, D, X_{AS}) \tag{3.12}$$

$$TD = TD(TC, AS, RL, WD, X_{TD}) \tag{3.13}$$

$$D = D(RL, AS, X_D) \tag{3.14}$$

$$RL = RL(TC, , X_{RL}) \tag{3.15}$$

In the long run, the simultaneous choice of location and travel decisions is assumed to affect density levels across a given urban area. This model best describes a long-run equilibrium, in which both location and travel decisions are optimized under constraint. Urban form is treated as endogenous to the process and is itself affected by household travel decisions and location behavior. Aspects of this relationship and its influences on transit patronage have been previously considered in the literature. For example, while modeling long-run transit demand responses to fare changes, Voith (1997) treats density as endogenous and being affected directly by transit patronage levels. In the long run, these levels are affected by supply-side changes. Voith (1997) assumes that as transit services improve, more people tend to live in proximity to transit stations, thus increasing the demand for transit services.

Ideally, empirical testing of this model would rely on panel data of individual travel diaries. Generally, however, panel data are unavailable and cross-section data are relied on. With cross section data, we can study changes in behavior by controlling for individual heterogeneity.

Comparative Static Analysis

Given the endogenous treatment of density, we can use this model to test the effects of policies geared at directly affecting density, such as policy interventions intended to increase density around transit stations. Assuming an exogenous shock, θ, positively affecting density, comparative statics can be obtained. The inclusion of two more equations complicates the calculations to derive the relevant comparative static results. The results are basically the same as Model II, although the expected magnitudes of impacts differ. To avoid cluttering the text, Appendix A reports the comparative statics results,

which we will use in the empirical work of Chapter 4. Table 3.1 reports a summary of the comparative statics highlighting the expected signs from changes in the most relevant variables affecting trip chaining, *TC*, activity space, *AS*, and transit demand, *TD*.

TABLE 3.1 Comparative Static Results

Exogenous Variable	D	RL	$AS_\varphi{}^*$	$TC_\varphi{}^*$	WD
Effect on Trip Chaining, *TC*	-	+	+	+	+/-
Effect on Activity Space, *AS*	-	+	+	-	+/-
Effect on Transit Demand, *TD*	+	+/-	-	-	+/-
*Shift parameters affecting *AS* and *TC*					

Conclusions

The analytical framework we presented in this chapter seeks to strike a balance between the complexity of activity-based modeling and the more traditional discrete-choice frameworks. The added complexity of the models introduced here is intrinsic to the explicit consideration of non-work travel behavior and its interrelationship with the spatial extent of non-work activities.

These analytical models are general and can be applied to data from any urban area. Empirical testing of the hypotheses of these models requires detailed travel beha-vior data at the individual level. The increased level of sophistication of activity-based travel diaries allows collecting information on activities conducted at home and out of home, as well as their spatial location. As we shall see, the contribution of geographic information system (GIS) modeling permits the measurement of the geographic dimen-sion of both activities and travel and relating them to the surrounding urban landscape.

Coupling GIS with econometric modeling allows conducting empirical tests of the relationships generated by the models of this chapter.

Chapter 4: Empirical Analysis

Introduction

In this chapter, we test all relevant hypotheses about the relationship between urban form and transit patronage introduced in Chapter 3. The objectives are:

1. to test the signs summarized by Table 3.1;
2. to assess the presence of endogeneity in the relationship between transit and urban form; and,
3. to assess the magnitude of this relationship.

The aim is to ascertain to what extent density matters in shaping the demand for transit, after accounting for any endogeneity or simultaneity that might be present. To test these hypotheses, we rely on a dataset that provides travel behavior information at the disaggregate level. First, we provide descriptive statistics for the models' dependent and independent variables. Then, we proceed to specify Model I through Model III and choose the appropriate multivariate regression method. We finally present the results of regression at the end of the chapter.

Data Sources

To test the models presented in the previous chapter, we must rely on travel-diary data. Travel diaries ask respondents to compile a log of activities and travel made during a selected time frame, usually one or two days, encompassing both weekday and weekend travel. In these surveys, respondents log in information on activities by purpose

(work, recreation, shopping, etc.). The new generation of activity-based travel surveys is characterized by travel diaries that provide a high level of activity detail, both at home and out-of-home, to obtain a comprehensive picture of all behavioral aspects at the individual and household levels affecting travel decisions. The main advantage of these new type of surveys, as highlighted by Davidson et al. (2007), is that they are based on tour structure of travel, with travel derived within a general framework of the daily activities undertaken by households and persons.

In this study, we use travel-diary data from the 2000 Bay Area Travel Survey (BATS2000). BATS2000 is a large-scale regional household travel survey conducted in the nine-county San Francisco Bay Area of California by the Metropolitan Transportation Commission (MTC). Completed in the spring of 2001, BATS2000 provides consistent and rich information on travel behavior of 15,064 households with 2,504 households that make regular use of transit.[1] BATS2000 used the latest applications of activity and time-based survey instruments to study travel behavior. The data from BATS2000 are accessible online and maintained as a set of relational data files and are available as comma-separated value (CSV) and American Standard Code for Information Interchange (AS-CII) text files (MTC 2008). Each data file has a corresponding statistical analysis system (SAS®) script to read the data file and act as the data dictionary for the data file (MTC 2007). In the dataset, 99.9 percent of home addresses and 80 percent of out-of-home activities were geocoded using geographic information systems (GIS) to the street address or street intersection level (99.5 percent to the street address level). This permits a precise geographic determination of non-work activities, job, and residential unit locations.

---

[1]MTC defines a transit household as one where one or more members used transit at least once during the two-day surveying period.

The choice of this dataset goes beyond its quality. Most of the relevant academic and practitioner work on the relationship between transit and urban form, research on the issue of residential self-selection, and the efficacy of transit-oriented development policies (TOD) made use of BATS2000. Most of the work we reviewed in Chapter 2 used this dataset. MTC also compiles a list of research papers that made use of the data (Gossen 2005).

Our dataset combines BATS2000 travel behavior data with geographical data from the Census Bureau. Census data are from Summary File 3, which consists of detailed tables of social, economic, and housing characteristics compiled from a sample of approximately 19 million housing units (about 1 in 6 households) that received the Census 2000 long-form questionnaire ("Census 2000 Summary File 3" 2007). We obtained these data at the Census block-group level. Thus, we measure housing and neighborhood characteristics at the block-group level where the residential unit is located.

The unit of observation is the household to reflect the higher hierarchical decision making process of both residential location and travel needs. Referring to MTC work on transit use and station proximity (MTC 2006), a transit household is defined as one where one or more members used transit at least once during the two-day surveying period.

Dependent Variables Descriptive Statistics

While in Chapter 3 we defined activity space, *AS*, residential location, *RL*, trip chaining, *TC*, walking distance to the nearest station (i.e., station proximity), *WD*, and density, *D*, we now provide some additional explanation on their measurement.

Measures of Activity Space, *AS*

   Activity space measures the spatial dispersion of non-work activity locations. Non-work activities consist of shopping, recreational activities (e.g. visiting friends or dining out), and non-recreational activities (doctor visits, child rearing, recurring activities). These activities can be located in proximity to the household residential unit or be located away from it. To measure the spatial extent of these activities across the urban landscape, we employ area-based geometric measures developed in transportation geography. Different metrics that describe the spatial extent of activity locations can be employed. The simplest measure is represented by the standard distance circle (SDC) (or standard distance deviation), which is essentially a bivariate extension of the standard deviation of a univariate distribution. It measures the standard distance deviation from a mean geographic center and is computed as

$$SDC = \sqrt{\frac{\sum(x_i - \bar{x})^2 + \sum(y_i - \bar{y})^2}{n}} \qquad (4.1)$$

where $\bar{x}$ and $\bar{y}$ represent the spatial coordinates of the mean center of non-work activities at the household level, and the i subscript indicates the coordinates of each non-work activity. The mean activity center is analogous to the sample mean of a dataset, and it represents the sample mean of the x and y coordinates of non-work activities contained in each household activity set. The coordinates represent longitude and latitude measurement of each activity and are reported in meters following the Universal Transverse Mercator (UTM) coordinate system. Household activity locations are those visited by surveyed household members during a specified time interval, in this case two representative weekdays. Thus, the standard distance of a household's activity pattern is estimated as the standard deviation (in meters or kilometers) of each activity location from the mean

center of the complete daily activity pattern. Interpretation is relatively straightforward, with a larger standard distance indicating greater spatial dispersion of activity locations. The area of the SDC is the area of a circle with a radius equal to the standard distance. The SDC provides a summary dispersion measure that can be used to explore systematic variations of activities subject to socio-demographic, travel patterns, and patterns of land-use.

As pointed out by Ebdon (1977), this measure is affected by the presence of outliers or activities that are located farthest from the mean center. As a result of the squaring of all the distances from the mean center, the extreme points have a disproportionate influence on the value of the standard distance. To eliminate dependency from spatial outliers, another measure of dispersion, called the standard deviational ellipse (SDE) is usually employed, which uses an ellipse instead of a circle. The advantages of the SDE with respect to the SDC have been discussed in the literature (Ebdon 1977). In addition to control for outliers, the SDE also allows accounting for directional bias of activities with respect to their mean center. The ellipse is centered on the mean center with the major axis in the direction of maximum activity dispersion and its minor axis in the direction of minimum dispersion (See Figure 4.1). In this study, we employ the standard distance ellipse (*SDE*), using the formula described in Levine (2005)

$$SDE = \sqrt{\frac{\sigma_x^2 + \sigma_y^2}{2}} \qquad (4.2)$$

where σ_x and σ_x represent the length of the major and minor axes of the ellipse.

60
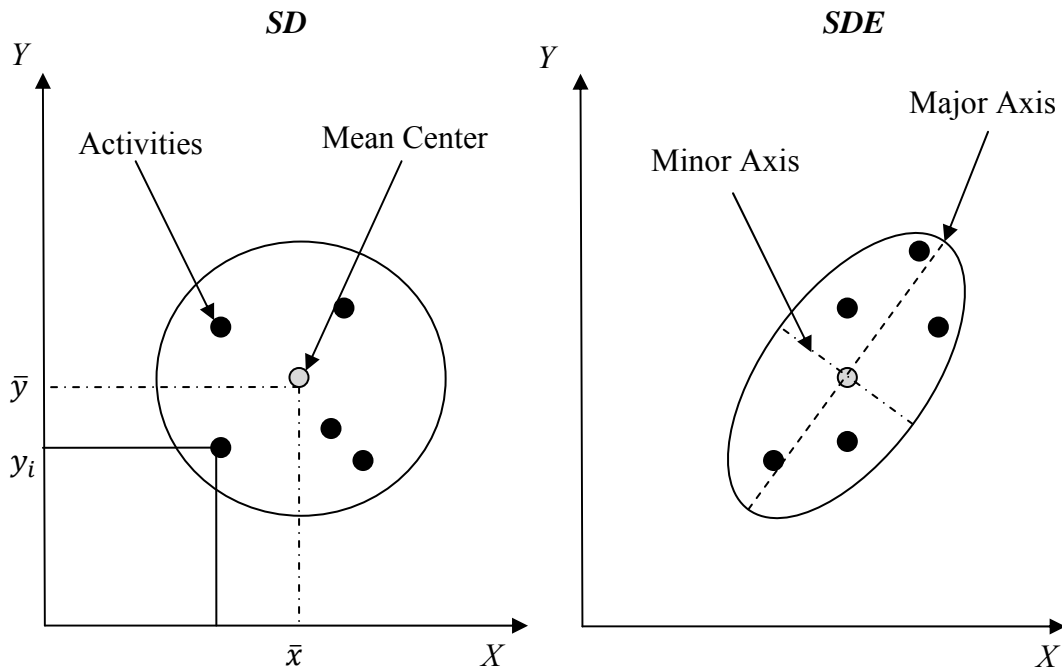
FIGURE 4.1 Standard Distance Circle and Standard Distance Ellipse

Measures of Residential Location, *RL*

We define residential location as the average distance of household employment

activities to the household residential unit

$$RL = \frac{\sum_{m=1}^{k} dist_{mj}}{k} \tag{4.3}$$

where $dist_{mj}$ is the Euclidean distance to the residential unit located at j, from a house-

hold member work location m, and k is the total number of employed household mem-

bers. An alternative specification only considers the distance between the household

head's work location and the residential unit. This assumes that the residential location

choice puts more relevance to the location of the household "breadwinner," as discussed

in detail later in this chapter.

Measures of Transit Station Proximity, *WD*

In this study, we treat transit proximity as a continuous variable measuring distance to the nearest transit station from the household residential unit. A 2006 publication from MTC made use of BATS2000 data to look at the relationship between transit use, population density, and characteristics of individuals living near transit stations (MTC 2006). An appendix to this study was recently published on the MTC website which reports an updated version of the household file containing an additional variable measuring network walking distance from each household residential unit to the nearest transit station (Purvis 2008). Using this file, we measure walking distance as actual distance based on network characteristics to take into consideration the existence of accessibility impediments.

Measures of Density, *D*

We measure the dependent variable density, *D*, as gross population density of the Census block group in which the household residential unit is located. The Census block-group area is measured in square miles. As discussed in Chapter 2, other studies on transit and urban form tend to utilize number of dwelling units per square mile. We also consider additional urban form measures, initially treated as exogenous to the model, which we describe under the exogenous variable section of this chapter.

Table 4.1 presents basic descriptive statistics of the dependent variables, split by different gross population density levels corresponding to the classification adopted by MTC to differentiate between urbanized and non-urbanized areas. As documented in Chapter 2, there exists an underlying correlation between density levels and travel behavior. This table shows how the activity space is slightly larger for transit households than

for non-transit households (19.1 versus 17.2 square miles) and contracts as density increases, while trip chaining does not follow this linear relationship. Walking distance to the nearest station noticeably decreases at higher density levels. To highlight the relevance of transit patronage, Table 4.2 compares sample transit trip averages to auto, walk and other trips. This table shows marked differences in terms of trip making and trip chaining behavior between transit and non-transit households, as well as in average travel times between home and work between transit and non-transit households (51.9 minutes versus 37.4 minutes).

Explanatory Variables Descriptive Statistics

Socio-Demographic Variables

We treat the following socio-demographic variables as exogenous explanatory variables:

- Household characteristics
- Householder gender
- Householder race
- Number of children of school age
- Number of persons employed full-time
- Household income
- Number of vehicles
- Number of licensed individuals
- Tenure (own versus rent)

These variables are available from the BATS2000 person file. Some of these socio-demographic variables have been included in the studies reviewed in Chapter 3 dealing with the influence of land use on transit patronage, while the most current literature

TABLE 4.1 Descriptive Statistics: Overall Sample Means

Density (per-sons/mile^2)	Activity Space (mile^2)	Residential Location, RL (miles)	Residential Location, RL (min)	Trip Chaining, TC (number)	Transit Trips (number)	Auto Trips (number)	Walk Trips (number)	Walking Distance, WD (mile)
0 to 499	27.84	14.12	43.40	2.96	0.14	9.00	0.50	2.33
500 to 5,999	19.31	11.82	40.97	3.04	0.27	8.78	0.72	0.45
6,000 to 9,999	15.69	10.02	38.70	2.98	0.29	8.40	0.80	0.23
>=10,000	13.70	8.56	39.41	3.01	0.73	5.98	1.33	0.14

Data Source: 2000 Bay Area Travel Survey (BATS2000) and 2000 Census Summary File 3, Census Bureau

TABLE 4.2 Descriptive Statistics: Sample Means of Dependent Variables and Selected Trip Measures

Transit House-hold		Gross Popu-lation Density (per-sons/mile^2)	House-hold Ac-tivity Space (mile^2)	Residen-tial Loca-tion, RL (miles)	Residen-tial Loca-tion, RL (min)	Trip Chain-ing, TC (num-ber)	Transit Trips (num-ber)	Auto Trips (num-ber)	Walk Trips (num-ber)	Walking Dis-tance, WD (mile)
No	Mean	7,910.51	17.16	10.33	37.36	2.87	-	8.32	0.73	0.49
	SD	8,752.95	38.40	10.07	33.32	1.77	-	6.14	1.62	1.44
	N	12,260	10,548	9,128	8,353	11,242	12,260	12,260	12,260	12,260
Yes	Mean	15,172.65	19.14	11.58	51.92	3.65	2.32	5.96	1.70	0.22
	SD	17,193.12	37.84	9.76	35.35	1.73	1.29	5.77	2.38	0.38
	N	2,503	2,176	2,138	1,918	2,446	2,503	2,503	2,503	2,503
Overall Sample	Mean	9,141.78	17.50	10.57	40.08	3.01	0.39	7.92	0.89	0.45
	SD	11,006.88	38.31	10.03	34.18	1.79	1.02	6.14	1.81	1.33
	N	14,763	12,724	11,266	10,271	13,688	14,763	14,763	14,763	14,763

Data Source: 2000 Bay Area Travel Survey (BATS2000) and 2000 Census Summary File 3, Census Bureau

on self-selection considers all of them. Table 4.3 provides a summary of these variables for the overall sample. As with the vast majority of travel survey, the white population is overly represented, as well as the higher income groups.

Travel Behavior Variables

We also created additional explanatory variables at the household level to control for factors affecting both the spatial extent of non-work activities and the ensuing travel behavior:

- Activity travel time
    - mean travel time to shopping trips starting at home
    - mean travel time to recreational trips starting at home
    - mean travel time to school trips starting at home
    - mean travel time to other trips not starting at home
    - mean travel time across all non-work activities
- Activity duration
    - mean time duration across all non-work activities

These variables are commonly used in the activity-based literature in modeling activity duration and scheduling (Bhat 1997, 1999, 2001) and activity travel patterns (Kuppam and Pendyala 2001). Transit households spend less time shopping compared to non-transit households (28.9.0 versus 30.3 minutes), they also spend less time on recreational activities (161.9 versus 175.9 minutes) and at home (181.8 versus 210.1 minutes). The time spent travelling to reach out-of-home activities also differs, with transit households spending an average of 15.7 minutes on the road versus 12.9 minutes for non-transit households. The trade-off between leisure and work is also reflected in less time spent sleeping (243.6 versus 249.6 minutes for non-transit households). These time-use

variations and the comparison between transit and non-transit households provided in

Table 4.2 are indicative of the trade-offs inherent to total time available, residential loca-

tion, and trip-chaining behavior discussed in Chapter 3.

Urban Form Variables

Although BATS2000 does not include land-use variables, it provides exact geo-

graphical information about the location of each of the 15,064 households. GIS coordi-

nates permit a precise allocation of each household residential unit within each Census

Bureau geographical unit of reference using GIS techniques. By linking each house-

holds' residential unit x and y geographic coordinates to GIS Census block-group maps of

the San Francisco Bay area, we merged a comprehensive set of land-use variables with

the travel diary dataset.[2] We obtained other variables related to non-residential land use

from the 2000 U.S. Census Bureau County Business Patterns (CBP) data file. Table 4.4

describes these variables and data sources.

We intend to use the last two variables of Table 4.4 as proxy measures of centrali-

ty (CBD distance) and polycentricity (distance from the nearest subcenter). As men-

tioned in Chapter 3, monocentric models only consider measures of the strength of the

relationship between CBD employment (and other activities located at the CBD) and tra-

vel behavior.

---

[2] Detailed GIS maps and other geographical data are available online at the MTC website <http://www.mtc.org>.

TABLE 4.3 Summary of Selected Demographic Variables

Variables	Frequency	% Share
Householder Gender		
Male	6,901	45.8%
Female	8,163	54.2%
Householder Race		
Asian	1,223	8.1%
Black	442	2.9%
Hispanic	647	4.3%
Other	674	4.5%
White	12,078	80.2%
Children, by age group		
< 6 year	1,539	10.2%
6 to 11 year	1,973	13.1%
12 to 18 year	2,202	14.6%
Employed, Full Time (persons)		
0	876	7.0%
1	7,214	57.8%
2	4,063	32.5%
>=3	335	2.7%
Household Income ($)		
Less than 10,000	225	1.7%
10,000 to 14,999	230	1.7%
15,000 to 19,999	322	2.4%
20,000 to 24,999	368	2.8%
25,000 to 29,999	464	3.5%
30,000 to 34,999	424	3.2%
35,000 to 39,999	514	3.9%
40,000 to 44,999	756	5.7%
45,000 to 49,999	833	6.3%
50,000 to 59,999	1,352	10.2%
60,000 to 74,999	1,660	12.6%
75,000 to 99,999	2,359	17.9%
100,000 to 124,999	1,620	12.3%
125,000 to 149,999	804	6.1%
>= 150,000	1,260	9.6%
Vehicles		
0	610	4.0%
1	4,938	32.8%
2	6,542	43.4%
3	2,238	14.9%
>=4	736	4.9%
Tenure		
Own	10,415	69.4%
Rent	4,597	30.6%

TABLE 4.4 Urban Form Variables

Variable	Definition	Source
Gross population density	Number of persons/Census block group area size (square miles)	U.S Census Bureau Summary File 3
Dwelling units	Number of owner occupied units	U.S Census Bureau Summary File 3
Dwelling density	Number of owner occupied units/ Census block group area size (square miles)	U.S Census Bureau Summary File 3
Number of retail establishments	Total number of retail establishments within a zip code	U.S Census County Business Patterns: 2000
Retail establishment density	Total number of retail establishments/zip code area	U.S Census County Business Patterns: 2000
Number of wholesale establishments	Total number of retail establishments within a zip code	U.S Census County Business Patterns: 2000
Wholesale establishment density	Total number of wholesale establishments/zip code area	U.S Census County Business Patterns: 2000
Distance from CBD	Distance from CBD	BATS2000-GIS derived
Distance from subcenter	Distance from the nearest subcenter	BATS2000-GIS derived

Through decades of decentralization, the urban landscape has taken a polycentric form, with a number of clustered employment centers affecting both employment and population distributions. The majority of these centers is subsidiary to an older CBD. Such centers are usually called subcenters or sub-regional centers (a more formal definition of subcenter is a set of contiguous tracts with significantly higher employment densities than surrounding areas). The transportation includes few studies of the influence of subcenters on travel behavior. One such study is Cervero and Wu (1998), who have examined the influence of subcenters in the San Francisco Bay Area on commute distances

to conclude that employment decentralization has lead to increased travel. Studies treating subcenters generally take subcenters as exogenously determined either by assumption or by an empirical determination that makes use of specific density thresholds. There are no established methods to determine the number of subcenters present in any urban area. Existing methods rely on rules of thumb based on knowledge about specific geographic areas (Giuliano and Small 1991), while others account for an endogenous determination based on their impact on agglomeration and employment (McMillen 2001).

To account for urban decentralization and its effect on transit use, we adopt the Census definition of cities and designated places to first identify subcenters and then produce a distance measure between a household residential unit and the nearest subcenter.[3] In addition to the above variables, we obtained a set of explanatory variables to control for household idiosyncratic preferences for location. The literature provides some insight on the choice of land-use variables as controls or instrumental variables (Boarnet and Crane 2001; Boarnet and Sarmiento 1998; Crane 2000; Crane and Crepeau 1998b).

Using the Summary 3 Census Bureau file, we obtained the following variables at the block-group level:

1. Stock of housing built before 1945 (number of housing units)

2. Housing median value (dollars; owner-occupied units)

3. Housing median age (years; non-rent units)

4. Housing size (median number of rooms; owner-occupied units)

5. House median monthly cost (owner-occupied units)

---

[3] According to the U.S. Census, a city is a type of incorporated place. A census designated place is a statistical entity consisting of a densely settled concentration of population that is not within an incorporated place, but is locally defined by a name.

6. Percent of household living below poverty line

7. Diversity index (0 = homogeneous; 1 = heterogeneous neighborhood)

The first variable has been used before as an instrumental variable in multivariate regression studies that considered travel behavior as endogenous to urban form (Boarnet and Crane 2001; Boarnet and Sarmiento 1998; Crane 2000; Crane and Crepeau 1998b), while the remaining ones are unique to this study. Additional controls for neighborhood characteristics have also been used elsewhere. For example, the proportion of block-group or census-tract population that is Black and the proportion Hispanic have been used as instruments by Boarnet and Sarmiento (1998) and the percent of foreigners by Vance and Hedel (2007).

In this study we use variables one through five to control for idiosyncratic preferences for housing characteristics not directly affecting travel behavior but directly affecting the residential choice decision at the household level. We use variables six and seven as controls for neighborhood characteristics. In particular, the percentage of households living below poverty levels (henceforth defined as poverty) serves as a proxy for crime, while the diversity index (henceforth called diversity) is used as a proxy for ethnic preferences (i.e., moving into a neighborhood with similar ethnic characteristics). The latter is an index of ethnic heterogeneity that varies from zero (only one race living in the neighborhood) to one (no race is prevalent), similar to Shannon's diversity index (Begon and Towsend 1996).[4] As discussed in further detail in Chapter 5, poverty and diversity serve

---

[4] The Shannon Index is a measurement used to compare diversity between habitat samples. The comparison is made by taking into account the proportion of individuals of a given species to the total number of individuals in the set.

a dual role as instrumental variables when we treat transit station proximity, *WD*, endogenous to the model.

Table 4.5 and Table 4.6 present relevant sample mean values split by households by mode choice. Transit households tend to live in highly populated areas characterized by higher than average poverty levels, as well as smaller and older housing units. We also generated one-way analysis of variance tables (not reported here) that include an interaction term between transit household and the transit station dummy variable. All variables exhibit a significant difference in means, indicating that housing price, housing age, room size, neighborhood diversity and poverty levels differ across households according to their location and mode choice. To gain additional insight on the trade-off between residential location and preference for transit, Table 4.7 and Table 4.8 report the same measures of Table 4.5 and Table 4.6, but differentiate between households living in proximity to a transit station. We measure proximity using a Euclidean half-mile buffer around a transit rail line in existence when the BATS2000 travel survey was being conducted.

Transit Supply Variables

We include the following measures of transit supply:

- Presence of a transit stop at workplace
- Supply of park-and-ride within a half-mile of transit stop
- Presence of a transit-oriented development (TOD) stop within a half-mile of residential unit

TABLE 4.5 Urban Form Variables by Household Type

Transit Household	Gross Population Density (persons/mile^2)	Dwelling density (dwellings/mile$^{2)}$)	Retail Establishments Density (number/mile^2)	Wholesale Establishment Density (number/mile2)
No	7,911	3,313	18.4	6.9
Yes	15,173	7,198	43.1	12.6
Overall	9,144	3,974	22.5	7.9

TABLE 4.6 Housing and Demographic Variables by Household Type

Transit Household	House Median Value ($)	House Median Age (years)	Housing Stock (% built before 1949)	Housing Size (rooms)	Households Median Income	Households Below Poverty	Diversity Index
No	399,819	34.18	0.20	5.97	74,189.52	0.06	0.57
Yes	399,374	41.77	0.36	5.92	67,140.84	0.08	0.62
Overall	399,591	35.47	0.23	5.92	72,994.44	0.06	0.58

TABLE 4.7 Urban Form Variables by Transit-Station Proximity

Within 1/2 mile of Transit Station	Gross Population Density	Dwelling density (dwellings/mile^2)	Retail Establishments Density (number/mile^2)	Wholesale Establishment Density (number/mile^2)
No	7,313.8	2,939.2	14.8	5.7
Yes	19,871.4	10,039.7	67.6	20.8
Overall	9,144.4	3,974.3	22.5	7.9

TABLE 4.8 Urban Form Variables by Transit-Station Proximity

Within 1/2 mile of Transit Station	House Median Value ($)	House Median Age (years)	Housing Stock (% built before 1949)	Housing Size (rooms)	Households Median Income	Households Below Poverty	Diversity Index
No	396,509.6	33.6	18.6%	6.0	75,050.4	5.4%	0.57
Yes	417,647.7	46.3	46.4%	5.2	60,501.5	8.9%	0.64
Overall	399,591.1	35.5	22.6%	5.9	72,994.4	5.9%	0.58

The relevance of transit station proximity to the workplace is confirmed by the literature, as seen in Chapter 3. For example, using BATS2000, Cervero (2007) showed that the presence of a station within one mile of a workplace (with good accessibility) strongly influences both residential choice decisions and transit use. The relationship gets stronger as distance to the station declines.

The presence of park-and-ride lots nearby transit stops also positively influences transit ridership by improving accessibility to those households located farther than the one-mile threshold. Furthermore, as highlighted by TCRP Report 95 (2007), the presence of park-and-ride lots provides increases opportunities to trip chain from the residence to the transit station on the way to work . The relevance of park-and-ride lots is measured by a dichotomous variable indicating the presence of a park-and-ride lot within a half-mile of a transit stop. To produce these transit-supply explanatory variables, the same GIS maps created by MTC as part of their transit station proximity study were used (MTC 2008) (a detailed discussion of the GIS methodology is provided in Appendix G of the MTC study).

Finally, to test the relevance of urban design policies on transit patronage, we introduce in the model a dichotomous variable qualifying a transit stop as having the characteristics of a TOD station. TOD stops are characterized by land development policies geared at facilitating transit use by improving transit station accessibility (by reducing physical barriers), and by promoting mixed land-use development (residential and commercial) in their immediate surroundings. For example Cervero (2007) used BATS2000 and census land-use data to evaluate transit-oriented development (TOD) impacts on ridership and self-selection. In his analysis, he notes that between 1998 and 2002 about

13,500 apartment and condominium units were built within a half-mile of urban stations of southern California and the San Francisco Bay Area, often using land previously occupied by park-and-ride lots; this makes the dataset suitable to also test the impact of TOD on ridership. We relied on the California Department of Transportation Transit-Oriented Database to identify these stations (CALTRANS 2008).

Table 4.9 summarizes the full set of endogenous and exogenous explanatory variables.

TABLE 4.9 List of Variables for Model Estimation

Variable	Definition	Use
inc	Household income	Socio-demographic
sch	Number of children of school age (pre-k to middle)	Socio-demographic
veh	Number of vehicles	Socio-demographic
own	Tenure (1 = owner; 0= renter)	Socio-demographic
licensed	Number of persons with driving license	Socio-demographic
tswork	Presence of a transit stop within 0.5 mile of workplace (1=yes, 0=otherwise)	Transit supply
prkride	Presence of a park-and-ride within 0.5 mile of a transit stop (1=yes, 0=otherwise)	Transit supply
ts_tod	Transit stop characterized as transit-oriented development stop (1=yes, 0=otherwise)	Transit supply
cbd_dist	Residential unit distance from CBD	Urban form/land use
subc_dist	Residential unit distance from nearest subcenter (cities and designated places)	Urban form/land use
r_est	Number of retail establishments, zip code level	Urban form/land use mix
w_est	Number of wholesale establishments, zip code level	Urban form/land use mix
hprice	Median house price, block group level	Residential/neighborhood characteristics
hage	Median house age, block group level	Residential/neighborhood characteristics
room	Median number of rooms owner occupied unit, block group level	Residential/neighborhood characteristics
inc_blkgrp	Median household income, block group level	Residential/neighborhood characteristics
pov	Proportion of households living below poverty line, block group level	Residential/neighborhood characteristics
div	Diversity index (ranges from 0 if block group level is ethnically homogenous to 1 if heterogeneous)	Residential/neighborhood characteristics
act_dur	Mean non-work activity duration	Travel behavior
act_tt	Mean travel time to non-work activities	Travel behavior
TC	Trip chain; number of non-work trip stops on the job-residence commute	Trip chaining behavior
AS	Household activity space; standard distance ellipse area ($mile^2$)	Spatial extent of non-work activities
RL	Residential location (home-work distance)	Household residential location
WD	Walking distance from the residential unit to the nearest transit station	Transit station proximity
D	Gross population density (persons/$mile^2$)	Urban Form

Method of Analysis

Given the structural framework of Chapter 3, the empirical test of the proposed hypotheses requires the use of structural equation modeling (SEM). SEM is used to capture the causal influences of the exogenous variables on the endogenous variables and the causal influences of the endogenous variables upon one another. The use of SEM in transportation research is linked to the development of activity-based modeling in travel behavior research, which explicitly points out the causal mechanisms underlying individuals' location and travel decisions. Furthermore, more recent developments in the literature studying the efficacy of urban design policies dealing with residential sorting effects try to sort out causality links between urban form and travel behavior. To uncover causality when travel behavior and urban form simultaneously affect each other, requires suitable econometric techniques. As the literature review of Chapter 2 highlighted, it is only recently that transportation researchers have recognized that causal relationships among travel behavior and urban form can be effectively represented in a structural equation framework (Cao, Mokhtarian, and Handy 2006, 2007; Guevara and Moshe 2006; Mokhtarian and Cao 2008; Peng et al. 1997). Available methods include maximum likelihood estimation (ML), generalized least squares (GLS), two-stage least squares (2SLS), three-stage least squares (3SLS), and asymptotically distribution-free estimation (ADF).

Before proceeding with the estimation, it is necessary to ensure that the model is identified. We subject each of the three models presented in Chapter 3 to the rank condition for identification prior to estimation. Detailed rank conditions are reported in Appendix B. We also discuss the inclusion and exclusion of relevant explanatory variables for each equation.

Model I Results

Using the set of variables summarized in Table 4.7, we specify the first model of Chapter 3 with exogenous residential location, *RL*, and density, *D*, as

$$TC = \alpha_0 + \alpha_1 AS + \alpha_2 RL + \alpha_3 WD + \alpha_4 veh + \alpha_5 act_{tt} + \alpha_6 act_{dur} + \alpha_7 sch +$$

$$\alpha_8 subc\_dist + \varepsilon_1 \tag{4.4}$$

$$AS = \beta_0 + \beta_1 TC + \beta_2 D + \beta_3 act\_dur + \beta_4 inc + \beta_5 r\_estd + \varepsilon_2 \tag{4.5}$$

$$TD = \gamma_0 + \gamma_1 TC + \gamma_2 AS + \gamma_3 WD + \gamma_4 RL + \gamma_5 tswork + \gamma_6 prkride + \gamma_7 ts\_tod +$$

$$\gamma_8 veh + \varepsilon_3 \tag{4.6}$$

Equation (4.4) describes trip-chaining behavior occurring on the commute trip to and from the work location. Trip chaining, jointly determined with the activity space, *AS*, is affected by vehicle availability (*veh*) and transit-station proximity, activity travel time and duration (*act_tt* and *act_dur*), and household structure (*sch*). Vehicle ownership and transit proximity, together with household characteristics (income and children), affect the capability of engaging in complex tours.

Equation (4.5) describes how the spatial extent of non-work activities responds to changes in urban form, being affected directly by density levels and retail establishment concentrations (*r_estd*). Drawing from the work of Anas (2007) on trip-chaining beha-vior and non-work travel, we assume that activity space is a result of utility maximizing behavior determining goods consumption and non-work travel. As income levels in-crease, so does the demand for (normal) goods and travel. We assume that individuals have preferences for heterogeneity in consumption (convexity of indifference curves in-dicates preference for balanced consumption bundles). As assumed by Anas (2007), in-

dividuals prefer to visit different locations, a behavior that positively affects the size of the activity space.

Equation (4.6) describes the demand for transit trips as brought about by the necessity to engage in non-work travel (directly affected by *AS* and *TC*) and by the relative distance of the residential unit to the work location, *RL*. We expect that transit supply directly affects transit ridership in terms of transit station accessibility both at origin and destination. We also wish to test the relevance of TOD policies in affecting ridership by including the dichotomous variable *ts_tod*, which measures the impact of a TOD station.

All three equations pass the rank condition for identification. Equation (4.4) is overidentified, and equation (4.5) and (4.6) are classified as just identified. The results of a three-stage least square regression (3SLS) are displayed in Table 4.10.

The results show that the joint determination of trip chaining and the spatial extent of non-work activities relate to transit patronage as hypothesized in Chapter 3. The presence of a transit stop at workplace (*tswork*) positively affects transit demand, as well as the presence of a TOD transit stop in proximity of the residence unit (*ts_tod*). The size of the activity space reduces as density increases, which, in turn, positively affects the demand for transit. This assumption, as stated in Chapter 3, relates more compact urban environments to increased transit patronage. As locations where non-work activities are more clustered, the need to engage in long and complex journeys requiring modes other than transit decreases, resulting in increased transit usage. The converse is also true, suggesting that policy interventions related to directly affect the clustering of non-work activity locations, such as mixed-land use policies, are likely to significantly affect ridership

levels. However, the relevance of this relationship is better appreciated in a context where residential location is also treated as a choice variable (i.e., endogenous).

To better appreciate the magnitude of these effects, Table 4.11 reports point elasticities of transit demand with respect to selected explanatory variables. For example, to obtain the elasticity of travel demand with respect to changes in density, we use

$$\varepsilon_{TD,D} = \frac{dTD}{dD} * \frac{D}{TD} \tag{4.8}$$

where dTD/dD is from equation (3.4) of Model I.

Table 4.11 shows that, for example, a 20-percent increase in gross population density, D, which is equal to about 1,830 persons per square mile, produces an approximate nine-percent increase in transit demand (linked trips at household level). Transit station proximity also plays a relevant role. A doubling of the average walking distance, *WD*, to the nearest transit station, or an increase from 0.3 miles to 0.6 miles, decreases transit demand by 14 percent; at about one mile, transit demand declines by 28 percent.

The presence of a transit station (*tswork*) within a half-mile of the workplace increases transit demand by 69 percent. Living in proximity to a TOD transit station (*tstod*) increases transit demand by about 28 percent. There seems to be a ridership bonus associated with proximity to a station characterized by accessibility features intended to promote transit use.

TABLE 4.10 3SLS Regression Results—Model I

Equation	Coefficient	Std. Error	P
Trip chaining, TC			
RL	0.0096	0.0040	0.0160
AS	0.0648	0.1658	0.6960
WD	-0.0570	0.0137	0.0000
veh	-0.0793	0.0308	0.0100
act_tt	0.0014	0.0004	0.0010
act_dur	-0.0022	0.0003	0.0000
subc_dist	0.0439	0.0068	0.0000
sch	0.0778	0.0144	0.0000
constant	1.2771	0.2611	0.0000
Activity space, AS			
TC	0.5863	0.0592	0.0000
D	-0.0974	0.0121	0.0000
act_dur	0.0001	0.0002	0.6880
inc	0.0299	0.0050	0.0000
r_estd	-0.0022	0.0003	0.0000
constant	1.7226	0.1351	0.0000
Transit demand, TD			
TC	0.6548	0.0732	0.0000
AS	-0.3002	0.0920	0.0010
WD	-0.0800	0.0124	0.0000
RL	0.0057	0.0021	0.0070
tswork	0.3848	0.0422	0.0000
prkride	-0.0737	0.0514	0.1510
ts_tod	0.2063	0.1097	0.0600
veh	-0.0456	0.0221	0.0390
constant	-0.1256	0.1014	0.2150

Note: N= 8,229; F_{TC}=49.3; F_{AS}=73.6; F_{TD}=122.1

TABLE 4.11 Elasticity Estimates—Model I

Elasticity	RL	WD	D	subc_dist	r_estd	tswork*	ts_tod*
TC	0.087	-0.007	-0.044	0.109	0.000	-	-
AS	0.100	-0.008	-0.066	0.125	0.000	-	-
TD	-0.157	-0.137	0.475	-0.388	0.001	0.687	0.279

Indicates a proportional change

The model reports a negative elasticity between residential location, (*RL*) and transit use. This is consistent with the assumption that households characterized by longer commutes engage in more complex trip chains, which positively affect the spatial extent of non-work activities. With exogenously fixed transit supply, as the activity space expands, transit demand declines.

The results also show that transit demand is sensitive to the presence of nearby subcenters (*subc_dist*), or, in general, to decentralization. The negative sign associated with the elasticities shows that increased polycentricity significantly affects transit demand adversely. The farther a household lives from a subcenter, the less it uses transit. A 50-percent increase in distance to a subcenter (from 2.9 to 4.3 miles) decreases transit demand by about 19.4 percent. This is so because households tend to rely more on other transport modes to carry out more complex trip chains. This result is consistent with the current literature on transit competitiveness and polycentric metropolitan regions. For example, in a study of transit services and decentralized centers, Casello (2007) finds that transit improvements between and within activity centers (i.e., subcenters) are necessary to realize the greatest improvements in transit performance.

Next, we extend Model I to ascertain the extent to which the above relationships are affected by treating residential location as a choice variable.

Model II Results

As discussed in Chapter 3, residential self-selection refers to individuals or households preferring certain residential locations due to idiosyncratic preferences for travel. In applied work, if residential self-selection is not accounted for, findings tend to overstate the importance of policies to increase transit use by mixed-used development.

To deal with this issue, Model II treats residential location as endogenous while retaining density as exogenous. Theoretical considerations inferred in Chapter 3 lead us to specify a model where individuals can locate anywhere within an urban area, choosing a utility-maximizing *job-residence pair*. This process is carried out in conjunction with the optimal choice of both consumption and non-work travel. A household optimally located at a distance to work engages in trip-chaining to benefit from time-savings gained by combining errands to and from work. Time savings can either be allocated to a move farther out or to engage in additional non-work travel.

We specify Model II as

$$TC = \alpha_0 + \alpha_1 AS + \alpha_2 RL + \alpha_3 WD + \alpha_4 veh + \alpha_5 act\_tt + \alpha_6 act\_dur + \alpha_7 sch +$$
$$\alpha_8 subc\_dist + \varepsilon_1 \tag{4.9}$$

$$AS = \beta_0 + \beta_1 TC + \beta_2 D + \beta_3 act\_dur + \beta_4 inc + \beta_5 r\_estd + \varepsilon_2 \tag{4.10}$$

$$TD = \gamma_0 + \gamma_1 TC + \gamma_2 AS + \gamma_3 WD + \gamma_4 RL + \gamma_5 tswork + \gamma_6 prkride + \gamma_7 ts\_tod +$$
$$\gamma_8 veh + \varepsilon_3 \tag{4.11}$$

$$RL = \delta_0 + \delta_1 TC + \delta_2 TD + \delta_3 hprice + \delta_4 hage + \delta_5 rooms + \delta_6 div + \delta_7 pov +$$
$$\delta_8 own + \varepsilon_4 \tag{4.12}$$

We consider housing characteristics (pricing, age, size) as relevant factors affecting residential location, as well as neighborhood characteristics (ethnicity, crime). In terms of exclusion restrictions, Equation (4.12) assumes that while residential location is affected by travel decisions (trip chaining and transit use), housing and neighborhood characteristics do not directly affect travel behavior at the disaggregate level. Other housing-characteristics variables, such as the stock of housing built before 1945, are not included in Equation (4.12) as they serve the same role of those just discussed (beside being highly correlated with pricing and size, thus potentially causing multicollinearity).

Equation (4.10) passes the rank condition for identification and is classified as just identified. Table 4.12 displays the results of the 3SLS regression.

TABLE 4.12 3SLS Regression Results—Model II

Equation	Coefficient	Std. Err.	P
Trip chaining, TC			
RL	0.0096	0.0118	0.4130
AS	0.0725	0.1980	0.7140
WD	-0.0573	0.0142	0.0000
veh	-0.0786	0.0316	0.0130
act_tt	0.0014	0.0005	0.0020
act_m	-0.0022	0.0003	0.0000
subc_dist	0.0435	0.0070	0.0000
sch	0.0778	0.0144	0.0000
constant	1.2604	0.2673	0.0000
Activity space, AS			
TC	0.2357	0.0538	0.0000
D	-0.0858	0.0107	0.0000
act_m	-0.0007	0.0002	0.0000
hhinc	0.0412	0.0045	0.0000
r_estd	-0.0014	0.0003	0.0000
constant	2.0943	0.1202	0.0000
Transit demand, TD			
TC	0.6964	0.0753	0.0000
AS	-0.2598	0.1157	0.0250
WD	-0.0669	0.0127	0.0000
RL	-0.0090	0.0088	0.3110
tswork	0.3716	0.0446	0.0000
prkride	-0.0669	0.0524	0.2020
ts_tod	0.1304	0.1147	0.2560
veh	-0.0365	0.0221	0.0990
constant	-0.1119	0.1020	0.2720
Residential location, RL			
TC	3.7324	0.5009	0.0000
TD	-1.2408	0.4660	0.0080
hprice	-2.8117	0.2722	0.0000
hage	-0.0849	0.0094	0.0000
rooms	1.1279	0.1468	0.0000
div	-2.6312	0.7238	0.0000
pov	-5.9629	2.4133	0.0130
own	0.4966	0.2658	0.0620
constant	39.1808	3.3743	0.0000

Note: $N= 8,212$; $F_{TC}=42.7$; $F_{AS}=72.5$; $F_{TD}=118.5$; $FRL=57.2$

The relevant signs and coefficient magnitudes of the first three equations are consistent with those of Model I. Table 4.12 reports a negative sign but statistically insignificant sign of the effect of residential location on transit demand (TD_{RL}). This might be due to the transit supply characteristics where the travel survey was conducted (e.g., fairly well-served commute routes). The parameter does not have a *ceteris paribus* interpretation as it changes concurrently with the other endogenous variables. Compared to Model I, changes in activity space negatively affect transit use. More dispersed activity-travel locations result in reduced transit patronage, although this effect is now less important.

As with Model I, we produce the relevant point elasticities, summarized by Table 4.13 (only reporting statistically significant estimates).

TABLE 4.13 Elasticity Estimates—Model II

Elasticity	WD	D	subc_dist	r_estd	tswork*
TC	-0.009	-0.036	0.108	-0.014	-
AS	-0.003	-0.069	0.041	-0.232	-
TD	-0.028	0.269	0.065	0.170	0.766
RL	0.002	-0.027	0.052	-0.017	-

* Indicates a proportional change

Compared to Model I, the endogenous treatment of residential location reduces the magnitude of the elasticity of travel demand with respect to density elasticity by 56 percent. When households can locate anywhere in an urban area and they adjust trip chaining and commuting costs, an exogenous 20-percent increase in density produces a 5.4-percent increase in the demand for transit (household linked trips). Transit station

proximity to the workplace, however, increases in importance. The presence of a transit stop within a half-mile of the workplace increases transit demand by about 76 percent.

Accounting for self-selection reduces the relevance of transit-station proximity indicated by an 80-percent decrease in magnitude in its point elasticity estimate with respect to Model I. An increase from 0.3 to 0.6 miles to the nearest transit station reduces transit demand by only 2.8 percent as opposed to the 14 percent reduction of Model I. This result shows that self-selection is more relevant than what noted by Cervero (2007), who found that self-selection accounts for about 40 percent of transit ridership for individuals residing near a transit station.

To understand the reasons for these changes, it is sufficient to look at the specification of Model II. Equation (4.12) assumes households optimally choose residential location and non-work activities, which also optimally define the spatial extent of non-work activities. Households locate their residences farther from the job locations, trading lower housing costs against increased commute distance. Trip chaining optimization is part of this trade-off process, which leads to an expansion of the activity space. This in turn reduces the opportunities to use transit to engage in non-work travel. This behavior is empirically validated by the statistical significance of all housing and neighborhood controls in equation (4.12).

Model III Results

Up to this point, we have treated urban form as exogenous. What happens if urban form, as measured by gross population density, is affected by travel decisions? To what extent is the relationship between density and transit in Model I and Model II affected by treating density as endogenous? The following model endogenizes density

$$TC = \alpha_0 + \alpha_1 AS + \alpha_2 RL + \alpha_3 WD + \alpha_4 veh + \alpha_5 act\_tt + \alpha_6 act\_dur + \alpha_7 sch +$$

$$\alpha_8 subc\_dist + \varepsilon_1 \tag{4.13}$$

$$AS = \beta_0 + \beta_1 TC + \beta_2 D + \beta_3 act\_dur + \beta_4 inc + \beta_5 r\_estd + \varepsilon_2 \tag{4.14}$$

$$TD = \gamma_0 + \gamma_1 TC + \gamma_2 AS + \gamma_3 WD + \gamma_4 RL + \gamma_5 tswork + \gamma_6 prkride + \gamma_7 ts\_tod +$$

$$\gamma_8 veh + \varepsilon_3 \tag{4.15}$$

$$RL = \delta_0 + \delta_1 TC + \delta_2 TD + \delta_3 hprice + \delta_4 hage + \delta_5 rooms + \delta_6 div + \delta_7 pov +$$

$$\delta_8 own + \varepsilon_4 \tag{4.16}$$

$$D = \vartheta_0 + \vartheta_1 RL + \vartheta_2 AS + subc\_dist + \vartheta_3 cbd\_dist + \varepsilon_5 \tag{4.17}$$

Equation (4.17) treats as endogenous population density at the residential unit location. This model introduces exogenous variables serving as a proxies for centrality dependence (*cbd_dist*) and for polycentricity (*subc_dist*). Compared to Model I and Model II, the joint endogenous treatment of residential location and density produces a model whose relevant hypotheses are confirmed.

Regarding Equation (4.17) both CBD and subcenter distance are statistically significant. The sign of the CBD measure of centrality (*cbd_dist*) is negative as expected. As distance to the CBD or the nearest subcenter increases, density decreases. This finding indicates the spatial attraction of the CBD relative to subcenters even within a polycentric urban area, such as the San Francisco Bay area.

The relevance of these two variables is better highlighted by the elasticities presented in Table 4.15.

The elasticity of travel demand with respect to walking distance is less than that of Model I, but greater (in absolute terms) than that of Model II. An increase from 0.3 to

0.6 miles to the nearest transit station reduces transit demand by 9 percent, compared to the 14-percent reduction of Model I and 2.4-percent reduction of Model II. The presence of a transit stop at the workplace almost doubles the demand for transit, substantially increasing the importance of that variable in this model as compared to the others.

The sign and statistical significance associated with the centrality measure (*cbd_dist*) confirms the relevance of the CBD as a generator of transit ridership. Treating density endogenously results in a more elastic travel demand with respect to distance to the nearest transit center. It is relevant to note that both *cbd_dist* and *subc_dist*, appear as explanatory variables but are treated as endogenous in the model. An initial specification treated these two variables as exogenous, but overidentification tests (discussed in the next chapter) revealed that this treatment led to weak instruments (a problem leading to inconsistent estimates).

The exogenous treatment of subcenters assumes that they directly affect density, *D*, without being affected by its changes. The literature on the formation of subcenters demonstrates that the exogenous treatment of subcenters presents problems related to their identification and to the role they play in affecting both employment and population density. Recent studies show that the formation of subcenters is endogenous to the process leading to urban development (i.e., subcenters are endogenous to changes in density) (McMillen 2001). Thus this study treats them as endogenous.

TABLE 4.14 3SLS Regression Results—Model III

Equation	Coefficient	Std.Error	P
Trip chaining, TC			
RL	0.0777	0.0158	0.0000
AS	1.0087	0.2241	0.0000
WD	-0.6626	0.0554	0.0000
veh	-0.0292	0.0316	0.3570
act_tt	-0.0009	0.0005	0.0560
act_m	-0.0004	0.0003	0.2650
subc_dist	0.1875	0.0317	0.0000
sch	0.0570	0.0132	0.0000
constant	-2.9369	0.3204	0.0000
Activity space, AS			
TC	0.5389	0.0654	0.0000
D	-0.2817	0.0002	0.0000
act_m	0.0000	0.0050	0.8390
hhinc	0.0182	0.0316	0.0000
r_estd	-0.0018	0.0010	0.0000
constant	3.5109	0.2583	0.0790
Transit demand, TD			
TC	0.2310	0.0782	0.0030
AS	0.2130	0.1103	0.0540
WD	-0.4740	0.0405	0.0000
RL	0.0162	0.0089	0.0700
tswork	0.4463	0.0414	0.0000
prkride	-0.0788	0.0457	0.0840
ts_tod	0.1280	0.0995	0.1980
veh	-0.0641	0.0204	0.0020
constant	-1.3114	0.1379	0.0000
Residential location, RL			
TC	2.4695	0.4889	0.0000
TD	1.1677	0.4700	0.0130
hprice	-2.7930	0.2491	0.0000
hage	-0.0961	0.0080	0.0000
rooms	1.3432	0.1071	0.0000
div	-6.1904	0.5571	0.0000
pov	-4.5075	1.6515	0.0060
own	1.3780	0.1901	0.0000
constant	40.3105	3.0969	0.0000
Density, D			
RL	-0.0091	0.0108	0.4000
AS	-0.5333	0.0710	0.0000
cbd_dist	-0.0401	0.0016	0.0000
subc_dist	-0.0707	0.0301	0.0190
constant	11.7688	0.1800	0.0000

Note: N= 8,212; χ_{TC}^2=2,512.8; χ_{AS}^2=611.2; χ_{TD}^2=1,712.7; χ_{RL}^2=646.3; χ_{D}^2=1,448.6

TABLE 4.15 Elasticity Estimates—Model III

Elasticity	WD	subc_dist	cbd_dist	r_estd	tswork*
TC	-0.067	-0.195	-1.066	0.014	--
AS	-0.060	-0.088	-0.102	-0.009	--
TD	-0.093	-0.522	-1.177	-0.366	0.961
RL	-0.023	-0.076	-0.301	0.011	--
D	-0.012	-0.153	-2.972	-0.002	--

Indicates a proportional change

The elasticity of transit demand with respect to distance to the CBD (– 1.17) is greater in absolute value than the elasticity with respect to distance to the nearest subcenter (– 0.52). In other words transit patronage is more responsive to a residential location near the CBD than near subcenters. This is probably due to differences in existing transit station locations near the CBD compared to suburban areas. This result is inconsistent with recent findings that found increased transit use in better served decentralized urban areas (Brown and Thompson 2008; Thompson and Brown 2006) and empirical findings showing that transit ridership is not affected by the CBD (Brown and Nego 2007).

Next, we subject the models to post-estimation testing to confirm their statistical validity. We also discuss additional factors that could potentially affect the validity of results.

Post Estimation Analysis

The validity of the empirical results hinges on factors associated with the quality of the data used and the statistical techniques employed. This section discusses some of the key factors that might affect the results of the empirical investigation, namely:

- Dataset issues

    o Measurement problems

    o Scaling

- Modeling issues

    o Use of cross-sectional data

    o Misspecification

    o Endogeneity not accounted for

    o Nonlinearities

Dataset Issues

The travel-behavior dataset relies on the travel-diary information from BATS2000. The geographic coordinates of households' residences and their travel destinations allow the calculations of residential location, *RL*, activity space, *AS*, and other measures, such distance from the household residential location to the CBD and the nearest subcenter. Land-use data from the Census 2000 Summary File 3 are measured at the block-group level, while data from the Census county business patterns survey (CBP) are measured at the zip-code level. The different geographical units can lead to scale measurement issues.

Measurement Problems

While we measure residential location as home-work distance, we could have considered other measures as well. For example, an alternative is represented by the average commute time between work and home. This measure has the advantage of accounting for spatial characteristics as well as network characteristics (such as street net-

work design and level of service). It also represents a measure of the opportunity cost of residing at a certain distance from work. This measure can be expressed as

$$RL_{min} = \frac{\sum_{m=1}^{k} dur_{mj}}{k} \tag{5.1}$$

where dur_{mj} is commute length (measured in minutes of travel) to the residential unit located at j from a household member work location m, and k is the total number of employed household members.

In the models of Chapter 4, residential location is measured by linear distance between work and home using geographical coordinates of the residence and the work location, thus providing a relatively accurate measure. In contrast, measuring residential location by travel time entails using the survey reported travel time, which is subject to measurement error (under or overstatement of actual travel time by the respondents) and unobserved factors related to the time the survey was conducted (unobserved, non-random, factors affecting traffic levels during the two-day data collection period).

Chapter 3 and Chapter 4 discussed the definition and measurement of activity space, *AS*, and the adoption of the standard distance ellipse (SDE) to measure the household spatial dispersion of non-work activities. In choosing SDE, we compared it to the second-best alternative, the standard distance circle (SDC). As discussed, the advantage of SDE over SDC is the diminished relevance of outliers. Indeed, sample descriptive statistics showed outlier influence that could not be eliminated without relevant loss of information. In addition, we normalized SDE using a log transformation.

The literature provides additional activity-space measures. For example, while Buliung and Kanaroglou (2006) use SDE, they also introduce the household activity space (HAS). HAS is an area-based geometry that defines a minimum convex polygon

containing activity locations visited by a household during a reference period (i.e., the travel-survey period). The advantage of HAS is that it weights the activity space by the relevance of activities, such as their type (recreational, maintenance, etc.) and their relative frequencies. Although HAS reports an accurate geographical measurement of the activity space, Buliung and Remmel (2008) show that the use of the minimum convex polygon algorithm provides similar results to SDE in terms of behavioral interpretation. Other research shows that the choice of an appropriate shape representing an individual's activity space is highly dependent on the spatial distributions and frequencies of the locations visited by the person in the given time period (Rai et al. 2007).

Scaling Issues

As described in detail in Chapter 4, land use and urban form are measured at two geographic levels. Gross population density is measured at the Census block-group level. This scale of measurement, besides being the level that corresponds closely to the neighborhood, is also consistent with the literature and allows comparison of findings. Retail establishment density, a proxy for land-use mix (commercial land uses) is measured at the zip-code level, which is a wider geographical area. As argued by Boarnet and Crane (2001), this scale is appropriate when investigating the role of non-work travel, as non-work trips usually involve distances more than a block from the residential unit. Also, in the sample dataset, geocodes coincide with traffic analysis zones (TAZ)[5]. In this study, retail establishment density directly affects the activity space. As summarized in Table 4.16, the average size of the activity space is much larger than the average size of a cen-

---

[5] According to the U.S. Census Bureau, a TAZ is a special area delineated by state and/or local transportation officials for tabulating traffic-related data, especially journey-to-work and place-of-work statistics (2008).

sus block group level, while the average size of a zip code is more is approximately the size of the activity space, at least for the household using transit.

TABLE 4.16 Land-Area Geographic Measures

Transit House-hold		Household Activity Space (mile^2)	Block Group Area (mile^2)	Zip Code Area (mile^2)
No	Mean	17.16	2.30	42.88
	SD	38.40	10.92	88.66
	N	10,548	12,260	12,260
Yes	Mean	19.14	0.87	18.37
	SD	37.84	4.18	51.33
	N	2,176	2,503	2,503
Overall Sample	Mean	17.50	2.06	38.72
	SD	38.31	10.11	84.02
	N	12,724	14,763	14,763

Modeling Issues

Post Estimation Tests

The models presented above explicitly deal with endogeneity of urban form and travel by applying simultaneous equation modeling. As seen, the first step requires correctly identifying a model. This step generates models that are either just identified or overidentified, based on the number of exclusion restrictions applied to each equation (See Appendix B for more details).

96

Tests of Endogeneity and Overidentification

A property of the 3SLS regression is its loss of efficiency if the explanatory variables treated as endogenous are, in fact, exogenous, making its use unnecessary when compared to OLS. It is thus useful to test the explanatory variables suspected to be endogenous to the model.

The null hypothesis of the endogeneity test is that an OLS estimator of the same equation would yield consistent estimates; that is, any endogeneity among the regressors would not have deleterious effects on the OLS estimates. A rejection of the null hypothesis indicates that endogenous regressors' effects on the estimates are meaningful, and instrumental variables are required. The test was first proposed by Durbin (1954) and later by Wu (1974) and Hausman (1978). The procedure to test endogeneity of multiple explanatory variables requires (i) estimating in reduced form each endogenous variable on all exogenous variables (including those in the structural equation and those used as instruments; i.e., the explanatory variable included in the other equations); (ii) adding the estimated error terms back into the structural equation; and, (iii) testing for the joint significance of these residuals in the structural equation. Joint significance indicates that at least one variable is endogenous to the model. Under the null hypothesis, the test statistic is distributed χ_q^2 (Chi-squared) with q degrees of freedom, where q is the number of regressors specified as endogenous in the original instrumental variables regression. The procedures to conduct this test are available in Stata® (the statistical package used in this study) using the *ivreg2* routine (specifically, by using the command *ivendog*) developed by Baum et al. (2007).

Furthermore, after verifying the presence of endogeneity, additional tests are needed to confirm the correct choice of the exclusion restrictions characterizing the system of equation. These tests are needed to confirm the proper choice of instruments and to eliminate doubts of a poor model performance (bias and inconsistency). The overidentification tests used here are conducted by regressing the residuals from a 3SLS regression on all exogenous variables (both included exogenous regressors and excluded instruments). Under the null hypothesis that all instruments are uncorrelated with the residuals, a Lagrangean multiplier (LM) statistic of the form NxR^2 (N = number of regressors, while R^2 is calculated from the residuals' regression), has a large sample Chi-squared distribution, χ_r^2, where r is the number of *overidentifying* restrictions (i.e., the number of excess instruments). If the hypothesis is rejected, there is doubt about the validity of the instrument set; one or more of the instruments do not appear to be correlated with the disturbance process. The Stata® procedure reports the Sargan (1958) overidentification test (using the *overid* command).

Finally, when dealing with a relatively large number of exclusion restrictions, a situation encountered in Model III, it has been shown that the power of the overidentification tests is reduced (Baum, Schaffer, and Stillman 2007). Furthermore, there is a need to be able to test subsets of instruments to identify weak ones, which would adversely affect validity of results. In this context, another test statistic can be used to test a subset of instruments; the *difference-in-Sargan* test, or *C* test. The statistic is computed as the difference between two statistics; one obtained by regression using the entire set of instruments and a second one obtained with the smaller set of restrictions (excluding the suspected variables). Under the null hypothesis that the variables are proper instruments,

the *C*-test statistics is distributed χ_k^2 with *k* degrees of freedom equal to the number of suspect instruments being tested.

Table 4.17 reports the results of the endogeneity and overidentification tests for the travel demand equation, *TD* (the same tests and same results were obtained for the other equations but are not reported here). The Durbin-Wu-Hausman (DWH) test is numerically equivalent to the standard Hausman endogeneity test. Results across the three models indicate the presence of endogeneity, confirming the appropriateness of 3SLS versus OLS regression.

Model III fails the overidentification test in its initial specification that treated the land-use measures (*cbd_dist*, *subc_dist*, *r_estd*) as exogenous to the system (Sargan test = 24.951; *p-value* = 0.0030). After their endogenous treatment, Model III passes the overidentification test, as signaled both by the Sargan (7.1540 with *p-value* of 0.3068) and *C* tests.

Overall, the tests indicate that SEM is an appropriate technique and that the equation specifications of Chapter 4 produce models that also pass the overidentification tests. The validity of the models allows making conclusions regarding the parameters of interest.

Other Issues

The use of SEM is best exploited in the context of panel datasets, which are better suited to uncover underlying causality among the relationships of interest. In the transportation literature there exist several applications of SEM using cross-sectional data. For example, Pendyala (1998) uses SEM to investigate the homogeneity of causal travel behavior across a population of interest; Fuji and Kitamura (2000) and Golob (2000) de-

velop models of trip generation developing models of activity duration and trip genera-

tion. Additional examples of applications of SEM using cross-sectional datasets are dis-

cussed by Golob (2003).

TABLE 4.17 Endogeneity and Overidentification Tests

Test	Model I	Model II	Model III
Wu-Hausman F test	78.073	83.369	13.000
p-value	0.000	0.000	0.000
Durbin-Wu-Hausman χ^2 test	153.423	243.059	90.217
χ^2 p-value	0.000	0.000	0.000
Anderson canon. corr. LR statistic (identification/IV relevance test):	42.137	27.137	33.524
χ^2 p-value	0.000	0.003	0.000
Sargan statistic (overidentification test of all instruments):	9.638	11.365	24.951
χ^2 p-value	0.057	0.252	0.003
Sargan statistic without suspect instruments*	-	-	7.154
χ^2 p-value	-	-	0.307
C statistic (exogeneity/orthogonality of suspect instruments)**			17.798
χ^2 p-value			0.001

Test conducted after endogenous treatment of: cbd_dist, subc_dist, r_estd
** Test conducted on exclusion of instruments: cbd_dist, subc_dist, r_estd

The models of this study require a substantial amount of information, not only in

terms of travel behavior data from travel diaries, but also on the spatial location of resi-

dences, work, and non-work activities.

The increased sophistication of communication systems that can easily track indi-

viduals' travel patterns in space and time makes the data-collection effort less daunting

than otherwise, allowing increased used of sophisticated models, such as the ones devel-

oped in this study. For example, the recent uses of GPS tracking devices reveals that

human behavior results in optimized patterns of travel based on socio-demographic characteristics. These methods not only allow tracking travel and non-work activity locations, they also provide more accurate measures of travel itself, such as actual travel-time speed based on network characteristics.

Transit-Station Proximity

Notwithstanding the validity of the above post-estimation tests, there still exists the possibility of endogeneity of some of the exogenous variables. This endogeneity, although confuted by statistical tests, is not ruled out by theoretical assumptions. For example, while this study treats vehicle ownership as exogenous and not directly influenced by the location decision, the literature contains studies that consider vehicle ownership as a discrete-choice variable endogenous to the residential location process and to density levels. One extension of this dissertation might include an endogenous treatment of this variable, while overcoming the limitations imposed by ad-hoc choice-set specifications.

Endogeneity also extends to transit supply measures. For example, measures of supply, such as the number of transit stations and frequency of service are treated as exogenous to the model. As discussed in several places throughout this dissertation, the implications of treating a variable as exogenous, while being endogenous to the process, are not trivial.

An additional consideration must be made regarding the use of walking distance as a measure of transit-station proximity that cannot be made when using the more traditional half-mile buffer. As density increases, the number of transit stops at the geographical unit (i.e., block group) increases. This reduces the average distance from any given household to its nearest transit station independently of location preferences. Further-

101

more, as shown in Figure 4.2, in densely populated areas, stations are located in neighborhoods characterized by higher than average poverty levels, and that are increasingly diverse (i.e., characterized by ethnic minorities). In other words, in higher urban density settings, a supply-side spatial bias is present and correlated with relevant instrumental variables that control for neighborhood characteristics. For this reason, Model III, which endogenously treats residential location and density, considers walking distance as endogenous.
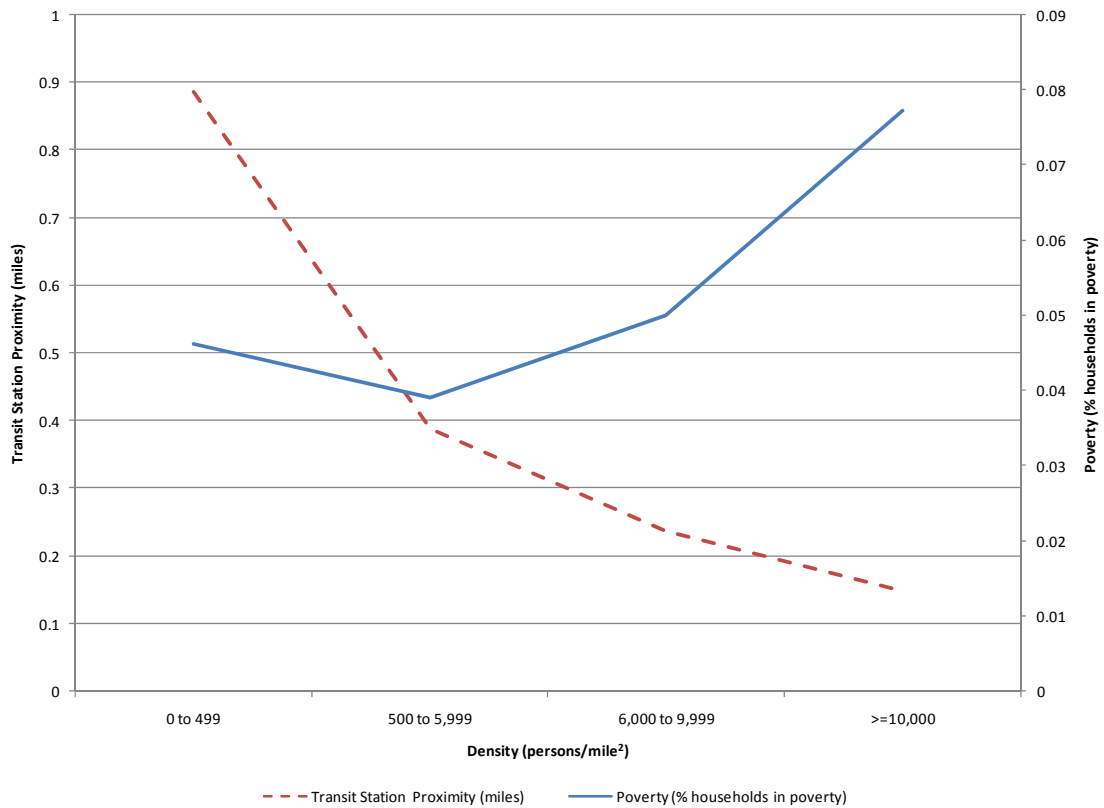

FIGURE 4.2 Poverty and Transit-Station Proximity

Implications

The models of Chapter 3 are innovative in many aspects, above all for its explicit incorporation of the links between consumption, travel, the spatial location of non-work activities, and the ensuing interrelationship with the surrounding built environment.

The empirical application of the behavioral model requires the use of simultaneous equation modeling. The biggest challenge when employing structural equation modeling lies in defining properly specified models. The necessary identification steps outlined in Chapter 4 and summarized in Appendix B are paramount to reliable estimates. The literature reviewed in this study revealed that none of the papers and studies formally follows this process. The result is the estimation and presentation of sets of parameters that are not unique, which make statistical inference unreliable. The validity of the empirical models of Chapter 4 is confirmed by the relevant endogeneity and overidentification tests presented in this chapter.

Chapter 5: Conclusions

Summary of Findings

  This dissertation research sought to overcome shortcomings of the empirical lite-rature modeling of the relationship between transit travel behavior and urban form. A review of the current state of empirical research on the subject uncovered the main weak-nesses of findings relating the built environment to travel behavior as well as noting the paradigm shift epitomized by the activity-based literature. The findings of this review show that there has been a shift from the study of density threshold levels that make tran-sit cost-feasible to an analysis of the effect of urban design and land-use mix on travel behavior, after controlling for density levels. The issue is no longer at what density thre-sholds it makes sense to implement transit, but what is the best set of policies affecting urban design and land-use mix that most influences the spatial arrangements of activity locations, so that individuals are more likely to utilize transit. This shift is reflected by an increasing number of studies that assess the relevance of transit-oriented development (TOD) to transit use when households or individuals prefer certain urban setting to oth-ers.

  While early work sought to provide a framework that made use of aggregate data, the more recent literature models the simultaneous decision of location and travel when individuals choose locations based on idiosyncratic travel preferences.

Finally, there is a lack of empirical work that examines the relationship between urban form and travel behavior within an analytical framework that takes into account the complexity of travel by considering trip chaining among other travel complexities. To avoid these shortcomings and to incorporate the activity-based approach, we developed and estimated a simultaneous equation model of transit usage and urban form.

Empirically Estimable Model of Transit and Urban Form

The models of Chapter 3 allow household travel to respond to changes in urban form, by considering trip-chaining for non-work travel. In the model, trip-chaining results from households' reductions in non-work travel time while accounting for constraints that the built environment imposes. Any travel-time saving is spent on additional non-work travel or provides inducement to reassess residential location decisions. These changes in travel behavior and residential location then affect the demand for travel.

The constraints imposed by the built environment are captured by the activity space. Empirical evidence in Chapter 4 shows that lower densities define a larger activity space, which, in turn, decreases transit use. Conversely, as density increases, the activity space contracts, as does the need to engage in complex trip chains. Idiosyncratic preferences for transit also affect transit demand. For example, in the absence of adequate transit, households that need to engage in complex trip-chain patterns, independent of changes in the surrounding built-environment, may use the automobile. In contrast, if adequate transit services are available to accommodate their travel patterns, households would choose transit, other things equal.

To facilitate a summary of Chapter 4's findings and for ease of comparison, Table 5.1 presents elasticities from the three estimated models (only statistically significant results are shown).

Exogenous density change does not have a large effect on transit demand, and the magnitude of the effect decreases when residential location becomes endogenous. A 20-percent increase in gross population density (1,830 persons per square mile) increases transit demand from a minimum of 5.4 percent to a maximum of 9.5 percent.

TABLE 5.1 Relevant Land-Use and Transit-Supply Elasticities of Transit Demand

Elasticity	Model I[a]	Model II[b]	Model III[c]
Density	0.475	0.269	n/a
Walking distance	-0.137	-0.028	-0.093
Transit station at workplace*	0.687	0.766	0.961
TOD station*	0.279	n/a	n/a
Distance to CBD	n/a	n/a	-1.177
Distance to nearest subcenter	-0.388	-0.065	-0.522
Retail establishments density	0.001	0.170	n/a
Residential location	-0.157	n/a	n/a

[a] residential location exogenous; density exogenous

[b] residential location endogenous; density exogenous

[c] residential location and density endogenous

n/a = not available

* Indicates a proportional change

The importance to transit demand of station proximity, as measured by walking distance, decreases after accounting for idiosyncratic preferences for location. In Model III, the elasticity of transit demand with respect to walking distance is about one-third

smaller than in Model I, in which residential location and density are exogenous. This decline in magnitude is due to allowing households to choose their residential location and by accounting for omitted-variable bias error. This contrasts with what found by Cervero (2007), who shows that self-selection accounts for about 40 percent of transit ridership for individuals residing near a transit station.

The presence of a transit station in proximity to a workplace also has a significant positive impact on ridership, as indicated by the magnitude of the proportional changes across all three models.

In Model I, transit-oriented development near transit stations has a positive impact on transit use; a TOD stop increases transit demand by about 28 percent. In conformity to the literature, a transit station near a workplace exerts a positive impact on ridership, as indicated by the magnitude of the proportional changes across all three models.

An established central business district (CBD) is still a relevant driver of transit use, as highlighted by an elasticity of transit demand with respect to distance to the CBD of -1.17. Although subcenters play a less important role, our findings support a policy of providing transit services in decentralized employment and residential areas to increase ridership.

The importance of mixed-use development to increase transit patronage is highlighted by the elasticity of travel demand with respect to retail establishment density. Model II shows that a 20-percent increase in retail establishment density (or about 28 establishments per square mile) increases transit demand by 3.4 percent.

Households living farther from work, as measured by residential location use less transit, which is due to trip-chaining behavior. Such households engage in complex trip

107

chains and have, on average, a more dispersed activity space, which requires reliance on more flexible modes of transportation. The results support policies that would reduce the spatial allocation of activities and improve transit accessibility at and around subcenters. Similar results can be obtained by policies that increase the presence of retail locations in proximity to transit-oriented households.

Research Contributions

The major contribution of this research effort is the development of a simultaneous equation model of transit patronage and land-use that acknowledges the interrelationship between travel behavior and urban form. In particular, the framework embraces the paradigm shift from trip generation to activity-based modeling by considering travel demand as a derived demand brought about by the necessity to engage in out-of-home activities. In addition, this framework presented in Chapter 3 departs from the monocentric models of residential location, which do not account for decentralized work places, by explicitly acknowledging both the presence and the relevance of subcenters. The models take into account for the trade-off between consumption and travel brought about by the finite nature of time and its allocation among household members.

Another contribution of this dissertation is the empirical treatment of density as an explanatory variable for trip-making behavior. As opposed to the current practice of regressing trip making behavior against density measures, we assume that density does not directly to affect the demand for travel. In our models, density first directly affects the spatial dispersion of goods and services, as measured by the activity space. It is only by affecting the size of the activity space that density affects both trip chaining and the de-

mand for transit services. The consequences introduced by this structure are not trivial and as demonstrated by the empirical results.

In addition, the empirical analysis shifts the analysis from individual travel behavior to household travel behavior, recognizing that travel decisions are taken jointly among individuals. The models

Finally, the empirical work takes advantage of the advances in geographic information systems (GIS) tools and geographic science contributions to the spatial analysis of the interactions of travel behavior and urban form.

Directions for Further Research

Notwithstanding the validity of the post-estimation tests performed in Chapter 5, there still exists the possibility that some of the variables treated as exogenous are, in fact, endogenous. For example, this study treats vehicle ownership as exogenous. The literature review, however, revealed studies that consider vehicle ownership endogenous to residential location and density. One extension to this research, therefore, would be to include an endogenous treatment of this and other mode-choice variables.

Another extension would be to include leisure time available to households. Indeed, the behavioral model of Chapter 3 assumes that households can save time by engaging in trip chaining. Time savings are then reallocated to either more non-work travel or to an extended commute. The model does not explicitly explain what happens to leisure time. The inclusion of total time constraints that includes all relevant time uses (in-home and out-of-home) would provide insight on time use and its effect on trip chaining.

Finally, in contrast to multiple linear regression analysis, nonparametric estimation methods would permit less restrictive assumptions. These methods can uncover the pres-

ence of nonlinearities among dependent and independent variables which could lead to a

better parameterization of equations of interest. Although nonlinearity in trip-chaining

formation and density levels is better captured by these methods than by more commonly

used techniques, being computationally challenging, they are rarely used in applied work,

especially in the field of travel behavior research and simultaneous equation modeling.

Further research that makes use of these methods is warranted.

References

Adler, Thomas and Moshe Ben-Akiva. 1979. "A Theoretical and Empirical Model of Trip Chaining Behavior." *Transportation Research Part B: Methodological* 13B, 243–257.

Alonso, William. 1964. *Location and Land Use*. Cambridge, Mass.: Harvard University Press.

Anas, Alex. 2007. "A Unified Theory of Consumption, Travel and Trip Chaining." *Journal of Urban Economics* 62, 162–186.

Anas, Alex, Richard Arnott and Kenneth A. Small. 1998. "Urban Spatial Structure." *Journal of Economic Literature* 36, 1426--1464.

Anas, Alex and Ikki Kim. 1996. "General Equilibrium Models of Polycentric Urban Land Use with Endogenous Congestion and Job Agglomeration." *Journal of Urban Economics* 40, 232–256.

Anas, Alex and Rong Xu. 1999. "Congestion, Land Use and Job Dispersion: A General Equilibrium Analysis." *Journal of Urban Economics* 45, 451–473.

Arnott, Richard. 1998. "Economic Theory and the Spatial Mismatch Hypothesis." *Urban Studies* 35, 1171--1185.

Badoe, Daniel A. and Eric J. Miller. 2000. "Transportation--Land-Use Interaction: Empirical Findings in North America, and Their Implications for Modeling." *Transportation Research Part D* 5, 235–263.

Bagley, Michael N. and Patricia L. Mokhtarian. 2002. "The Impact of Residential Neighborhood Type on Travel Behavior: A Structural Equations Modeling Approach." *The Annals of Regional Science* 36, 279–297.

Baum-Snow, Nathaniel and Matthew E. Kahn. 2000. "The Effects of New Public Projects to Expand Urban Rail Transit." *Journal of Public Economics* 77, 241–263.

Baum, Christopher F., Mark E. Schaffer and Steven Stillman. 2007. "Enhanced Routines for Instrumental Variables/Generalized Method of Moments Estimation and Testing." *Stata Journal* 7, 465–506.

Begon, M. J. and C.R. Towsend. 1996. *Ecology: Individuals, Populations, and Communities*. Cambridge, MA: Blackwell Science Ltd.

Ben-Akiva, Moshe and John L. Bowman. 1998. "Integration of an Activity-Based Modeling System and a Residential Location Model." *Urban Studies* 35, 1131–1153.

Bento, Antonio M., Maureen L. Cropper, Ahmed Mushfiq Mobarak and Katja Vinha. 2005. "The Effects of Urban Spatial Structure on Travel Demand in the United States." *The Review of Economics and Statistics* 87, 466–478.

Bertaud, Alain. 2003. "Clearing the Air in Atlanta: Transit and Smart Growth or Conventional Economics?" *Journal of Urban Economics* 54, 379–400.

Bhat, Chandra. 1997. "An Econometric Framework for Modeling the Daily Activity-Travel Pattern of Individuals," paper presented at Activity-Based Workshop National Institute of Statistical Sciences, North Carolina

———. 1999. "An Analysis of Evening Commute Stop-Making Behavior Using Repeated Choice Observations from a Multi-Day Survey." *Transportation Research Part B: Methodological* 33, 495–510.

———. 2001. "Modeling the Commute Activity-Travel Pattern of Workers: Formulation and Empirical Analysis." *Transportation Science* 35, 61–79.

Bhat, Chandra and Jessica Guo. 2004. "A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modeling." *Transportation Research Part B: Methodological* 38, 147-168.

———. 2004. "A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modeling." *Transportation Research Part B: Methodological* 38, 147–168.

Bhat, Chandra and Frank Koppelman. 1999. "Activity-Based Modeling of Travel Demand," in R. W. Hall (ed.),*The Handbook of Transportation Science* Massachussetts: Kluwer Academic Publishers. pp. 35–61.

Boarnet, Marlon and Randall Crane. 2001. "The Influence of Land Use on Travel Behavior: Specification and Estimation Strategies." *Transportation Research Part A* 35, 823–845.

Boarnet, Marlon G. and Sharon Sarmiento. 1998. "Can Land-Use Policy Really Affect Travel Behaviour? A Study of the Link between Non-Work Travel and Land-Use Characteristics." *Urban Studies* 35, 1155–1169.

Boarnet, Marlon Gary and Randall Crane. 2001. "Travel by Design : The Influence of Urban Form on Travel," in,*Spatial Information Systems* Oxford ; New York: Oxford University Press. pp. 33–59.

———. 2001. *Travel by Design : The Influence of Urban Form on Travel*. Oxford ; New York: Oxford University Press.

Boarnet, Marlon Gary and Sharon Sarmiento. 1996. *Can Land Use Policy Really Affect Travel Behavior : A Study of the Link between Non-Work Travel and Land Use Characteristics*. Berkeley, Calif.: University of California Transportation Center.

Bollinger, Christopher R. 1997. "The Impact of Rapid Rail Transit on Economic Development: The Case of Atlanta's Marta." *Journal of Urban Economics* 42, 179–204.

Bowes, David R. and Keith R. Ihlanfeldt. 2001. "Identifying the Impacts of Rail Transit Stations on Residential Propery Values." *Journal of Urban Economics* 50, 1–25.

Brown, Jeffrey R. and Dristi Nego. 2007. "Urban Structure and Transit Ridership: A Reexamination of the Relationship in the United States," paper presented at 87th Annual Meeting of the Transportation Research Board Washington, D.C. .

Brown, Jeffrey R. and Gregory L. Thompson. 2008. "The Relationship between Transit Ridership and Urban Decentralization:Insights from Atlanta." *Urban Studies* 45, 1119–1139.

BTS. 2001. "Highlights of the 2001 National Household Travel Survey," in: Bureau of Transportation Statistics.

Buliung, Ron and Pavlos Kanaroglou. 2006. "Urban Form and Household Activity-Travel Behaviour." *Growth and Change* 37, 174–201.

Buliung, Ron and Tarmo K. Remmel. 2008. "Open Source, Spatial Analysis, and Activity-Travel Behavior Research: Capabilities of the *Aspace* Package." *Journal of Geography Systems* 10, 191–216.

Burchell, Robert W., Naveed A. Shad, David Listokin and Hilary Phillips. 1998. "The Cost of Sprawl - Revisited," *TCRP,* 39. Washington, Transportation Research Board.

CALTRANS. 2008. "California Transit-Oriented Development (Tod) Database," in CALTRANS (ed.): California Department of Transportation.

Cao, Xinyu, Patricia L. Mokhtarian and Susan L. Handy. 2006. "Examining the Impacts of Residential Self-Selection on Travel Behavior: Methodologies and Empirical Findings." California, University of California.

———. 2007. "Do Changes in Neighborhood Characteristics Lead to Changes in Travel Behavior? Structural Equations Modeling Approach." *Transportation* 35, 535–556.

Casello, Jeffrey M. 2007. "Transit Competitiveness in Polycentric Metropolitan Regions." *Transportation Research Part A* 41, 19–40.

"Census 2000 Summary File 3." 2007. *2000 Census of Population and Housing - Technical Documentation*. U.S. Census Bureau.

Cervero, Robert. 2007. "Transit Oriented Development's Ridership Bonus: A Product of Self Selection and Public Policies." *Environment and Planning A* 39, 2068–2085.

Cervero, Robert and Kara Kockelman. 1997. "Travel Demand and the 3ds: Density, Diversity, and Design." *Transportation Research Part D* 2, 199–219.

Cervero, Robert and John Landis. 1997. "Twenty Years of the Bay Area Rapid Transit System: Land Use and Development Impacts." *Transportation Research Part A* 31, 309–333.

Cervero, Robert, Steven Murphy, Christopher Ferrel, Natasha Goguts and Yu-Hsin Tsai. 2004. "Transit-Oriented Development in the United States: Experiences, Challenges, and Prospects," *TCRP* TCRP Report 102. Washington, Transportation Research Board.

Cervero, Robert and Kang-Li Wu. 1998. "Sub-Centring and Commuting: Evidence from the San Francisco Bay Area, 1980-90." *Urban Studies* 35, 1059–1076.

Chu, You-Lian. 2003. "Empirical Analysis of Commute Stop-Making Behavior." *Transportation Research Record* 1831, 106–113.

Committee on Physical Activity, Health, Transportation, and Land Use. 2005. "Does the Built Environment Influence Physical Activity? Examining the Evidence," ISBN 0-309-09498-4. Transportation Research Board, Institute of Medicine of the National Academies.

Concas, Sisinnio and Philip L. Winters. 2007. "The Impact of Carpooling on Trip Chaining Behavior and Emission Reductions." *Transportation Research Record* 2010, 83–91.

Crane, Randall. 2000. "The Influence of Urban Form on Travel: An Interpretive Review." *Journal of Planning Literature* 15, 1–23.

Crane, Randall and Richard Crepeau. 1998a. *Does Neighborhood Design Influence Travel? A Behavioral Analysis of Travel Diary and Gis Data*. Berkeley: University of California Transportation Center.

———. 1998b. "Does Neighborhood Design Influence Travel?: A Behavioral Analysis of Travel Diary and Gis Data." *Transportation Research Part D* 3, 225–238.

Davidson, D. 1991. "Impact of Suburban Employee Trip Chaining on Transportation Demand Management." *Transportation Research Record* 1321, 82–89.

Davidson, William, Robert Donnelly, Peter Vovsha, Joel Freedman, Steve Ruegg, Jim Hicks, Joe Castiglione and Rosella Picado. 2007. "Synthesis of First Practices and

Operational Research Appraoches in Activity Based Travel Demand Modeling." *Transportation Research Part A* 41, 464–488.

DOE. 1996. "Policy and Procedure Guidelines," in D. o. Environment (ed.).

Domencich, Thomas A. and Daniel McFadden. 1975. *Urban Travel Demand : A Behavioral Analysis : A Charles River Associates Research Study*. New York: North-Holland.

Dong, Xiaojing, Ben-Akiva Moshe, John L. Bowman and Joan L. Walker. 2006. "Moving from Trip-Based to Activity-Based Measures of Accessibility." *Transportation Research Part A* 40, 163–180.

Durbin, J. 1954. "Errors in Variables." *Review of the International Statistical Institute* 22, 23–32.

Ebdon, David. 1977. *Statistics in Geography: A Practical Approach*. Oxford: Blackwell Publishing.

Evans, John E., Richard H. Pratt, Andrew Stryker and J. Richard Kuzmyak. 2007. "Chapter 17 - Transit Oriented Development," *TCRP Report 95 - Traveler Response to Transportation System Changes,* 17. Washington, Transit Cooperative Research Program.

Ewing, R. 1997. "Is Los Angeles-Style Sprawl Desirable? ." *Journal of the American Planning Association* 63, 107–126.

Ewing, Reid and Robert Cervero. 2001. "Travel and the Built Environment - a Synthesis." *Transportation Research Record*, 87–114.

Fuji, S. and R. Kitamura. 2000. "Evaluation of Trip-Reducing Effects of New Freeways Using a Structural Equations Model System of Commuters' Time Use and Travel." *Transportation Research Part B: Methodological* 34, 339–354.

Giuliano, Geneveve and Kenneth A. Small. 1991. "Subcenters in the Los Angeles Region." *Regional Science and Urban Economics* 21, 163–182.

Golob, Thomas F. 2000. "A Simultaneous Model of Household Activity Participation and Trip Chain Generation." *Transportation Research Part B: Methodological* 34, 355–376.

———. 2003. "Structural Equation Modeling for Travel Behavior Research." *Transportation Research Part B: Methodological* 37, 1–25.

Gossen, Rachel. 2005. "Research Papers Using San Francisco Bay Area Survey Data Sets." Metropolitan Transportation Commission.

Greenwald, Michael J. and Marlon Boarnet. 2001. "Built Environment as Determinant of Walking Behavior: Analyzing Nonwork Pedestrian Travel in Portland, Oregon." *Transportation Research Record* 1780, 33–42.

Guevara, Cristian Angelo and Ben-Akiva Moshe. 2006. "Endogeneity in Residential Location Models." *Transportation Research Record* 1977, 60–66.

Hägerstrand, T. 1970. "What About People in Regional Science? ." *Papers of the Regional Science Association*, 7–21.

Handy, Susan, Xinyu Cao and Patricia Mokhtarian. 2005. "Correlation or Causality between the Built Environment and Travel Behavior? Evidence from Northern California." *Transportation Research Part D* 10, 427–444.

Hausman, J.A. 1978. "Specification Tests in Econometrics." *Economics Letters* 8, 239–245.

Jara-Diaz, Sergio R. and Francisco J. Martinez. 1999. "On the Specification of Indirect Utility and Willingness to Pay for Discrete Residential Location Models." *Journal of Regional Science* 39, 675–688.

Jones, P, Frank Koppelman and J. P. Orfeuil. 1990. "Activity Analysis: State-of-the-Art and Future Directions," in,*Developments in Dynamic and Activity-Based Approaches to Travel Analysis* Aldershot: Aveburey. pp. 34–55.

Kahn, Matthew. 2007. "Gentrification Trends in New Transit-Oriented Communities: Evidence from 14 Cities That Expanded and Built Rail Transit Systems." *Real Estate Economics* 35, 155–182.

Kain, John F. 1968. "Housing Segregation, Negro Employment, and Metropolitan Decentralization." *Quarterly Journal of Economics* 82, 175–197.

Kondo, Katsunao and Ryuichi Kitamura. 1987. "Time-Space Constraints and the Formation of Trip Chains." *Regional Science and Urban Economics*, 49–65.

Kuby, Michael, Anthony Barranda and Christopher Upchurch. 2004. "Factors Influencing Light-Rail Station Boardings in the United States." *Transportation Research Part A* 38, 223–247.

Kuppam, Arun K. and Ram M. Pendyala. 2001. "A Structural Equation Analysis of Commuters' Activity and Travel Patterns." *Transportation* 28, 35–54.

Levine, Ned. 2005. "Crimestat Iii: A Spatial Statistics Program for the Analysis of Crime Incident Locations ". Washington D.C. , National Institue of Justice.

Maat, Kees and Harry J.P. Timmermans. 2006. "Influence of Land Use on Tour-Based Travel in the Netherlands." *Transportation Research Record*, 234–241.

Maoh, H and Pavlos Kanaroglou. 2007. "Geographic Clustering of Firms and Urban Form: A Multivariate Analysis." *Journal of Geography Systems*, 229–252.

McDonald, John F. and Clifford I. Osuji. 1995. "The Effect of Anticipated Transportation Improvement on Residential Land Values." *Journal of Regional Science and Urban Economics* 25, 261–278.

McFadden, Daniel. 1973. *Conditional Logit Analysis and Qualitative Choice Behavior*. New York: Academic Press.

McGuckin, Nancy and Elaine Murakami. 1999. "Examining Trip-Chaining Behavior: Comparison of Travel by Men and Women." *Transportation Research Record* 1693, 79–85.

McMillen, Daniel P. 2001. "Nonparametric Employment Subcenter Identification." *Journal of Urban Economics* 50, 448–473.

Mills, Edwin S. 1972. *Urban Economics*. Glenview, Ill.,: Scott Foresman.

Misra, R and Chandra Bhat. 2002. "Activity Travel Patterns of Non-Workers in the San Francisco Bay Area: Exploratory Analysis." *Transportation Research Record* 1718, 43–51.

Misra, R, Chandra Bhat and S Srinivasan. 2003. "Continuous Time Representation and Modeling Framework for Analysis of Nonworker Activity-Travel Patterns: Tour Episodes and Attributes." *Transportation Research Record* 1831, 11–20.

Modarres, Ali. 2003. "Polycentricity and Transit Service." *Transportation Research Part A* 37, 841–864.

Mokhtarian, Patricia L. and Xinyu Cao. 2008. "Examining the Impacts of Residential Self-Selection on Travel Behavior: A Focus on Methodologies." *Transportation Research Part B: Methodological* 42, 204–228.

Moshe, Ben-Akiva and John L. Bowman. 1998. "Integration of an Activity-Based Model System and a Residential Location Model." *Urban Studies* 35, 1131–1153.

MTC. 2006. "Characteristics of Rail and Ferry Station Area Residents in the San Francisco Bay Area: Evidence from the 2000 Bay Area Travel Survey." San Francisco, CA, Metropolitan Transportation Commission.

———. 2008. "Gis Maps and Data," in M. P. Commission (ed.),*Maps and Data* San Francisco, CA: Metropolitan Planning Commission.

———. 2008. "Household Travel Survey Data and Reports," in M. T. Commission (ed.),*Maps and Data* San Francisco, CA: Metropolitan Transportation Commission.

Muth, Richard F. 1969. *Cities and Housing; the Spatial Pattern of Urban Residential Land Use*. Chicago,: University of Chicago Press.

Nelson, Peter, Andrew Baglino, Winston Harrington, Elena Safirova and Abram Lipman. 2007. "Transit in Washington, D.C.: Current Benefits and Optimal Level of Provision." *Journal of Urban Economics* 62, 231–251.

NHTS. 2007. "Congestion: Who Is Traveling in the Peak?," in,*NHTS Brief*: U.S. Department of Transportation.

Nishii, K, K Kondo and R Kitamura. 1988. "Empirical Analysis of Trip Chaining Behavior." *Transportation Research Record*, 48–59.

Noland, Robert and John V. Thomas. 2007. "Multivariate Analysis of Trip-Chaining Behavior." *Environment and Planning B: Planning and Design*, 953–970.

Oster, C.V. 1978. "Second Role of Work Trip - Visiting Nonwork Destinations." *Transportation Research Record*, 79–89.

Pass, Eric I. 1985. " State of the Art and Research Opportunities in Travel Demand: Another Perspective." *Transportation Research Part A* 19, 460–464.

Pendyala, Ram. 1998. "Causal Analysis in Travel Behavior Research: A Cautionary Note," in S. Jara-Diaz (ed.),*Travel Behaviour Research: Updating the State of Play* Oxford: Pergamon. pp. 35–48.

Peng, Zhong-Ren, Kenneth J. Duecker, James Strathman and Janet Hopper. 1997. "A Simultaneous Route-Level Transit Patronage Model." *Transportation* 24, 159–181.

Pinjari, Abdul Rawoof, Ram M. Pendyala, Chandra R. Bhat and Paul A. Waddel. 2007. "Modeling Residential Sorting Effects to Understand the Impact of the Built Environment on Commute Mode Choice." *Transportation* 34, 557–573.

Purvis, Chuck. 2008. "Bats2000_Hh_Proximity_Nearest_Transit.Zip," in: Metropolitan Transportation Commission.

Pushkarev, Boris S. and Jeffrey M. Zupan. 1977. *Public Transportation and Land Use Policy*. Ontario, Canada: Fitzhenry & Whiteside.

———. 1982. *Urban Rail in America: An Exploration of Criteria for Fixed-Guideway Transit*. Ontario, Canada: Fitzhenry & Whiteside.

Pushkarev, Boris, Jeffrey M. Zupan and Regional Plan Association (New York N.Y.). 1977. *Public Transportation and Land Use Policy*. Bloomington: Indiana University Press.

Rai, R.K., Michael Balmer, Marcel Rieser, V.S. Vaze, Stefan Schonfelder and Kay W. Axhausen. 2007. "Capturing Human Activity Spaces." *Transportation Research Record* 2021, 70–80.

Reilly, Michael and John Landis. 2002. "Title": Institute of Urban and Regional Development.

Sargan, J. 1958. "The Estimation of Economic Relationships Using Instrumental Variables." *Econometrica* 26, 393–415.

Schimek, Paul. 1996. "Household Motorvehicle Ownership and Use: How Much Does Residential Density Matter?" *Transportation Research Record*, 120–125.

Shiftman, Yoram. 1998. "Practical Approach to Trip Model Chaining " *Transportation Research Record* 1645, 17–23.

Strathman, James G. 1995. "Understanding Trip Chaining, Special Report on Trip and Vehicle Attributes," *1990 NPTS Report Series*. Washington, D.C., U.S. Department of Transportation.

Strathman, James G., Kenneth J. Dueker and Judy S. Davis. 1994. "Effects of Household Structure and Selected Travel Characteristics on Trip Chaining." *Transportation*, 23–45.

Taylor, Brian D. and Douglas Miller. 2003. "Analyzing the Determinants of Transit Ridership Using a Two-Stage Least Square Regression on a National Sample of Urbanized Areas," in,*2004 Transportation Research Board 83rd Annual Meeting* Washington, D.C. .

TCRP. 1995. "An Evaluation of the Relationships between Transit and Urban Form," in,*Research Results Digest* Washington: Transit Cooperative Research Program. pp. 1-55.

Thompson, Gregory L. and Jeffrey R. Brown. 2006. "Explaining Variation in Transit Ridership in U.S. Metropolitan Cities between 1990 and 2000: A Multivariate Analysis," in. pp. 172–181.

Vance, Colin and Ralf Hedel. 2007. "The Impact of Urban Form on Automobile Travel: Disentangling Causation from Correlation." *Transportation* 34, 575–588.

Voith, Richard. 1991. "The Long-Run Elasticity of Demand for Commuter Rail Transportation." *Journal of Urban Economics* 30, 360–372.

———. 1997. "Fares, Service Levels, and Demographics: What Determines Commute Rail Ridership in the Long Run?" *Journal of Urban Economics* 41, 176–197.

Wallace, Brett, Jannifer Barnes and G. Scott Rutherford. 2000. "Evaluating the Effects of Traveler and Trip Characteristics on Trip Chaining, with Implications for Transportation Demand Management Strategies. Transportation Research Record." *Transportation Research Record* 1718, 97–106.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass.: MIT Press.

———. 2002. *Introductory Econometrics: A Modern Approach*. Mason, OH: Thomson South-Western.

Wu, D.M. 1974. "Alternative Tests of Independence between Stochastic Regressors and Disturbances: Finite Sample Results." *Econometrica* 42, 529–546.

Zegras, P Christopher. 2004. "The Influence of Land Use on Travel Behavior: Empirical Evidence from Santiago De Chile," in. pp. 175–182.

Zheng, Siqi and Matthew E. Kahn. 2008. "Land and Residential Property Markets in a Booming Economy: New Evidence from Beijing." *Journal of Urban Economics* 63, 743–757.

Zupan, Jeffrey, Robert Cervero, Inc Parsons Brinckerhoff Quade & Douglas and Inc Howard/Stein-Hudson Associates. 1996. "Commuter and Light Rail Transit Corridors: The Land Use Connection," *TCRP Project H-1 Transit and Urban Form* TCRP 16. Transportation Research Board National Research Council.

Appendices

Appendix A: Comparative Static Analysis

In this appendix, we derive the most relevant comparative static results of Model I through Model III. In Model I, we consider the impact of changes in exogenous density, D, and exogenous residential location, RL, on trip chaining, TC, activity space, AS, and travel demand, TD. Starting from an equilibrium state, we consider the impact of an increase in density and residential location. To conduct comparative static analysis, we first introduce a set of basic assumptions related to residential location, trip chaining behavior, activity space, and urban form. Also, although trips are integers in reality, we treat them as a continuous non-negative variable for analytical purposes.

We begin with some definitions and assumptions followed by a detailed discussion of the behavioral equations of the model. We first show how we derive the residential location, RL, trip chaining, TC, transit demand, TD, and activity space, AS, equations by relaying on a generalized form of a model of consumption, travel and trip chaining first exposited by Anas (2007).

Definitions and Assumptions

We assume the urban area is divided into zones, which are linked via a transport network. To focus on consumer decisions, we consider a partial equilibrium model in which urban density and firm location are predetermined. In this context, we derive the demands for goods consumption and non-work travel. Consumers are price takers in all

Appendix A (Continued)

markets and take as given all transport costs and travel times. They supply labor hours

after allocating time to work and non-work travel, and leisure.

The following notation is used:

$i = 1, \dots, I$ the number of stores a consumer visits of a period of time

$x_0, \dots, x_I$ goods consumption, x_0 with being housing

$p_0, \dots, p_I$ goods prices, with p_0 being housing price

$c_0, \dots, c_I$ individual round trip costs

$t_0, \dots, t_I$ individual round trip times, with t_0 being commuting time

n number of chained trips

c chained-trip costs

t chained-trip times

W work time

T travel time

L leisure time

M total time

Assumptions

- $p_0 = p_0(t_0)$, i.e., housing price is a function of time distance between home and work. We also assume that housing price falls with distance from the CBD and that residential locations are more decentralized than job locations. Hence $p_0'(t_0) < 0$.

Appendix A (Continued)

- $T = \sum_{i=0}^{I} n_i t_i + nt$, i.e., total travel time includes commuting, individual shopping trips, and trip chains.

- t_0 represents (time) distance between home and work, not necessarily the CBD. Note that the choice of t_0 will determine $t_i + c_i$ $(i \neq 0)$, so $t_i = t_i(t_0)$ and $c_i = c_i(t_0)$ $(i \neq 0)$. Also $c_0'(t_0)$, while $c_i'(t_0) \lesseqgtr 0$, $t_i'(t_0) \lesseqgtr 0$, and $c_i'(t_0) \lesseqgtr 0$ for $i \neq 0$. Furthermore $t_0(t_0) \equiv t_0$

Following Anas (2007), we assume that chained trips involve trips to all places selling $x_0, \dots, x_I$ and that individual trips are in addition to trip chains. This could be restrictive, but Anas argues that not all trips may involve a chained trip, i.e., that there are corner solutions. In our empirical model we assume that chained trips occur as part of the commute. This is consistent with empirical findings (see reference) on trip-chaining formation.

Consider a consumer who visits each of $i = 1,...,I$ stores over a period of time. Although the number of trips may be an integer, we treat it as continuous, for over a period of time involving many trips, the per-unit-time number of trips can be continuous, e.g., five trips per week is 0.71 trips per day. The consumer buys a quantity z_i $(i = 1,...,I)$ from each store per trip. The utility function is

$$U = U(x_0, \dots, x_I; n_0, \dots, n_I; n; L)$$

The budget constraint is

$$wW = p_0(t_0)x_0 + \sum_{i=1}^{I} n_i p_i x_i + \sum_{i=0}^{I} n_i c_i + n \sum_{i=1}^{I} p_i x_i + nc \qquad (1)$$

Considering the identity

$$M = L + W + T \qquad (2)$$

126

Appendix A (Continued)

where

$$T = \sum_{i=0}^{I} n_i t_i + nt \tag{3}$$

with

$$wW = p_0(t_0)x_0 + \sum_{i=1}^{I} n_i p_i x_i + \sum_{i=0}^{I} n_i c_i + n \sum_{i=1}^{I} p_i x_i + nc \tag{4}$$

The Lagrangean objective function is then given by

$$\Lambda = U(x_0, \dots, x_I; n_0, \dots, n_I; n; L) + \lambda\{w[M - \sum_{i=0}^{I} n_i t_i(t_0) - nt(t_0) - L] - p_0(t_0) -$$

i=1Inipixi−i=0Inicit0−ni=1Ipixi−nct0 (5)

from which we obtain the following first-order conditions

$$\Lambda_{x_0} = u_{x_0} - \lambda p_0(t_0) = 0 \text{ (housing consumption)} \tag{6}$$

$$\Lambda_{x_i} = u_{x_i} - \lambda(n_i + n)p_i = 0, i \neq 0 \text{ (consumption of non-housing goods)} \tag{7}$$

$$\Lambda_{n_0} = u_{n_0} - \lambda[wt_0 + c_0(t_0)] = 0 \text{ (number of commuting trips)} \tag{8}$$

$$\Lambda_{n_i} = u_{n_i} - \lambda[wt_i(t_0) + p_i x_i + c_i(t_0)] = 0, \ i \neq 0 \text{ (number of non-commuting individ-}$$

ual trips) (9)

$$\Lambda_n = u_n - \lambda[wt(t_0) + \sum_{i=1}^{I} p_i x_i + c(t_0)] = 0 \text{ (number of non-commuting chained}$$

trips) (10)

$$\Lambda_{t_0} = u_{t_0} - \lambda(wn_0 + p_0'(t_0)x_0) = 0 \text{ (commuting time)} \tag{11}$$

$$\Lambda_{t_i} = u_{t_i} - \lambda wn_i = 0, i \neq 0 \text{ (non-commuting individual travel time)} \tag{12}$$

$$\Lambda_t = u_t - \lambda wn = 0 \text{ (non-commuting chained trip travel time)} \tag{13}$$

$$\Lambda_L = u_L - \lambda w = 0 \text{ (leisure choice)} \tag{14}$$

Appendix A (Continued)

$$\Lambda_\lambda = w\left(M - \sum_{i=0}^{I} n_i t_i - L\right) - p_0(t_0) - \sum_{i=1}^{I} n_i p_i x_i - \sum_{i=0}^{I} n_i c_i - n\sum_{i=1}^{I} p_i x_i - nc = 0$$

(budget constraint) (15)

From the first-order conditions, we derive solutions for the equations of interest.

The demand for commuting trips plus non-commuting trips is

$$n_i = n_i(w; p_1, \dots, p_I; c_0, \dots, c_I; c), i = 0, \dots I \tag{16}$$

The demand for non-commuting chained trips

$$n = n(w; p_1, \dots, p_I; c_0, \dots, c_I; c), i = 0, \dots I \tag{17}$$

The optimal work-residence travel time is

$$t_0 = t_0(w; p_1, \dots, p_I; c_0, \dots, c_I; c), i = 0, \dots I \tag{18}$$

From this, we obtain the demand for chained trips, the demand for individual trips and

the optimal commuting time, which are related to the demand equations of Chapter 3 as

$$TC = n(w; p_1, \dots, p_I; c_0, \dots, c_I; c) \text{ (trip chaining equation)} \tag{19}$$

$$TD = \sum_{i=0}^{I} n_i = TD(w; p_1, \dots, p_I; c_0, \dots, c_I; c) \text{ (transit demand equation)} \tag{20}$$

$$RL = t_0(w; p_1, \dots, p_I; c_0, \dots, c_I; c) \text{ (residential location equation)} \tag{21}$$

Next, we introduce some additional assumptions to carry out the comparative static analysis.

Assumption A.1. We assume that as the distance defining the *job-residence pair* increases, then the need to chain non-work trips increases

$$TC_{RL} \equiv \frac{\partial TC}{\partial RL} > 0 \tag{22}$$

The partial equilibrium model of trip chaining and consumption, Anas (2007),

shows that trip chaining saves time, which, in turn can be allocated to more consumption

128

Appendix A (Continued)

and discretionary travel. In this study, we assume that individuals can allocate the travel time savings of trip chaining to either more commute time, a move farther from work (more commute time), or more non-work travel time. Empirical evidence linking complex trip chaining to the work commute is found in Oster (1978), Kondo, and Kitamura (Kondo and Kitamura 1987), Nishii et al. (1988), and Strathman et al. (1994). These studies find that the propensity to link non-work travel to the work commute increases with distance from work. Oster (1978) shows that the probability of adding non-work trips to the commute increases with the distance to household members' employment destinations. Adopting Hägerstrand 's (1970) concept of space-time prisms, Kondo and Kitamura (1987) model the formation of trip chains and empirically show that under diminishing marginal benefits ($\partial^2 TC / \partial RL^2 < 0$), households living farther from work tend to chain non-work trips to the work commute.

Assumption A.2. If density, D, increases, then non-work activity locations, such as shopping or recreational locations, tend to be more clustered together, thus reducing the activity space

$$AS_D \equiv \frac{\partial AS}{\partial D} < 0 \tag{23}$$

Although this assumption seems intuitively acceptable, it is obtained theoretically from the generalization of the partial equilibrium model of trip chaining developed by Anas (2007) and reported in Appendix A. Empirical evidence on this assumptions is found in Noland and Thomas (2007) who, in a multivariate analysis of trip chaining behavior, show a positive relationship between lower densities and the complexity of trip

Appendix A (Continued)

chaining behavior. Noland and Thomas (2007) find that low density leads to both a greater reliance upon trip chaining and tours that involve more stops, thereby expanding the activity space..

Assumption A.3. As the activity space gets more dispersed (increases) then trip chaining increases

$$TC_{AS} \equiv \frac{\partial TC}{\partial AS} > 0 \tag{24}$$

As with assumption (a.2), this assumption seems intuitively acceptable, but it is justified from generalizing the partial equilibrium model of trip chaining developed by Anas (2007) and reported in Appendix A. Also, empirical work supports this assumption (Noland and Thomas 2007).

Assumption A.4 As trip chaining increases, the household activity space decreases

$$AS_{TC} \equiv \frac{\partial AS}{\partial TC} < 0 \tag{25}$$

The optimization of a trip chaining sequence results in a reduction of the activity space. The rate of reduction decreases as trip chaining increases. This is so because the activity space contracts up to a point where the location of activities is close enough to make the individual indifferent between chaining the additional trip and making a separate trip to a given store

Appendix A (Continued)

Model I Comparative Static Results

Now, consider Model I. Equations (3.1), (3.2), and (3.3) can be written as impli-

cit functions in the form $F^j(TC, AS, TD, RL, WD, D, X_{TC}, X_{AS}, X_{TT})$, where $j = 1, .., 3$.

With continuous partial derivatives and with the relevant assumptions (A.3) and (A.4),

the *Jacobian determinant* is

$$|J| = \begin{vmatrix} \frac{\partial F^1}{\partial TC} & \frac{\partial F^1}{\partial AS} & \frac{\partial F^1}{\partial TD} \\ \frac{\partial F^2}{\partial TC} & \frac{\partial F^2}{\partial AS} & \frac{\partial F^2}{\partial TD} \\ \frac{\partial F^3}{\partial TC} & \frac{\partial F^3}{\partial AS} & \frac{\partial F^3}{\partial TD} \end{vmatrix} = \begin{vmatrix} 1 & -TC_{AS} & 0 \\ -AS_{TC} & 1 & 0 \\ -TD_{TC} & -TD_{AS} & 1 \end{vmatrix} = 1 - \overset{(-)}{\widetilde{AS_{TC}}} \overset{(+)}{\widetilde{TC_{AS}}} > 0 \qquad (26)$$

Therefore, *TC*, *AS*, and *TD* [no comma] can be considered implicit functions of

$(RL, D, WD, X_{TC}, X_{AS}, X_{TD})$ at and around any point that satisfies Equations (3.1), (3.2),

and (3.3). Hence the implicit function theorem justifies writing

$$TC = f^1(RL, D, WD, X_{TC}, X_{AS}, X_{TD}) \qquad (27)$$

$$AS = f^2(RL, D, WD, X_{TC}, X_{AS}, X_{TD}) \qquad (28)$$

$$TD = f^3(RL, D, WD, X_{TC}, X_{AS}, X_{TD}) \qquad (29)$$

indicating that the equilibrium values of the endogenous variables are implicit

functions of the exogenous variables and parameters. The partial derivatives of the im-

plicit functions provide the comparative-static results.

Next, we obtain the comparative static results of changes in density, residential

location and transit station proximity.

Effects of an Increase in Density, D

The general form for the comparative static analysis of Model I for changes in *D*

is given by

Appendix A (Continued)

$$
\begin{vmatrix} 1 & -TC_{AS} & 0 \\ -AS_{TC} & 1 & 0 \\ -TD_{TC} & -TD_{AS} & 1 \end{vmatrix} \begin{bmatrix} \partial TC/\partial D \\ \partial AS/\partial D \\ \partial TD/\partial D \end{bmatrix} = \begin{bmatrix} 0 \\ AS_D \\ 0 \end{bmatrix} \tag{30}
$$

Density Effect on Trip Chaining, *TC*

The effect of density on trip chaining is

$$
dTC/dD = \frac{\begin{vmatrix} 0 & -TC_{AS} & 0 \\ AS_D & 1 & 0 \\ 0 & -TD_{AS} & 1 \end{vmatrix}}{|J|} = \frac{\overset{(-)}{\overbrace{\overset{+}{TC_{AS}}\overset{-}{AS_D}}}}{\underset{(+)}{\underbrace{1 - \underset{(-)}{AS_{TC}}\underset{(+)}{TC_{AS}}}}} < 0 \tag{31}
$$

An increase in density causes a clustering of activities which contracts the activity space, which, in turn, reduces the need to engage in trip chaining. This outcome has been confirmed in the literature on trip chaining behavior, which shows that lower density environments increase the need to engage in trip chaining (Wallace, Barnes, and Rutherford 2000; Noland and Thomas 2007).

Density Effect on Activity Space, *AS*

The effect of an increase in density on the activity space is

$$
dAS/dD = \frac{\begin{vmatrix} 1 & 0 & 0 \\ -AS_{TC} & AS_D & 0 \\ -TD_{TC} & 0 & 1 \end{vmatrix}}{|J|} = \frac{\overset{(-)}{\overbrace{AS_D}}}{(+)} < 0 \tag{32}
$$

Appendix A (Continued)

Note, by assumption (A.2), we have $\partial AS/\partial D < 0$. An increase in density contracts the activity space both directly and indirectly through feedback effect coming by way of $(AS_{TC}TC_{AS})$ in the denominator of (32).

Density Effect on Transit Demand, *TD*

The effect of an increase in density on transit demand is

$$\frac{dTD}{dD} = \frac{\begin{vmatrix} 1 & -TC_{AS} & 0 \\ -AS_{TC} & 1 & AS_D \\ -TD_{TC} & -TD_{AS} & 0 \end{vmatrix}}{|J|} = \frac{\overset{(+)}{\overbrace{\underset{(-)}{\widehat{\alpha}}}} \overset{(+)}{\overbrace{\underset{(-)}{\widehat{\beta}}}}}{|J|} > 0 \tag{33}$$

where the product $\alpha = TD_{AS}AS_D$ give the increase in transit demand caused by a contraction in the activity space as a result of increased density, and $\beta = TD_{TC}TC_{AS}AS_D$ gives the increase in transit demand caused by decreasing trip chaining as a result of increased density.

First, increased density reduces the extent of the activity space, which increases the demand for transit trips. Second, higher densities reduce the activity space, which reduces the need to chain trips (as time savings opportunities decrease) and increases the demand for transit trips. Thus transit demand increases since it is sensitive to changes affecting the spatial allocation of non-work activities and to changes affecting trip chaining behavior.

This result relies on the assumption that demand for transit trips decreases as trip chaining increases

$$\frac{\partial TD}{\partial TC} < 0 \tag{34}$$

Appendix A (Continued)

An increase in the number of chained trips decreases the demand for transit as the need to rely on more flexible modes of transport increases. This is also reflected by the following assumption on the relationship between transit demand and the size of the activity space

$$\frac{\partial TD}{\partial AS} < 0 \qquad (35)$$

That is, the increased spatial dispersion of non-work activities cannot be accommodated by additional transit trips. Given the characteristics of transit service supply (being fixed at least in the short to medium run), increased spatial dispersion is accommodated by substituting transit travel with other, more flexible, modes, such as auto travel. Auto is a more flexible mode in terms of allowing serving a more dispersed activity space.

Effect of a Change in Residential Location, *RL*

Next, we look at the effect of a change in exogenous residential location, *RL*. Applying Cramer's rule to Model I, we have

$$\begin{vmatrix} 1 & -TC_{AS} & 0 \\ -AS_{TC} & 1 & 0 \\ -TD_{TC} & -TD_{AS} & 1 \end{vmatrix} \begin{bmatrix} \partial TC/\partial RL \\ \partial AS/\partial RL \\ \partial TD/\partial RL \end{bmatrix} = \begin{bmatrix} TC_{RL} \\ 0 \\ TD_{RL} \end{bmatrix} \qquad (36)$$

Residential Location Effect on Trip Chaining, *TC*

From assumption (A.1), as distance between home and work increases, trip chaining increases. The new equilibrium results in a higher number of trips per chain

$$dTC/dRL = \frac{\begin{vmatrix} TC_{RL} & -TC_{AS} & 0 \\ 0 & 1 & 0 \\ TD_{RL} & -TD_{AS} & 1 \end{vmatrix}}{|J|} = \frac{\overset{(+)}{\overbrace{TC_{RL}}}}{|J|} > 0 \tag{37}$$

When testing this hypothesis with using cross-sectional data, individuals with a more living farther from work are expected to engage in a higher number of trips per chain (or in more complex tours characterized by more stops). In a longitudinal context, a move farther out entails more time spent commuting, which increases the propensity to engage in trip chaining to save overall time.

Residential Location Effect on Activity Space, *AS*

The effect of an increase in *RL* on the activity space is given by

$$dAS/dRL = \frac{\begin{vmatrix} 1 & TC_{RL} & 0 \\ -AS_{TC} & 0 & 0 \\ -TD_{TC} & TD_{RL} & 1 \end{vmatrix}}{|J|} = \frac{\overset{(-)}{\overbrace{\overset{(-)}{\overbrace{AS_{TC}}}\,\overset{(+)}{\overbrace{TC_{RL}}}}}}{\underset{(+)}{}} < 0 \tag{38}$$

A move farther away from work increases trip chaining, which in turn decreases the activity space.

Residential Location Effect on Transit Demand, *TD*

The change in transit demand caused by a change in residential location is given by:

$$dTD/dRL = \frac{\begin{vmatrix} 1 & -TC_{AS} & TC_{RL} \\ -AS_{TC} & 0 & 0 \\ -TD_{TC} & -TD_{AS} & TD_{RL} \end{vmatrix}}{|J|} = \frac{\overset{(\pm)}{\overbrace{TD_{RL}}}+\overset{(-)}{\overbrace{AS_{TC}}}\left(\overset{(+)}{\overbrace{TC_{RL}}}\overset{(-)}{\overbrace{TD_{AS}}}-\overset{(-)}{\overbrace{TC_{AS}}}\overset{(\pm)}{\overbrace{TD_{RL}}}\right)+\overset{(+)}{\overbrace{TC_{RL}}}\overset{(-)}{\overbrace{TD_{TC}}}}{\underset{(+)}{}} \gtreqless 0 \tag{39}$$

The sign is ambiguous as the overall effect on transit demand hinges on the sign of TD_{RL}. We posit that to the extent that an urban area is well served by transit, then the

relationship between transit demand and residential location is positive. A positive rela-

tionship is observed in older, more monocentric-type cities, with existing transit services

supporting major work commute travel routes. On the other hand, if supply constraints

exist, transit demand declines as the job-residence distance increases. Therefore, the

overall effect on transit demand due to a change in location depends on both the sign and

magnitude of TD_{RL}.

Effect of a Change in Walking Distance, *WD*

Effect of Walking Distance on Transit Demand. We now look at the effect on

transit demand from an increase in distance from the nearest transit station. The empiri-

cal literature provides unequivocal evidence of a negative relationship between distance

to transit stops and the demand for transit services (Cervero 2007; Cervero and

Kockelman 1997). The debate is mostly centered on the magnitude of this relationship,

as high-lighted by the growing body of literature on residential self-selection. All else

equal, being located farther away from a transit station results in a change in transit de-

mand as

$$\frac{dTD}{dWD} = \frac{\overset{(-)}{\overbrace{TD_{WD}}} + \overset{(-)}{\overbrace{TD_{TC}}}\overset{(-)}{\overbrace{TC_{WD}}} + \overset{(-)}{\overbrace{AS_{TC}}}\left(\overset{(-)}{\overbrace{TC_{WD}}}\overset{(-)}{\overbrace{TD_{AS}}} - \overset{(+)}{\overbrace{TC_{AS}}}\overset{(-)}{\overbrace{TD_{WD}}}\right)}{1 - \underbrace{\underset{(-)}{AS_{TC}}\underset{(+)}{TC_{AS}}}_{(+)}} \lesseqgtr 0 \quad (40)$$

The overall effect of an increase in walking distance is ambiguous. An increase

in distance to the nearest station directly reduces transit demand ($TD_{WD} < 0$). At the

same time, reduced accessibility impacts and the ability to engage in trip chaining using

transit, producing an ambiguous effect on transit demand. The sign hinges on the rela-

Appendix A (Continued)

tionship between trip chaining and distance to the nearest transit station, $(TC_{WD} \gtrless 0)$,

which is undetermined.

Appendix B: Equation Identification

Identification

In the context of simultaneous equation modeling, the validity of results hinges on the determination of the exclusion restrictions. That is, the researcher must a priori determine what explanatory variables are to be included and excluded from each equation. The determination of the exclusion restrictions defines a model that is correctly specified in the sense that the matrix of the reduced form parameters to be estimated is unique in its representation of the more primitive structural matrix. Exclusion restrictions need to be drawn outside of the variables a researcher has available from a given dataset (i.e, they should be based on sound behavioral theory).

A necessary and sufficient condition for identification of a structural equation is provided by the rank condition. The rank condition assures that the exclusion restrictions are sufficient and are unique. The following steps are required to obtain the rank condition for a given structural equation:

1) Let Δ be a matrix of all the structural parameters

$$\Delta = \begin{bmatrix} B \\ \Gamma \end{bmatrix} \qquad (1)$$

and let R_i be the matrix of exclusion restrictions defining structural equation i

Appendix B (Continued)

$$R = [1 \quad \dots \quad 0] \tag{2}$$

2) Premultiply (c.1) by (c.2) to obtain the list of variables excluded from equation

i

$$R\Delta = [1 \quad \dots \quad 0] \begin{bmatrix} \beta_{1i} \\ \vdots \\ \gamma_{ki} \end{bmatrix} = \beta_{1i} + \cdots \tag{3}$$

3) Compute the rank of $R\Delta$

4) Equation (*i*) is identified (overidentified) if the rank is equal (greater) to G-1; where G is equal to the number of endogenous variables

Next, each of the four models presented next is subject to the rank condition for identification prior to estimation and results are reported below. Note that the size of R depends on the number of exogenous and endogenous structural parameters excluded by each equation. The following notation is used to denote exogenous and endogenous appearing or being excluded by each equation

G= total number of endogenous variables

K= total number of exogenous variables

g_i= number of endogenous variables included in equation *i*

g_i^*= number of endogenous variables excluded from equation *i*

k_i= number of exogenous variables included in equation *i*

k_i^*= number of exogenous variables excluded from equation *i*

Appendix B (Continued)

Model I

Following the equation specifications of Chapter 3, the following rank conditions for identification are obtained. Given the dimensions of the matrices involved, we used the mathematica software package to compute the rank conditions.

Trip Chaining Equation, *TC*

Inclusions/Exclusions	*Number*
G	3
K	13
g_i	1
g_i^*	1
k_i	7
k_i^*	6

The rank condition is

$$\underset{7x16}{R}\ \underset{16x3}{\Delta} = 2;\ (G-1) = 2 \text{ (just identified)} \tag{4}$$

Activity Space Equation, *AS*

Inclusions/Exclusions	*Number*
G	3
K	13
g_i	1
g_i^*	1
k_i	4
k_i^*	9

The rank condition is

$$\underset{10x16}{R}\ \underset{16x3}{\Delta} = 3;\ (G-1) = 2\ \text{(overidentified)} \tag{5}$$

Appendix B (Continued)

Transit Demand Equation, *TD*

	Inclusions/Exclusions	Number
G		3
K		13
g_i		2
g_i^*		0
k_i		4
k_i^*		7

The rank condition is

$$\underset{7\text{x}16}{R} \ \underset{16\text{x}3}{\Delta} = 2; \ (G-1) = 2 \ \text{(just identified)} \tag{6}$$

Model II

Following the specification of Chapter 4, the following rank conditions for identification are obtained.

Trip Chaining Equation, *TC*

	Inclusions/Exclusions	Number
G		4
K		18
g_i		2
g_i^*		1
k_i		6
k_i^*		12

The rank condition is given by

$$\underset{13\text{x}22}{R} \ \underset{22\text{x}4}{\Delta} = 3; \ (G-1) = 3 \ \text{(just identified)} \tag{7}$$

141

Appendix B (Continued)

Activity Space Equation, *AS*

Inclusions/Exclusions	*Number*
G	4
K	18
g_i	1
g_i^*	2
k_i	4
k_i^*	14

The rank condition is given by

$$\underset{16\text{x}22}{\underset{\smile}{R}} \; \underset{22\text{x}4}{\underset{\smile}{\Delta}} = 3; \; (G-1) = 3 \; (\text{ just identified}) \tag{8}$$

Transit Demand Equation, *TD*

Inclusions/Exclusions	*Number*
G	4
K	18
g_i	3
g_i^*	0
k_i	5
k_i^*	13

The rank condition is given by

$$\underset{13\text{x}22}{\underset{\smile}{R}} \; \underset{22\text{x}4}{\underset{\smile}{\Delta}} = 3; \; (G-1) = 3 \; (\text{ just identified}) \tag{9}$$

Appendix B (Continued)

Residential Location Equation, *RL*

Inclusions/Exclusions	Number
G	4
K	18
g_i	2
g_i^*	1
k_i	6
k_i^*	12

The rank condition is given by

$$\underset{13\text{x}22}{R}\ \underset{22\text{x}4}{\Delta}\ = 3;\ (G-1) = 3 \text{ (just identified)} \tag{10}$$

Model III

Following the specification of Chapter 4, the following rank conditions for identification are obtained.

Trip Chaining Equation, *TC*

Inclusions/Exclusions	Number
G	5
K	18
g_i	2
g_i^*	2
k_i	6
k_i^*	12

The rank condition is given by

$$\underset{14\text{x}23}{R}\ \underset{23\text{x}5}{\Delta}\ = 4;\ (G-1) = 4 \text{ (just identified)} \tag{11}$$

143

Appendix B (Continued)

Activity Space Equation, *AS*

Inclusions/Exclusions	Number
G	5
K	18
g_i	2
g_i^*	2
k_i	3
k_i^*	15

The rank condition is given by

$$\underset{17\text{x}23}{R}\ \underset{23\text{x}5}{\Delta} = 4;\ (G-1) = 4\ \text{(just identified)} \tag{12}$$

Transit Demand Equation, *TD*

Inclusions/Exclusions	Number
G	5
K	18
g_i	3
g_i^*	1
k_i	5
k_i^*	13

The rank condition is given by

$$\underset{14\text{x}23}{R}\ \underset{23\text{x}5}{\Delta} = 4;\ (G-1) = 4\ \text{(just identified)} \tag{13}$$

144

Appendix B (Continued)

Residential Location Equation, *RL*

	Inclusions/Exclusions	Number
G		5
K		18
g_i		2
g_i^*		2
k_i		6
k_i^*		12

The rank condition is given by

$$\underset{14\text{x}23}{\underbrace{R}}\ \underset{23\text{x}5}{\underbrace{\Delta}} = 4;\ (G-1) = 4\ \text{(just identified)} \tag{14}$$

Density Equation, *D*

	Inclusions/Exclusions	Number
G		5
K		18
g_i		2
g_i^*		2
k_i		2
k_i^*		16

The rank condition is given by

$$\underset{18\text{x}23}{\underbrace{R}}\ \underset{23\text{x}5}{\underbrace{\Delta}} = 4;\ (G-1) = 4\ (\text{ just identified}) \tag{15}$$

.

About the Author

Sisinnio Concas is a transportation economist with broad experience in urban and regional economic impact analysis. He currently serves as senior research associate for the Center for Urban Transportation Research (CUTR) at the University of South Florida, Tampa, USA. He has extensive expertise in evaluating transportation infrastructure investments for state and local transportation agencies. His research interests include the study of the linkages between transportation infrastructure investment and economic development, the application of nonparametric estimation and inference techniques to transportation research, and the theoretical underpinnings of the linkages between household activity-travel patterns and land-use in urban areas. A native of Sardinia, Italy, he holds a Doctoral degree in Political Sciences, with a field specialization in Political Economy from the University of Sassari, Italy, and a Master of Arts in Economics from the University of South Florida.