1. Report No. SWUTC/06/167860-1	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle A Comprehensive Analysis of Built Environment Characteristics on Household Residential Choice and Auto Ownership Levels		5. Report Date September 2006	
		6. Performing Organization Code	
7. Author(s) Chandra R. Bhat and Jessica Y. Guo		8. Performing Organization Report No. Report 167860-1	
9. Performing Organization Name and Address Center for Transportation Research The University of Texas at Austin 3208 Red River, Suite 200 Austin, Texas 78705-2650		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No. 167860	
12. Sponsoring Agency Name and Address Southwest Region University Transportation Center Texas Transportation Institute Texas A&M University System College Station, Texas 77843-3135		13. Type of Report and Period Covered	
		14. Sponsoring Agency Code	
15. Supplementary Notes Supported by general revenues from the State of Texas.			

16. Abstract

In this report, we identify the research designs and methodologies that may be used to test the presence of "true" causality versus residential sorting-based "spurious" associations in the land-use transportation connection. The report then develops a methodological formulation to control for residential sorting effects in the analysis of the effect of built environment attributes on travel behavior-related choices. The formulation is applied to comprehensively examine the impact of the built environment, transportation network attributes, and demographic characteristics on residential choice and car ownership decisions. The model formulation takes the form of a joint mixed multinomial logit-ordered response structure that (a) accommodates differential sensitivity to the built environment and transportation network variables due to both demographic and unobserved household attributes and (b) controls for the self-selection of individuals into neighborhoods based on car ownership preferences stemming from both demographic characteristics and unobserved household factors.

| 17. Key Words

Built environment; residential sorting; self-selection; auto ownership; integrated land use-transportation modeling. | | 18. Distribution Statement

No restrictions. This document is available to the public through NTIS:
    National Technical Information Service
    5285 Port Royal Road
    Springfield, Virginia 22161 | |
| 19. Security Classif.(of this report)
Unclassified | 20. Security Classif.(of this page)
Unclassified | 21. No. of Pages
58 | 22. Price |

Form DOT F 1700.7 (8-72)   **Reproduction of completed page authorized**

A Comprehensive Analysis of Built Environment Characteristics on Household Residential Choice and Auto Ownership Levels

by

Dr. Chandra R. Bhat
The University of Texas at Austin
Dept of Civil, Architectural & Environmental Engineering

and

Dr. Jessica Y. Guo
University of Wisconsin – Madison
Department of Civil and Environmental Engineering

Research Report SWUTC/06/167860-1

September 2006

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program in the interest of information exchange. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

ABSTRACT

There has been an increasing interest in the land use-transportation connection in the past decade, motivated by the possibility that design policies associated with the built environment can be used to control, manage, and shape individual traveler behavior and aggregate travel demand. In this line of research and application pursuit, it is critical to understand whether the empirically observed association between the built environment and travel behavior-related variables is a true reflection of underlying causality or simply a spurious correlation attributable to the intervening relationship between the built environment and the characteristics of people who choose to live in particular built environments.

In this report, we identify the research designs and methodologies that may be used to test the presence of "true" causality versus residential sorting-based "spurious" associations in the land-use transportation connection. The report then develops a methodological formulation to control for residential sorting effects in the analysis of the effect of built environment attributes on travel behavior-related choices. The formulation is applied to comprehensively examine the impact of the built environment, transportation network attributes, and demographic characteristics on residential choice and car ownership decisions. The model formulation takes the form of a joint mixed multinomial logit-ordered response structure that (a) accommodates differential sensitivity to the built environment and transportation network variables due to both demographic and unobserved household attributes and (b) controls for the self-selection of individuals into neighborhoods based on car ownership preferences stemming from both demographic characteristics and unobserved household factors.

The analysis in the report represents, to our knowledge, the first instance of the formulation and application of a unified mixed multinomial logit-ordered response structure in the econometric literature. The empirical analysis in the report is based on the residential choice and car ownership decisions of San Francisco Bay area residents.

ACKNOWLEDGEMENTS

EXECUTIVE SUMMARY

This research develops a methodological framework to control for residential sorting effects (also referred to as self-selection effects) in the analysis of the effect of built environment attributes on travel behavior-related choices. The formulation is applied to examine the impact of the built environment, transportation network attributes, and demographic characteristics on residential choice and car ownership decisions. The model formulation takes the form of a joint mixed multinomial logit-ordered response structure that (a) accommodates differential sensitivity to the built environment and transportation network variables due to both demographic and unobserved household attributes and (b) controls for the self-selection of individuals into neighborhoods based on car ownership preferences. To our knowledge, the analysis in this research represents the first instance of the formulation and application of such a unified mixed multidimensional-ordered response structure in the econometric literature. The framework can be used to control for residential self-selection for any kind of travel behavior variable and directly provides the correct standard errors regarding the effect of the built environment attributes. It is geared toward cross-sectional analysis, recognizing that almost all existing data sources available for analysis of BE effects are cross-sectional in nature. Unlike earlier studies, the methodology also explicitly considers and models the residential location choice decision jointly with the travel behavior choice of interest.

The empirical analysis in the research is based on the residential choice and car ownership decisions of San Francisco Bay area residents. The data is drawn from several sources, including the (a) 2000 San Francisco Bay area travel survey, (b) land-use/demographic, network level-of-service, and GIS-based bicycle facility data files for the Bay area obtained from the Metropolitan Transportation Commission (MTC), and (c) the US census data, the US 2000 Tiger files, and the Public Use Microdata Sample (PUMS) data. A whole range of zonal size and density measures, land-use structure variables, regional accessibility indices, commute-related characteristics, local transportation network measures, zonal demographic and housing variables, zonal ethnic composition characteristics, household demographics, and interactions of these variables are considered in the analysis.

There are several important findings from our study. <u>First</u>, BE attributes do affect residential choice decisions as well as car ownership decisions. Thus, policy decisions regarding

changes in BE characteristics have to be evaluated in the joint context of both decisions, so that spatial relocation patterns as well as car ownership changes can be analyzed. Such a complete picture enables a comprehensive assessment of potential travel-related changes due to BE policies. Second, our findings support the notion that the commonly used population and/or employment density measures are actually proxy variables for such BE measures as street block density and transit accessibility. Third, in the context of car ownership decisions, both household demographics and BE characteristics are influential. However, household demographics have a more dominant effect. Fourth, there is variation in sensitivity to BE attributes due to both demographic and unobserved factors, in both residential choice as well as car ownership decisions. But, while the study examined a suite of demographic interactions and allowed random variations in sensitivity to several BE characteristics, most of these did not turn out to be statistically significant. Among demographics, income is a key variable in affecting the sensitivity to BE attributes and related variables. Unobserved household-specific factors also play an important role in the sensitivity to commute time and street block density (in the residential choice model) and employment density and street block density (in the car ownership model). Ignoring such systematic and random variations in sensitivity to BE attributes will, in general, lead to inconsistent results regarding the effect of BE attributes on travel behavior decisions, which can, in turn, lead to inappropriate policy decisions. Fifth, household income is the dominant factor in residential sorting. Specifically, low income households consciously choose to (or are constrained to) locate in neighborhoods with low commute costs, long commute times, and high employment density compared to their high income counterparts. Such low income households also intrinsically choose to own fewer cars. Thus, ignoring income effects in car ownership (and by extension, other travel decisions) can lead to an inflated effect of the built environment and related variables on travel behavior decisions. Other demographic factors that impact residential sorting based on car ownership preferences correspond to the presence of senior adults in the household and whether or not a person lives alone. Finally, and rather surprisingly, our results did not support the notion of residential sorting in car ownership propensity based on unobserved household factors. This result implies that independent models of residential choice and car ownership choice (after accommodating the residential sorting effects of demographics) are adequate to examine BE effects on car ownership choice, in the current empirical context (see also the important caveats related to this issue in Section 4.3).

But, in general, it is important to consider the methodology developed in this research to control for the potential presence of self selection due to both observed and unobserved household factors.

To summarize, the model in the research can be used to assess the impacts of changing demographics, built environment characteristics, and transportation network attributes for land-use planning and transportation public policy analysis. The study, to our knowledge, represents the first formulation and application of a comprehensive modeling framework to consider residential self-selection effects, as well as observed and unobserved variations in sensitivity to the built environment, in a joint model of residential location choice and car ownership decisions.

x

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER 1. INTRODUCTION

Transportation engineers and planners have routinely assumed for several decades now that there is an association between land-use development patterns and the travel behavior of individuals. This is reflected in the different trip generation rates and (sometimes) mode shares attributed to different land-use development patterns. However, there is no rigorous attempt to explain the causal thread or mechanism that generates the association between land use and travel demand in such transportation planning practice. One reason for this is that the primary goal of traditional transportation planning has been to predict, in a reactive manner, the travel demand corresponding to a particular future land-use scenario, so that adequate transportation supply can be provided to meet the projected future travel demand. In such a reactive planning process, the difference between an association and the causal thread in land use-transportation interaction may be relatively mute.

Increasingly, however, a number of different forces, including high capital costs of new infrastructure, dwindling land space to build additional transportation infrastructure, air quality deterioration, and public opposition to the potential adverse side-effects of new infrastructure construction, have combined to extend the emphasis of travel demand analysis from the reactive, supply-enhancing, prediction-oriented focus to include a proactive, demand-reducing, policy-oriented focus. As part of this expanded focus of transportation planning, there has been interest in the land use-transportation connection in the past decade, motivated by the possibility that land-use and urban form design policies can be used to control, manage, and shape individual traveler behavior and aggregate travel demand. In this line of research and application pursuit, however, the difference between an association and a casual thread in land use-transportation interactions is no more a mute issue; rather it takes the center stage. Only by clearly establishing whether a causal thread actually exists to explain associations between the built environment and travel behavior, or whether these associations are generated through intervening variables, can researchers make credible, persuasive, policy recommendations.

To be sure, there has been an expanding and lively body of literature debating the causal versus the associative nature of the relationship between the built environment and travel behavior (we will use the term built environment or BE in this report to refer to land-use, urban

form, and street network attributes). Another dimension of the debate is whether any causal effect of the built environment on travel behavior is of adequate magnitude to actually cause a discernible shift in travel patterns. These issues are at the heart of the potential effectiveness of design policies manifested in "new urbanism" and "smart growth" concepts (see Pickrell, 1999; Ewing and Cervero, 2001; and Ewing, 2005). On the one side of the debate, proponents of the new urbanism and smart growth concepts claim that the association between the built environment and travel behavior represents a causal effect, and is of a sufficient enough magnitude to lead to tangible reductions in motorized vehicle use. In addition, according to these proponents, car dependence-reducing BE strategies will also lead to friendlier, and socially vibrant, neighborhoods. Several state, regional, and local governments have embraced the new urbanism and smart growth concepts, and have responded with land use planning proposals targeted toward reducing travel demand and improving air quality (see Transportation Research Board Conference Proceedings on Smart Growth and Transportation, 2005, for a review of agencies that have adopted such land use policy mechanisms). On the other side of the debate, opponents of the new urbanism and smart growth movement contend that any association between the BE and travel behavior is due to the intervening relationship between the BE and the demographic/other characteristics of people choosing to live in particular built environments. Further, opponents indicate that the increasingly isolated and auto-dependent orientation of the population is simply a manifestation of demographic shifts and lifestyle preferences, rather than any consequence of BE designs that do not subscribe to smart growth and new urbanism concepts (see Audirac and Shermyen, 1994; Guiliano, 1995; and Gordon and Richardson, 1997).

Between the polarized groups of ardent proponents and opponents of the new urbanism/smart growth concepts is a body of scholarly and applied works that is at best mixed and inconclusive. A review by Ewing and Cervero (2001) describes several studies that found reasonably significant elasticity effects of the BE attributes on travel demand variables. Some more recent studies have also found significant effects of the BE on one or more dimensions of activity/travel behavior (see Rajamani *et al.*, 2003; Krizek, 2003; Shay and Khattak, 2005; Bhat *et al.*, 2005; Bhat and Singh, 2000; and Rodriguez *et al.*, 2005). However, several studies reviewed by Crane (2000) and some other works (see, for example, Boarnet and Sarmiento, 1998; Boarnet and Crane, 2001; Bhat and Lockwood, 2004; Bhat *et al.*, 2005; and Bhat and Zhao, 2002) have found that BE measures have little to no impact on such dimensions of travel

behavior as activity/trip frequency and non-motorized mode use. Further, because of the widely varying estimation techniques, units of analysis, empirical contexts, travel behavior dimensions, and BE characteristics and their scales used across the studies, it is difficult to compare and contrast results. The net result is that there is reasonable agreement in the academic field that, despite the explosion of empirical studies in the past 15 years, it is still premature to draw any conclusive evidence regarding the impacts of the BE on activity-travel behavior. Further, two major inter-related problems need to be carefully addressed and recognized as we move forward in improving our understanding of the relationship between the BE and travel behavior: (1) The relationship between the BE and travel behavior can be very complex, and (2) The "true" causal impact of the BE on travel behavior can be assessed only if the spurious association due to residential sorting based on demographics and other characteristics is controlled for. Each of these two issues is discussed in turn in the next two sections (see also Boarnet and Crane, 2001; Crane, 2000; Krizek, 2003; and Handy, 1996).

1.1 Complex Nature of the Built Environment-Travel Behavior Relationship

There are at least three elements characterizing the complex relationship between the BE and travel, as discussed below.

1.1.1 Multidimensional Nature

The first element of the complex relationship between the BE and travel is that both of these are multidimensional in nature. That is, there are many aspects to the BE, including accessibility to transit stops, presence and connectivity of walk and bike paths, land-use mix, street network density (such as average length of links and number of intersections per unit area), block sizes, and proportion of street frontage with buildings. Similarly, there are many dimensions of travel, including car ownership, number of trips, time-of-day, route choice, travel mode choice, purpose of trips, and chaining of trips. A fundamental question then is what dimension of the BE impacts what dimension of travel, a seemingly innocuous, but very complex, question to address. Many earlier research works have focused on the impact of selected BE characteristics on selected travel dimensions (for example, see Bhat and Singh, 2000; Dunphy and Fisher, 1996; Pozsgay and Bhat, 2002; Cervero, 2002; Greenwald and Boarnet, 2001; Kitamura *et al.*, 1997; and Handy and Clifton, 2001). Such analyses provide only

3

a limited picture of the many interactions leading up to travel impacts. In particular, the use of a narrow set of BE measures potentially renders the measures as proxies for a suite of other BE measures, making it difficult to identify which element of the multidimensional package of BE measures is actually responsible for the travel impact. A similar problem arises when studies compare activity/travel behaviors of individuals across judgmentally pre-defined neighborhoods (such as conventional neighborhood and neo-urbanist neighborhoods; see, for example, Shay and Khattak, 2005; Saelens *et al.*, 2003; Handy *et al.*, 2005; Rodriguez *et al.*, 2005; and Schwanen and Mokhtarian, 2005). To the extent that neighborhoods are different across many different BE measures, it is not possible to isolate the individual effects or interaction effects of specific sets of BE variables. Similar to the use of a narrow set of BE attributes, the focus on the impacts of the BE on narrow dimensions of travel does not provide the overall effect on travel. For instance, a denser environment may be associated with less of pick up/drop off activity episodes, but more of recreational episodes (see Bhat and Srinivasan, 2005). The net impact on overall travel will depend on the "aggregation" across the effects on individual travel dimensions. Finally, most empirical analyses consider a trip-based approach to analysis, ignoring the chaining of activities and the resulting intricate interplay of the effect of BE attributes on the many dimensions characterizing activity participation and travel.

1.1.2 Moderating Influence of Decision-Maker Characteristics

The second element of the complex relationship between BE measures and travel is the moderating influence of the characteristics of decision makers on travel behavior (individuals and households). These characteristics may include sociodemographic factors (such as gender, income, and household structure), travel-related and environmental attitudes (such as preference for non-motorized/motorized modes of transportation and concerns about mobile source emissions), and perceptions regarding the BE attributes (that is, cognitive filtering of the objective built environment attributes). The decision maker characteristics may have two kinds of moderating influences: (1) a direct influence on travel behavior (for example, higher income households are more likely to own cars; see Bhat and Pulugurta, 1998, and Shay and Khattak, 2005), and (2) an indirect influence on travel behavior by modifying the sensitivity to BE characteristics (for example, it may be that high income households, wherever they live, own several cars and use them more than low income households; this creates a situation where high

income households are less sensitive to BE attributes in their car ownership and use patterns than low income households). Almost all individual and household-level analyses of the effect of BE characteristics on travel behavior recognize and control for the direct influence of decision-maker attributes by incorporating sociodemographic characteristics as determinants of travel behavior. A handful of studies also control for the direct impact of attitudes and perceptions of decision-makers on travel behavior (see Schwanen and Mokhtarian, 2005; Kitamura *et al.*, 1997; Handy *et al.*, 2005; and Lund, 2003). However, while there has been recognition that the sensitivity to BE attributes can vary across decision-makers (see Badoe and Miller, 2000), most previous empirical studies have not examined the indirect effect of demographics on the sensitivity to BE attributes. And, to our knowledge, no earlier study has recognized the potential effect of unobserved decision-maker characteristics on the response to BE attributes. On the other hand, it is possible that the varying levels and sometimes non-intuitive effects of BE attributes on travel behavior found in earlier empirical studies (for example, in Bhat and Gossen, 2004 and TRB, 2003) is, at least in part, a manifestation of varying BE attribute effects across decision-makers in the population.

1.1.3 Spatial Scale of Analysis

The third element characterizing the complex relationship between the built environment and travel is the "neighborhood" shape and scale used to measure the BE measures. Most studies use predefined spatial units based on census tracts, zip codes, or transport analysis zones as operational surrogates for neighborhoods because urban form data is more readily available and easily matched to travel data at these scales. However, it is anything but clear as to how individuals perceive the "neighborhood" space and scale, and how they filter spatial information when making spatial choice decisions (see Golledge and Gärling, 2003; Krizek, 2003; and Guo and Bhat, 2004, 2006, for detailed discussions of this issue). Further, it is possible that different BE attributes have different spatial extents of influence on travel choices, as empirically illustrated by Guo and Bhat (2006) and Boarnet and Sarmiento (1998).

1.2 Residential Sorting Based on Travel Behavior Preferences

The second major issue in the BE-travel behavior relationship is residential sorting based on travel behavior preferences. A fundamental assumption in almost all earlier research efforts is

that there is a one-way causal flow from the BE characteristics to travel behavior. Specifically, the assumption is that households and individuals locate themselves in neighborhoods and then, based on neighborhood attributes, determine their travel behaviors. Thus, on the basis of these studies, if good land-use mixing has a negative influence on the number of motorized trips, the implication would be that building neighborhoods with good land-use mix would result in decreased motorized trips in the population, with a concomitant decrease in traffic congestion levels. A problem with the above line of reasoning is that it does not take a comprehensive view of how individuals and households make residential choice and travel decisions. In reality, households and individuals who are auto-disinclined, because of their demographics, attitudes, or other characteristics, may search for locations with high residential densities, good land-use mix, and high public transit service levels, so they can pursue their activities using non-motorized travel modes. If this were true, urban land-use policies aimed at, for example, increasing density or land-use mix, would not stimulate lower levels of auto use in the overall population, but would simply alter the spatial residence patterns of the population based on motorized mode use desires. Ignoring this self-selection in residence choices can lead to a "spurious" causal effect of neighborhood attributes on travel, and potentially lead to misinformed BE design policies.[1]

Disentangling the "spurious" and "true" causal effects of neighborhood BE attributes is critical to understanding the causal relationships between the BE and travel, and contributes to the discussions regarding the effectiveness of new urbanism and smart growth strategies to reduce auto use. Several earlier authors, including Boarnet and Crane (2001), Cervero and Duncan (2003), and Krizek (2003), have raised the issue of self selection in the assessment of BE attribute impacts on travel choices. Suggestive evidence of self-selection has been noted in empirical studies by Kitamura *et al.*, (1997), Handy and Clifton (2001), and Krizek (2000).

The literature that has considered the self-selection issue (also refereed to as the residential sorting issue) in assessing the impact of BE attributes on travel choices has done so in one of three ways: (1) Controlling for decision-maker attributes that jointly impact residential and travel choices, (2) Using instrumental variable methods to econometrically accommodate the

---

[1] A caveat here. The above discussion assumes that there is an adequate supply of neighborhoods to choose from for persons who are auto-disinclined. If there is an undersupply, then building neighborhoods that promote alternatives to auto use would lead to a reduction in auto use in the population even if the only process at work is residential sorting. However, in this scenario, the policy questions shift from impacting travel behavior to providing a better balance between the demand for non-auto oriented neighborhoods and the supply of such neighborhoods (see also Crane, 2000).

potential endeneity of residential choice decisions, or (3) Using before-after household move data that potentially controls for household travel desires and attitudes.

1.2.1 Controlling for Decision-Maker Attributes

The first approach is to control for demographic and other travel-related attitudes/perceptions of decision-makers that may impact the neighborhood type individuals choose. This can be accomplished by incorporating decision-making characteristics as explanatory variables in models of travel behavior. For instance, households with small children might locate in neighborhoods with easy-to-access park facilities and pursue several non-motorized recreation trips to nearby parks. By including "households with small children" as a variable in a model of non-motorized recreation trips, one controls for neighborhood selection and obtains the "true" impact of park accessibility on recreational trip generation. As indicated earlier in Section 1.1.2, most disaggregate-level studies accommodate demographics in modeling travel choices. However, it is likely that factors other than the typically collected demographic data on decision-makers are at play in residential sorting and travel choices. As an example, Lund (2003) includes three attitudinal variables (in addition to demographic and perception variables) in a study of BE effects on weekly frequency of strolling trips and utilitarian trips by walk. The three attitudinal variables are (1) importance of walking to daily activities, (2) interacting with one's neighbors, and (3) feeling "at home" in the neighborhood. The first one of these is statistically significant, indicating that, if this variable was not controlled for, it would have potentially led to an overestimation of the effect of BE characteristics on walk trips (because individuals who value walking are likely to locate themselves in neighborhoods with a walk-conducive BE). Other studies that have included travel-related attitudes to, in part, alleviate the residential sorting issue are Kitamura *et al.* (1997), Bagley and Mokhtarian (2002), Schwanen and Mokhtarian (2004, 2005), Handy *et al.* (2005), and Khattak and Rodriguez (2005). The basic reasoning in all these studies is that after controlling for demographic and attitudinal factors that are likely to affect residential sorting, the remaining effect of BE measures is closer to the "cleansed and true" causal effect of the BE measures on travel. This is a creative, and simple, way of tackling the self-selection problem, but its use in practice is limited by the fact that most travel survey data sets do not collect attitudinal data. Further, it is unlikely that all the demographic and travel lifestyle attitudes that have any substantive impact on residential

sorting can be collected in a survey, because of which it becomes difficult to gauge how close the estimated BE effects are to the "true" causal effect.

1.2.2 Instrumental Variables Approach

The second approach to alleviate the residential sample selection effect is to use a two-stage instrumental variable approach where the endogenous "explanatory" BE attributes are first regressed on instruments that are related to the BE attributes, but have little correlation with the randomness in the primary travel behavior of interest. The predicted values of the BE attributes from this first regression are next introduced as independent variables (along with other demographic attributes of the individual) in the travel behavior relationship of interest. For example, Boarnet and Sarmiento (1998) and Boarnet and Crane (2001) select four non-transportation neighborhood amenities as instruments, and use the predicted values of various density measures on these instrumental variables to estimate the effect of BE measures on non-work automobile trips.

A problem with the instrumental variable approach as discussed above, however, is that it is not applicable to the case where the travel behavior equation of interest has a non-linear structure, such as a discrete choice or a limited/truncated variable (this is the reason that Boarnet and Sarmiento switch from an ordered response model to a simple linear regression model within the same paper when using the instrumental variable approach). There are control function and related approaches today to deal with the case of endogenous "explanatory" variables in the context of discrete choice and other non-linear models (see Berry *et al.*, 1995; Lewbel, 2004; Louviere *et al.*, 2005), but these methods need rather tedious computations to recognize the sampling variation in the predicted value of the endogenous BE attributes to obtain the correct standard errors in the main equation of interest. The alternative of ignoring the sampling variance in the predicted values of the BE attributes, as done by Boarnet and Sarmiento, can lead to incorrect conclusions about the statistical significance of the effects of the BE attributes.

1.2.3 Using Before-After Household Move Data

The third approach to alleviate the residential sorting effect is to examine the travel patterns of households immediately before and immediately after a household relocation. The potential advantage of examining the same household in two different neighborhoods is that one can

ostensibly control for the overall travel desires and attitudes of the members of a household, so that the before-after relocation changes in travel behavior may be attributed to the different built environments in the two neighborhoods. The essential idea in this approach is to consider the relocation as a "treatment", with the associated travel behavior changes being the response variable. The assumption one would make in such an analysis is that relocating households are in equilibrium in their pre-move neighborhood in terms of BE attributes, and moved because of factors unrelated to their preference of BE attributes (such as to upgrade the physical housing stock in response to higher incomes or a change in lifecycle).

A longitudinal before-after relocation study of the type discussed above is undertaken by Krizek (2003), who examines the changes in travel behavior between two consecutive years for relocating households using the Puget Sound Transportation Panel. While such a longitudinal approach is one way of alleviating the self-selection problem, a problem with the approach is that the relocating households are themselves a self-selected group, and may have moved because of dissonance in the pre-move neighborhood BE attributes vis-à-vis their desired configuration of BE attributes.

1.3 The Current Report

In this report, we contribute to the literature on BE-travel behavior interactions by addressing some of the challenges discussed in the previous two sections. In particular, and organized by the order of points raised in the earlier two sections, our research may be characterized as follows. First, we develop and use a whole gamut of BE attributes in our analysis of BE effects. On the travel side, however, we narrow our analysis to BE effects on car ownership choice (see further discussion of this issue toward the end of this section). Second, we consider a range of demographic variables, and their direct as well as indirect effects (through interactions with BE attributes), in our car ownership choice model. In addition, we explicitly recognize the presence of unobserved heterogeneity (that is, sensitivity variations due to unobserved household/individual factors) in examining the effect of BE attributes on car ownership choice. Ignoring such unobserved heterogeneity when present will, in general, lead to inconsistent conclusions about the impact of BE attributes (see Chamberlain, 1980; Bhat, 1998). Third, we propose and apply a general methodology to control for the self-selection of individuals into neighborhoods in an effort to obtain a "cleansed" and "true" causal effect of BE

measures on travel behavior. The methodology can be used to control for self-selection for any kind of travel behavior variable and directly provides the correct standard errors regarding the effect of BE attributes. It is geared toward cross-section analysis, recognizing that almost all existing data sources available for analysis of BE effects are cross-sectional in nature. Unlike earlier studies, the methodology also explicitly considers and models the residential location choice decision jointly with the travel behavior choice of interest. Such a joint model provides a valuable tool for policy analysis, since it can predict how residential choices would change due to urban form design policies as well as estimate the travel behavior change of individuals. Thus, in combination, the joint model provides a complete picture of the spatial pattern of travel changes in response to BE design policies.

Three limitations of our study are also worth noting here. <u>First</u>, we do not consider travel-related attitudinal variables or perceptions of neighborhood BE attributes in our analysis. Some earlier studies discussed above have shown that attitudes and subjective perceptions can play an important role in travel behavior choices. Further, controlling for such attitudes/perceptions can serve to lessen the impact of the residential sorting issue. However, most travel surveys do not collect such attitudinal/perception information. Besides, our methodology is general and is readily applicable for use with attitude and perception data, should such data become available. <u>Second</u>, we use rather aggregate and pre-defined spatial units (*i.e.*, traffic analysis zones) as surrogates for neighborhoods that people choose from in making their residential choices. Also, our computations of BE measures are based on these aggregate spatial units (as opposed to having a sliding neighborhood definition for computation of BE measures, as advocated by Guo and Bhat, 2004, 2006, and others). Again, however, our overall methodology can be extended for use with higher resolutions of spatial unit definitions. <u>Third</u>, we assume that employment location is pre-determined for all employed individuals in the household. Extending the analysis framework to consider a joint model of work location and residential location would be a fruitful avenue for further research.

As indicated earlier, the research in the report uses household car ownership level as the travel-related variable of focus to assess the impact of BE attributes. This choice was based on three considerations. First, car ownership is an important intervening variable in the effect of BE attributes on travel decisions. After all, car ownership and residential choice decisions may be characterized as medium-term decisions as opposed to the shorter-term day-to-day travel

decisions of individuals. For example, Messenger and Ewing (1996) indicate that density affects the use of the transit mode of travel only through its effect on car ownership. If the impact of BE measures on car ownership is ignored, and car ownership is included directly as an exogenous variable along with BE attributes in a travel choice model, the coefficients on the BE attributes will underestimate the true cumulative impact (directly and indirectly through car ownership changes) on the travel choice of changes in BE attributes. Second, there has been relatively less empirical attention on the effect of BE characteristics on car ownership than on other choices. As indicated by TRB (2003), "most researchers have not isolated effects of land-use mix or site design on auto ownership". A few recent attempts at shedding more light on this issue include the research of Hess and Ong, 2001, Bhat and Pulugurta, 1998, Holtzclaw *et al.*, 2002, and Shay and Khattak, 2005). Third, auto ownership has been found to impact almost all aspects of daily activity-travel patterns, including the number of out-of-home activity episodes that individuals participate in, the location of out-of-home participations, and the travel mode and time-of-day of out-of-home activity participations (see, for example, Bhat and Lockwood, 2004; Pucher and Renne, 2003; Bhat and Castelar, 2002).

The rest of this report is structured as follows. The next chapter presents the econometric framework. Chapter 3 discusses the data sources and variables considered in the analysis. Chapter 4 focuses on the empirical results. Chapter 5 concludes the report by summarizing the important findings of the research.

CHAPTER 2. ECONOMETRIC FRAMEWORK

2.1 Mathematics of the Model Structure

In the following presentation, we will use the index i to represent the spatial unit of residential choice ($i = 1, 2,\ldots, I$), index k to represent the number of cars in a household ($k = 0, 1, 2,\ldots, K$), and the index q to represent the qth household ($q = 1, 2,\ldots, Q$). Let x_i be a vector of BE attributes characterizing residential spatial unit i. The equation system for the joint residential choice and car ownership model is then as follows:

$$u_{qi}^* = \beta_q' z_i + \gamma_q' x_i + \varepsilon_{qi}, \text{ spatial unit } i \text{ chosen if } u_{qi}^* > \max_{\substack{j=1,2,\ldots I \\ j \neq i}} u_{qj}^* \tag{1}$$

$$c_{qi}^* = \alpha' y_q + \delta_q' x_i + \xi_{qi}, \; c_{qi} = k \text{ if } \psi_{k-1} < c_{qi}^* < \psi_k.$$

The first equation above is the usual indirect utility function for the choice of residence. Specifically, u_{qi}^* is the indirect (latent) utility that the qth household derives from locating itself in spatial unit i, z_i is a vector of non-BE attributes of spatial unit i affecting residential choice (for example, quality of schools, average cost of homes, racial composition, commute time, *etc.*), and x_i is a vector of BE attributes impacting residential choice (land-use mix, density, transit-accessibility, *etc.*). β_q is a household-specific coefficient vector capturing the sensitivity to the attributes in vector z_i. β_q can vary based on observed (to the analyst) household characteristics as well as unobserved (to the analyst) household characteristics. For instance, the sensitivity to the average cost of homes in zone i may be moderated by the income earnings of household q, as well as household unobserved characteristics (such as money consciousness and risk taking nature). In the rest of this section, and only for notational convenience, we will ignore the varying nature of the sensitivity to non-BE attributes of spatial units across households and write $\beta_q = \beta$. γ_q is a household-specific coefficient vector capturing the sensitivity to BE attributes in vector x_i. We parameterize each element of γ_q as follows: $\gamma_{ql} = (\gamma_l + \theta_l' w_{ql} + v_{ql} + \omega_{ql})$, where w_{ql} is a vector of observed household-specific factors affecting sensitivity to the lth BE attribute in vector x_i, and v_{ql} and ω_{ql} are household-specific unobserved factors impacting

household q's sensitivity to the lth BE attribute. v_{ql} includes only those household-specific unobserved factors that influence sensitivity to residential choice, while ω_{ql} includes only those household-specific unobserved factors that impact both residential choice and car ownership choice. For example, consider a household's sensitivity to street block density. The household may have a higher sensitivity (than its observationally equivalent peer group) to street block density because members of the household are social extroverts and perceive higher street block density as providing a more socially vibrant setting conducive to their social outlook. The socially extroverted nature, however, may not have an impact on car ownership. This would be captured in v_{ql}. Now, another unobserved household factor may be overall auto disinclination due to environmental concerns. This is likely to impact the sensitivity to street block density in residential choice (because higher street block densities may be more conducive to non-motorized and transit forms of travel) and also influence car ownership propensity. This would be included in ω_{ql} (more on this later). Finally, in the first equation of the model system in Equation (1), ε_{qi} is an idiosyncratic error term assumed to be identically and independently standard extreme-value distributed across alternatives i and households q.

The second equation in equation system (1) corresponds to an ordered-response structure for car ownership decisions (see Bhat and Pulugurta, 1998 and Hess and Ong, 2001 for examples of earlier applications of the ordered-response structure for car ownership). The ordered response structure explicitly recognizes the ordinal and discrete nature of car ownership. c_{qi}^{*} represents the latent car ownership propensity of household q should the household choose to locate in spatial unit i, y_q is a set of household characteristics (such as income and number of children) that influences car ownership levels, and x_i is the vector of BE attributes corresponding to spatial unit i.[2] α, in the car ownership equation, is a coefficient vector representing the impact of sociodemographics on car ownership propensity, and δ_q is a household specific coefficient vector capturing the impact of BE attributes on car ownership

---

[2] Note that we are introducing the full vector x_i of BE attributes in both the residential choice and car ownership equation for notational ease. In general, some of the BE attributes will not impact residential choice (the corresponding element in γ_q is zero for all q) and some will not influence car ownership choice (the corresponding element in δ_q is zero for all q). Additionally, it is possible that BE attributes have a mean effect of zero across households for residential choice and/or car ownership, but have a statistically significant distribution around the zero mean.

decisions. We parameterize the elements of δ_q as follows: $\delta_{ql} = (\delta_l + \Delta_l' s_{ql} + \eta_{ql})$, where s_{ql} is a vector of observed household-specific factors influencing sensitivity to the lth BE attribute in x_i, Δ_l is a corresponding vector of coefficients, and η_{ql} is a term capturing the impact of household-specific unobserved terms associated with different sensitivities to BE attributes in car ownership decisions. ξ_{qi} is an error term that we partition into two components: $\sum_l (\pm \omega_{ql}) x_{il} + \zeta_{qi}$. The $\pm \omega_{ql} x_{il}$ terms are the common error components relating to the sensitivity to BE attributes in residential choice and car ownership propensity, while ξ_{qi} is an idiosyncratic term assumed to be identically and independently standard logistic distributed across individuals. The car ownership propensity c_{qi}^* is mapped to the observed car ownership level c_{qi} using the ψ thresholds in the usual ordered-response fashion.

With the notational and parameterizations discussed above, the equation system of (1) can be written as:

$$u_{qi}^* = \beta' z_i + \sum_l \left(\gamma_l + \theta_l' w_{ql} + v_{ql} \right) x_{il} + \left(\sum_l \omega_{ql} x_{il} + \varepsilon_{qi} \right) \tag{2}$$

$$c_{qi}^* = \alpha' y_q + \sum_l \left(\delta_l + \Delta_l' s_{ql} + \eta_{ql} \right) x_{il} + \left(\sum_l \pm \omega_{ql} x_{il} + \zeta_{qi} \right)$$

2.2 Intuitive Discussion of Model Structure

The reader will note that the self-selection into neighborhoods based on seeking out those neighborhoods that are compatible with car ownership desires is accommodated in the joint model system in Equation (2) in two ways. <u>First</u>, we are controlling for the effect of systematic sociodemographic differences among households in their resident location patterns. Consider, for example, that high income households stay away from high density neighborhoods, and prefer exclusive, sprawling, residential enclaves for their residences. This can be reflected by including income as a variable in the w_{ql} vector in the residential choice equation. High income households are also likely to own more cars than low income households. The residential sorting based on income can then be controlled for when evaluating the effect of density on car

ownership by including income as a variable in the y_q vector. <u>Second</u>, unobserved attributes (such as travel attributes, lifestyle, and environmental considerations) may influence both residential choice and car ownership propensity. For example, households who intrinsically like to use non-motorized forms of transport or transit may locate themselves in high density neighborhoods that are conducive to their preferences, and also own fewer cars. Such unobserved preferences are captured in the common $\omega_{ql} x_{il}$ terms in the two equations. The '$\pm$' in the front of the $\omega_{ql} x_{il}$ terms in the car ownership propensity equation indicates that the correlation in unobserved factors may be positive or negative. If the sign is '+', it implies that households who intrinsically prefer the BE characteristic represented by x_{il} are also more likely to own cars, while a '–' sign implies that households who prefer the BE characteristic captured by x_{il} are less likely to own cars. If the x_{il} measures are defined in the context of promoting smart growth and neo-urbanism concepts (such as high density and increased land use diversity), then there may be an expectation that the appropriate sign should be negative. In our analysis, we are able to test which one of the two signs is appropriate. If the sign were to be indeed negative (that is, households who have an intrinsic preference for neo-urbanist neighborhoods also have a lower preference for cars due to unobserved attributes such as auto-disinclination), ignoring these $\omega_{ql} x_{il}$ terms while estimating the car ownership propensity equation leads to an artificial inflation of the negative sign on the corresponding neo-urbanist BE attributes (*i.e.*, an artificial inflation of the negative sign on the δ_l terms).

2.3 Model Estimation

The parameters to be estimated in the equation system of (2) include the β and α vectors, the ψ thresholds in the ordered response car ownership model, the γ_l, δ_l, θ_l, and Δ_l vectors (some of the elements of these vectors will be zero because not all BE attributes will have an effect on residential choice and/or car ownership), and the variances of $v_{ql} (= \sigma_{vl}^2)$, $\eta_{ql} (= \sigma_{\eta l}^2)$, and $\omega_{ql} (= \sigma_{\omega l}^2)$ for those BE attributes with random taste heterogeneity. The reader will note that the random heterogeneity in sensitivity to a particular BE attribute l may occur only in residential choice ($\sigma_{vl}^2 \neq 0$, $\sigma_{\eta l}^2 = 0$, $\sigma_{\omega l}^2 = 0$), only in car ownership propensity

16

($\sigma_{vl}^2 = 0$, $\sigma_{\eta l}^2 \neq 0$, $\sigma_{\omega l}^2 = 0$), independently in residential choice and car ownership propensity ($\sigma_{vl}^2 \neq 0$, $\sigma_{\eta l}^2 \neq 0$, $\sigma_{\omega l}^2 = 0$), and combinations of the above patterns with a common effect on both residential choice and car ownership propensity ($\sigma_{\omega l}^2 \neq 0$). Of course, there may not be any random heterogeneity at all for some BE attributes ($\sigma_{vl}^2 = 0$, $\sigma_{\eta l}^2 = 0$, and $\sigma_{\omega l}^2 = 0$).

Let Ω represent a vector that includes all the parameters to be estimated, and let $\Omega_{-\sigma}$ represent a vector of all parameters except the variance terms. Also, let d_q be a vector that stacks the v_{ql}, η_{ql}, and ω_{ql} terms across all BE attributes and let Σ be a corresponding vector of standard errors. In the current application, we will assume independence across (a) the elements of the d_q vector, (b) the d_q vector, and the ε_{qi} and ζ_{qi} terms for all i, and (c) all unobserved and observed elements. Define $a_{qi} = 1$ if household q resides in spatial unit i and 0 otherwise. Similarly, define $b_{qk} = 1$ if household q owns k cars and 0 otherwise. Then, the likelihood function for a given value of $\Omega_{-\sigma}$ and d_q may be written for household q as:

$$L_q(\Omega_{-\sigma}) \mid d_q = \left[\frac{\exp(\beta' z_i + \gamma_q' x_i)}{\sum_j \exp(\beta' z_j + \gamma_q' x_j)} \right]^{a_{qi}} \cdot \left[G(\psi_k - \alpha' y_q - \delta_q' x_i) - G(\psi_{k-1} - \alpha' y_q - \delta_q' x_i) \right]^{b_{qk}}, \quad (3)$$

where $G(.)$ is the cumulative distribution of the standard logistic distribution. Finally, the unconditional likelihood function can be computed for household q as:

$$L_q(\Omega) = \int_{d_q} \left(L_q(\Omega_{-\sigma}) \mid d_q \right) dF(d_q \mid \Sigma), \quad (4)$$

where F is the multidimensional cumulative normal distribution. The log-likelihood function is $L(\Omega) = \sum_q \ln L_q(\Omega)$.

In the current report, we apply simulation techniques to approximate the multidimensional integral in Equation (4), and maximize the resulting simulated log-likelihood function. Specifically, we use the scrambled Halton sequence to draw realizations of c_q from its

population normal distribution. Details of the Halton sequence and the procedure to generate this sequence are available in Bhat (2003) and Sivakumar *et al.* (2005).

CHAPTER 3. DATA SOURCES AND VARIABLE SPECIFICATIONS

3.1 Data Sources

The area selected for this study is the Alameda County in the San Francisco Bay area. The county contains 233 transport analysis zones, and the residential choice and car ownership levels of households within this county were the focus of the current empirical analysis.

The primary data source used in the analysis is the 2000 San Francisco Bay Area Travel Survey (BATS). This survey was designed and administered by MORPACE International, Inc. for the Bay Area Metropolitan Transportation Commission. The survey collected information on the car ownership levels and residential location from over 15,000 households in the Bay Area (see MORPACE International Inc., 2002 for details on survey, sampling, and administration procedures). Further, data on individual and household demographics, as well as individual employment-related characteristics (including employment location if employed), were obtained.

In addition to the activity survey, six other data sets associated with the San Francisco Bay area were used in the current analysis: land-use/demographic coverage data, zone-to-zone travel level-of-service (LOS) data, a GIS layer of bicycle facilities, the Census 2000 Tiger files, and the Census 2000 Population and Housing Data Summary Files. These data are discussed in the next three paragraphs.

Both the land-use/demographic and LOS data files were obtained from the Metropolitan Transportation Commission (MTC) in the San Francisco Bay area. The land-use/demographic file provided, for each of the Traffic Analysis Zones (TAZ), data on (1) area of coverage by land-use purpose, (2) number of housing units by dwelling type, (3) employment levels by sector, (4) population, income and age distribution of the population, and (5) area type by zone. The MTC also provided zone-to-zone travel level of service (LOS) data that included inter-zonal distances, as well as peak and off-peak travel times and costs. The land-use/demographic and LOS files were used to characterize the demographic characteristics of households in each zone, the urban environment, and accessibility to work and other activity opportunities, as discussed in the next section.

Another data source obtained from MTC was a GIS line layer describing all existing bicycle facilities in the Bay Area region. It included class 1 facilities (separate paths for cyclists and pedestrians), class 2 facilities (painted lanes solely for cyclists), and class 3 facilities (signed routes on shared roads). A fourth source of data was the Census 2000 TIGER files, from which two GIS line layers were extracted for the Bay Area region: one is the highway network (including interstate, toll, national, state, and county highways) and the other is the local roadways network (including local, neighborhood, and rural roads).

Finally, the Census 2000 Population and Housing Data Summary File 1 (SF1) was used to compute the ethnic composition, average household size, average household income, and average housing cost of each zone. The census block-group level data were aggregated to the TAZ level using a spatial overlay process to obtain the requisite zonal-level measures. The final sample for analysis comprised 2,954 households.

3.2 Variable Specification

The six spatial data sources discussed in the previous section provide a rich set of variables for consideration in model specification. Almost all the non-demographic variables derived for use in the current analysis are related to the built environment. These variables may be broadly classified into five groups, and are briefly discussed in the subsequent sections (Sections 3.2.1 through 3.2.5). Section 3.2.6 and 3.2.7 present the zonal demographics, housing cost, and ethnic composition variables used in the analysis. Section 3.2.8 provides an overview of the household demographic factors considered in the model. Finally, Section 3.2.9 discusses interactions of the household demographic characteristics with the variables presented in sections 3.2.1 through 3.2.8.

3.2.1 Zonal Size Density Measures

These variables relate to the size of the zone (population, number of housing units, *etc.*) and the density of the zone (# of households per acre, employment per acre by sector, *etc.*). These measures are included to examine the influence of the residential environment on residential choice and car ownership choice.

3.2.2 Zonal Land-Use Structure

These variables include land-use composition measures (percentages of zonal area in residential, commercial, and other land-uses), housing type measures (fractions of single family, multifamily, duplex, and other housing units), and a land-use mix diversity index computed from the land-use composition measures as:

$$\text{Land - use mix diversity} = 1 - \left\{ \frac{\left| \dfrac{r}{L} - \dfrac{1}{3} \right| + \left| \dfrac{c}{L} - \dfrac{1}{3} \right| + \left| \dfrac{o}{L} - \dfrac{1}{3} \right|}{4/3} \right\}, \tag{5}$$

where $L = r + c + o$, and r is the zonal acreage in residence use, c is the acreage in commercial/industrial use, and o is the acreage in other uses.

3.2.3 Regional Accessibility Measures

The regional accessibility measures are of the Hansen-type (Fotheringham, 1983) and are computed separately for the drive and transit modes, using the land-use/demographic and level-of-service files obtained from MTC. Three accessibility measures are developed for each mode m as follows:

$$A_{im}^{emp} = \left[\sum_{j=1}^{N} \left(\frac{E_j}{t_{ijm}} \right) \right] \Big/ N, \ A_{im}^{shop} = \left[\sum_{j=1}^{N} \left(\frac{R_j}{t_{ijm}} \right) \right] \Big/ N, \text{ and } A_{im}^{rec} = \left[\sum_{j=1}^{N} \left(\frac{V_j}{t_{ijm}} \right) \right] \Big/ N, \tag{6}$$

where A_{im}^{emp}, A_{im}^{shop}, and A_{im}^{rec} denote the employment, shopping, and recreational accessibility, respectively, for zone i by mode m; E_j, R_j, and V_j are the number of basic employees, number of retail employees, and vacant land acreage, respectively, in zone j; t_{ijm} is the travel time from zone i to zone j by mode m; and N is the total number of TAZs.

The computation of the accessibility measures is straightforward for the drive mode, since each zone is connected to each other zone by the highway network. However, some zones are not serviced by transit from a particular origin zone (these are identified in the MTC transit network). In such cases, we generated the transit accessibility measures only over those zones with service availability (the denominator N in Equation (6) changes to the total number of zones

reachable by transit from zone *i*). But, in addition, we also generated a dummy variable indicating if a zone *i* is not connected to one or more other zones, and a count variable indicating the number of zones serviced by transit from zone *i*.

3.2.4 Commute-Related Variables

The commute variables include times and costs by each of the drive and transit modes. These are computed from the level-of-service file obtained from MTC, assuming that the employment location is predetermined. The variables are obtained as aggregate values across all workers in the household. For the transit mode, the variables are computed over all workers in the household who have transit available from a candidate residential zone to their work zone. In addition, a dummy variable indicating if transit is not available for the commute of one or more workers (from each candidate residential zone to the employment zone) is created.[3]

3.2.5 Local Transportation Network Measures

The local transportation network measures relate to bicycle facilities, highway facilities, and transit access within a zone. Specifically, we computed bikeway density (miles of bikeway facility per square mile), street block density (# of blocks per square mile), highway density (miles of highway per square mile), transit availability in zone, and access time to transit stop if transit is available. These measures are included because they represent local measures of non-motorized, transit, and auto levels of service, which can impact both residential choice decisions and car ownership choices.

3.2.6 Zonal Demographics and Housing Cost Variables

These variables refer to average household size, household income, and housing cost in each zone, and are derived from the Census 2000 population and housing data summary file 1 (SF1).

---

[3] The joint model in the report recognizes the possibility that individuals select their commute level of service by locating in zones that are attractive based on their preferred car ownership level (for example, a household that prefers transit would locate in zones with good transit access and also be less likely to own cars; failure to accommodate this jointness would result in an inflated negative influence of good transit access on car ownership).

3.2.7 Zonal Ethnic Composition

These variables are considered to accommodate ethnic preferences in household residence location, and are also obtained from the Census SF1 file. The ethnic composition for each zone is available in the following categories: non-Hispanic Caucasians, African-Americans, Asians, Hispanics, and others.

3.2.8 Household Demographic Variables

A number of household demographic variables are considered in the joint model of residential choice and car ownership. These include household size, presence and number of active adults, presence and number of senior adults (> 65 years of age), presence and number of employed individuals, presence and number of children (< 16 years of age), presence and number of physically challenged individuals, race of household, household structure (nuclear family, couple family, single person, single parent family, and other family types), household income, whether or not the household owns or rents its residence, and the housing type (*i.e.*, whether the household lives in a single family detached housing or multifamily housing). The demographic variables are introduced by interacting with the non-household demographic variables (see next section). In addition, in the car ownership model, these variables are also directly introduced to explain car ownership decisions.

3.2.9 Interaction of Household Demographics with Attributes

An important focus of the current effort is to examine the variations in sensitivity across households to zonal built environment characteristics, zonal demographics, regional accessibility, and commute characteristics. For instance, lower income individuals may be likely to choose to locate in zones with good transit service more so than higher income individuals. Further, it is possible that good transit availability has a higher negative impact on owning fewer cars for lower income households than for higher income households. Clearly, demographics can moderate the influence of the built environment and other variables on residential choice and car ownership levels. Thus, we consider several interactions of such household demographics as income, size, household structure, presence of children in household, and number of employed individuals, with the different groups of variables identified in the earlier sections.

The final specification of variables was based on a systematic process of eliminating variables found to be statistically insignificant and parsimony in representation. The specification was additionally guided by intuitive considerations, and the results from earlier studies.

CHAPTER 4. EMPIRICAL RESULTS

The joint model estimates both the residential location choice parameters as well as the car ownership parameters simultaneously, as discussed in Chapter 2. However, for ease in presentation, we discuss the estimation results separately in the next two sections (Sections 4.1 and 4.2). Section 4.3 discusses residential sorting effects.

4.1 Residential Location Choice Model Results

The residential location model is estimated with 233 alternative zonal locations to choose from. We do not undertake a sampling scheme to reduce the number of alternatives because our overall residential choice model does not correspond to a simple multinomial logit model (rather there is mixing in the model). Nerella and Bhat (2004) indicate that sampling of alternatives in the presence of mixing distributions leads to inaccurate parameter estimates.

The residential choice parameter estimates are presented in Table 1. The coefficient on the logarithm of number of households in a zone has the expected positive sign, indicating that households are more likely to locate in zones with a large number of housing opportunities. The parameter on this size variable is close to, and statistically insignificantly different from, the value of 1, suggesting a well-specified model that is not dependent on the zonal configuration used in estimation (see Bhat *et al.*, 1998; Ben-Akiva and Lerman, 1985, p. 249; Daly, 1982). Among the density measures, we found both household density (number of households per acre) and total employment density (number of employed individuals per acre) to have an impact on residential choice. Households without seniors locate in zones with high household density, perhaps due to better housing availability in these zones or simply due to population clustering. However, households with seniors shy away from high housing density developments. The effect of total employment density indicates that middle and high income households (households not in the lowest income quartile) prefer zones with a low employment density, while low income households are indifferent to employment density in their residential choices.

Table 1. Estimation Results of the Residential Location Choice Model

Variables	Parameter	t-statistic
Zonal size and density measures (including demographic interactions)		
Logarithm of number of households in zone	0.9105	19.931
Household density (#households per acre x 10^{-1})	0.4159	4.697
interacted with presence of seniors in household	-0.8017	-4.653
Employment density (#employment per acre x 10^{-1})	-0.2063	-3.646
interacted with household income in the lowest quartile	0.2230	3.657
Zonal land-use structure variables (including demographic interactions)		
Fraction of residential land area	-0.6038	-4.936
interacted with presence of seniors in household	1.1581	4.888
Fraction of single family housing interacted with household living in single family detached housing	2.2445	15.291
Regional accessibility measures (including demographic interactions)		
Recreation accessibility (by drive mode)	0.1687	2.077
interacted with household income in the lowest quartile	-0.3199	-1.963
Commute-related variables (including demographic interactions)		
Drive commute time (minutes x 10^{-1})	-1.2667	-27.521
interacted with household income in the second highest quartile	-0.0557	-3.172
interacted with household income in the third and fourth highest quartiles	-0.1170	-4.408
Standard deviation	0.7524	15.145
Drive commute cost ($ x 10^{-1}) interacted with household income in the lowest quartile	-5.0371	-3.103
Local transportation network measures (including demographic interactions)		
Street block density (number of block per square mile x 10^{-1})	-0.1790	-2.717
interacted with single person household	0.2027	3.082
Standard deviation	0.4911	5.879
Bicycle facility density (miles per square mile x 10^{-1})	0.4030	6.369
Transit availability	0.4462	2.528
Transit access time to stop (minutes x 10^{-1})	-0.2113	-3.217
Zonal demographics and housing cost (including demographic interactions)		
Absolute difference between zonal median income and household income ($ x 10^{-5})	-1.8610	-12.279
Absolute difference between zonal average household size and household size	-0.5445	-9.196
Average housing value	-0.0893	-4.250
Zonal ethnic composition measure		
Fraction of Caucasian population interacted with Caucasian dummy variable	2.5211	14.378
Fraction of African-American population interacted with African-American dummy variable	4.2856	9.642
Fraction of Hispanic population interacted with Hispanic dummy variable	3.3652	6.474

Among the zonal land-use structure variables, the results indicate that households with no senior adults tend to stay away from zones with a high fraction of land invested in residential land use, though the reverse holds for households with senior adults. Further, households who live in single-family detached housing are drawn to zones with a large share of single family housing units, as one would expect. Interestingly, land-use mix diversity does not influence residential location after controlling for other variables.

The regional accessibility measures do not appear to be important determinants of residential location choice. The only statistically significant variables correspond to recreational accessibility. The parameter estimates on the two recreational accessibility variables in Table 2 indicate that middle and high income households (those in the top three quartiles of income earnings) locate themselves in good recreational accessibility areas, while low income households tend to reside in areas of low recreation accessibility. The latter finding is interesting, and presumably reflects market and other forces that make it difficult for low income households to be close to good recreational opportunities. This is an issue that warrants additional attention and equity analysis.

Table 2. Estimation Results of the Car Ownership Model

Variables	Parameter	t-statistic
Zonal size and density measures (including demographic interactions)		
Household density (#households per acre x 10^{-1})	-0.0562	-0.301
interacted with presence of seniors in household	-0.3123	-1.162
Employment density (#employment per acre x 10^{-1})	-0.2148	-1.554
interacted with household income in the lowest quartile	-0.3411	-2.075
Standard deviation	-0.4235	-2.586
Commute-related variables (including demographic interactions)		
Drive commute time (minutes x 10^{-1})	0.1975	2.478
Drive commute cost ($ x 10^{-1})	-2.9501	-2.254
Local transportation network measures (including demographic interactions)		
Street block density (number of block per square mile x 10^{-1})	-0.3416	-3.376
Standard deviation	0.4602	5.339
Transit availability	-0.4623	-1.262
interacted with household income in the lowest quartile	-0.5693	-3.658
Transit access time to stop (minutes x 10^{-1})	0.1994	1.475
Household demographic variables		
Number of active adults	1.4189	14.197
Number of senior adults	1.4101	10.219
Number of employed individuals	0.3415	3.961
Number of physically challenged individuals	-1.0679	-6.954
Household income ($ x 10^{-5})	0.4428	3.545
Single parent household	-1.2949	-4.167
Single individual household	-1.3040	-8.221
Residing in a multi-family housing unit	-0.8343	-6.538
Non-Caucasian non-African-American household	-0.4578	-3.459
Own household dwelling	0.7995	6.361

The commute-related variables are important determinants of residential choice. Households, in general, locate to reduce their drive commute time. Further, high income households, in particular, choose residential locations that are closer to their work place, perhaps because they are more able to afford housing at locations close to work locations (this result lends credence to earlier urban planning research studies that indicate a high level of jobs-housing imbalance for low income families; see, for example, Ihlanfeldt and Sjoquist, 1998). The results also indicate the presence of significant unobserved variations in the sensitivity to the drive commute time. Specifically, the results show that, among low income individuals, while the average effect of commute time is negative, a small percentage (5%) have a preference to locate farther from their workplace than closer. This percentage of individuals who prefer to locate farther rather than closer to the work place reduces to 4% for middle income households (those in the second income quartile), and 3% for high income households (those in the third and fourth income quartiles). Also, as expected, low income households locate themselves in zones with low driving costs to work. Interestingly, all the commute related variables are associated with the relatively ubiquitous drive mode and not the transit mode, suggesting that, at least when making residential choice decisions, households make decisions based on driving times and costs to work.

Several local transportation network measures are found to be important determinants of residential location. The results show that multiple individual households are generally reluctant to locate in areas with a high street block density, while single individual households prefer to locate in such neighborhoods. This may be a reflection of the more socially and physically active lifestyles of single individual households, which make them seek neighborhoods that are walking-friendly. However, there is a high degree of variation in the sensitivity to street block density due to unobserved factors (some households may prefer neighborhoods with high street block density because of, say, potentially more social interactions, while others may stay away from such neighborhoods because of a preference for privacy or a sprawling neighborhood). Specifically, the standard error on the variable, along with the mean effects, indicates that about 35% of multiple person households prefer high street block density, while 65% of such households prefer low street block density. The corresponding numbers for single-person households are 48% and 52%, respectively. The effect of the next variable, the bicycle facility density variable, indicates that, without exception, households prefer zones with a dense bicycle

network. Finally, among the local transportation network measures, it is clear that households prefer zones with transit availability and with smaller access times to transit stops. Surprisingly, we did not find any income-related or other demographic differences in the sensitivity to transit availability and transit stop access time.

The group of zonal demographics and housing cost variables are also important determinants of residential choice. The income dissimilarity measure, captured by the absolute difference between zonal median income and household income, confirms the income segregation phenomenon observed in earlier studies (Waddell, 1993; Bhat and Guo, 2004). Similarly, there is clustering of households based on household size. The zonal average housing value has the expected negative, and statistically significant, effect on residential choice.

Finally, the zonal ethnic composition measures, when interacted with the race of the household, reveal a clear racial segregation trend in household residential location choices.

4.2 Car Ownership Model Parameter Estimates

The results of the ordered-response car ownership model are provided in Table 2. The coefficient estimates represent the impact of the variables on the car ownership propensity of households.

The effects of the zonal size and density measures show marginally significant impacts of household density and employment density on car ownership propensity. An important reason for this is the high correlation between the density measures and the local transportation network measures (street block density, transit availability, and transit access time). In fact, when the local transportation network measures are removed, we obtain the usual negative and strongly significant effects of household and employment density on car ownership propensity (as obtained in earlier car ownership studies). Thus, our results indicate that the density measures used in earlier studies are partial proxies for local transportation network measures. Overall, the results suggest marginally significant to insignificant negative influences of household and employment density on car ownership propensity, though low income households residing in high total employment density zones are significantly less likely to own cars than other households. Further, there is a high level of sensitivity variation due to unobserved factors across households in response to total employment density, as reflected by the statistically significant standard deviation on this variable.

The commute-related variables indicate a high car ownership propensity among households whose members have a long drive commute time, and a low car ownership propensity among households whose members have a high drive commute cost, both of which are consistent with intuitive considerations.

The local transportation network variables show the highly significant mean negative influence of street block density on car ownership propensity. However, there is significant unobserved heterogeneity in the responsiveness to street block density, with about 23% of households responding to an increase in street block density by increasing car ownership. Households residing in zones with transit availability are less likely to own cars than those residing in zones without transit availability, and this effect is particularly pronounced for households with low income earnings. Also, a longer transit access time at the residence end leads to higher car ownership propensity.

Finally, there are strong demographic and housing tenure effects on car ownership propensity. Specifically, households with a high number of active and senior adults, employed individuals, income, and who live in owned dwellings, have a high car ownership propensity, while single-parent households, single-individual households, households with several physically challenged individuals, households residing in multi-family housing units, and households of non-Caucasian and non-African American races have a low car ownership propensity. These demographics are consistent with those of earlier car ownership studies (see Bhat and Pulugurta, 1998 and Holtzclaw *et al.* 2002).

4.3 Residential Self-Selection Effects

The residential sorting effect in the response to the built environment and commute variables can be due to observed demographic effects or due to unobserved correlations, as discussed earlier in Section 1.2. The results in Sections 4.1 and 4.2 indicate the presence of demographic-based residential sorting of households based on car ownership preferences. Specifically, the results show that (1) Households with senior adults stay away from high density areas and those same households have a high preference for cars (relative to households with small children and no senior adults), (2) Households with low income earnings choose to (or are constrained to) locate in neighborhoods with long drive commutes, low drive commute costs, and high employment densities, and these same households own fewer cars, and (3) Single

31

individual households have a strong preference to locate in areas with high street block density and also own fewer vehicles. Thus, failure to control for these demographic effects in a car ownership model would lead to inflated effects of BE attributes. For instance, an urban policy directed toward high employment density developments, according to the model estimated in the report, would draw a disproportionately large fraction of low income households into the area of policy implementation. These households, intrinsically, also own fewer cars. Similarly, a transportation policy to increase street block density would draw a large fraction of single individuals into the neighborhood. These single individuals are, by nature, also likely to own fewer cars. Ignoring these residential self-selection effects would then lead to the misinformed result that the low car ownership in areas with high density development or high street block density is solely due to the urban policy.

The joint model formulated in Chapter 2, in addition to recognizing observed demographic self-selection effects, also accommodates the potential presence of residential sorting effects due to unobserved household factors through the $\omega_{ql}x_{il}$ terms. These sorting effects would be manifested in statistically significant $\sigma^2_{\omega l}$ estimates. However, none of these terms turned out to be significant in our estimations. This suggests that, at least in the current empirical context, the significant impacts of the built environment and other variables on car ownership are "true" effects rather than "corrupted" effects. However, the lack of sorting effects due to unobserved household factors may also be the result of (1) Measurement errors in accessibility indices and other BE measures (that is, the measurement errors on these attributes are so large that they swamp correlations in residential choice and car ownership propensity due to common unobserved sensitivity effects to these attributes) and/or the (2) Non-inclusion of important neighborhood measures actually considered by households (even though we have made a concerted effort in this research to include a comprehensive set of neighborhood measures based on data we were able to assemble).

While the absence of unobserved residential sorting effects collapses the joint model into independent models of residential choice and car ownership, it is important to note that the joint model formulated in this report needs to be estimated before one can conclude about whether to use independent models in any particular empirical setting.

4.4 Overall Likelihood-Based Measures of Fit

The log-likelihood value at convergence of the final joint model (which collapsed to independent models corresponding to a mixed multinomial logit residential location model and a mixed ordered-response logit car ownership model because of the absence of self-selection due to unobserved factors) is -16,050. The corresponding value for the model with no allowance for unobserved variations in sensitivity to the BE and commute attributes is -16,135. The likelihood ratio test for testing the presence of unobserved variations in sensitivity is 170, which is larger than the critical chi-square value with 4 degrees of freedom at any reasonable level of significance (the 4 degrees of freedom correspond to the standard deviations on the drive commute time and street block density coefficients in the residential location model, and on the employment density and street block density coefficients in the car ownership model). Further, the log-likelihood value corresponding to equal probability for each of the 233 zonal alternatives in the residential location model and sample shares in the car ownership model (corresponding to the presence of only the threshold parameters) is -19912.0. The likelihood ratio index for testing the presence of exogenous variable effects and unobserved taste variations is 7724, which is substantially larger than the critical chi-squared value with 54 degrees of freedom at literally any level of significance.

4.5 Elasticity Effects of Exogenous Variables

The parameters on the exogenous variables in Table 3 do not directly provide the magnitude of the effects of variables on car ownership levels (we do not focus on the elasticity effects of the residential choice model since the alternatives correspond to a large number of spatial units). To examine the magnitude of variable impacts on car ownership choice, we compute the aggregate level "elasticity effects" of variables on the expected aggregate car ownership level. To do so, we first define the expected car ownership level for any household q residing in spatial unit i as:

$$E(c_{qi}) = \sum_k P_{qik} \times k, \tag{7}$$

where P_{qik} is the probability that household q in spatial unit i will choose to own k cars and is given by

$$P_{qik} = \int_{\eta_q} \left[G(\psi_k - \alpha'y_q - \delta_q'x_i) - G(\psi_{k-1} - \alpha'y_q - \delta_q'x_i) \right] dF(\eta_q). \tag{8}$$

The multidimensional integral above is taken over the multivariate normal η_q vector that corresponds to the random elements embedded in δ_q (see Sections 2.1 through 2.3). The expected aggregate car ownership level across the entire sample is then computed by summing across the expected car ownership levels of individual households.

Table 3. Elasticity of Variables in Car Ownership Model

Variables	Elasticity Effect
Zonal size and density measures (including demographic interactions)	
Household density (#households per acre x 10^{-1})	-0.0004
Employment density (#employment per acre x 10^{-1})	-0.0009
Commute-related variables (including demographic interactions)	
Drive commute time (minutes x 10^{-1})	0.0039
Drive commute cost ($ x 10^{-1})	-0.0033
Local transportation network measures (including demographic interactions)	
Street block density (number of block per square mile x 10^{-1})	-0.0027
Transit availability	-0.0461
Transit access time to stop (minutes x 10^{-1})	0.0023
Household demographic variables	
Number of active adults	0.0934
Number of senior adults	0.0928
Number of employed individuals	0.0239
Number of physically challenged individuals	-0.0818
Household income ($ x 10^{-5})	0.0022
Single parent household	-0.1004
Single individual household	-0.1012
Residing in a multi-family housing unit	-0.0630
Non-Caucasian non-African-American household	-0.0337
Own household dwelling	0.0545

With the preliminaries above, one can compute an aggregate-level "elasticity" of an ordinal exogenous variable (such as the number of active adults in the household) by increasing the value of the ordinal variable by 1 unit for each household and obtaining the relative change in expected aggregate car ownership level. Thus, the "elasticities" for the ordinal exogenous variables can be viewed as the proportional change in expected aggregate car ownership level due to an increase of 1 unit in the ordinal variable across all households.

Next, to compute an aggregate-level "elasticity" of a dummy exogenous variable (such as transit availability), we change the value of the variable to one for the subsample of observations (*i.e.*, households) for which the variable takes a value of zero and to zero for the subsample of observations for which the variable takes a value of one. We then sum the shifts in expected aggregate car ownership level in the two subsamples after reversing the sign of the shifts in the second subsample and compute an effective proportional change in expected aggregate car ownership level in the entire sample due to a change in the dummy variable from 0 to 1. The aggregate-level elasticity of a continuous exogenous variable (household and employment density, drive commute time and cost, street block density, transit access time, and household income) is computed by increasing the continuous variable by a uniform 10% across all individuals and obtaining the proportional change in the expected aggregate car ownership level.

The elasticity effects are presented in Table 3 by variable category. The table presents only the effects of the non-interaction variables from Table 2, since the effects of the interaction variables (such as employment density interacted with the low household income dummy variable) is accommodated by increasing the interaction variable appropriately whenever a component variable is increased. In general, the results in Table 3 indicate the strong effect of demographic and housing tenure variables, all of which except household income are ordinal or dummy variables. However, the transit availability dummy variable (under local transportation network measures) is also an important determinant of car ownership. Among the various continuous variables, the results show the important influence of household income, street block density, transit access time, and the commute variables (drive time and cost). As alluded to earlier, the effects of the density variables (household density and employment density) are small relative to other variables. This is because of the inclusion of the transportation network measures. Overall, the elasticity effects indicate the important, though inelastic, changes in car ownership levels due to changes in the built environment.

CHAPTER 5. CONCLUSION

This report develops a methodological framework to control for residential sorting effects (also referred to as self-selection effects) in the analysis of the effect of built environment attributes on travel behavior-related choices. The formulation is applied to examine the impact of the built environment, transportation network attributes, and demographic characteristics on residential choice and car ownership decisions. The model formulation takes the form of a joint mixed multinomial logit-ordered response structure that (a) accommodates differential sensitivity to the built environment and transportation network variables due to both demographic and unobserved household attributes and (b) controls for the self-selection of individuals into neighborhoods based on car ownership preferences. To our knowledge, the analysis in this report represents the first instance of the formulation and application of such a unified mixed multidimensional-ordered response structure in the econometric literature. The framework can be used to control for residential self-selection for any kind of travel behavior variable and directly provides the correct standard errors regarding the effect of the built environment attributes. It is geared toward cross-sectional analysis, recognizing that almost all existing data sources available for analysis of BE effects are cross-sectional in nature. Unlike earlier studies, the methodology also explicitly considers and models the residential location choice decision jointly with the travel behavior choice of interest.

The empirical analysis in the report is based on the residential choice and car ownership decisions of San Francisco Bay area residents. The data is drawn from several sources, including the (a) 2000 San Francisco Bay area travel survey, (b) land-use/demographic, network level-of-service, and GIS-based bicycle facility data files for the Bay area obtained from the Metropolitan Transportation Commission (MTC), and (c) the US census data, the US 2000 Tiger files, and the Public Use Microdata Sample (PUMS) data. A whole range of zonal size and density measures, land-use structure variables, regional accessibility indices, commute-related characteristics, local transportation network measures, zonal demographic and housing variables, zonal ethnic composition characteristics, household demographics, and interactions of these variables are considered in the analysis.

There are several important findings from our study. First, BE attributes do affect residential choice decisions as well as car ownership decisions. Thus, policy decisions regarding changes in BE characteristics have to be evaluated in the joint context of both decisions, so that spatial relocation patterns as well as car ownership changes can be analyzed. Such a complete picture enables a comprehensive assessment of potential travel-related changes due to BE policies. Second, our findings support the notion that the commonly used population and/or employment density measures are actually proxy variables for such BE measures as street block density and transit accessibility. Third, in the context of car ownership decisions, both household demographics and BE characteristics are influential. However, household demographics have a more dominant effect. Fourth, there is variation in sensitivity to BE attributes due to both demographic and unobserved factors, in both residential choice as well as car ownership decisions. But, while the study examined a suite of demographic interactions and allowed random variations in sensitivity to several BE characteristics, most of these did not turn out to be statistically significant. Among demographics, income is a key variable in affecting the sensitivity to BE attributes and related variables. Unobserved household-specific factors also play an important role in the sensitivity to commute time and street block density (in the residential choice model) and employment density and street block density (in the car ownership model). Ignoring such systematic and random variations in sensitivity to BE attributes will, in general, lead to inconsistent results regarding the effect of BE attributes on travel behavior decisions, which can, in turn, lead to inappropriate policy decisions. Fifth, household income is the dominant factor in residential sorting. Specifically, low income households consciously choose to (or are constrained to) locate in neighborhoods with low commute costs, long commute times, and high employment density compared to their high income counterparts. Such low income households also intrinsically choose to own fewer cars. Thus, ignoring income effects in car ownership (and by extension, other travel decisions) can lead to an inflated effect of the built environment and related variables on travel behavior decisions. Other demographic factors that impact residential sorting based on car ownership preferences correspond to the presence of senior adults in the household and whether or not a person lives alone. Finally, and rather surprisingly, our results did not support the notion of residential sorting in car ownership propensity based on unobserved household factors. This result implies that independent models of residential choice and car ownership choice (after accommodating the residential sorting

effects of demographics) are adequate to examine BE effects on car ownership choice, in the current empirical context (see also the important caveats related to this issue in Section 4.3). But, in general, it is important to consider the methodology developed in this report to control for the potential presence of self selection due to both observed and unobserved household factors.

To summarize, the model in the report can be used to assess the impacts of changing demographics, built environment characteristics, and transportation network attributes for land-use planning and transportation public policy analysis. The study, to our knowledge, represents the first formulation and application of a comprehensive modeling framework to consider residential self-selection effects, as well as observed and unobserved variations in sensitivity to the built environment, in a joint model of residential location choice and car ownership decisions.

REFERENCES

Audirac, I. and Shermyen, A. (1994). An Evaluation of Neotraditional Design's Social Prescription: Postmodern Placebo or Remedy for Suburban Malaise? Journal of Planning Education and Research, 13, 1161-173.

Badoe, D.A. and Miller, E. J. (2000). Transportation-Land-Use Interaction: Empirical Findings in North America, and Their Implications for Modeling. Transportation Research Part D, 5(4), 235-263.

Bagley, M. N. and Mokhtarian, P. L. (2002). The Impact of Residential Neighborhood Type on Travel Behavior: A Structural Equations Modeling Approach. Annals of Regional Science, 36(2), 279-297.

Ben-Akiva, M., and Lerman, S. R. (1985). Discrete Choice Analysis, MIT Press, Cambridge, MA.

Berry, S., Levinsohn, J. and Pakes, A. (1995). Automobile Prices in Market Equilibrium. Econometrica, 63, 841-889.

Bhat, C. R. (1998). Accommodating Variations in Responsiveness to Level-of-Service Variables in Travel Mode Choice Modeling. Transportation Research Part A, 32(7), 495-507.

Bhat, C. R. (2003). Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences. Transportation Research Part B, 37(9), 837-855.

Bhat, C. R. and Castelar, S. (2002). A Unified Mixed Logit Framework for Modeling Revealed and Stated Preferences: Formulation and Application to Congestion Pricing Analysis in the San Francisco Bay Area. Transportation Research Part B, 36(7), 593-616.

Bhat, C. R., and Gossen, R. (2004). A Mixed Multinomial Logit Model Analysis of Weekend Recreational Episode Type Choice. Transportation Research Part B, 38(9), 767-787.

Bhat, C. R. and Guo, J.Y. (2004). A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modeling. Transportation Research Part B, 38(2), 147-168.

Bhat, C. R. and Lockwood, A. (2004). On Distinguishing Between Physically Active and Physically Passive Episodes and Between Travel and Activity Episodes: An Analysis of Weekend Recreational Participation in the San Francisco Bay Area. Transportation Research Part A, 38(8), 573-592.

Bhat, C. R. and Pulugurta, V. (1998). A Comparison of Two Alternative Behavioral Mechanisms for Car Ownership Decisions. Transportation Research Part B, 32(1), 61-75.

Bhat, C. R. and Singh, S.K. (2000). A Comprehensive Daily Activity-Travel Generation Model System for Workers. Transportation Research Part A, 34(1), 1-22.

Bhat, C. R. and Srinivasan, S. (2005). A Multidimensional Mixed Ordered-Response Model for Analyzing Weekend Activity Participation. Transportation Research Part B, 39(3), 255-278.

Bhat, C. R. and Zhao, H. (2002). The Spatial Analysis of Activity Stop Generation. Transportation Research Part B, 36(6), 557-575.

Bhat, C. R., Govindarajan, A., and Pulugurta, V. (1998). Disaggregate Attraction-End Choice Modeling. Transportation Research Record, 1645, 60-68.

Bhat, C. R., Srinivasan, S., and Axhausen, K.W. (2005). An Analysis of Multiple Interepisode Durations Using a Unifying Multivariate Hazard Model. Transportation Research Part B, 39(9), 797-823.

Boarnet, M. G. and Crane, R. (2001). The Influence of Land Use on Travel Behavior: Specification and Estimation Strategies. Transportation Research Part A, 35, 823-845.

Boarnet , M. G. and Sarmiento, S. (1998). Can Land-use Policy Really Affect Travel Behavior? A Study of the Link Between Non-work Travel and Land-Use Characteristics. Urban Studies, 35(7), 1155-1169.

Cervero, R. (2002). Built Environments and Mode Choice: Toward a Normative Framework. Transportation Research Part D, 7(4), 265-284.

Cervero, R. and Duncan, M. (2003). Walking, Bicycling, and Urban Landscapes: Evidence from the San Francisco Bay Area. American Journal Of Public Health, 93(9), 1478-1483.

Chamberlain, G. (1980). Analysis of Covariance with Qualitative Data. Review of Economic Studies, 47, 225-238.

Crane, R. (2000). The Influence of Urban Form on Travel: An Interpretive Review. Journal of Planning Literature, 15(1), 3-23.

Daly, A. (1982). Estimating Choice Models Containing Attraction Variables. Transportation Research Part B, 16, 5-15.

Dunphy, R. T. and Fisher, K. (1996). Transportation, Congestion, and Density: New Insights. Transportation Research Record, 1552, 89-96.

Ewing, R. (2005). Building Environment to Promote Health. Journal of Epidemiology and Community Health, 59(7), 536-537.

Ewing, R. and Cervero, R. (2001). Travel and the Built Environment – Synthesis. Transportation Research Record, 1780, 87-114.

Fotheringham, A. S. (1983). Some Theoretical Aspects of Destination Choice and Their Relevance to Production-Constrained Gravity Models. Environment and Planning, 15A, 1121-1132.

Golledge, R. and Garling, T. (2003). Cognitive Maps and Urban Travel. Research Paper No. 601, University of California Transportation Center.

Gordon, P. and Richardson, H. W. (1997). Are Compact Cities a Desirable Planning Goal? Journal of the American Planning Association, 63(1), 95.

Greenwald, M. J. and Boarnet, M. G. (2001). Built Environment as Determinant of Walking Behavior - Analyzing Nonwork Pedestrian Travel in Portland, Oregon. Transportation Research Record, 1780, 33-42.

Guiliano, G. (1995). The Weakening Transportation-Land Use Connection. Access, 6, 3-11.

Guo, J. Y. and Bhat, C. R. (2004). Modifiable Areal Units: Problem or Perception in Modeling of Residential Location Choice? Transportation Research Record, 1898, 138-147.

Guo, J. Y. and Bhat, C. R. (2006). Operationalizing the Concept of Neighborhood: Application to Residential Location Choice Analysis. Forthcoming, Journal of Transport Geography.

Handy, S. L. (1996). Methodologies for Exploring the Link Between Urban Form and Travel Behavior. Transportation Research Part D, 1(2), 151-165.

Handy, S. L. and Clifton, K. J. (2001). Local Shopping as a Strategy for Reducing Automobile Travel. Transportation, 28(4), 317-346.

Handy, S., Cao, X., Buehler, T. J., and Mokhtarian, P. L. (2005). The Link Between the Built Environment and Travel Behavior: Correlation or Causality? Presented at the 84th Annual Meeting of the Transportation Research Board, Washington, D.C.

Hess, D. and Ong, P. (2001). Traditional Neighborhoods and Auto Ownership. Working Paper, Lewis Center for Public Policy Studies, UCLA, July. (www.sppsr.ucla.edu/lewis/WorkingPapers.html)

Holtzclaw, J., Clear, R., Dittmar, H., Goldstein, D., and Haas, P. (2002). Location Efficiency: Neighborhood and Socio-Economic Characteristics Determine Auto Ownership and Use? Transportation Planning and Technology, 25, 1-27. (www.tandf.co.uk/journals/online/0308-1060.html)

Ihlanfeldt, K. R. and Sjoquist, D. L. (1998). The Spatial Mismatch Hypothesis: A Review of Recent Studies and Their Implications for Welfare Reform. Housing Policy Debate, 9(4), 849-892.

Khattak, A. J. and Rodriguez, D. (2005). Travel Behavior in Neo-traditional Neighborhood Developments: A Case Study in USA. Transportation Research Part A, 39(6), 481-500.

Kitamura, R., Mokhtarian, P. L., and Laidet, L. (1997). A Micro-Analysis of Land Use and Travel in Five Neighborhoods in the San Francisco Bay Area. Transportation, 24, 125-158.

Krizek, K. J. (2000). Pretest-Posttest Strategy for Researching Neighborhood-Scale Urban Form and Travel Behavior. Transportation Research Record, 1722, 48-55.

Krizek, K. J. (2003). Residential Relocation and Changes in Urban Travel - Does Neighborhood-Scale Urban Form Matter? Journal Of The American Planning Association, 69(3), 265-281.

Lewbel, A. (2004). Simple Estimators For Hard Problems: Endogeneity in Discrete Choice Related Models. Working Papers in Economics No. 604, Department of Economics, Boston College.

Louviere, J., Train, K., Ben-Akiva, M., Bhat, C.R., Brownstone, D., Cameron, T.A., Carson, R.T., DeShazo, J.R., Fiebig, D., Greene, W., Hensher, D., and Waldman, D. (2005). Recent Progress on Endogeneity in Choice Modeling. Marketing Letters, Special Issue: Sixth Invitational Choice Symposium, 16(3-4), 255-265.

Lund, H. (2003). Testing the Claims of New Urbanism. APA Journal, 69(4), 414-429.

Messenger, T. and Ewing. R. (1996). Transit-Oriented Development in the Sunbelt. Transportation Research Record, 1552, 145-152.

MORPACE International, Inc. (2002). Bay Area Travel Survey Final Report, March. ftp://ftp.abag.ca.gov/pub/mtc/planning/BATS/BATS2000/

Nerella, S. and Bhat, C. R. (2004). Numerical Analysis of Effect of Sampling of Alternatives in Discrete Choice Models. Transportation Research Record, 1894, 11-19.

Pickrell, D. (1999). Cars and Clean Air: A Reappraisal. Transportation Research Part A, 33(7), 527-547.

Pozsgay, M. A. and Bhat, C. R. (2002). Destination Choice Modeling for Home-Based Recreational Trips: Analysis and Implications for Land-Use, Transportation, and Air Quality Planning. Transportation Research Record, 1777, 47-54.

Pucher, J., and Renne, J. (2003). Socioeconomics of Urban Travel: Evidence from the 2001 NHTS. Transportation Quarterly, 57(3), 49-78.

Rajamani, J., Bhat, C. R., and Handy, S. L. (2003). Assessing Impact of Urban Form Measures on Nonwork Trip Mode Choice After Controlling for Demographic and Level-of-Service Effects. Transportation Research Record, 1831, 158-165.

Rodriguez, D. A., Khattak, A. J., and Evenson, K. R. (2005). Can Neighborhood Design Encourage Walking and Bicycling? – Physical Activity on a New Urbanist and a Conventional Suburban Community. Presented at the 84th Annual Meeting of the Transportation Research Board, Washington D.C.

Saelens, B. E., Sallis, J. F., and Frank, L. D. (2003). Environmental Correlates of Walking and Cycling: Findings from the Transportation, Urban Design, and Planning Literatures. Annals of Behavioral Medicine, 25, 80-91.

Schwanen, T. and Mokhtarian, P. L. (2005). What Affects Commute Mode Choice: Neighborhood Physical Structure or Preferences Toward Neighborhoods? Journal of Transport Geography, 13(1), 83-99.

Schwanen, T. and Mokhtarian, P. L. (2004). The Extent and Determinants of Dissonance Between Actual and Preferred Residential Neighborhood Type. Environment And Planning B, 31(5), 759-784.

Shay, E., and Khattak, A. J. (2005). Auto Ownership and Use in Neo-Traditional and Conventional Neighborhoods. Presented at the 84th Annual Meeting of the Transportation Research Board, Washington, D.C.

Sivakumar, A., Bhat, C. R., and Ökten, G. (2005). Simulation Estimation of Mixed Discrete Choice Models with the Use of Randomized Quasi-Monte Carlo Sequences: A Comparative Study. Transportation Research Record, 1921, 112-122.

Transportation Research Board (2003). TCRP Report 95: Traveler Response to Transportation System Changes, Chapter 15 - Land Use and Site Design, Washington, D.C.

Transportation Research Board (2005). Smart Growth and Transportation: Issues and Lessons Learned. Conference Proceedings 32, Washington, D.C.

Waddell, P. (1993). Exogenous Workplace Choice in Residential Location Models: Is the Assumption Valid? Geographical Analysis, 25, 65-82.