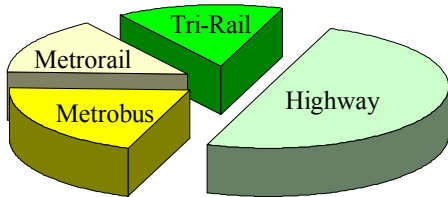


FSUTMS Mode Choice Modeling: Factors Affecting Transit Use and Access

Final Report



1. Report No. NCTR 392-07, 416-03		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle FSUTMS Mode Choice Modeling: Factors Affecting Transit Use and Access				5. Report Date July 2002	
				6. Performing Organization Code	
7. Author(s) Fang Zhao, Min-Tang Li, Lee-Fang Chow, Albert Gan, L. David Shen				8. Performing Organization Report No.	
9. Performing Organization Name and Address National Center For Transit Research (NCTR) University of South Florida CUT 100 4202 East Fowler Avenue, Tampa, FL 33620				10. Work Unit No.	
				11. Contract or Grant No. DTRS98-G-0032	
12. Sponsoring Agency Name and Address Office of Research and Special Programs Florida DOT U.S. Department of Transportation 605 Suwannee Street Washington, DC 20590 Tallahassee, Florida 32399				13. Type of Report and Period Covered Final Report	
				14. Sponsoring Agency Code	
15. Supplementary Notes Supported by a Grant from the USDOT Research and Special Programs Administration, and the Florida Department of Transportation; conducted at the Lehman Center for Transportation Research (LCTR) at the Florida International University (FIU), Miami. Mr. Ike Ubaka was the project manager.					
16. Abstract This study utilized the transit onboard survey data from the 2000 Southeast Florida Travel Characteristic Study to investigate the needs to incorporate additional variables other than travel time and costs into the modal split model of the Florida Standard Urban Transportation Model Structure (FSUTMS) to improve its transit forecasting capability. Factors including land use, demographics, the socioeconomic characteristics of the population, and transit supply were studied using regression analysis. GIS was used extensively for both data compilation and spatial analysis. The study identified a number of transit supply variables as well as population, employment density, and land use mix as strong indicators of transit use. Analysis of transit walk accessibility indicated that transit use beyond one half mile was minimum. A methodology was developed for estimating the percentage of transit service population in a zone, which proved to be a better indicator of transit use than the commonly used buffer method. Regression models that may be used for estimating and forecasting, for a future model year, the percentage of population and workers served by transit were also developed.					
17. Key Words Public Transit, Modal Split, Transit Accessibility, Transit Supply, Land Use, Demographics, Socioeconomic.			18. Distribution Statement Available to the public through the National Technical Information Service (NTIS), 5285 Port Royal Road, Springfield, VA 22161, (703) 487-4650, http://www.ntis.gov/ , and through the NCTR web site at http://www.nctr.usf.edu/ .		
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of pages 169	
				22. Price	

**FSUTMS MODE CHOICE MODELING:
FACTORS AFFECTING TRANSIT USE AND ACCESS**

Final Report

Prepared for

National Center for Transit Research (NCTR)
University of South Florida
Tampa, Florida 33620

in Cooperation with

Florida Department of Transportation
U.S. Department of Transportation
Contract No.: BC137
Subcontract No.: 2117-392-LO-A

Prepared by

Fang Zhao, Min-Tang Li, Lee-Fang Chow, Albert Gan, and L. David Shen
Lehman Center for Transportation Research
Florida International University
10555 W Flagler Street
Miami, Florida 33174
Phone: (305) 348-3821
Fax: (305) 348-2802
E-mail: fang@eng.fiu.edu

The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the State of Florida Department of Transportation or the U.S. Department of Transportation.

July 2002

ACKNOWLEDGEMENTS

We would like to thank Mr. Bob McCullough, Mr. Wilson Hernandez, and Mr. Shi-Chiang Li, who helped initiate, conceptualize, and shape the study. We are also grateful to Mr. Shi-Chiang Li, Mr. Frank Baron, and Mr. Mike Neidhart for their careful review of the final report and for their helpful comments. Dr. Jill Strube was the technical editor of the report while she was a Ph.D. candidate in the College of Health and Urban Affairs, Florida International University. Special thanks go to Mr. Ike Ubaka, the project manager, for his guidance, support, and help throughout the study.

This project was funded by the Research Center of the Florida Department of Transportation through the National Center for Transit Research (NCTR). Mr. Dennis Hinebaugh's assistance in managing the project at NCTR is also acknowledged.

TABLE OF CONTENTS

TABLE OF CONTENTS.....	i
LIST OF TABLES.....	iii
LIST OF FIGURES	v
EXECUTIVE SUMMARY	1
1. INTRODUCTION	10
2. LITERATURE REVIEW	13
2.1 Transit Level of Service.....	13
2.2 Accessibility to Opportunities and Pedestrian Accessibility	15
2.3 Transit Ridership Forecasting Models	18
2.3.1 State-of-the-Art General Purpose Travel Demand Models	18
2.3.2 Models Designed for Transit Ridership Forecast	19
2.3.3 Dynamic Structural Model.....	30
2.4 Mode Choice and Urban Forms	30
2.5 Neotraditional Neighborhood and Urban Design	37
2.6 Summary	39
3. STUDY AREA SELECTION.....	42
3.1 Selection of Geographic Analysis Units	42
3.2 Household Travel Characteristics Survey.....	43
3.2.1 Sample Locations at Municipality Level	43
3.2.2 Sample Locations at Neighborhood Level.....	44
3.2.3 Summary	44
3.3 Transit Onboard Survey.....	45
3.4 Existing Transit Services in Miami-Dade County	50
3.5 Study Area Selection.....	53
4. DATA COLLECTION AND PROCESSING	55
4.1 Introduction.....	55
4.2 Transit Level of Service.....	55
4.3 Land Use	59
4.4 Accessibility.....	64
4.5 Transit Users' Socioeconomic Characteristics	69
4.6 Demographics	70
4.7 Data for Broward County.....	71
4.8 Data for Palm Beach County	73
5. TRANSIT ACCESSIBILITY	76
5.1 Introduction.....	76
5.2 Transit Walk Accessibility.....	78
5.2.1 Estimation of Population Accessibility to Transit	78
5.2.2 Estimation of Workers' Accessibility to Transit	85
5.2.3 Estimation of Percentages of Trips with Transit Accessibility.....	87
5.3 Forecast of Transit Accessibility	91
5.3.1 Forecast of Transit Accessibility for Production Trips.....	92
5.3.2 Forecast of Transit Accessibility for Attraction Trips	98
5.3.3 Auto Accessibility.....	100

6.	REGRESSION ANALYSIS OF FACTORS CONTRIBUTING TO TRANSIT USE ..	107
6.1	Multiple Linear Regression Analysis Procedure	107
6.2	Regression Results for Miami-Dade County	108
6.3	Regression Results for Broward and Palm Beach Counties	118
6.4	Regressions with Demand Variables	122
	6.4.1 Data Compiled Using GIS Land Use Method	122
	6.4.2 Data Compiled Using Property Method	127
	6.4.3 Data Compiled Using Buffer Method.....	132
	6.4.4 Data Compiled Using Tract Method.....	135
6.5	Summary	149
7.	SUMMARY AND CONCLUSIONS	152
8.	RECOMMENDATIONS	155
	REFERENCES	156

LIST OF TABLES

Table 2-1.	Summary of Accessibility Measures	16
Table 2-2.	Elasticity of Model Variables	29
Table 3-1.	Dwelling Units (DUs) with Bus Trips as a Percent of Total Samples in Miami-Dade Municipalities	44
Table 3-2.	Dwelling Units (DUs) with Bus Trips as a Percent of Total Samples in Miami-Dade Neighborhoods	45
Table 3-3.	Geocoded Results for Transit Trips Ends	46
Table 3-4.	Number of Transit Trips Interchanged between Tri-County Areas.....	46
Table 3-5.	Number of Samples for Transit On-Board in Municipalities of Miami-Dade.....	47
Table 3-6.	Number of Trips with Origins in Neighborhoods of Miami-Dade	48
Table 3-7.	Number of Trips with Destinations in Neighborhoods of Miami-Dade	49
Table 3-8.	Percentages of Areas Overlapped by Transit Route Buffers	52
Table 3-9.	Neighborhoods Selected for Study Areas	53
Table 4-1.	Transit LOS Variables	57
Table 4-2.	Land Use Variables.....	60
Table 4-3.	Tract Level Highway Accessibility Variables	66
Table 4-4.	Tract Level Transit Accessibility Variables	67
Table 4-5.	Tract Level Transit-Highway Accessibility Differences	68
Table 4-6.	Socioeconomic Variables.....	69
Table 4-7.	Demographic Variables	70
Table 4-8.	Transit LOS Variables for Broward County	71
Table 4-9.	Land Use Variables for Broward County	71
Table 4-10.	Tract Level Highway Accessibility Variables for Broward County.....	71
Table 4-11.	Tract Level Transit Accessibility for Broward County	72
Table 4-12.	Tract Level Transit-Highway Accessibility Differences for Broward County.....	72
Table 4-13.	Socioeconomic Variables for Broward County	73
Table 4-14.	Demographic Variables for Broward County	73
Table 4-15.	Land Use Variables for Palm Beach County	73
Table 4-16.	Tract Level Highway Accessibility Variables for Palm Beach County	74
Table 4-17.	Tract Level Transit Accessibility Variables for Palm Beach County	74
Table 4-18.	Palm Beach Tract Level Transit-Highway Accessibility Differences.....	74
Table 4-19.	Socioeconomic Variables for Palm Beach County	74
Table 4-20.	Demographic Variables for Palm Beach County.....	75
Table 5-1.	Comparison of Transit Service Population (1/4-Mile Distance)	85
Table 5-2.	Conversion between Dade County and FSUTMS Land Use Categories.....	86
Table 5-3.	Cumulative Percentages of Transit Trips by Walking Distance.....	86
Table 5-4.	Comparison of Percentage of Population and Percentage of Trips Served by Transit within a 0.5-Mile Distance	91
Table 5-5.	Regression Model For Predicting Percentage of Access trips to BART Stations by Walking, All Trip Purposes, 1992 (Cervero 2001).....	92
Table 5-6.	Comparison of Five Accessibility Models.....	97
Table 5-7.	Recommended Values for Intersection Density for Different Types of Street Structures	98
Table 5-8.	Comparison of Four Models for HBW Trips.....	100

Table 5-9.	Home-Based Transit Linked Trips.....	103
Table 6-1.	Transit Production Trip Models for Miami-Dade (GIS Land Use Method)*	109
Table 6-2.	Transit Attraction Trip Models for Miami-Dade (GIS Land Use Method)*	110
Table 6-3.	Transit Production Trip Models for Miami-Dade (Property Method)*	111
Table 6-4.	Transit Attraction Trip Models for Miami-Dade (Property Method)*	112
Table 6-5.	Transit Production Trip Models for Miami-Dade (Buffer Method)*	113
Table 6-6.	Transit Attraction Trip Models for Miami-Dade (Buffer Method)*	113
Table 6-7.	Transit Production Trip Models for Miami-Dade (Tract Method)*	114
Table 6-8.	Transit Attraction Trip Models for Miami-Dade (Tract Method)*	115
Table 6-9.	Adjusted R^2 's for the <i>DECAY_POP_</i> and <i>ServiceArea%</i> Models.....	117
Table 6-10.	Transit Production Trip Models for Broward (Buffer Method)*.....	118
Table 6-11.	Transit Attraction Trip Models for Broward (Buffer Method)*	119
Table 6-12.	Transit Production Trip Models for Broward (Tract Method)*	119
Table 6-13.	Transit Attraction Trip Models for Broward (Buffer Method)*	120
Table 6-14.	Transit Production Trip Models for Palm Beach (Tract Method)*	121
Table 6-15.	Transit Attraction Trip Model for Palm Beach (Tract Method)*	122
Table 6-16.	Models for Miami-Dade (GIS Land Use Method, Low Accessibility)	123
Table 6-17.	Models for Miami-Dade (GIS Land Use Method, Medium Accessibility)	123
Table 6-18.	Models for Miami-Dade (GIS Land Use Method, High Accessibility).....	123
Table 6-19.	Models for Miami-Dade (Property Method, Low Accessibility)	127
Table 6-20.	Models for Miami-Dade (Property Method, Medium Accessibility)	128
Table 6-21.	Models for Miami-Dade (Property Method, High Accessibility).....	128
Table 6-22.	Models for Miami-Dade (Buffer Method, Low Accessibility).....	132
Table 6-23.	Models for Miami-Dade (Buffer Method, Medium Accessibility)	133
Table 6-24.	Models for Miami-Dade (Buffer Method, High Accessibility)	133
Table 6-25.	Models for Miami-Dade (Tract Method, Low Accessibility).....	135
Table 6-26.	Models for Miami-Dade (Tract Method, Medium Accessibility)	136
Table 6-27.	Models for Miami-Dade (Tract Method, High Accessibility)	137
Table 6-28.	Variables for Intercept and No-Intercept Models	150

LIST OF FIGURES

Figure 3-1.	0.25- and 0.5-Mile Buffer Zones along Miami-Dade Bus Routes.....	51
Figure 3-2.	Neighborhoods Included in the Choice Set for Study Area Selection.....	52
Figure 3-3.	Neighborhoods Selected for Study Areas.....	54
Figure 4-1.	Sample of the Bus Stop Information Obtained from MDTA	56
Figure 4-2.	Example of the Headway Database (Partial)	57
Figure 4-3.	Cells Used for Calculation of Mean Entropy.....	62
Figure 5-1.	Streets That Do Not Allow Access to Transit from Properties.....	79
Figure 5-2.	Transit Catchment Area, Streets, and Property Distribution	80
Figure 5-3.	Frequency Distribution of Transit Trips versus Walking Distance	82
Figure 5-4.	Estimation of the Decay Function.....	83
Figure 5-5.	Transit Catchment Area, 0.25-mile Buffer, Streets, and Property Distribution ...	84
Figure 5-6.	Transit Buffer and Distribution of Nonresidential Land Use	87
Figure 5-7.	Value Ranges for INT_DEN and Corresponding Mostly Grid Street Structures.	94
Figure 5-8.	Ranges of Value of INT_DEN and Corresponding Mostly Curvilinear Street Structures	94
Figure 5-9.	Street Intersection Density Distribution by TAZ in Study Area.....	95
Figure 5-10.	Metrorail Alignment, Bus 95 Routes, an Park-n-Ride Sites.....	102
Figure 5-11.	Transit Tours.....	103
Figure 5-12.	Distribution of Transit Tours by Auto Access Distance.....	104
Figure 5-13.	Comparison of Auto Access Distance and Linked Transit Trip Distance	105
Figure 5-14.	Distribution of Linked Transit Trips by Length (miles)	106
Figure 6-1.	Distance decayed Transit Service Population and Percentage of Service Area in a Tract	117
Figure 6-2.	Through Origin Relationships between <i>SFDUDensity1</i> and Transit Production by Transit Accessibility Level (GIS Land Use Method)	124
Figure 6-3.	Non-Zero Intercept Relationships between <i>SFDUDensity1</i> and Transit Production by Transit Accessibility Level (GIS Land Use Method)	125
Figure 6-4.	Intercept and No-Intercept <i>SFDUDensity1</i> Models for Transit Productions (GIS Land Use Method)	125
Figure 6-5.	Through Origin Relationships between <i>TotalDensity1</i> and Transit Production by Transit Accessibility Level (GIS Land Use Method)	126
Figure 6-6.	Non-Zero Intercept Relationships between <i>TotalDensity1</i> and Transit Production by Transit Accessibility Level (GIS Land Use Method)	126
Figure 6-7.	Intercept and No-Intercept <i>TotalDensity1</i> Models for Transit Productions (GIS Land Use Method)	127
Figure 6-8.	Through Origin Relationships between <i>AvgEntropy</i> and Transit Production by Transit Accessibility Level (Property Method)	129
Figure 6-9.	Non-Zero Intercept Relationships between <i>AvgEntropy</i> and Transit Production by Transit Accessibility Level (Property Method)	130
Figure 6-10.	Intercept and No-Intercept <i>AvgEntropy</i> Models for Transit Production (Property Method).....	130
Figure 6-11.	Through Origin Relationships between <i>TotalDensity</i> and Transit Production by Transit Accessibility Level (Property Method)	131

Figure 6-12.	Non-Zero Intercept Relationships between <i>TotalDensity</i> and Transit Production by Transit Accessibility Level (Property Method)	131
Figure 6-13.	Intercept and No-Intercept <i>TotalDensity</i> Models for Transit Production (Property Method).....	132
Figure 6-14.	Through Origin Relationships between <i>AvgEntropy3</i> and Transit Production by Transit Accessibility Level (Buffer Method).....	134
Figure 6-15.	Non-Zero Intercept Relationships between <i>AvgEntropy3</i> and Transit Production by Transit Accessibility Level (Buffer Method).....	134
Figure 6-16.	Intercept and No-Intercept <i>AvgEntropy3</i> Models for Transit Productions (Buffer Method).....	135
Figure 6-17.	Through Origin Relationships between <i>%HH0Child</i> and Transit Production by Transit Accessibility Level (Tract Method).....	138
Figure 6-18.	Non-Zero Relationships between <i>%HH0Child</i> and Transit Production by Transit Accessibility Level (Tract Method).....	138
Figure 6-19.	Intercept and No-Intercept <i>%HH0Child</i> Models for Transit Production (Tract Method).....	139
Figure 6-20.	Through Origin Relationships between <i>AvgWrkrChild</i> and Transit Production by Transit Accessibility Level (Tract Method).....	140
Figure 6-21.	Non-Zero Intercept Relationships between <i>AvgWrkrChild</i> and Transit Production by Transit Accessibility Level (Tract Method).....	140
Figure 6-22.	Intercept and Non-Intercept <i>AvgWrkrChild</i> Models for Transit Productions (Tract Method).....	141
Figure 6-23.	Through Origin Relationships between <i>AvgWrkr0Child</i> and Transit Production by Transit Accessibility Level (Tract Method).....	142
Figure 6-24.	Non-Zero Intercept Relationships between <i>AvgWrkr0Child</i> and Transit Production by Transit Accessibility Level (Tract Method)	142
Figure 6-25.	Intercept and Non-Intercept <i>AvgWrkr0Child</i> Models for Transit Productions (Tract Method).....	143
Figure 6-26.	Through Origin Relationships between <i>AvgPChild</i> and Transit Production by Transit Accessibility Level (Tract Method).....	144
Figure 6-27.	Non-Zero Intercept Relationships between <i>AvgPChild</i> and Transit Production by Transit Accessibility Level (Tract Method).....	144
Figure 6-28.	Intercept and No-Intercept <i>AvgPChild</i> Models for Transit Productions	145
Figure 6-29.	Through Origin Relationships between <i>AvgP0Child</i> and Transit Production by Transit Accessibility Level (Tract Method).....	146
Figure 6-30.	Non-Zero Intercept Relationships between <i>AvgP0Child</i> and Transit Production by Transit Accessibility Level (Tract Method).....	146
Figure 6-31.	Intercept and No-Intercept <i>AvgP0Child</i> Models for Transit Productions	147

EXECUTIVE SUMMARY

Introduction

In the Florida Standard Urban Transportation Model Structure (FSUTMS), the mode choice utility functions typically include variables like travel times by automobile and by transit, terminal times, parking costs, and transit fares. Other factors may also influence the decision of transit use than travel costs alone. It has been widely recognized that socioeconomic characteristics of the population have an impact on the decision to use transit. Such characteristics, however, are generally not considered in modal split, as is true of FSUTMS. Other important factors such as land use may be missing from the models, preventing them from producing valid and reliable transit ridership forecasts.

The purpose of this study was to improve the estimates of transit accessibility and to identify and analyze factors that have a significant effect on transit ridership. The goal was to recommend a set of variables that may be incorporated into the FSUTMS modal split procedure to improve its transit forecasting capability. The specific objectives were:

- (1) To obtain a good understanding of the current state-of-the-art and state-of-the-practice in transit ridership forecasting;
- (2) To identify available GIS data resources for improving the accuracy of analysis, and determine appropriate geographic units for data analysis and for modeling;
- (3) To develop a standard procedure that can be used to determine transit accessibility by pedestrians and automobiles;
- (4) To identify additional variables that further capture the underlying causes of transit use; and
- (5) To develop practical recommendations with regards to incorporating improvements into FSUTMS.

This report presents the research methodologies, findings and conclusions, including factors that affect transit use and accessibility, and recommendations for improving the FSUTMS mode choice model.

Literature Review

A literature review of publications related to transit ridership models, transit accessibility evaluation, mode choice and urban forms, transit use, and urban design was conducted. The review indicated that the most common analysis approaches found were multiple regression and time-series analysis. While structural models are capable of modeling interrelated variables and are appropriate if the calibration of a model is necessary to predict the modal split, few were employed. Nevertheless, the objective of this research was to determine the relevance of different variables and their relative importance in transit use, which should be adequately served with regression methods. The weaknesses of the multiple regression approach, such as autocorrelation, non-normal distributions and invalidity of the independent, identically distributed (IID) normal assumption, were discussed. Although flawed in theory, multiple regression analysis is still widely applied since its concept is easy to understand. Thus, from a practitioner's point of view, this approach may be worth further investigation.

The factors that affect transit use can generally be classified into the following categories:

- (1) Transit level of service (LOS);
- (2) Accessibility;
- (3) Land use/urban design; and
- (4) Transit users' socioeconomic/demographic characteristics.

Transit service quality may be considered in terms of rider comfort, cleanliness and appearance, safety and security, pedestrian environment, amenities, headway, hours of service, parking spaces, reliability, service coverage, transfer, cost, etc. Some of these are traditional LOS variables, such as transit service frequency, route coverage, and fares, others have been recently included as transit LOS factors (Kittleson & Associates, Inc. 1999a, 1999b). The literature generally supports the ability of transit systems with high-quality services to attract more users. However, many of the LOS factors affecting transit use often cannot be easily quantified. For example, most of the studies point out the conceptual importance of the LOS measures of effectiveness (MOEs) without further suggesting the appropriate approaches to quantify many if not all of them. Some examples of such MOEs include passenger comfort in vehicle and in stations, security, and pedestrian environment. Consequently, it is still difficult to formulate LOS variables in models for estimating transit share. As a result, statistics such as Person Per Minute Served were developed and implemented to determine the quality of transit service. Although tremendous efforts have been devoted to exploring the LOS factors that significantly affect transit use, contradictory findings were cited by different researchers and practitioners in different study areas (e.g., Ulberg (1982) and Vaziri *et al.* (1990) did not find service hours to be significant in contrast to other researchers) indicating that some of the transit LOS MOEs may not be transferable from one urban area to the other.

Similar to the LOS MOEs, some accessibility variables were also demonstrated to significantly affect transit use by the research studies such as Metro (2000), Parsons Brinckerhoff (2000), Sun *et al.* (1998), and others. Numerous models are available for measuring accessibility, e.g., the accessibility index proposed by Richardson and Young (1982) and Kockelman (1997). Lately, with the advancements in geographic information system (GIS) techniques and more readily available data sources, more rigorous and detailed analyses (such as considering natural and manmade barriers that prevent access from residential areas to public transit stops), have been developed to better quantify accessibility. Nevertheless, the procedures that measure accessibility in a GIS environment are not yet simple menu-driven processes, requiring that someone trained in GIS acquire the necessary information and manipulate the related data. In addition, disaggregated accessibility variables are still not adopted in the traditional aggregated travel demand forecasting models and their effects on demand models' outcome are not yet verified.

Land use/urban design variables may include population density, employment density, land use mix, land use balance, etc. Although some of the literature consider the land use/urban design factor as an important component affecting the travel mode selected by road users, its effects are not as significant as individual characteristics such as gender, ethnicity, and age (Loutzenheiser, 1997). Thus, how to properly define and measure the effects of land use/urban design on transit ridership appears to be a relatively new research topic that has attracted the attention of many

researchers and practitioners. Different forms of entropy¹, e.g., those implemented in Frank and Pivo (1994) and Kockelman (1997), have been adopted by different studies in different study areas. Further research is needed to focus on land use factors and their links to travel behavior because their relationships have not yet been thoroughly investigated.

According to our literature review, research on socioeconomic/demographic characteristics of transit users results in significantly contradictory conclusions. For example, income was found to not significantly affect transit use as stated in seven research papers, while the majority of literature concluded otherwise. Conflicting findings can also be observed for age, gender, and vehicle availability variables. In addition, a household's or an individual's socioeconomic/demographic characteristics are usually highly correlated, such as in the case of vehicle ownership and income. Additional precaution, such as applying the Principal Components method to reduce the number of variables (Cambridge Systematics, 1994), may be required before incorporating correlated factors.

One of the problems with many studies on the links between various factors and transit use is that the conclusions are based on claims of statistical significance, while in reality the contribution or impact of these factors is insignificant. For instance, some models presented in the literature only achieved a small R^2 (i.e., 0.1 or less) while all the model variables are statistically significant. From a modeling perspective, a contributing factor should be selected based not only on its statistical significance, but also on its power to explain the variations in transit ridership.

Study Area Selection

The study was limited to a selected area due to time constraints and the significant effort required to collect, compile, and verify the quality of the necessary data for the entire southeast Florida Region or one or more of its counties. The decision was made to select a random sample from Miami Dade County for statistical analysis because this county had the largest share of transit users and more transit types compared to other counties. Data from Broward and Palm Beach counties were used to test whether the same factors identified by the models developed for Miami-Dade County correlated with transit uses in each of those counties.

This study used the transit onboard (TOB) survey data, collected as part of the Southeast Florida Regional Travel Characteristics Study (SEFRTCS). The information from the household survey was not used because it was not adequate to allow statistically significant analysis on factors that impact transit usage due to the small transit user response size.

In selecting the analysis units, Traffic Analysis Zones (TAZ's) were considered too small to be appropriate because the TOB sample sizes were too small to ensure adequate data points in individual TAZs. Instead, Year 2000 census tracts were considered appropriate because they were not so small that too few samples would be available in each tract or too large to lose too much of the variability of important land use characteristics. One hundred census tracts were selected following a random sampling procedure.

¹ A term in physics, entropy measures disorder and randomness. Here it measures land use dissimilarity or land use mix.

Multiple regression analysis was chosen as the methodology for determining the factors affecting transit usage. Although for the SEFRTCS the TOB survey samples were randomly collected from the transit systems based on their ridership and did not reflect the true transit share at either trip end, the number of samples at both transit trip ends were highly related to the area's transit services and thus theoretically the transit demand. Because the purpose of this study was to search for the factors that had impacts on transit use and not to estimate the true transit mode share in an area, the number of TOB samples at tract level may be used as the dependent variable for regression analysis.

Data Collection and Processing

Data from a transit-on-board (TOB) survey conducted in the southeast Florida area and data on transit level of service, accessibility, land use, and socioeconomic characteristics of transit users at aggregate levels were used to develop the models. More than 170 variables were compiled using GIS. The independent variables may be classified generally into the following categories:

- Transit level of service (LOS). Examples include average bus headway, total number of bus runs in a census tract, percentage of a census tract covered by 0.25-mile buffer around Metrorail stations, and others.
- Land use. Some of the land use variables include population density, dwelling unit density, employment density, population plus employment density, land use mix, parcel size, and jobs-housing balance.
- Aggregate tract level socioeconomic/demographic characteristics of the population. Examples of this group of variables include property values, racial makeup, birth place (U.S. or foreign born), percentage of household without a car, and some lifestyle variables.
- Regional Accessibility. This group of variables includes a set of measures reflecting opportunities available to transit users throughout the county and the ease of travel by transit to those opportunities. They were computed using a gravity model-like formula that incorporates zonal employment opportunities and a disutility function reflect travel time between zones.

Transit Accessibility

To improve the estimation of transit walk accessibility, which was one of the objectives of this study, the research team developed a new methodology for estimating the percentage of production trips served by transit. This method takes into consideration walking distance to transit stops, population distribution, and existence of barriers to pedestrians. Data included the detailed street network, bus stop locations, bus routes, population and dwelling unit information by TAZ, property locations as represented by their label points, and property tax database, which provided information on number of bedrooms, used as a proxy for number of persons, at each property location. The transit onboard survey data were used to determine the effect of walking distance on transit use; results suggest that transit use deteriorates exponentially with walking distance to transit stops. A decay function was determined based on the survey data reflecting this deteriorating trend in transit use with respect to walk distance, and transit walk accessibility was measured by the percentage of the population, weighted by the decay function, in a zone that was within 0.5 mile of walking distance from transit stops. This population is referred to as

“distance decayed transit service population.” Increasing the limit of walking distance beyond 0.5 mile produced no noticeable increase in accessibility based on the survey data. Because of the decay in transit use due to increases in walking distance, transit accessibility is much lower than the traditional buffer method or the network ratio method will estimate.

Results from analysis also showed that there was no significant difference between the percentage of population with transit access and the percentage of production trips with transit access in a given zone. This means that the percentage of population with transit access can be used directly in place of percentage production trips with transit access as required as input by FSUTMS modal split model.

To allow the use of the transit accessibility measure for forecast purposes, where detailed information on street configuration and population distribution may be lacking for new TAZs, regression models have been developed using easily obtainable data to predict transit accessibility. One of the variables used in predicting transit accessibility for production trips was the number of streets in a TAZ intersecting TAZ boundary per 1,000 feet. This variable in fact measures how well the street configuration provides walk accessibility, for example, in traditional neighborhoods with grid street patterns and small blocks versus suburban neighborhoods where streets are often curvilinear with cul-de-sacs. The value of this variable must be determined based on the anticipated type of the communities in new TAZs. This may be used as a planning tool to examine the impact of different alternatives of development in a new area in terms of its transit accessibility.

Employment accessibility to transit was defined as the percentage of employees in a zone within 0.25 mile air distance of transit stops. This choice was because most commercial developments are located along arterials and thus are rather accessible to transit, and because the spatial distribution of employees in a zone is difficult to determine. Employing land use data improves the information on the spatial distribution of employees in a zone as opposed to assuming employees are evenly distributed across the entire zone or they are evenly distributed along all arterials, although on a few occasions land use data have not been consistent with employment data from the ZDATA file.

Forecasting future employment accessibility to transit services may be accomplished using the regression model developed in this study. Application of the model only requires the calculation of bus route density in a zone and the service and commercial employment ratio; both are easily obtainable.

Auto access of transit was analyzed using the TOB data of transit trips that involved accessing a transit station/stop either by park-and-ride or kiss-and-ride. Analysis results show that while there was no relationship between auto access trip distance and the transit trip length, most transit trips were longer than the auto access trips. The auto access distance for 92% of the sampled transit trips was shorter than 10 miles; the longest was 14.6 miles. Therefore, we conclude that auto access distance in a zone may be assumed to be up to the longest transit trips likely from that zone (by considering premium transit modes and major activity centers) up to 14 miles. This upper limit depends on the route length of the rapid transit services and may change if the route configuration or length changes.

Regression Analysis of Factors Contributing to Transit Use

Other transit factors impacting transit use were also investigated. Multiple regression analysis was employed using data compiled for Miami-Dade, Broward, and Palm Beach counties. Four methods were used to compile the model data, which ranged in their levels of detail in the land use and demographic information. Regional accessibility variables, especially those weighted by population, appeared to have the most significant impact on both transit productions and attractions, regardless of which method was used to compile the data. Regional accessibility measures were calculated based on zonal population, employment, zonal trip time, and impedance functions calibrated based on travel survey data for both transit and highway networks. The impedance functions are unlikely to be transferable and may need to be calibrated for individual urban areas, which will require a moderate effort.

For Miami-Dade County, the average number of bus runs in a day per bus stop in a given tract, the percentage of tract area falling into a 0.25-mile transit buffer, and percentage of distance decayed transit service population were also relatively good indicators for predicting transit use at both production and attraction ends. The first two kinds of data are easier to compile and are more reliable than regional accessibility and may be used as alternatives to regional accessibility measures. The latter two variables, i.e., percentage of transit service areas and percentage of distance decayed transit service population in a tract, are strongly correlated. Further analysis revealed that percentage of transit service areas did not perform well in areas that have low service coverage. Therefore, percentage of distance decayed transit service population in a tract is considered a better predictor of transit use.

No significant linear relationship was found consistently between transit use and the demographic, socio-economic, and other transit LOS variables for either production or attraction trips beside the regional accessibility measure, the number of bus runs, transit service coverage, and distance decayed transit service population, which are all related to transit supply. This may be due to the loss of spatial variation in these variables when large spatial analysis units (such as census tracts) are used to compile the data. While smaller spatial units will preserve the characteristics of an area or population better, this research has determined that the TOB survey cannot be used due to the limited number of samples available resulting in a significant number of spatial units with no samples.

In comparison with transit production trip models, the goodness-of-fit for attraction models are relatively poor. The R^2 s for these models are consistently lower than those for the production trip models. This may be the result of less reliable data on zonal employment.

Although the models for Broward and Palm Beach counties show similar effects of the regional accessibility variables on transit use, no other variables can be identified as the supplement/substitute of the accessibility variables.

To investigate the effect of demand variables by controlling for transit LOS, the data were divided into three groups based on their regional accessibility levels: low, medium, and high. Variables that were consistently included in models for three regional transit accessibility groups with consistent and correct signs were considered to be candidate predictors for transit demand. Total density (*TotalDensity*) compiled with the GIS Land Use Method, and entropy measures

(*AvgEntropy*) compiled with the Property Method and the Buffer Method appeared to be such predictors. However, entropy is considered a characteristic of transit-oriented neighborhoods rather than the cause of transit use. Its predictive power is therefore uncertain.

Findings and Conclusions

This research demonstrated by using GIS that more realistic estimates of transit walk accessibility could be obtained. Transit accessibility was much lower than that estimated by traditional buffer method when man-made and natural barriers and the effect of long walking distance on transit use are considered. Transit walk accessibility can also be forecast based on data that are typically used by FSUTMS models. The only variable that needs to be determined for forecast is a policy variable that determines the type of urban design to be adopted or envisioned. The study concluded the following:

- Analysis of auto access to transit shows that auto access distance in a zone may be assumed to be up to the longest transit trips likely from that zone (by considering premium transit modes and major activity centers) up to 14 miles. This result, obtained based on data from Miami-Dade County, however, may not be applicable to areas of a different urban structure, premier transit alignment, and parking availability.
- Results from the regression analysis to identify the most influential factors on transit use suggest that transit supply variables, such as regional accessibility and number of daily bus runs, dominate other factors in contributing to transit use, which in turn is the result of transit service supply being determined based on demand. Some demand variables such as density and land use mix also appear to be significant. Additional research efforts are necessary to estimate the coefficients of the potential variables identified in this research for each travel mode's utility function utilized in the modal split process.
- For Miami-Dade County, the average number of bus runs per bus stop in a given tract and percentage of tract area falling into a 0.25-mile transit buffer were also relatively good indicators for predicting transit use at both production and attraction ends. Because these two kinds of data are easier to compile and are more reliable than regional accessibility, they may be used as alternatives to regional accessibility measures. However, the newly developed *DECAY_POP* variable, i.e., percentage of transit service population seems to be a stronger indicator for transit production than the percentage of transit service area.
- No significant linear relationships were found consistently between transit use and the demographic, socio-economic, and other transit LOS variables for either production or attraction trips beside the regional accessibility measures, the number of bus runs, and transit service coverage, which are all related to transit supply. This may be due to the loss of spatial variation in these variables when large spatial analysis units (such as census tracts) are used to compile the data.
- Compared to transit production trip models, the goodness-of-fit for attraction models were relatively poor. This may be the result of less reliable data on zonal employment. Furthermore, the models for Broward and Palm Beach counties showed similar effects of

the accessibility variables on transit productions. Again, no other variables can be identified as the supplement/substitute of the accessibility variables.

- This research has used GIS extensively to compile data for various variables created for the accessibility analyses and the transit use analyses. Transit accessibility analyses require bus stop locations and property locations. As it is becoming more common for counties to possess GIS parcel data, the availability of property location data is not foreseen as a problem for the application of the methodology developed in this research. Property tax records should include information on number of bedrooms for each residential property and information on the type of the property to indicate if it is a single- or multi-family dwelling. If number of bedroom information is unavailable, then an average household size will have to be assigned to dwelling units that are of the same type (single- or multi-family). In recent years, some of the Florida MPOs have begun to use lifestyle models, which do not require data on dwelling types and zonal population by dwelling types. For the purpose of obtaining better estimates of transit service population and possibly other applications, it is suggested that MPOs continue to maintain dwelling type and population information, which will not require significant efforts.
- Finally, the transit accessibility analysis can be automated with a specially designed GIS program. This program will automate the process of matching properties to streets, assign household size, create transit service network, calculate the percentage of population with transit access in a TAZ, estimate the percentage of workers with transit access in a zone using land use information, and so on.

Recommendations

Based on the research results described in this report, the following recommendations are provided for consideration for future effort in improving FSUTMS modal split models:

- (1) The long transit walk (assuming a one-mile walking distance) file is unnecessary since transit use will be extremely low beyond one-half mile from transit stops. Instead, a single transit walk file can be used that is based on one-half mile walking distance and the decay function calibrated in this research. The data may also be used for transit service planning as it can provide much more accurate information regarding service population at the transit stop level.
- (2) The percentage of population served by transit in a zone may be estimated using the transit service population forecast model developed in this research for any reasons such as lack of data or skilled GIS staff. The forecast models can also be used for forecasting transit service population for future year models.
- (3) A GIS application should to be implemented that will calculate percentage of transit service population in a zone based on street network data, ZDATA1 and ZDATA2 files, property tax database, and parcel GIS data. In the case that GIS parcel data are unavailable, the regression model will be applied to estimate the percentage.

- (4) If a buffer method is to be used to estimate transit service population, caution should be used when the calculated service population percentage from the buffer method exceeds 50%, a level rarely exceeded by the distance decayed transit service population. Only when population is actually clustered around transit stops, a percentage higher than 50% may be justified.
- (5) Regional accessibility may be considered for the inclusion in the modal split model. This requires that accessibility to be estimated for the future, which is possible given forecast population, employment, and transportation improvements.
- (6) Density as measured by total employment plus population per acre may serve as a measure for demand.
- (7) To ensure that adequate data are available to support a more conclusive study, carefully designed surveys are desired in the future to obtain adequate number of observations in areas that reflect the different spectrum of socioeconomic and land use characteristics. Future household surveys should target transit households so that studies can be carried out to directly determine the modal split instead of using an indirect approach as in this study. If a community leadership sees transit development as necessary, then the development of information, data, and forecasting methods to better place investments is worthwhile.

1. INTRODUCTION

While transit accounts for only a small percentage of all trips in Florida, it is the primary means of travel for certain segments of the population who may not have access to jobs, health services, and social and recreational facilities by other means. Additionally, when adequately and properly provided, transit offers a travel alternative that may help to alleviate roadway congestion. To determine the appropriate level of current transit services and future transit investments for specific service areas, accurate information on the number of trips by transit, trip origins and destinations, and transit modes and routes are needed. There are two basic approaches to obtain such information. One is to apply a comprehensive regional travel demand model, typically a four-step model, and the other is to develop special models or analyses, usually designed specifically for transit-only purposes.

In four-step demand models, transit ridership is estimated during the modal split step. This is generally based on a multinomial logit model or, more recently and frequently, a nested logit model, although cross-classification similar to that for trip generation has also been used. The multinomial logit model has the following form:

$$P_n(i) = \text{prob}(Y_n = i) = \frac{e^{V_{ni}}}{\sum_{j \in C_n} e^{V_{nj}}}$$

where

- $P_n(i)$ = the probability with which person n will choose mode alternative i ;
- Y_n = the value of the response variable for individual n ;
- C_n = the set of alternatives in person n 's choice set; and
- V_{ni} = the measurable component of the utility of alternative i for individual n .

The measurable component of the utility may be expressed in matrix form as:

$$V_{nj} = A' X_n + B' Z_j + C' W_{nj} \quad j \in C_n$$

where

- X_n = a vector of characteristics of individual n ;
- Z_j = a vector of attributes of alternative j ;
- W_{nj} = a vector of interactions between characteristics of individual n and attributes of alternative j ;
- C_n = the choice set faced by individual n ; and
- A, B, C = vectors of model parameters, which are determined during model calibration.

In the Florida Standard Urban Transportation Model Structure (FSUTMS), the utility functions typically include variables like travel times by automobile and by transit, terminal times, parking costs, and transit fares. It is widely recognized that more factors influence the decision of transit use than travel costs alone, such as socioeconomic characteristics of the population. Such characteristics other possibly other factors, however, are generally not considered in modal split, as is true of FSUTMS. Other important factors such as land use may be missing from the models, reducing their ability to produce valid and reliable transit ridership forecasts.

There are several other possible causes for the fact that it is difficult to obtain accurate transit forecasts from four-step models, resulting in unrealistic expectations:

- (1) Transit ridership information is usually scantily available compared with the large amount of data collected for highway users. The transit component validation of the models is often for aggregate market shares and for a few screen lines (Dehgani and Harvey, 1994).
- (2) Analyses are typically based on traffic analysis zones (TAZs) that often border arterials spaced one-half or one mile apart. In FSUTMS, the aggregate level of TAZs for modal split implies that within the same TAZ, characteristics of travelers and land uses are uniform, which is far from reality. For instance, people are generally not willing to walk much farther than 0.25 mile to a transit stop. Even within 0.25-mile air distance of transit stops, the actual walking distance may be significantly longer if the streets are winding or simply not accessible due to natural or man-made barriers. The land-use type and intensity may also vary within a TAZ. Such variances are masked when aggregated data are used.
- (3) Currently in FSUTMS, transit auto access trips are assumed to be of a maximum length of ten miles. This assumption has not been studied or validated. Additionally, auto access trips are not based on highway skims and therefore are not affected by highway congestion. The effect of highway congestion on use of transit needs to be evaluated.

In recent years, various studies designed to estimate transit ridership have been undertaken, but the results have not been considered for incorporation into FSUTMS. Additionally, GIS data resources have become readily available, allowing for more rigorous and detailed analyses than were possible before.

The traditional travel demand models can greatly benefit from utilizing more disaggregate data to better reflect the spatial patterns of population, employment, land use, access, etc. More explanatory variables that have been known to significantly influence transit ridership may also be incorporated into the existing FSUTMS modal split module to improve the predictive power of the model.

In an assessment of transit system modeling in Florida (Pendyala, 1996), it was recommended that trip makers' characteristics should be considered for inclusion as explanatory variables. More specifically, such variables may include income, employment type and status, vehicle availability, household structure, age, and gender. Another recommendation was to account for the time of day factor in transit modeling as a consideration, such as safety at night contributing to automobile use.

The purpose of this study was to improve the estimates of transit accessibility and to identify and analyze factors that have a significant effect on transit ridership. The goal was to recommend a set of variables that may be incorporated into the FSUTMS modal split procedure to improve its transit forecasting capability. The specific objectives were as follows:

- (1) Obtain a good understanding of the current state-of-the-art and state-of-the-practice in transit ridership forecasting;
- (2) Identify available GIS data resources for improving the accuracy of analysis, and determine appropriate geographic units for data analysis and for modeling;

- (3) Develop a standard procedure that can be used to determine transit accessibility by pedestrians and automobiles;
- (4) Identify additional variables that further capture the underlying causes of transit use; and
- (5) Develop practical recommendations with regards to incorporating improvements into FSUTMS.

The remainder of this report is divided into sections. The results of the literature review are summarized in Chapter 2. Chapter 3 details the procedure and results of selecting the study areas. Chapter 4 describes the data collected and processed for this project. Chapter 5 presents the methodology for estimating transit walk accessibility and analysis of auto accessibility. Regression analyses to identify potential variables that may be incorporated into FSUTMS modal split model are discussed in Chapter 6. Finally, conclusions and recommendations are provided in Chapter 7.

2. LITERATURE REVIEW

Many factors affect transit use. The best understood perhaps are fare and travel time. Both have been incorporated into existing travel demand models. In this report, the literature review will focus on studies that address issues beyond transit fare and travel time. We first survey the literature on transit level of service, followed by accessibility; both are important determinants in mode choice. Transit ridership forecasting models are then presented. Studies on the impact of land use form and neighborhood design on travel behavior including mode choice are also reviewed.

2.1 Transit Level of Service

Level of service (LOS) standards for highways were established based on the mobility of vehicles, not the mobility of people, as in the case of transit, and are easier to measure because they are based on vehicle speed, density, and delays. Transit service quality, on the other hand, is more difficult to measure because most of the time services are provided with a fixed schedule and routes, and an array of factors ranging from ease of access/egress and transfer time, wait time to ride, comfort, and security, to name a few, may affect the decision of a traveler to use transit or not. Transit service quality varies from one urban area to another, from mode to mode, even from route to route. Worldwide evidence establishes that high-quality transit services are able to attract more transit users and reduce automobile use; however, the reverse is also true—poor transit service encourages more automobile use and discourages transit use. Transit LOS, therefore, needs to be controlled in any studies that attempt to establish a relationship between transit use and land use, policy, and demographic and socioeconomic factors, as well as other relevant issues.

Cleland *et al.* (1997) reported the results of a survey on transit users' satisfaction with the largest transit systems in six urban areas in Florida. The survey was conducted for the purpose of identifying existing and potential future problems. More than 14,500 surveys were collected. Among the 22 factors included in the survey, transit users identified hours of service, location of routes, and headways as the biggest concerns. Additionally, bus rider comfort, printed schedules, safety, and cleanliness were identified as future potential problems.

A recent Transit Cooperative Research Program (TCRP) project established a means to measure a set of service quality factors that influence the decision to use transit (Kittelsohn & Associates, 1999a). The service quality factors include service coverage, pedestrian environment, scheduling, amenities, transit information, transfer, total trip type, cost, safety and security, passenger loads, appearance and comfort, and reliability. Service coverage includes spatial availability at trip origins and destinations. Pedestrian environment is defined by presence and conditions of sidewalks, street lights, topography, signalized crosswalks on busy arterials, crosswalks with pedestrian refuges in the medium, curb cuts and bus stop loading areas, etc. Scheduling is concerned with the days and hours for which services are provided. Service frequency is also important. Amenities may include such things as benches, shelters, informational signage, trash receptacles, telephones, vending facilities, and air conditioning onboard transit vehicles or inside transit stations. Transit information may be provided through printed maps, posted information onboard vehicles or at stops/stations, onboard announcements of stops, telephone information, and World Wide Web sites or electronic mailing lists. High

passenger loads reduce comfort due to overcrowding for passengers taking long trips, as they are less likely to find a seat. The loading factor (ratio of number of seats taken to number of seats available) is usually at or lower than 1.0 for long-distance service and may approach 2.0 for inner-city service. Appearance of transit facilities and equipment (including vehicles) can affect transit attractiveness, especially to non-riders. Ride comfort (e.g. comfortable seats) and quality (smooth ride) may also affect users' decision to use transit. Finally, but not the least, service reliability, in terms of schedule adherence and regular headway, is an extremely important aspect of service quality. In addition to these measures, Kittelson & Associates (1999a) defined a number of quality of service measures for selected variables including service frequency, service hours (service span), transit supportive area covered² (percent of transit supportive area within 0.25 mile of transit facilities), passenger loading, on-time performance, headway adherence, and difference between transit and automobile travel times.

Besides the quality of service measures defined above, other measures have been developed by transit properties. One example is the percent person-minutes served, developed by Ryus (1998) for the Florida Department of Transportation's Transit Level of Service (TLOS) software. The concept is based on the fact that transit service coverage is determined not only by the spatial separation between transit stops/stations to trip origins or destinations, but also by the frequency and hours of services. Therefore, even if street configuration and population or employment distribution remains constant, transit availability will improve with more frequent service and extended hours as increased opportunities encourage riders to use the service. The service availability from a user's point of view is thus measured by the percentage of cumulative population and employment person-minutes during which a transit vehicle is within the 0.25 mile (or 5 minutes) walking distance out of all the population and employment person-minutes (e.g. for one service-hour and a community with a population of 1,000, the person-minutes will be 60,000). For any moment, the population and employment that have transit available to them may be determined by calculating the number of people and employment within a 0.25-mile ring centered at the current transit vehicle location. This approach accounts for the amount of service provided to a large extent. It may be further enhanced by taking into consideration of the uneven demand distribution during a day.

A software package is designed to analyze fixed-route service utilizing the concept of percent of person-minutes served (PPMS) (Kittelson & Associates, 1999b). The Transit Level-Of-Service (TLOS) software output is to be used to adjust mode split value by multiplying the number of non-transit trips by the PPMS value before being inserted in the mode split equation, as follows:

$$\text{Adjusted Mode Split} = \frac{\# \text{transit trips}}{\# \text{transit trips} + PPMS \times \# \text{auto trips}}$$

In the equation, the transit trips are linked trips. The adjusted mode split reflects the mode share for the area and time for which transit service is available, and can be used to evaluate transit service improvements.

² Transit supportive area is defined as one that has a household density of three (3) units per gross acre or an employment density of four (4) jobs per acre or higher.

Rood (1997) developed the Local Index of Transit Availability (LITA) to measure transit availability through a letter grade system. Using this system, scores are computed for service frequency (weekly average of number of vehicles per day), capacity (seat-miles divided by a combined population and employment), and route coverage (transit stops per square mile) for areas served by transit. The variance of the raw score of each category is divided by the standard deviation of that category. The letter grade is then determined based on the average of the three scores for each area. This measure is designed to compare transit availability in different areas instead of being an absolute measure.

Vaga and Shortreed (1982) studied the negative impact of transfers. They used a logit model for transit mode choice and assumed the disutility function with the following form:

$$U = \text{constant} + 0.03 (\text{difference in transit and auto in-vehicle-times}) + 0.08 (\text{difference in transit and auto out-of-vehicle time}^3) + \beta (\text{transfer time})$$

The authors illustrated that with commonly calibrated β values (0.08, 0.1, and 0.2) and an initial mode split of 8 percent, a five-minute transfer time would result in 31.2 percent decrease in transit ridership if the typical transfer time to in-vehicle time of 3.0 was assumed. The same transfer time would result in a 28.5 percent decrease in transit demand if the initial mode split was assumed to be 20 percent, and a 20 percent decrease if the initial mode split was 50 percent.

2.2 Accessibility to Opportunities and Pedestrian Accessibility

Accessibility has been recognized as one of the most important factors that affect both land use and travel behavior. How to define and measure accessibility has attracted the attention of many researchers and many forms of accessibility measures have been developed, which Richardson and Young (1982) classified into a spectrum of accessibility measures as shown in Table 2-1.

The logit model logsum term is given by:

$$A_i = \ln \sum_{k=1}^m e^{c(B_k - C_{ik})}$$

where A_i = accessibility index;
 B_k = benefits gained by participating in activity at site k ;
 C_{ik} = cost of travel between sites i and k ; and
 c = a sensitivity coefficient.

Richardson and Young considered the above measures of accessibility to have one major deficiency: in the calculation of accessibility of a point within a region, it was assumed that all trips that contribute to the accessibility of that point start from that single point. Instead, these authors proposed that the choice of a destination does not depend on the travel cost between that destination and the origin for linked trips, but depends instead on the cost of travel between that destination and the immediately preceding destination, and so on. For a linked trip with two destinations, the linked accessibility of a site o is given by

³ The transfer time is excluded from the out-of-vehicle time.

Table 2-1. Summary of Accessibility Measures

Accessibility Measures	Features
<i>Topological</i>	Indicates if two points are connected by a transportation link
Modal accessibility	The degree of connectivity of two places depending on the modes available.
Temporal accessibility	Accessibility varying during different time periods (e.g. transit service is available only part of a day). (PPMS utilizes this concept.)
Legal accessibility	Limitations or restrictions to accessibility by legal or regulatory rules (e.g. special permits issued to allow access to a certain area, one-way traffic rules, and denial of access to the transportation system to certain population groups).
<i>Relative accessibility</i>	Ease of travel between two points (e.g. a residential location and an employment center) based on travel time or cost
<i>Integral accessibility</i>	Ease of travel between one point and multiple different points based on travel time or cost
Place accessibility	Only spatial separation between one place and other places accounted for
Activity accessibility	Activities at destinations accounted for explicitly
<i>Cumulative opportunity index</i>	Number of opportunities (e.g. jobs) reachable from the origin within a predefined travel time or cost
<i>Gravity type measures</i>	Sum of opportunities weighted by travel time or cost
<i>Logit model logsum term</i>	Based on logit model; log sum of expected value of the maximum utility to be gained in destination choice situation

$$LA_o = \ln 2 + (B_x + B_y + B_o) - (C_{ox} - C_{oy} + C_{xy})$$

where LA_o = the linked accessibility of site o ;
 B_i ($i = x, y, o$) = benefit to be gained by participating in activity at site i ; and
 C_{ij} ($i = o, x$ and $j = x, y$) = travel cost between sites i and j .

It was demonstrated that in the case of two-destination linked trips, accessibility calculated as the logit model logsum term would be significantly underestimated when the origin was far from the center point between the two destinations. In other words, as the distance between the destinations and the origin increases, the linked accessibility will better reflect the benefit of making a linked trip, which reduces the travel time as compared to two unlinked trips. One important implication is that the accessibility of a suburban resident may be improved by linking trips and thus long distance from the urban core may not be as large a deterrent to urban sprawl as expected if unlinked accessibility is used.

Allen *et al.* (1993) considered that the relative or integral accessibility in its original form or modified forms did not reflect the overall accessibility in an area. Consequently, they developed an area accessibility measure that was based on the average of the integral accessibility of a set

of random points to other points in the area, and showed that if a rectangular area of dimensions X and Y was divided such that there were I and J equally spaced internal points in the rectangle, respectively, then the average accessibility, E , would be

$$E = \frac{X}{3} + \frac{X}{3I} + \frac{Y}{3} + \frac{Y}{3J}$$

When J and I become large, E may be approximated by $(X + Y)/3$. Using this accessibility measure, Allen *et al.* studied the employment growth rates in major U.S. metropolitan areas using regression, showing that the accessibility index was significant at 0.02 level (p-value) based on the regression results.

An application of a gravity type accessibility measure to study travel behaviors is described in Kockelman (1997). The accessibility index was defined as the sum of all attractions (e.g. employment) weighted by friction terms that reflect the ease of travel between a location and those activity centers. Zonal attractiveness may be measured by total employment or commercial and service employment. The friction term $f(t_{ij})$ often assumes an exponential form with coefficients estimated by Levinson and Kumar (1995).

Pedestrian accessibility to transit has been long recognized as important in determining ridership. Transit use decreases as the walk distance to a transit stop or station increases. It will sharply drop in most areas after the first 0.06 mile (100 meters), and will diminish beyond 0.36 mile (600 meters) (Lam and Morrall, 1982; Levinson and Brown-West, 1984). Loutzenheiser (1997) examined pedestrian access to Bay Area Rapid Transit (BART) stations in order to investigate the characteristics of walking trips and determine why people choose to walk to BART. A travel survey of BART users was used to develop disaggregate discrete choice logit models of walk access. Walking distance, gender, ethnicity, age, and car availability were identified as the most important characteristics affecting the choice of walk mode. Men were found to be more likely to walk than women; safety concerns were cited as prevalent among female riders. Blacks and Asians were less likely to walk. While Asians seemed to prefer transit to walking, blacks were more likely to drive than walk. Senior citizens, who comprised a small percentage of the surveyed riders, were less likely to walk. Availability of transit to access the BART station had a positive impact on transit use, which the author identified as an indirect result of transit availability due to transit oriented developments. Income was not a significant factor among riders who chose to take transit instead of walking. However, high income and car ownership were identified as disincentives to walking. By analyzing areas around BART stations, density alone was found to be significant, but was insignificant after individual characteristics were accounted for. The author concluded that individual characteristics such as gender, ethnicity, and age were more significant than urban design variables. Additionally, for every additional 0.3 mile from a station, the probability of an individual choosing to walk to BART station decreases by 50 percent.

Zonal residential population and number of employees within walking distance of transit stops or routes have long been considered significant contributing factors of transit use. Traditional methods are based on the buffer zone analysis, which assumes that population and employment are evenly distributed throughout a traffic analysis zone (TAZ). Several studies have shown that the buffer analysis method usually overestimates access to transit, resulting in inaccurate

forecasts. O'Neill *et al.* (1995) developed a GIS procedure (referred to as the “network-ratio method” hereafter) for performing transit service area analysis using the street network map, census block, and bus stop table databases. The method assumes an evenly distributed population along streets, as opposed to throughout the TAZ. Therefore, for each TAZ, it estimates the portion of population within transit service coverage according to the ratio of the total length of streets that are within 1/4-mile walking distance to that of all streets in the TAZ. The results showed that the network-ratio method was more accurate in determining the household units in residential areas that had developed along a modified grid street network. The same method was also used by Hsiao *et al.* (1997) to link transit ridership to land use, pedestrian accessibility, and demographic characteristics. The proportions of population with access to transit in areas with different street configurations were compared using the network-based method and buffer method. The network-based method yielded lower estimates of accessibility than the buffer method, and more importantly, the accessibility in areas with primarily gridded streets was higher than in areas with primarily irregular street patterns. A comparison of the percentage of transit users among workers that were in a low income group (income < \$25,000) and higher income group (income > \$25,000) indicated that workers with lower income were more likely to use transit. However, for both income groups, better transit accessibility increased the likelihood of transit use. A similar comparison between groups of different levels of automobile ownership yielded similar results. The study, however, did not control other factors that may have impacts on transit use.

Zhao (1998) modified the network-based method by utilizing additional information on land uses such as single-family and multi-family population, and by incorporating natural and manmade barriers that prevent access from residential areas to public transit stops. It was demonstrated that the modified network-ratio method was able to account for better density variations among different land uses, particularly for residential developments, and that barriers could have a significant negative impact on transit walk access, especially in areas with new developments surrounded by community walls.

2.3 Transit Ridership Forecasting Models

This section reviews literature on transit ridership analysis and forecasting models. Innovative, state-of-the-art general travel demand models that incorporate variables other than the common travel time and travel time typically included in modal split models are summarized first. Literature on special purpose transit ridership modeling is then reviewed.

2.3.1 State-of-the-Art General Purpose Travel Demand Models

The Sacramento regional travel demand model possesses some interesting features (DKS Associates 1994). Some of the important features include: (1) the feedback of assigned travel impedances to the trip distribution module; (2) the utilization of accessibility variables in automobile ownership and trip generation steps; (3) a joint destination and mode choice model for work trips; (4) a mode choice model including walk-and-bike, walk-and-drive transit access, and two-car pool modes; and (5) the inclusion of land use and household attribute variables in the mode choice models in addition to travel costs and time.

A travel demand model is currently being developed for the San Francisco County Transportation Authority (SFCTA) (Parsons Brinckerhoff, 2000). Two distinct sets of mode choice models have been calibrated: (1) the tour mode choice models determine the primary mode for the tour; (2) the trip mode choice models determine the mode for each individual trip made on that tour, based on the mode chosen for the tour. There is one of each of the tour and trip models for each tour purpose (Work, School, Other, and Work-Based). The modes defined for trip mode choice model estimation include: Drive-Alone, Shared-Ride (2 person carpools), Shared-Ride (3+ person carpools), Walk, Bike, Walk-Local-Walk (WLW), Walk-MUNI-Walk (WMW), Walk-Premium-Walk (WPW), Walk-Premium-Auto (WPA), Auto-Premium-Walk (APW), Walk-BART-Walk (WBW), Walk-BART-Auto (WBA), Auto-BART-Walk (ABW). The variables included in each nested mode choice model, either for tours or trips, are: (1) “traditional” level-of-service variables (e.g., in-vehicle time, first and second wait time for transit, walk-access time); (2) tour chain type variables (number of stops); (3) pedestrian environment factor variables; (4) and household variables. The tour mode choice model structure nests auto modes together (Auto Driver and Auto Passenger), non-motorized modes (Walk and Bicycle), and transit modes (Walk-Access and Drive-Access), respectively. The nested model structure for trip mode choice is consistent with that of the tour mode choice models. However, the access mode in the transit nest (walk vs. drive) is higher in the nesting structure than are the transit sub-modes.

The Metro of Portland Oregon is undertaking an effort in the estimation of a new mode choice model, which is an element of the first major update of Portland’s trip-based model structure in the last six years. A multinomial logit procedure is applied to estimate the mode choice models (Metro, 2000). Eight trip purposes are classified: HBW (Home-Based Work), Hbshop (Home-Based Shop), Hbrec (Home-Based Recreation), Hboth (Home-Based Other), NHBW (Non-Home-Based Work), NHBW (Non-Home-Based Non-Work), HBsch (Home-Based School), and HBcoll (Home-Based College). In this study, eight discrete modes are available, except for HBcoll trips, in the mode choice set: drive alone, drive with passenger, auto passenger (e.g., car pool), bus by walk access, MAX⁴ (with or without bus) by walk access, transit (bus and/or MAX) by park & ride access, bike, and walk. For HBcoll trips, the mode choice set is limited to auto, transit, and bike/walk modes. HBcoll bike/walk trips are removed from total HBcoll trips prior to separating the auto and transit mode choices. Probabilities are applied to distribute trips for each purpose to determine the number of trips by each mode. The variables included in the utility models for mode choices of different trip purposes are varied. The variables are generalized as follows: travel time, travel cost, income, employment, accessibility (defined by number of intersections within ½ mile from home), land use mix (based on the household and employment numbers within a certain distance of a zone), number of workers, car ownership, and household structure. The final calibration of the model will be completed shortly later in 2001.

2.3.2 Models Designed for Transit Ridership Forecast

While most urban areas use general-purpose four-step models that forecast both automobile and transit travel demands, models for forecasting transit trips have been developed for several reasons. One reason is that a regional travel model requires a large amount of data collection and

⁴ MAX is the Portland light rail system.

is complex to calibrate, validate, and apply. Another reason is that such a model lacks accuracy at the corridor or route level, making results unsuitable for determining needed adjustments for local service. Furthermore, most four-step models do not adequately account for the variables that affect transit ridership.

Horowitz (1985) developed a transit ridership forecast model based on the four-step model concept. The mode split is determined using a logit model and is based on measures of trip dissatisfaction, which are expressed as a combination of trip time and trip cost:

$$\text{for car } a'_{ij} = a_{ij} + [c_j + (a_{ij} + 2e)M] / V$$

$$\text{for bus } t'_{ij} = t_{ij} + F / V$$

where a_{ij} = automobile travel time between zone i and zone j ;

c_j = automobile cost at zone j ;

M = automobile cost per minute of travel

V = value of time

t_{ij} = transit travel time (including weighted walk time, weighted waiting time, wait penalty, weighted transfer time and transfer penalty); and

F = transit fare.

Koppelman (1983) developed a simplified form of the multinomial logit model and applied it to the prediction of travel mode shares for a range of transit service changes. Changes in ridership due to new service changes are determined by an incremental logit equation in which utility is a function of the change in in-vehicle travel time, out-of-vehicle travel time, and out-of-pocket cost. No model calibration was performed. Variables and their coefficients in the logit model for determining transit mode share were “judgmentally selected” by the author based on model coefficients from seven urban areas. The utility model for transit service in the paper is illustrated as follows:

$$UT = -0.016 (\text{In-vehicle travel time}) - 0.17 / (\text{Distance in Mile}) (\text{Out-of-vehicle travel time}) - 0.0044 (\text{Out-of-pocket cost})$$

Nickesen *et al.* (1983) developed another simple transit ridership estimation model system for short-range planning. A sequence of simple trip generation, trip distribution and modal split models generate trip-purpose specific transit trip table, denoted as “trial” trip tables. A linear programming model converted the trial trip table into the final transit trip table, which nearly replicated the observed link volumes and was as close as possible to the trial trip table. The effort was aimed at refining the results from a logit modal split model with borrowed coefficients.

In Levinson (1985), the ridership potential of various public transportation options was estimated based on the following information: (1) corridor population and employment growth; (2) changes in service levels resulting from the various options; and (3) effects of changes in gasoline price, parking costs, and increased traffic congestion. General growth trends were derived from an analysis of actual experience and agency forecasts, which were then modified as appropriate to reflect results of the 1980 Census and likely development in the study area. Bus ridership was

estimated for 1985 and 2000 by applying both average and sectional growth factors. In both cases, it was assumed that sufficient service adjustments would be made to enable ridership to keep pace with corridor population and economic growth and that there would be no fare increases, relative to the base year, in real dollars. This author concluded that bus ridership would not keep pace with population and economic growth unless service was improved.

Eash *et al.* (1993) developed a spreadsheet version of the Chicago Area Transportation Study's mode choice model to estimate how service changes for the west-side rail lines affect ridership. The data utilized include a zonal trip table, existing line-haul and access time and costs, and service change scenarios.

In Dehghani and Harvey (1994) described a fully incremental logit model for the Seattle-area transit forecast. Forecasts were made based on the differences from an existing situation instead of an entirely new set of variables. The data used in the model included transit onboard survey, transit travel time and cost, automobile travel time and cost, transit walk access, and automobile access. These variables are commonly used in the modal split in traditional four-step models.

To deal with short-term transit ridership fluctuations, Seattle's Metro Transit developed a simple statistical model that forecast short-term ridership based on gasoline price, gasoline supply, service changes, fare changes, and employment (Ulberg, 1982). Other factors, such as service quality, population, CBD employment, parking prices, fuel efficiency of cars, and data disaggregated by time of day, route, or region served, were also considered but rejected because of lack of data or predicting power. The analysis is a time series based on the dependent variable: monthly change in seasonally adjusted weekly ridership. Ridership data were estimated from revenue data and adjusted for seasonal changes to obtain average monthly ridership. By studying the time lag effect of gasoline price on ridership, two months' lag time was found to result in the lowest residual, which was converted to the equivalent monthly change. Transit service was measured by monthly service hours. Employment was obtained monthly from the Washington State Employment Security Department for the entire service area. Employment was adjusted for seasonal changes and a time lag of three months was determined. A multiple regression model ($R^2 = 0.694$) was estimated using the above predictors and eight years of data. All predictors except service hours were significant. The authors offered explanations for the insignificance of service hours as being too gross a measure of service quality, the possible long lag between addition of service hours and new ridership in some areas, the Seattle transit riders being mostly marginal riders, on whom economic factors had more influence, and the fact that service changes resulted in ridership changes instead of the other way around. The model was able to predict ridership for 1980 and 1981 with an error of 0.9 and 1.7 percent, respectively. An examination of the stability of the model coefficients by estimating regression models using a subset of data and comparing the coefficients revealed that all coefficients except that for employment were stable. The instability in the employment coefficient was considered to be a result of the saturation of transit services that came later that could no longer attract ridership from existing employment, and new ridership that had to come from new employment.

Another short-term transit ridership forecasting model was reported in Nelson and O'Neil (1982). The multiple regression model has ten predictors for the dependent variable: home-based transit trips per thousand zonal population. The home-based transit trips were estimated based on a 1981 onboard survey; responses represent nearly 25 percent of total boardings. The survey data

were adjusted and expanded for 298 residential zones. 121 additional zones were not included because either the zones had too few dwelling units to ensure unbiased data or accurate socioeconomic data were unavailable. The predictors were grouped into two categories: (1) level of service (LOS) and (2) socioeconomic and land use. The LOS variables include composite in-vehicle travel time, composite transit wait time, composite number of transfers, number of routes serving a zone, and a dummy variable assigned to zones at the end of major transit corridors to control external trip assignments in those zones. Socioeconomic and land use variables include percentage of single-family homes in a zone (population or dwelling unit density were found to be correlated with LOS variables), commercial and service employment density, ratio of industrial employment to number of households (negatively impacting transit use), percentage of people aged 18 and over that were estimated to be employed, and average zonal household size. Elderly population was considered to be an important factor but was not used because of lack of data. Income and auto ownership also were not found to be negative factors as expected. The least square multiple regression technique was applied to calibrate the model. Independent variables for LOS measures were constructed by taking the weighted average derived from the estimated total daytime population of each major destination area in order to reduce the multicollinearity effect. The model achieved an adjusted R^2 of 0.7289. While the model appeared to perform well based on a comparison of the predicted ridership by the model to the actual ridership, it is unclear from the paper if the test data set was part of the model development data or not. The model did produce accurate results when additional services were added, increasing the number of routes in nine zones. The model predicted 74 to 81 new transit trips (the model prediction adjusted for non-home trips, which accounted for about 17 percent system-wide) compared to the actual 81 trips observed over a four-month period after the implementation of the service change.

In a study of public transit alternatives in the Forest Hills-Needham corridor, factors that were considered to produce the most significant mode shift included levels of service offered by different transit alternatives, out-of-pocket costs of using alternative modes, personal income, and automobile ownership (Harrington and Carakatsane, 1984). The transit alternatives included two that would combine commuter rail service and different local bus feeder services and one that would consist of express and local bus services with no commuter rail service. Data related to population, automobile ownership, mean income, percent of population living below poverty level, employment by types (private, manufacturing, and service), transit access modes (drive-alone park-n-ride, carpool, kiss-n-ride, walk, and other), transit ridership, transit levels of service, fuel price, travel time were compiled. Forecasting of ridership for the different alternatives involve three steps. Step one: intermediate total transit trip production in the forecast year was obtained by summing the zonal base year transit ridership, which was adjusted for changes in population, fuel price, per capita income, and per capita automobile registration. The elasticities used for these adjustment terms were 0.35, -0.29 (for the rail alternative, 0.25 for the bus alternatives), and -0.6 , respectively; Step two: the intermediate total ridership among production and attraction zones were allocated according to the ratio of zonal total trip production to area-wide total trip production, both adjusted by changes in employment, fuel price, and per capita income. The elasticity of total trip production with respect to employment was assumed 1.0. Step three: ridership for different alternatives was estimated by considering levels of service factors such as fare, between-station travel time, and access time.

Black (1993) tested a regression model for estimating rural transit ridership based on population over age 55, transit system size as measured by size of transit network in kilometers, and local (county) monthly rents. The R^2 was 0.986, but no additional information about the model was provided.

Multisystems⁵ developed a route-level transit ridership forecast model, the “Period Route Segment Model,” for the Southern California Rapid Transit District (Batchelder *et al.*, 1983). The model estimates the morning peak and the midday boarding in each direction for every segment of a route based on segment-specific demographics, attractions, and level of service. The model takes the following form:

$$BOARD_i^d = PROD_i \times OPP_i^d \times LOS_i^d$$

where $BOARD_i^d$ = boarding count on segment i in direction d ;
 $PROD_i$ = production factor in area surrounding segment i ;
 OPP_i^d = trip opportunity factor in direction d from zone i ; and
 LOS_i^d = level-of-service factor in segment i in direction d .

The transit trip production factor was computed for each segment of a route based on average monthly rent, total population and adult population (a surrogate for number of workers) that lived within walking distance (0.25 mile) of transit, number of riders on other transit routes that crossed or fed the route to be a proxy for transfers, and the route length in the TAZ. The population and adult population were estimated using a buffer zone method assuming their distributions are uniform throughout the TAZ. The opportunity factor for morning peak, e.g., OPP_i^d , was computed as follows:

$$OPP_i^d = \left(EMPLOP_i^d + 0.75POPOP_i^d \right)^{0.296}$$

$$EMPLOP_i^d = \frac{ORIGBD_i \times EMPL_{6-35}^d + 0.339 \times XRIDER_i \times EMPL_{3-15}^d}{ORIGBD_i + 0.339 \times XRIDER_i}$$

$$POPOP_i^d = \frac{ORIGBD_i \times POP_{6-35}^d + 0.339 \times XRIDER_i \times POP_{3-15}^d}{ORIGBD_i + 0.339 \times XRIDER_i}$$

$$ORIGBD_i = AFAC_i \times ADULT_i$$

where $AFAC_i$ = an income adjustment factor;
 $ADULT_i$ = the adult population near route i ;
 $EMPL_{6-35}^d$ = the employment size within 0.25 mile of the route, in direction d , and within 6 to 35 minutes of bus ride;
 $EMPL_{3-15}^d$ = the employment size within 0.25 mile of the route, in direction d , and within 3 to 15 minutes of bus ride;
 POP_{6-35}^d = the population size within 0.25 mile of the route, in direction d , and within 6 to 35 minutes of bus ride;

⁵ Multiplications Inc., Multisystems, the Consulting Division, 1050 Massachusetts Avenue, Cambridge, MA 02138.

- POP_{3-15}^d = the population size within 0.25 mile of the route, in direction d , and within 3 to 15 minutes of bus ride; and
- $XRIDER_i$ = the riders on routes that cross or feed into segment i .

The transit level-of-service factor was calculated as the product of the wait time and seat availability at segment i in direction d . The wait time depends on the headway, and seat availability depends on the cumulative length of the route from the origin to segment i , assuming the end of the route is a CBD. This model is interesting because it includes some factors, such as route level service quality and opportunity along a route, absent from other models that may contribute to transit ridership. Azar and Ferrira (1994) was later applied this model by using data from the Boston area to demonstrate the integration of the model with a GIS tool to allow transit planners to determine necessary service changes.

One important fact about transit ridership is that transit demand and supply are not independent of each other. Transit services are typically adjusted periodically (e.g., once a year) based on the past ridership levels and trends. Service changes in turn will have an impact on future ridership, although the impact is not immediate. To account for this interaction between the transit demand and supply, a simultaneous route-level ridership model that considers the demand, supply, and inter-route relationship was developed by Peng *et al.* (1997). The model consists of three equations to describe the demand, supply, and competing routes. The basic unit of observation is the route segment within a fare zone. For instance, when the model was applied to the Tri-Met service area, transit routes were segmented by four fare zones. Ten models developed for five periods (morning peak, mid-day, afternoon peak, evening, and night) and for two directions (inbound and outbound). The models were estimated using a three-stage least squares for all transit routes in the Tri-Met service area, and have the following form:

Demand

$$R_{iz} = -190.01 + 0.303S_{iz} - 0.121 \sum_j (R_{ijz} \cdot OVPOPPC_{ijz}) + 0.162 \sum_j R_{ijz} + 0.013POP_{iz} \\ + 0.0107Inc_{iz} + 0.502PARK_{iz} - 0.936CROSTWND_i - 20.324FEEDERD_i \\ + 111.04FAREZN1D + 128.45FAREZN2D - 43.435FAREZN3D$$

Supply

$$S_{iz} = 127.67 + 0.155R_{iz} + 0.510R_{-li} - 0.0045POP_{iz} + 0.327EMPDEN_{iz} \\ - 127.70CROSTWND_i + 94.499FEEDERD_i$$

Competing Routes

$$\sum_j R_{ijz} = 29.738 - 0.177S_{iz} \cdot \sum_j OVPOPPC_{ijz} + 9.077 \sum_j FRQ_{ijz} - 0.0188 \sum_j OVPOP_{ijz} \\ + 0.0180 \sum_j POP_{jz}$$

where

$$S_{iz} = \frac{\text{Service Time}_{iz}}{\text{Headway}_{iz}} \cdot \text{Seats per Bus}_{iz}$$

R	=	boarding rides;
R_{kz}	=	alighting from complementary routes k in zone z ;
Service Time	=	total minutes of service in a service period (e.g., AM peak, PM peak, off-peak)
S	=	service supply as measured by the total number of seats supplied during the service period;
POP	=	population within a quarter-mile buffer of a bus route or around a light rail station (referred to as transit service area);
INC	=	number of households with income less than \$25,000 in 1990 in a transit service area;
$EMPDEN$	=	number of employees per acre in 1990;
$PARK$	=	total number of parking spaces in Tri-Met park-n-ride lots within service area of route i ;
R_{-1i}	=	previous year ridership on route i ;
$OVPOPPC$	=	percentage of population in the overlapping portion of the transit service areas of two competing routes;
FRQ	=	number of buses per hour;
$CROSTWND$	=	crosstown route typology dummy variable;
$FEEDERD$	=	feeder route typology dummy variable;
$FAREZN1D$	=	fare zone dummy variable (fare zone 1);
$FAREZN2D$	=	fare zone dummy variable (fare zone 2); and
$FAREZN3D$	=	fare zone dummy variable (fare zone 3).

The models confirm the interactions between demand and supply, as well as the negative effects of competing routes. Additionally, the model indicates that inbound demand is mainly determined by population size at places of residence, while outbound demand by employment density. Parking availability at Park-and-Ride lots influences the morning peak and mid-day inbound demand. Results indicated that income was a strong factor contributing to inbound demand in the mid-day, afternoon-peak, and evening periods, but was not significant in the morning peak period. The authors contended that individuals who used transit during the morning peak hours were more likely to be choice riders while those during the other periods were the captive riders. The models also show that demand decreases as the distance from downtown increases.

Pendyala (1999) also developed a transit supply-demand model that consisted of a set of simultaneous equations. The Integrated Transit demand and SUPply model (ITSUP) was developed using demographic and socioeconomic data of Volusia County and transit system route data from Volusia County transit agency (VOTRAN). ITSUP is intended for planning service adjustments in the short-term as well as estimating the impact of service changes on ridership. It consists of three equations: one for estimating transit demand; one for estimating transit supply; and one for representing inter-route relationships. Using the three equations, the model interactively computes ridership and service supply on each route segment, defined as the transit link between two bus stops, until convergence is achieved. The default equations are:

Ridership Equation

$$RIDER = 23 + 0.0107 * HH_LT_25K + 0.013 * POPULATION + 0.303 * SRVC - 0.121 * CMPT_EQN * CMPT_RATIO + 0.162 * CMPL_EQN$$

Supply Equation

$$SRVC = 127.67 + 0.327 * WRKR_DNSTY - 0.0045 * POPULATION + 0.155 RIDERS$$

Inter-Route Equations

Competing Routes

$$CMPT_EQN = 29.738 - 0.177 * SRVC * CMPT_RATIO + 9.077 * CMPT_FREQ - 0.0188 * CMPT_OVPOP + 0.018 * CMPT_POP$$

Complementary Routes

$$CMPL_EQN = 29.738 - 0.177 * SRVC * CMPL_RATIO + 9.077 * CMPL_FREQ - 0.0188 * CMPT_OVPOP + 0.018 * CMPL_POP$$

where	<i>RIDERS</i>	=	daily 24-hour ridership on a route segment;
	<i>SRVC</i>	=	total daily seats supply on route segment;
	<i>HH_LT_25K</i>	=	number of households residing in 1/4-mile buffer around route segment with annual 1990 income less than or equal to \$25,000;
	<i>POPULATION</i>	=	total number of persons residing in 1/4-mile buffer around route segment;
	<i>WRKR_DNSTY</i>	=	employment density in 1/4-mile buffer around route segment;
	<i>CMPT_EQN</i>	=	number of riders on competing route segments;
	<i>CMPL_EQN</i>	=	number of riders on complementary route segments;
	<i>CMPT_RATIO</i>	=	ratio of the population in the overlap area of two competing route buffers over the total population in the subject route and competing route buffers;
	<i>CMPT_POP</i>	=	population in the competing route buffers;
	<i>CMPL_POP</i>	=	population in the complementary route buffers;
	<i>CMPT_OVPOP</i>	=	population in the overlap area of competing route buffers;
	<i>CMPT_FEEQ</i>	=	frequency of service on competing routes; and
	<i>CMPL_FREQ</i>	=	frequency of service on complementary routes.

Users may specify new equations appropriate for their own areas. This approach to a transit ridership model recognizes that transit service quantity and quality are not exogenous to ridership. Rather than assuming that ridership will increase with route density, the model also considers the competition among different routes for transit ridership, as ridership is unlikely to grow at the same rate at which route density increases.

A bus service planning and marketing model that utilizes a GIS approach is described in Hunt *et al.* (1986). The model consists of two regression models, one for residential locations and the

other for workplaces. The model by residence incorporates variables representing population density, proportion of white population in a census tract, proportion of zero car households, and percentage of census tract area falling in the 0.25-mile buffer around transit route multiplied by the total number of buses passing through the service area. The workplace model includes variables representing employment density in the destination census tract, proportion of workforce that is white, proportion of households in workforce with no automobile available, and combined bus frequency. The residence model has a R^2 of 0.84, while the workplace model has a R^2 of 0.41. The model may be used to predict ridership at route level and examine the effects of demographic and transit service changes.

Nelson *et al.* (1997) analyzed the characteristics of the population in Atlanta's northern affluent tier, including their demographic, socioeconomic, and trip characteristics. The authors found that riders were encouraged to use transit when they had easy access to stations and accessibility by rail to major employment centers even if they were in a high-income bracket. On the other hand, disincentives to rail service use included long walking distance or the need to ride a bus to access rail stations. Affluent suburban workers also tended to drive to rail stations, making ample park-n-ride facilities necessary to ensure good ridership.

Based on regression analyses, Tri-Met in Portland determined that housing density, employment density, and retail employment density were the most significant variables in the Portland area in determining transit ridership, accounting for 81% of the variations (Nelson\Nygaard Consulting Associates, 1997). A regression model estimated future year ridership levels, which were then converted to a transit orientation index (TOI) to measure transit ridership potential. The index values were then used to determine which areas needed to be given priority for transit service improvements to maximize the ridership. This index alone, however, does not reflect the deficiency or the need of transit services in an area. To maximize ridership, improvements should be made where the greatest potential to increase ridership exist.

A GIS-based transit forecasting approach to modeling transit ridership in tourist corridors is described in Preslar (1998). The corridor was the 6.0-mile International Drive with a high concentration of tourist attractions (Universal Studio, Sea World, Convention Center) and 18,000 hotel rooms. Data were aggregated using GIS at the parcel level, which included information such as land use, dwelling units, hotel rooms, square footage of retail space, and number of employees. The front door locations of buildings were also geocoded to allow accurate estimates of walking distance. A four-step model was used for ridership forecasting.

Some studies have focused on answering specific questions on how certain factors influenced transit use. Spillar and Rutherford (1990) examined the relationship between income, density, and transit ridership at the route level in western American cities. Per capita transit ridership data were derived from the 1980 Census, calculated by dividing total zonal transit ridership by total zonal population. The Census data also provided information on income distribution. The percentage of all families in each tract earning less than \$10,000 was calculated for each tract. The census tracts were then divided into subgroups according to this calculated income characteristic. One group, the 18% group, consisted of census tracts with fewer than 18% of low income families. The second group, the 82% group, consisted of census tracts containing 18% or more families with low income. A non-linear regression technique was used to regress against the primary variable of zonal gross population density for the following three household groups:

total, 18% income households, and 82% households. The authors concluded that the relationship between per capita transit ridership and gross population density was a second order polynomial type function, increasing rapidly as very low densities increase, and then curving at some maximum value once a specific density level was achieved. Income groupings displayed slightly different per capita ridership characteristics.

According to a study on bus route demand in Cleveland by Krechmer and Lantos (1983), bus-to-rail transfer is mainly determined by the in-vehicle travel between trip origin to rail station, while bus-to-bus transfer is minimum and is related to the frequency of the two bus routes, as follows:

$$PT_{ab} = \begin{cases} 0.498 - 0.1242 \left[\ln(CH_a + CH_b) \right] & \text{if } (CH_a + CH_b) \leq 55 \\ 0 & \text{if } (CH_a + CH_b) > 55 \end{cases}$$

where PT_{ab} = percentage of passengers transferring from bus route “a” to bus route “b”;
 CH_a = combined headway (minutes) for bus route “a”; and
 CH_b = combined headway (minutes) for bus route “b”.

The R^2 of the model was 0.55. The model was developed using data from Cleveland and may not be applicable to other areas, which have different transit services and demographic and socioeconomic characteristics. A report prepared by the Central Planning Staff for the Federal Transit Administration (1997) states that transfer penalties may be quantified and are equivalent to 12~15 minutes per transfer depending on the particular model specification.

When transit planners use a simpler model specially design for transit ridership estimation in place of a regional four-step model, they utilize longitudinal data, especially ridership trend data. Such an example may be found in McLeod Jr. *et al.* (1991). Two models were constructed using the statistical technique of least squares multiple regression: an annual passenger revenue-trips (R-TRIPS) model and an annual linked trips (L-TRIPS) model. Historical data on a small number of economic, demographic, and transportation variables from 1958 to 1986 were used. The final models are given as follows:

$$\begin{aligned} R-TRIPS &= -118.9 + 52.2 * \ln(JOBS) - 60.9 * \ln(INCOME) - 27.8 * \ln(FARE) + 7.9 * \\ &\quad \ln(BUSES) - 4.4 * STRIKES \\ L-TRIPS &= -118.3 + 38.2 * \ln(JOBS) - 44.1 * \ln(INCOME) - 36.0 * \ln(FARE) + 10.6 * \\ &\quad \ln(BUSES) - 4.1 * STRIKES \end{aligned}$$

where $JOBS$ = number of civilian jobs;
 $INCOME$ = per capita income in 1982 dollars (\$million);
 $FARE$ = fare in 1982 dollars (\$million);
 $BUSES$ = number of buses in the fleet; and
 $STRIKES$ = dummy variable for occurrence of strikes.

The following table summarizes the elasticities for all the variables in the two models.

Table 2-2. Elasticity of Model Variables

Variable	R-TRIPS Model	L-TRIPS Model
JOBS	1.04	0.64
INCOME	-0.98	-0.59
FARE	-0.56	-0.61
BUSES	0.25	0.28

According to the authors, the models explained 97 to 98 percent of the variation in the bus ridership.

Caution is advised, however, as direct application of trend data in standard regression models may cause statistical problems. Kyte *et al.* (1985) identify some of the problems as follows:

- A high degree of correlation was found among the input variables;
- The residuals were highly correlated and not independent as required for regression models;
- The delay in the response to service level changes would have been missed if only contemporaneous correlations were included in the model; and
- The biased standard errors from the regression model would have erroneously led to including variables that were not significant.

To properly utilize trend data, time series analysis is necessary and more appropriate and powerful than the traditional regression analysis. An example of application of this approach is presented in Kyte *et al.* (1985), who employed a statistical approach developed by Box and Jenkins (1976) for time-series data known as autoregressive-integrated moving average (ARIMA) models (also known as univariate ARIMA models). Four input variables were used: (1) transit service level (platform hours, platform miles, and route miles); (2) transit fare; (3) gasoline price; and (4) employment as a measure of the travel market size. Sixteen transit ridership models were developed using data for Portland, Oregon covering 1971 through 1982: one for the system as a whole, six representing distinct geographic sectors of the Portland region, and nine for individual routes in the Portland transit system. Kyte *et al.* determined that the Box-Jenkins time-series models were appropriate for evaluation and forecasting of transit ridership changes. The lag structure of the market response to the factors that influenced transit ridership were identified as follows:

- (1) Service level delays ranged from 1 to 10 months for the system model and 0 to 3 quarters for the sector and route models;
- (2) Fare delays ranged up to 2 quarters.
- (3) Response to gasoline prices and employment level changes are more rapid, though lag effects have been found at the route level for up to 3 quarters for gasoline price change.

Vaziri *at al.* (1990) studied temporal variation of specialized transportation monthly ridership and demonstrated the superiority of intervention models ridership, or multivariate ARIMA models, for time-series analysis over regression models based on the results of analyses of data collected over a period of time in Lexington/Fayette County, Kentucky. One model was developed using the entire data set, while ten other models were estimated using one of the subsets of the data: passenger type (elderly, handicapped), passenger's ability to walk

(ambulatory, semi-ambulatory, nonambulatory), and trip purpose (medical, employment, shopping, education, other). In sum, 11 regression and 11 intervention models were developed. The variables used in the model calibration for both approaches include: the service expansion in January 1981 (X_{1t}), Saturday service availability (X_{2t}), the fare increase of July 1984 (X_{3t}), and the fare increase of July 1985 (X_{4t}). Variable X_{2t} was found not significant in all of the 11 intervention models, while X_{3t} and X_{4t} were not significant in the majority of the models. The authors claim that these models are able to accurately predict the months that have minimum and peak ridership.

2.3.3 Dynamic Structural Model

Altinoglu and Smith (1992) proposed in a working paper the application of a Covariance Structural Model (CSM), a dynamic structural model, for modeling transit demand using travel survey data from 1971, 1981, and 1991. The CSM has two components: (1) a structural model that describes the “unassumed causal structure” and is in the form of simultaneous equations of unobserved (latent) variables; and (2) a measurement model that is a simultaneous equation relating unobserved (latent) variables to measured variables observed (or measured) variables. The authors suggested that the model parameters be estimated by minimizing the difference between the estimated covariance and the calculated covariance. The authors identified three types of latent factors: (1) mobility (endogenous); (2) socioeconomic characteristics (exogenous); and (3) individual land use characteristics (endogenous). The observed indicators for these three latent variables are public transit trips, car trips, public transit travel distance, car travel distance, car ownership, and public transit and car attributes for mobility; income, household size, life cycle of household, number of workers, number of drivers, age, gender, education level, company car use, employment category, employment status, ethnic background for socioeconomic characteristics; and residential area per individual, employment area per individual, residential location, home business/teleworking for land use characteristics. No information on the final outcome of the project is available. Data for most of the proposed variables are, however, available from travel survey data. Some of the variables are known to have direct impact on transit use, such as car ownership, life cycle of household, age, company car use, residential area per individual, etc., while the impact of others (such as ethnic background and employment category) is uncertain.

2.4 Mode Choice and Urban Forms

The need to understand how urban forms may affect travel behavior has become urgent due to recent policy initiatives at the federal, state, and local levels to look for ways to improve mobility and reduce congestion without building new highways. These policy initiatives are motivated by the Intermodal Surface Transportation Efficiency Act of 1991 (ISTEA), which provided new funding opportunities for transportation improvement projects not targeting single-occupancy-vehicle (SOV) mobility, the Transportation Equity Act for the 21st Century (TEA-21), which initiated a new sustainable development pilot program to help state and local governments plan environmentally friendly development, the Clean Air Act Amendments of 1990 (CAAA), which sets vehicle miles traveled (VMT) as a form of mitigation to meet air quality attainment, rising public concerns about petroleum consumption in the U.S. and global warming, and political pressure to reduce fuel consumption. One of the approaches to reduce VMT is to change travel behavior via policies such as taxation, pricing, and land use planning. The question is therefore

whether land use policies that encourage “transit/pedestrian friendly” neighborhoods will be effective. Researchers have been attempting to answer these questions by looking into land use factors and their links to travel behavior.

One of the most influential works may be that by Pushkarev and Zupan (1977) who investigated the impact of land use, spatial separation, and transit service quality on transit ridership. The land use variables are the suburban residential housing unit density and central business district (CBD) floor space, which is used as a proxy of jobs. Spatial separation is measured by distance between the CBD and the residential areas. By comparing different bus routes, the authors found that there was a significant correlation (0.75) between transit use and density. There is a four percent increase of workers using transit for every doubling of density. The results of their analysis lead to several interesting findings: a density of seven to thirty dwelling units per acre is the threshold of significant transit use: “high residential density by itself does little for transit if there is no dominant place to go to.” They point out, however, that the higher transit ridership was not induced by density per se, but was also due to increased availability of employment and other opportunities, as well as higher parking cost and more congested roads that have limited capacity to accommodate automobiles.

In another study of the 1979 New York Urban Region survey data, Pushkarev and Zupan conclude that “there is no statistically significant effect of income on driving once other variables (density, household size, number of adults, etc.) are held constant” (Holtzclaw, 1990).

By simple regression, Newman and Kenworth (1989) also find high correlation between automobile use (measured by petroleum consumption) and density by studying major cities around the world. Specifically, they found a correlation of -0.74 between urban density and private car use, +0.74 between density and transit passenger trips, and -0.76 between density and auto ownership. The correlation between density in central business districts (CBDs) and private car use is, however, much lower at -0.14; this might be explained by the fact that other important factors such as culture, government policy, gasoline prices, transportation system, transit service level, income, etc., were not controlled. These factors vary significantly in different countries and may have an important influence on travel behavior.

Thompson and Frank (1995) argue in their report that transit trip production and automobile trip production are independent to a certain degree. They point out that introduction of a new mode tends to cause increased travel in addition to causing mode shift. On the other hand, inferior modes may also have their own market. Traditional four-step models do not account for the independency of trip productions by different modes and tend to ignore the fact that trip rates may be affected by transit service and quality and roadway congestion level. The authors believe that studies supporting the claim that urban form and socioeconomic variables determine transit success or failure do not adequately control transit level of service, and studies concluding transit level of service is important do not control adequately for socioeconomic variables. In their study, Thompson and Frank attempt to control both sets of variables by studying transit ridership between two points as a function of mobility between the two points by transit and automobile, population and employment density, income and transit dependence characteristics, and design features of the two points. The model assumes a general form of the double constrained gravity model, in which the “production” and “attraction” terms interact with each other. The authors

argue that while the model needs significant improvements, it is potentially less cumbersome and theoretically more accurate than the traditional four-step models.

In a 1996 study, Gray and Thompson (1996) test the interaction between urban form, transit route configuration and transit demand. A set of four models are established: a single-occupancy auto trip model, a car-pool trip model, a transit trip model, and a biking and walking model. The explanatory variables include zonal population density, employment density, service employment, proportion of population of age 16 or younger, proportion of population of age over 65, ratio of male population, zero-car households, total employment, ratio of employment to population in origin zone, ratio of service employment to total employment, proportion of population in households with income lower than \$15,000, congested auto travel time, tolls, transit travel time, and straight line distance between origin and destination zones. Data for the dependent variables were taken from the Census Transportation Planning Package (CTPP) for Orange County, Florida for 1990. For independent variables, socioeconomic data came from CTPP while network, tolls, and travel times were from the regional demand model. The models were estimated using poisson regression. While some of the variables were statistically significant, the models generally had little explanatory powers. The SOV model had a R^2 of 0.12, with destination employment density, parking fees, and origin mixed use (employment to population ratio) being significant. The R^2 values for the other three models were 0.05 or less. One possible cause for the weak models was considered by the author to be the zero-inflation problem: too many dependent variables (90% ~ 98%) had zero values.

An empirical study was performed by Frank and Pivo (1994) to determine if density was a proxy of other factors or itself caused a difference in mode choice, with the purpose of discovering ways to implement urban forms that promote accessibility in urban areas. By analyzing mode choice for work and shopping trips based on land use variables such as population density, employment density, and land use mix at census tract level, life-style variables such as age distribution within a surveyed household and mean age of survey participants per census tract, and other non-urban-form variables including proportions of survey participants with a driver's license, mean number of vehicles for survey participants ending trips in a census tract, and proportions of transit trip ends made by survey participants employed outside home, by those participants who had a bus pass, and by those who had access to less than one vehicle. The land use mix was measured by an entropy index defined as follows:

$$\begin{aligned}
 entropy = & - \left[singlefamily \times \log_{10}(singlefamily) \right] \\
 & + \left[multifamily \times \log_{10}(multifamily) \right] \\
 & + \left[retail\ and\ services \times \log_{10}(retail\ and\ services) \right] \\
 & + \left[office \times \log_{10}(office) \right] \\
 & + \left[entertainment \times \log_{10}(entertainment) \right] \\
 & + \left[institutional \times \log_{10}(institutional) \right] \\
 & + \left[industrial/manufacturing \times \log_{10}(industrial/manufacturing) \right]
 \end{aligned}$$

Multivariate regression analyses showed that urban-form variables entered after including significant non-urban-form variables in the models did contribute to mode choice, with positive impact on transit use and walk and negative impact on SOV use. The analyses also suggested that employment density at both trips ends should be used to explain the variation in mode choice instead of using the density at one trip end. Additionally, land use mix best explains why individuals choose to walk. The authors also investigate the property of the functions that relate the urban-form variables to mode choice, suggesting that such functions are non-linear in nature. Plotting mode choice versus gross employment per acre was created and from the plots the authors determined that significant shifts from SOV to transit use and walking occur between an employment density of 20 and 75 employees per acre and again when density exceeded 125.

Kockelman supports Pushkarev's and Zupan's conclusion in a study on the relative effect of population density and income on modal split (Kockelman, 1995). She shows that density (or other factors proxied by density such as land prices, parking fees, transit service frequency, and congested roadways), not income, was the influential factor on modal split. The study analyzed three different levels of data covering 108 San Francisco Bay Area (SFBA) census tracts, 41 SFBA cities, and 35 U.S. metropolitan areas. Due to data limitations, the author examines work trips only. By analyzing census tract data using single variable regression, the percent of workers not driving alone was found to be significantly related to density (correlation 0.891 and R^2 0.794), but not so to income level (correlation -0.289 and R^2 0.084). Density and income are not significantly correlated. In multiple regression analyses, a destination index serves as a coarse proxy for transit level-of-service to and at the workplace and the regional importance of that destination for employment. The index was constructed as the weighted sum of percent of workers that commute to different cities. The multiple regression results again show that an index of density and destination is more important than income level in determining mode choice. The elasticity of percent of workers not driving alone is +0.35 with respect to residential population density, -0.10 with respect to income, and 0.2 with respect to the destination index. Not included in the model are working place parking policies, congestion along traveled routes, access to alternative modes, land use mix, trip length and cost, and transit service supply (destination index is a crude estimate of transit service availability), and non-work trips, all affecting mode choice and overall impact of these factors on travel behavior.

Similar analyses performed at the city level for the San Francisco Bay Area include a dummy variable for access to the rail rapid transit system, BART. The regression models suggest an elasticity of +0.35 for density, -0.25 for income, and +0.17 for BART access. Although BART access appears to have a significant impact on single vehicle occupancy, Kockelman concedes that the measure at the city level is coarse, pointing out that a study by Robert Cervero (1994) suggests that workplace parking policies, destination relative to station locations, and vehicle ownership are important factors in determining the mode choice for residents near the BART stations.

Kockelman (1997) investigates the link between urban form and travel behaviors, concluding that accessibility, land use mix, and land use balance are all statistically significant and influential to travel behaviors, including mode choice. In addition to the accessibility index described in the accessibility section previously, other measurements used are briefly introduced below:

$$\text{Entropy (Land Use Balance)} \quad \text{Entropy} = - \sum_j \frac{P_j \times \ln(P_j)}{\ln(J)}$$

Where P_j is the proportion of land development of the j^{th} type and J is the number of different types of land development, which include residential, commercial, public, offices and research sites, industrial, and parks for analysis of work trips, and residential, commercial, public, and parks for analysis of non-work trips. To avoid bias against small census tracts that do not have adequate area to allow a variety of land use types, a mean entropy is used:

$$\text{Mean entropy} = - \sum_k \frac{\sum_j \frac{P_{jk} \times \ln(P_{jk})}{\ln(J)}}{K}$$

where K is the number of actively developed hectares in a census tract, and P_{jk} the proportion of land use type j within a 0.8-km radius of developed area surrounding the k^{th} hectare. In addition, a “Dissimilarity Index” is another measure of land use mix, with a larger value indicating more types of land uses in a tract:

$$\text{Dissimilarity Index (Land Use Mix)} \quad \text{mix index} = \sum_k \frac{1}{K} \sum_i \frac{X_{ik}}{8}$$

where K is the number of actively developed hectares in the census tract, and X_{ik} is a dummy variable that assumes 1 if the central active hectares’ use type is different from that of a neighboring hectare, and 0 otherwise.

Linear regression models relating vehicle kilometers traveled (VKT) per household and different sets of predictors showed that the inclusion of the accessibility, entropy, and land use mix indicators significantly increase the value of R^2 when compared with models that only included household size, income per household member, and auto ownership. In the logit mode choice model, the inclusion of accessibility, population density, and employment density (all measured at both the origin and destination zones) also increased the psuedo- R^2 compared to models that only include trip distance, gender, age, race, number of workers, number of drivers, number of professional workers, auto ownership, household size, and member income as explanatory variables. Analysis of the elasticities of independent variables with respect to household VKT (total and non-work home-based) and mode choice shows that these variables are highly sensitive to accessibility (e.g. with an elasticity of -0.35 for non-work home-based VKT and 0.22 for walk/bike choice). Land use mix and mean entropy are also influential. The study concludes that accessibility is a far better predictor of VKT than density. While capable of identifying statistical correlation among travel behaviors and variables used in this study, the limitations of simple regression or logit models in determining the direction of causation have been recognized by the author, who contended that a structural model may be able to better explain the causation.

A study by Sun *et al.* (1998) takes a similar approach. Using the 1994 Portland Travel Survey data, density (population, employment, dwelling units), land use mix, accessibility, annual household income, household size, dwelling type, number of phone lines in a household, presence of a car phone, auto ownership, home ownership and year in current residence, number

of activities, and proximity to light rail are analyzed to determine their impact on household trip rates and VMT. Transit mode choice was not studied. The accessibility measure is given as:

$$Accessibility = e^{\ln(Y) - 0.175T + 0.009T^2 + 0.000009T^3}$$

Accessibility is computed for home-to-job and job-to-home trips and analyzed through ANOVA, linear regression, and sensitivity analysis. The authors proved that dwelling type was independent of household income. The authors utilized histograms to plot density and land use mix against income levels; results show that low-income households have a slightly higher concentration in high density areas and areas with better land use mix, and there is no fundamental difference in household income distribution in different types of neighborhoods. Regression analysis showed that density and land use balance make little difference in the number of daily trips but has a significant impact on house VMT. High density and high entropy both contribute to a reduction of VMT (by 19 percent and 45 percent, respectively).

In a study of Miami-Dade County in Florida, Messenger and Ewing (1996) establish two sets of simultaneous equations by place of residence and by place of work. The first set of equations establish three relationships: (1) transit share by place of residence to zero or one automobile households, land use mix/balance, and bus peak frequency; (2) zero or one automobile households to household income, logarithm of residential density (residential and employment), morning peak bus run time to downtown; and (3) logarithm of residential density to zero or one automobile households, logarithm of overall density, a variable rating street network resemblance to a grid system, and a dummy variable indicating proximity to the rail rapid transit. The second set of equations establish two relationships: (1) transit share to morning peak bus run time to downtown and zonal average seven-hour parking cost; and (2) the parking cost to the logarithm of overall density, a dummy variable indicating a zone is part of the downtown, and proportion of jobs in commercial and service sectors. The equations are simultaneously estimated by a full-information maximum likelihood method. The first set of equations (based on place of residence) has a better explanatory power (R^2 values ranging from 0.34 to 0.49) than the second set (based on place of work) (R^2 values ranging from 0.11 to 0.38). From the estimated equations, it was decided that the density needed to support a 25-minute bus headway was 8.4 dwelling units per acre (1.4 higher than that proposed by Pushkarev and Zupan) at the transit operator's minimum productivity and 19.4 dwelling units per acre at the system-wide average productivity. Additionally, different factors affect transit use at different trip ends. Bus mode share at trip origins is primarily a function of low automobile ownership, and secondarily of job-housing balance and transit service level, although job-housing balance has a small effect. This study finds that street configuration has no apparent effect on transit use, which disagrees with results from several other studies (Cervero and Gorham, 1995; Handy, 1992; Hsiao et al., 1997; Kockelman, 1997). This study finds that bus mode share at trip destinations is primarily a function of parking cost, overall density, and access to downtown; in addition, the trip end transit mode share models only explain a small portion of the variation in the data, indicating that other factors need to be identified.

In an attempt to determine if land use truly has a causal relationship with travel behavior or whether other socioeconomic, demographic, and transportation supply characteristics (also associated with land use) are the real determinants of travel behavior, Kitamura *et al.* (1997) conducted a household survey (including a three-day travel diary) in five neighborhoods in the

San Francisco Bay Area (SFBA) and investigated the travel behavior variables and a wide array of variables that are objectively or subjectively measured. The authors hold medium zonal income at a uniform level to control the effect of income on travel in the five neighborhoods, each approximately one square mile in size, while other characteristics such as land use density and mix are chosen to represent extreme conditions. Travel behavior is measured by number of trips, number of transit trips, number of non-motorized trips per person per day, and mode share. Data about the sites were collected, which included street characteristics (width, sidewalk, bike lanes, speed limits and other traffic control devices), public transit service (bus stops, service frequencies, etc.), location and type of commercial developments, parks and other public facilities, and general neighborhood characteristics (for detail see Kitamura *et al.* 1994). Dummy variables were used to represent access to rail transit, mixed land use, high density, presence of sidewalk, presence of bike lanes, backyard, available parking space, house ownership, sex, homemaker, student, professional, low education level, college education, graduate degree, high and medium personal income, respectively, apartment, single-family home, and responses to an array of questions related to reasons for staying in the area (no reason to move, streets pleasant to walk, cycling pleasant, good local transit, enough parking, and congestion problem). Measured variables include distances to nearest bus stop, rail station, grocery store, gas station, and park, respectively, and household size, number of persons over age 16, number of vehicles, number of vehicles per persons over age 16, household income, age, driver's license holding. Results of the regression models indicated that the variables had weak power to explain mode choice (R^2 values for all models are less than 0.14). Nonetheless, these results led the authors to conclude that parking availability negatively impacts the total number of person trips, and high density, proximity to parks and bus stops, access to rail transit stations, and presence of sidewalks encourages non-motorized travel. Furthermore, attitudes (pro-environment, pro-transit, suburbanite, automotive mobility, time pressure, urban villager, TCM, and workaholic) were determined to have a more significant impact on travel behavior than socioeconomic and land use characteristics. In fact, land use characteristics were found to be the weakest predictors. This is of particular interest because current modal split models do not include them as determinants of mode choice. These variables may also account for some of the unexplained variability in transit mode choice since we know people are not always as rational as assumed in logit models in which a trip maker is supposed to make a mode choice by maximizing the utility of the trip, which involves comparing the generalized costs for a trip via different means. On the other hand, it is impractical to include such attitudinal information in the models as such information is difficult, if not entirely impossible, to forecast.

The many facets of the relationship between urban form and transit were re-examined, explained, evaluated, and documented in a TCRP project for the purpose of helping make effective public transportation investment (Seskin, 1996). The TCRP project attempts to determine how urban form influences the demand for light rail and commuter rail transit and how transit influences land use. This study finds that urban structure, employment and residential densities, land use mix and urban design influence transit use. However, although land use mix and urban design are significant in explaining transit use, individual land use and design are not. Also, density is more powerful than land use mix and urban design in explaining transit use. On the other hand, the influences of transit on urban form were described by using the following four factors: property value, intensity of development, urban structure, and timing of development. First, accessibility to rail transit typically results in higher residential and commercial property values and rents. Second, although rail transit impacts transit station areas where transit confers a

distinct accessibility advantage on a location, the effects are varied among different networks. Third, both CBDs and subregional centers benefit from station-area development. Finally, major rail investments can accelerate development in station areas.

2.5 Neotraditional Neighborhood and Urban Design

Neotraditional neighborhoods are characterized by closely spaced street grids, high density, and location, often near street car tracks. Such neighborhoods are often older, built before the end of World War II. Whether urban design has any impact on transit use is a topic for debate; some argue that neotraditional neighborhood design encourages walking and transit use, while others disagree. Many studies have been conducted to determine the effect of urban design variables.

Handy (1992) studied shopping trips in the San Francisco Bay Area based on regional and local accessibility indices. The indices are based on the gravity model and are proportional to local attractions (or regional centers) and inversely proportional to an exponential function of travel time. Data from the 1980 Census and a regional travel survey of 7,235 households were aggregated at “superdistrict” level (34 in total) and used for analyses. Handy found that two to four more bicycle and walk trips were made by residents in the two districts that more closely resemble a neotraditional neighborhood than by those living in areas that are automobile oriented. She did not address the question of whether the non-motorized trips actually replaced some of the automobile trips or if the neotraditional neighborhood simply encourages more walk and bicycle trips. Her analytical approach (based on accessibility indices) has several weaknesses. Firstly, the use of superdistricts may mask the variability of accessibility in different parts of a zone. Secondly, local accessibility is easily affected by the choice of zonal boundaries, which are somewhat arbitrary. Finally, trip data do not distinguish convenience shopping (happening mostly locally) and comparison shopping (often at regional centers). Therefore it is impossible to evaluate how local and regional accessibility affect the travel patterns individually. Furthermore, socioeconomic and other factors that may affect travel patterns are not controlled in the study.

Cervero (1994) contends that many comprehensive studies on the relationship between built environment and travel behavior do not adequately control income and other extraneous factors. In his study of travel characteristics comparison using data from the San Francisco Bay Area and Los Angeles, he carefully paired “transit neighborhoods” and “auto neighborhoods” by a set of selection criteria (Cervero, 1994). The “transit neighborhoods” are defined as those that were built before 1945 along streetcar lines or a rail station on a grid street network. In contrast, the “auto neighborhoods” built after 1945 are not designed for transit and have no transit services, consisting of random street patterns (over 50% of intersections being “T” intersections or cul-de-sacs). To match the auto neighborhoods with the transit neighborhoods, criteria controlling income, transit services, topography, and size are used. For an auto neighborhood to match a transit neighborhood, there can be no more than ten percent variation of medium household income from that of the transit neighborhood; there should be transit services (type and density) comparable to the transit neighborhood; it should have similar topographic and natural conditions; and it should be located no more than four miles from the transit neighborhood. Additionally, an auto neighborhood also has to have a significantly lower percentage of four-way intersection cross roads and the net residential density lower than or equal to that of the transit

neighborhood. By applying these specific criteria, seven neighborhood pairs in SFBA and six in Los Angeles were identified.

A comparison of the SFBA paired neighborhoods revealed that while other demographic characteristics (such as mean vehicles per household, percent of white households, and mean age of residents) of the neighborhood pairs did not differ significantly, most auto neighborhoods had a higher auto ownership, produced many more drive-alone trips, had a lower transit use, and had much lower walk trip rates than transit neighborhoods. On average, transit neighborhoods generated about 70 percent more transit trips and 120 pedestrian/bicycle trips. This may be in part attributed to the fact that transit neighborhoods tend to have better transit service supplies (measured by daily VMT per acre). By comparison, the transit neighborhoods in Los Angeles did not enjoy the same significant amount of transit use or reduction of single occupancy driving. Cervero attributes this phenomenon to the overall strong auto orientation in Los Angeles such that the positive effects of transit neighborhoods are limited. To take his conclusion one step further, however, one may argue that the inconclusive relationship between transit neighborhoods and transit use in Los Angeles may be a result of inadequate transit services, which is affected by the built environment. In SFBA, transit services were much more concentrated in transit neighborhoods than in Los Angeles, perhaps due to the higher percentage of neighborhoods that qualified as transit neighborhoods. Not only did this attract people who desired to use transit to these neighborhoods, but it also allowed the transit providers to provide a good level of service in a large area and to increase overall accessibility via transit. In contrast, the dominance of auto-oriented neighborhoods in Los Angeles made it difficult to provide good transit services even to transit neighborhoods with the same efficiency and level of accessibility.

In addition to the comparison at the neighborhood level, Cervero ran a regression using data of the entire Los Angeles area on the percent of transit trips against variables such as gross residential density (households per acre), natural logarithm of household income, neighborhood type (auto or transit), and density interaction (product of residential density and neighborhood type). The R^2 is 0.55. According to the model, all variables are significant at a significance level of $p < 0.001$. In Los Angeles, everything else held constant, transit neighborhoods would generate 1.4 percent transit trips per every 1,000 households while those in SFBA would generate 5.1 percent transit trips. In Los Angeles, density had a stronger correlation to transit use than neighborhood type. Increasing density by one dwelling unit per acre would increase transit trips by two to four percent. The density-neighborhood type interaction term had a stronger effect in the SFBA than in Los Angeles. Work trips by transit averaged 8 percent more if density was 10 units per acre and 13.5 percent more when density was 30 units per acre. Cervero speculated that congestion might influence mode choice, though he did not control for it.

Cambridge Systematics, Inc. (1994) investigated the interactive effects of land use/urban design characteristics and transportation demand management (TDM) strategies on the transportation choices made by commuters in the Los Angeles metropolitan area. They suggested that employment site characteristics had an important influence on a person's willingness to commute by modes other than driving alone, and that employers' TDM strategies should be a function having a positive interactive effect in influencing an employee's choice of commute travel mode. Land use factors considered included accessibility to services, particularly the mix and intensity of services within walking distance of the workplace, and employment density. Urban design characteristics included architecture, streetscape, and site layout, and other factors contributing to

feelings of comfort and safety. The Principal Components method was applied to reduce a large number of land use/urban design variables into five composite variables: the perception of safety, the accessibility of services for walking on midday trips, the availability of convenience services, the mix of surrounding land uses, and the aesthetics of the area surrounding the work site. In addition, five categories were created for TDM strategies: financial incentives, flexible work schedules, assistance programs, award programs, and other strategies. ANOVA techniques are applied to test the interactive impacts of the five composite variables and the five categories of TDM strategies on mode choice. Some of the relevant results from the study include the following:

- When financial incentives are present, the greatest reduction in the drive alone share is realized in areas with an aesthetically pleasing urban character.
- When individually considered, TDM strategies have a larger influence on reducing the drive alone mode share than do land use characteristics.
- When jointly considered, there is a positive cumulative impact on increasing average vehicle ridership and reducing drive alone mode share when both financial incentives and one of the five land use component variables is present.
- Employer-provided transportation assistance programs have a small but statistically significant impact on reducing the drive alone modal share and increasing the average vehicle ridership at sites having a mix of convenience-oriented services.

A study of six Austin, Texas neighborhoods by Clifton and Handy (1998) also support the inconclusive effects of various urban form variables on travel behaviors, particularly on reducing automobile dependency. The study explores the motivations for travel as well as the patterns of travel. Travel surveys and focus groups were used to study the travel choices of residents of the six case study neighborhoods. The results suggest that the role urban form plays in travel behavior is not entirely straightforward, sometimes influencing travel choices directly, sometimes indirectly, sometimes influencing choices in the short term, sometimes in the long term, and sometimes not having any measurable influence on choices at all. In the end, it appears that certain land use policies can help to provide alternatives to driving, but that the reduction in driving is likely to be small.

2.6 Summary

Based on the literature review, the most common analysis approaches are multiple regression and time-series analysis. Structural models are capable of modeling interrelated variables and are appropriate if the calibration of a model is necessary to predict the modal split. Nevertheless, the objective of this research was to determine the relevance of different variables and their relative importance in transit use, which should be adequately served with regression methods. The weakness of the multiple regression approach, such as autocorrelation, non-normal distributions and invalidity of the IID normal assumption, is discussed. Although flawed in theory, multiple regression analysis is still widely applied since its concept is easy to understand. Thus, from a practitioner's point of view, this approach may be worth further investigation.

The factors that affect transit use can generally be classified into the following categories:

- (1) Transit level of service (LOS);
- (2) Accessibility;
- (3) Land use/urban design; and
- (4) Transit users' socioeconomic/demographic characteristics.

Transit service quality factors may be considered in terms of rider comfort, cleanliness and appearance, safety and security, pedestrian environment, amenities, headway, hours of service, parking spaces, reliability, service coverage, transfer, cost, etc. The literature generally supports the ability of transit systems with high-quality services to attract more users. However, some of the LOS factors affecting transit use often cannot be easily quantified. For example, most of the studies point out the conceptual importance of the LOS measures of effectiveness (MOEs) without further suggesting the appropriate approaches to quantify them. Some examples of such MOEs include passenger comfort in vehicle and in stations, security, and pedestrian environment. Consequently, it is still difficult to formulate LOS variables in models for estimating transit share. As a result, statistics such as PPMS were developed and implemented to determine the quality of transit service. Although tremendous efforts have been devoted to exploring the LOS factors that significantly affect transit use, contradictory findings were cited by different researchers and practitioners in different study areas (e.g., Ulberg (1982) and Vaziri *et al.* (1990) did not find service hours to be significant in contrast to other researchers) indicating that some of the transit LOS MOEs may not be transferable from one urban area to the other.

Similar to the LOS MOEs, some accessibility variables were demonstrated to significantly affect transit use by the research studies such as Metro (2000), Parsons Brinckerhoff (2000), Sun *et al.* (1998), and others. Numerous models are available for measuring accessibility, e.g., the accessibility index proposed by Richardson and Young (1982) and Kockelman (1997). Lately, with the advancements in GIS techniques and the more readily available data sources, more rigorous and detailed analyses (such as considering natural and manmade barriers that prevent access from residential areas to public transit stops), have been developed to better quantify accessibility. Nevertheless, the procedures that measure accessibility in a GIS environment are not yet simple menu-driven processes, requiring that someone trained in GIS acquire the necessary information and manipulate the related data. In addition, disaggregated accessibility variables are still not adopted in the traditional aggregated travel demand forecasting models and their effects on demand models' outcome are not yet verified.

Land use/urban design variables may include population density, employment density, land use mix, land use balance, etc. Although some of the literature considers the land use/urban design factor an important component affecting the travel mode selected by road users, its effects are not as significant as individual characteristics such as gender, ethnicity, and age (Loutzenheiser, 1997). Thus, how to properly define and measure the effects of land use/urban design on transit ridership appears to be a relatively new research topic that has attracted the attention of many researchers and practitioners. Different forms of entropy, e.g., those implemented in Frank and Pivo (1994) and Kockelman (1997), have been adopted by different studies in different study areas. Further research is needed to focus on land use factors and their links to travel behavior because their relationships have not yet been thoroughly investigated.

According to our literature review, research on socioeconomic/demographic characteristics of transit users (including age, education, ethnicity, gender, household structure, vehicle availability, and so on), results in significantly contradictory conclusions. For example, income was found to not significantly affect transit use by Sun *et al.* (1998), Loutzenheiser (1997), Gray and Thompson (1996), Spillar and Rutherford (1990), Holtzclaw (1990), and Nelson and O'Neil (1982) while the majority of literature concluded otherwise. Conflicting findings can also be observed for age, gender, and vehicle availability variables. In addition, a household's or an individual's socioeconomic/demographic characteristics are usually highly correlated, such as in the case of vehicle ownership and income. Additional precaution, such as applying the Principal Components method to reduce the number of variables (Cambridge Systematics, 1994), may be required before incorporating correlated factors.

One of the problems with many studies on the links between various factors and transit use is that the conclusions are based on claims of statistical significance, while in reality the contribution or impact of these factors is insignificant. For instance, some models presented in the literature only achieved a small R^2 (i.e., 0.1 or less) while all the model variables are statistically significant. From a modeling perspective, a contributing factor should be selected based not only on its statistical significance, but also on its power to explain the variations in transit ridership.

3. STUDY AREA SELECTION

This chapter presents the methodology by which the study area is selected. Data have been compiled for the statistical analyses to determine the contributing factors to transit ridership. The study is limited to a selected area due to time constraints and the significant effort required to collect, compile, and verify the quality of the necessary data for the entire southeast Florida Region or one or more of its associated counties. For instance, there were over 10,000 bus stops in Miami-Dade County, close to 4,000 of which were geocoded based on a printed list of bus stops, making it infeasible to use the entire data set. The decision was made to select a random sample from Miami Dade County for statistical analysis because this county has the largest share of transit users and more transit types compared to other counties. Data from Broward and Palm Beach counties were also used to test whether the same factors identified by the models developed for Miami-Dade County correlated with transit uses in each of those counties.

The rest of this chapter is organized as follows. Section 3.1 explains the rationale behind the selection of census tracts as the spatial analysis units. Sections 3.2 and 3.3 illustrate the statistics summarized from the household travel characteristics survey and transit onboard (TOB) survey. Section 3.4 addresses the elimination of neighborhoods based on existing transit service coverage in Miami-Dade County. Finally, Section 3.5 describes which neighborhoods are selected as the study areas and how they were selected.

3.1 Selection of Geographic Analysis Units

Traffic Analysis Zones (TAZ's) are considered too small to be used as the geographic units of analysis because the sample sizes are relatively small (see the discussion on survey data that follows). Instead, Census 2000 tracts are considered to be more appropriate because they are not so small that too few samples would be available in each tract or too large to lose too much of the variability of important land use characteristics.

Using census tracts as the analysis units has three advantages. First, census tract data are readily available and may be used to supplement existing data. Second, census data are updated every decade with great detail. This temporal information will allow longitudinal analysis to determine the effect of demographic, socioeconomic, and land use changes on transit use. Third and finally, as a commonly used unit of analysis, they allow the approach developed in this project to be applied to other urban areas.

Because sample units in different municipality and/or neighborhood boundaries are usually associated with different physical characteristics, 30 neighborhoods in Miami-Dade County have been randomly selected⁶. One hundred (100) census tracts were then selected randomly from those in the 30 neighborhoods. As will be seen later, since the census tracts are the actual analysis units, the 100 tracts falling within the 30 neighborhoods were selected for analysis.

Neighborhood and census tract boundaries were jointly considered in the process of study area selection to reflect different spatial characteristics within the county. The selection process has

⁶ The number 30 is commonly chosen as the minimum size of random samples that assure that the samples represent statistically the population to be studied.

involved careful examination of the transit service availability and survey data, which include the household survey and transit onboard (TOB) survey data from the Southeast Florida Regional Travel Characteristics Study (SEFRTCS). The procedure used in selecting the study areas is described as follows.

- First, auto dependent neighborhoods that had limited accessibility to existing transit services were identified. Since transit is not likely to be a feasible mode of choice for the people living in or commuting to these auto dependent areas, they were excluded. Transit accessibility for this purpose was determined based on the existing Miami-Dade Transit Agency (MDTA) bus route coverage (without jitney services since accurate information is not easily available). A quarter-mile buffer along the transit routes was created, which covered the potential transit service areas with walk accessibility to transit services. Neighborhoods that have either none or only a small proportion of their areas falling within the buffer zone were excluded.
- Next, 30 neighborhoods were randomly selected from the choice set created in the first step. Random selection ensures that the samples selected for analysis is not biased (e.g., based on the tracts with the highest number of samples). In other words, causes behind low transit use are of equal interest as those behind high transit use. The census tracts that were completely within the boundaries of 30 neighborhoods were then selected and included in the analysis. For the census tracts that were located partially inside the neighborhood boundaries, researchers' judgment determined if the tracts were to be included or not. The number of samples from the TOB survey as part of the SFRTC study will be used to estimate the transit share within each census tract. The TOB survey data provide detailed information of socio-economic attributes and travel characteristics of transit riders. The locations of sampled transit trips were geocoded for both trip ends. Since areas with more sampled transit trips at either the origin or destination ends reflect higher transit demand, the number of transit trip samples may be used as an indicator.

3.2 Household Travel Characteristics Survey

The primary purpose of the household travel characteristics survey was to collect data that could be utilized in travel demand model development. This survey collected data that characterize the demographic characteristics of households and travel patterns of household members. A total of 5,067 households with valid addresses completed the household survey of the SEFRTCS study. The detailed survey results categorized by household demographics and travel patterns may be found in Technical Report No. 1 of the SEFRTCS study. The following subsections describe the spatial distribution of the sampled households at the municipality and neighborhood levels in Miami-Dade County.

3.2.1 Sample Locations at Municipality Level

In Miami-Dade County, there are 31 incorporated municipalities. In total, 1,687 households (HHs) were sampled in Miami-Dade County in the household survey of the SFRTC study. Among these, only 88 HHs were recorded with transit trips. For the tri-county area, only 158 HHs were recorded with transit trips. Table 3-1 shows the number of samples recorded with transit trips, the total number of samples, and the ratio between these two statistics in each

municipality. The results in Table 3-1 indicate that transit trips are a very small proportion of total trips when data are aggregated to the municipality level.

Table 3-1. Dwelling Units (DUs) with Bus Trips as a Percent of Total Samples in Miami-Dade Municipalities

Municipality Name	Sampled DUs			Municipality Name	Sampled DUs		
	Total	With Bus Trips	%		Total	With Bus Trips	%
Unincorporated	875	35	4.00	Bay Harbor ISL	8	0	0.00
Miami	294	27	9.18	Sunny Isles	6	0	0.00
Hialeah	163	3	1.84	Surfside	6	1	16.67
Miami Beach	69	9	13.04	El Portal	5	0	0.00
North Miami	42	3	7.14	Sweetwater	5	0	0.00
Coral Gables	40	2	5.00	West Miami	4	0	0.00
Homestead	24	0	0.00	Medley	3	0	0.00
Pinecrest	23	0	0.00	N Bay Village	3	0	0.00
Opa Locka	17	0	0.00	Virginia Gdns	2	0	00.0
N Miami Beach	16	3	18.75	Florida City	2	0	0.00
Miami Springs	16	1	6.25	Biscayne Park	2	0	0.00
Aventura	14	1	7.14	Golden Beach	2	0	0.00
Miami Shores	13	1	7.69	Bal Harbour	0	0	0.00
South Miami	13	1	0.08	Islandia	0	0	0.00
Hialeah Gdns	10	0	0.00	Indian Creek	0	0	0.00
Key Biscayne	10	1	10.00				

3.2.2 Sample Locations at Neighborhood Level

Miami-Dade County agencies define the boundaries of 163 major neighborhoods. Neighborhood boundaries may or may not coincide with the municipality boundaries (i.e., some neighborhoods straddle two or more cities). Table 3-2 illustrates the number of household survey samples in the various neighborhoods of Miami-Dade County. Only the neighborhoods that have sampled households are included in Table 3-2. As shown in Table 3-2, of the 54 neighborhoods sampled, only nine have three or more households that recorded transit trips. Six out of the nine neighborhoods are located within the City of Miami. Additionally, Little Havana, with seven transit trips recorded, has the largest number of samples. Thus, with such small sample sizes within limited number of neighborhoods, the household survey data do not suit the purposes of this study.

3.2.3 Summary

The information from the household survey is not adequate to perform statistically significant analysis on factors that impact transit usage due to the small transit user response size. In fact, of the total 5,159 households surveyed, only 191 recorded transit trips, of which 113 were in Miami-Dade County; only 28 households recorded transit trips in the 100 selected census tracts that make up the study area. Such limited information is inadequate to perform statistical

analysis on the demographic, socioeconomic, and other lifestyle characteristics of the transit users. The SEFRTCS household survey data are therefore not useful for this study.

Table 3-2. Dwelling Units (DUs) with Bus Trips as a Percent of Total Samples in Miami-Dade Neighborhoods

Neighborhood Name	Sampled DUs			Neighborhood Name	Sampled DUs		
	Total	With Bus Trips	%		Total	With Bus Trips	%
Kendale Lakes	73	3	4.11	Hialeah – Area 6	14	1	7.14
Carol City	45	2	4.44	North Shore	13	1	7.69
West Flagler	40	5	12.50	Cutler	13	1	7.69
Little Havana	39	7	17.95	Norland	12	1	8.33
Kendall	39	1	2.56	Westwood Lakes	12	1	8.33
Tamiami	38	1	2.63	North Miami East	12	1	8.33
Richmond	36	4	11.11	Saga Bay E. Cutler	12	1	8.33
Westchester	36	1	2.77	Scott Lake	11	1	9.09
Civic Center	29	2	6.90	Granada	10	1	10.00
Hialeah – Area 1	29	1	3.45	West – Miami Shores	10	1	10.00
West Little River	27	5	18.52	Westview	10	1	10.00
Cutler Ridge	25	1	4.00	South Gables	8	1	12.50
North Bayfront	23	3	13.04	West Ave	7	3	42.86
South Miami Heights	22	2	9.09	South – N. Miami Beach	7	2	28.56
Buena Vista	21	2	9.52	West – N. Miami	7	1	14.28
Brickell	21	1	4.76	Miami Industrial	7	1	14.28
Hialeah – Area 5	21	1	4.76	Miami Springs – Area 1	7	1	14.28
Douglas Park	21	1	4.76	Surfside	6	1	16.67
Eastern Shores	21	1	4.76	West South Miami	6	1	16.67
West Kendall	19	2	10.52	Wynwood	6	1	16.67
Kendall North	19	1	5.26	Venetian Islands	5	1	20.00
Flamingo	17	3	17.65	Bunche Park	5	1	20.00
Sunset West	17	1	5.88	Interama	4	1	25.00
Grapeland	17	1	5.88	Central Downtown	4	1	25.00
South Naranja	16	2	12.50	Biscayne Point	4	1	25.00
Liberty City East	15	3	20.00	Key Biscayne – Bay Area	2	1	50.00
Aventura	14	1	7.14	Lake Lucerne	1	1	100.00

3.3 Transit Onboard Survey

Transit trips usually compose a small percentage of total person trips and data collected in a traditional household travel survey may not provide enough samples to adequately represent the trip patterns of transit users. In order to collect more transit ridership information, a TOB survey was completed as part of the SEFRTCS study. A total of 11,173 transit surveys were collected. Samples were collected from the fixed-schedule, fixed-route transit systems in Broward, Miami-

Dade, and Palm Beach counties. The routes and trips to survey were randomly selected from each system's weekday service schedule based on route level ridership. The detailed surveyed results categorized by household demographics and travel patterns may be found in Technical Report No. 2 of the SEFRTCS study.

The transit trip origins and destinations from the TOB survey have been geocoded. Table 3-3 shows the number of geocoded transit trip ends by origin, destination, or their combinations. After records without both trip ends geocoded were eliminated, 7,341 transit trips were found to have both valid origin and destination locations. Table 3-4 gives the number of transit trip interchanges between and within each of the three counties. As expected, Miami-Dade County had considerably higher transit ridership at both trip ends in comparison with Broward and Palm Beach counties. Among the trips with both ends geocoded, 3,884 trips started and ended within Miami-Dade County. Overall, 4,152 (56.6%) trips had their origins and 4,159 (56.7%) had their destinations in Miami-Dade County. Next, we discuss the spatial distribution of the samples from the TOB survey in Miami-Dade County by municipality and neighborhood.

Table 3-3. Geocoded Results for Transit Trips Ends

Trip End	O¹	D²	O / D³	O & D⁴
Number of Geocoded Trips	8,746	8,731	10,136	7,341
Percentage of Geocoded Trips	78.3	78.1	90.7	65.7

- 1 Origin
2 Destination
3 Origin or Destination
4 Origin and Destination

Table 3-4. Number of Transit Trips Interchanged between Tri-County Areas

Origin County	Destination County			Total
	Miami-Dade	Broward	Palm Beach	
Miami-Dade	3884	231	37	4152
Broward	209	2065	89	2363
Palm Beach	66	103	657	826
Total	4159	2399	783	7341

Table 3-5 shows the number of samples from the TOB survey of the SEFRTCS study by their geocoded locations in each municipality of Miami-Dade County. The origins and destinations of the sampled transit trips are located in 28 and 30 cities (including unincorporated areas), respectively. While the TOB samples were collected from the transit systems, the spatial locations for the origins and destinations of sampled transit trips reflect the transit usage in the sampled areas. For example, areas such as City of Miami, Miami Beach, Hialeah, North Miami, and North Miami Beach, where more transit services were provided, resulted in more transit trips from the TOB survey than the other cities. On the other hand, the areas with nearly no transit services at all resulted in zero or a small number of samples.

Table 3-5. Number of Samples for Transit On-Board in Municipalities of Miami-Dade

Municipality Name	Samples		Municipality Name	Samples	
	TOB_O¹	TOB_D²		TOB_O	TOB_D
Miami	1,431	1,280	Biscayne Park	5	1
Miami Beach	391	384	El Portal	1	5
Coral Gables	72	61	Golden Beach	0	1
Hialeah	163	179	Pinecrest	21	11
Miami Springs	10	4	Indian Creek	0	1
North Miami	120	138	Medley	2	3
N Miami Beach	166	176	N Bay Village	2	1
Opa Locka	9	12	Key Biscayne	6	12
South Miami	24	17	Sweetwater	44	49
Homestead	3	5	Virginia Gdns	2	3
Miami Shores	14	30	Hialeah Gdns	10	4
Bal Harbour	6	6	Aventura	70	137
Bay Harbor ISL	1	2	Islandia	0	0
Surfside	9	34	Unincorporated	1,534	1565
West Miami	9	5	Sunny Isles	24	26
Florida City	3	7			

¹ Origin location of transit trip

² Destination location of transit trip

Table 3-6 lists the number of sampled trips from the TOB survey with their origin trip ends located in the neighborhoods of Miami-Dade County. In total, 4,152 origin trip ends were sampled from 138 neighborhoods in Miami-Dade County. As in Table 3-6, Table 3-7 presents the number of sampled transit trips in the neighborhoods of Miami-Dade County for destination trip ends. A total of 4,159 destination trip ends were sampled from 142 neighborhoods. The neighborhoods including Central Downtown, Civic Center, West Flagler, Little Havana, North Bayfront, Omni-Boulevard, Liberty City East, Buena Vista, Auburdale, Shenandoah, Brickell, Wynwood, and Overtown were within the City of Miami boundary. Because these areas were known to have higher transit travel demand, MDTA provided them with more frequent service. Tables 3-6 and 3-7 illustrate that the number of samples from the TOB survey reflects the same tendency.

Table 3-6. Number of Trips with Origins in Neighborhoods of Miami-Dade

Name of Neighborhood	#	Name of Neighborhood	#	Name of Neighborhood	#
Central Downtown	209	Hammocks	24	South Miami Heights	8
Civic Center	189	Lake Lucerne - West	24	Cutler	8
Little Havana	146	Sunny Isles	24	South Miami City East	7
Liberty City East	138	East Kendall	23	South Miami City West	7
Brownsville & Liberty City W.	136	Hialeah - Area 3	23	Westwood Lakes	7
Flamingo	128	Richmond	23	Bal Harbor	6
West Flagler	116	Hialeah - Area 7	22	Dadeland	6
Buena Vista	108	North Grove	22	Granada	6
South - N. Miami Beach	104	Westview	22	Hialeah Gardens	6
North Shore	98	West - Miami Shores	22	Medley	6
Omni - Boulevard	98	West South Miami	21	South Naranja	6
Norland	95	Goulds - East	19	West Cutler Area	6
Miami Industrial	91	Hialeah - Area 4	19	Biscayne Park	5
Westchester	87	Ives Estate	17	Biscayne Point	5
Eastern Shores	85	Kendale Lakes	17	Miami Springs - Area 1	5
Shenandoah	82	North Miami Northwest	17	North Gables	5
Auburdale	79	West Sweetwater Est	17	Sunset East	5
Kendall	75	Miami Lakes	17	Sunset West	5
Aventura	70	Country Club Of Miami	16	Miami Springs - Area 3	4
Flagler Westside	69	Golden Glades - West	16	Naranja	4
West Kendall	67	Normandy Isle	16	East South Miami	3
North Bayfront	66	Scott Lake	16	Florida City	3
West Little River	65	So. Coral Terrace	16	Homestead	3
North Miami South	57	University	15	Key Biscayne - Bay Area	3
Sweetwater	53	Hialeah - Area 2	14	Key Biscayne - Oceanfront	3
Tamiami	52	Bay Shore	13	Miami Lakes West	3
NE Little River	48	Blue Lagoon & NW Tamiami Area	13	Miami Springs - Area 2	3
Oceanpoint	47	Calusa Area Or Lindgren Ac.	13	Saga Bay E. Cutler	3
Ojus	46	Douglas Park	13	Venetian Islands	3
Brickell	45	Metro-Lindgren Ac.	13	East Turnpike Area	2
Hialeah - Area 6	43	No. Coral Terrace	13	Gables Bayfront	2
Wynwood	41	South - Golden Glades	13	La Gorce Island	2
Nautilus	39	Bird Drive Basin	12	Leisure City Area	2
Central Gables	38	West Ave	12	NE Transitional Area	2
Opalocka North	36	West Miami	12	No Bay Village	2
West - N. Miami	35	Hialeah - Area 1	11	Olympia Heights	2
Golden Glades	34	North Miami East	11	Perrine	2
North Miami Northeast	33	Kendall North	10	West Tamiami	2
Carol City	30	Miami Shores	10	Bay Harbor Island	1
South Point	30	Catalina Lakes	9	El Portal	1
Hialeah - Area 5	29	Cutler Ridge	9	Palm Springs North	1
South Grove	28	Surfside	9	South Gables	1
Tamiami - Lindgren Ac.	28	Interama	9	West Homestead	1
Doral Area	26	Bunche Park	8	West Quail Roost	1
Grapeland	26	Key Biscayne - Co Parks	8	Homestead Base	1
Overtown & Spg Gardens	26	Opalocka City	8	Naranja - Priceton -East	1

Table 3-7. Number of Trips with Destinations in Neighborhoods of Miami-Dade

Name of Neighborhood	#	Name of Neighborhood	#	Name of Neighborhood	#
Central Downtown	244	Golden Glades - West	21	Metro-Lindgren Ac.	6
Civic Center	183	Nautilus	21	Perrine	6
Aventura	137	West - Miami Shores	21	Saga Bay E. Cutler	6
Flamingo	135	Ives Estate	21	Bal Harbor	6
Brownsville & Liberty City W	127	North Miami East	21	South Naranja	6
West Flagler	113	Hialeah - Area 3	18	Hialeah - Area 1	5
Westchester	112	Goulds - East	17	University	5
Eastern Shores	112	South Miami Heights	17	La Gorce	5
Miami Industrial	108	North Miami Northeast	16	Biscayne Point	5
Little Havana	107	South - Golden Glades	16	Sunset West	5
South - N. Miami Beach	104	Normandy Isle	16	Homestead	5
Norland	102	North Grove	15	El Portal	5
North Bayfront	97	Country Club Of Miami	15	Key Biscayne - Oceanfront	5
Omni - Boulevard	94	West Sweetwater East	15	Hialeah Gardens	4
West Little River	91	Westwood Lakes	14	Medley	4
Liberty City East	86	Miami Lakes	14	Tamiami - Lindgren Ac.	4
North Shore	84	Scott Lake	14	Calusa Area Or Lindgren Ac	4
Buena Vista	82	Kendale Lakes	13	Gables Bayfront	4
West Kendall	76	West Ave	13	East Turnpike Area	4
Kendall	75	Douglas Park	13	South Gables	4
North Miami South	72	No. Coral Terrace	13	Miami Shores	3
Auburdale	63	Cutler	12	Miami Springs - Area 3	3
Carol City	58	Granada	12	Naranja	3
Opalocka North	57	North Miami Northwest	12	Miami Lakes West	3
Shenandoah	56	Hialeah - Area 2	12	Miami Springs - Area 2	3
Sweetwater	54	Bunche Park	12	NE Transitional Area	3
Hialeah - Area 6	54	South Miami City West	12	Fisher Island	2
NE Little River	53	East Kendall	12	South Miami City East	2
Flagler Westside	52	Opalocka City	11	Venetian Islands	2
Oceanpoint	52	Bird Drive Basin	10	West Tamiami	2
Ojus	46	Catalina Lakes	10	Bay Harbor Island	2
South Point	44	North Gables	10	La Gorce Island	1
Tamiami	43	Interama	10	C-9 Basin Area	1
Hialeah - Area 5	42	Grapeland	10	No Bay Village	1
Brickell	36	Olympia Heights	10	Palm Springs North	1
Golden Glades	34	Blue Lagoon & NW	9	West Homestead	1
Surfside	34	So. Coral Terrace	9	West Quail Roost	1
West - N. Miami	32	Key Biscayne - Co Parks	9	Indian Creek Village	1
Central Gables	32	Hammocks	8	Port of Miami	1
Doral Area	30	West South Miami	8	Golden Beach	1
Wynwood	29	West Miami	7	North Redlands	1
Overtown & Spg Gardens	28	Cutler Ridge	7	Biscayne Park	1
Sunny Isles	26	Key Biscayne - Bay Area	7	Miami Springs - Area 1	1
Richmond	26	Lake Lucerne - West	7	Horse Country	1
Hialeah - Area 7	26	West Cutler Area	7	Dadeland	1
Westview	25	East South Miami	7	Palm Springs North	1
South Grove	23	Florida City	7		
Hialeah - Area 4	22	Bay Shore	6		

The TOB survey was performed to provide information on transit trip characteristics as well the socioeconomic attributes of transit users. The data collected may be used to formulate, calibrate, and validate existing and planned travel demand model structures. Since the number of households sampled with transit trips in the household survey of the SEFRTCS study was not

large enough, the results from the TOB survey had to be used in order to analyze the factors that impact transit use. Although the samples have been randomly collected from the transit systems based on their ridership and do not reflect the true transit share at either trip end, the number of samples at both transit trip ends is highly related to the area's transit services thus the transit demand. Because the purpose of this study is to search for the factors that have impacts on transit use and not to estimate the true transit mode share in an area, the number of TOB samples at tract level may be used as the dependent variable for regression analysis.

3.4 Existing Transit Services in Miami-Dade County

Figure 3-1 illustrates the 70 bus routes in Miami-Dade County that were operated by MDTA. Jitney transit services were not included. The 0.25- and 0.50-mile buffer zones covering the areas with possible walking distance to the existing transit systems are also shown. The buffer was created along the transit routes in the county, which included the Metrobus, Metrorail, and Metromover services.

In general, walking was the main mode for accessing transit. Based on the TOB survey results, 3,314 (79.8%) of the 4,152 trips involved walking to transit stops/stations from the origins, while 3,128 (75.2%) of the 4,159 trips involved walking to destinations after leaving the transit systems. Even for the people with trip origins or destinations outside of the 0.25-mile buffer, walking was still the dominant mode for accessing transit. For example, there were 258 and 274 trips with origins and destinations, respectively, located outside the 0.25-mile buffer. Among them, 193 (74.8%) and 206 (75.2%) trips respectively accessed transit by walking.

It is commonly assumed that for the majority of the transit users who walk to access transit services, the distance from either the trip origin or destination to transit stop or station is within 0.25 mile. According to the TOB survey results, 3,894 (93.8%) of the 4,152 trips had their origins located in the 0.25-mile buffer zone of the existing bus routes. Additionally, 3,885 (93.4%) of the 4,159 trips had their destinations located in the 0.25-mile buffer zone. By extending the buffer to 0.50-mile, more than 98% of the samples had trip ends located within the buffer area.

Transit, therefore, is not likely to be a competitive mode for areas outside the 0.25-mile buffer around transit routes. Overlaying the buffer layer and city and neighborhood boundary layers allows us to calculate the percentage of areas overlapped with the transit buffer zone. Table 3-8 shows the results of the overlay. Areas such as the City of Miami and the City of South Miami almost completely fall within the 0.25-mile transit buffer. For those municipalities that are not well covered, transit services are only available to some neighborhoods within each city. Figure 3-2 illustrates the neighborhood boundaries and the 0.25-mile transit buffer. The areas with 25% or more areas covered by the 0.25-mile transit buffer are highlighted. A total of 128 neighborhoods are included in the study area selection.

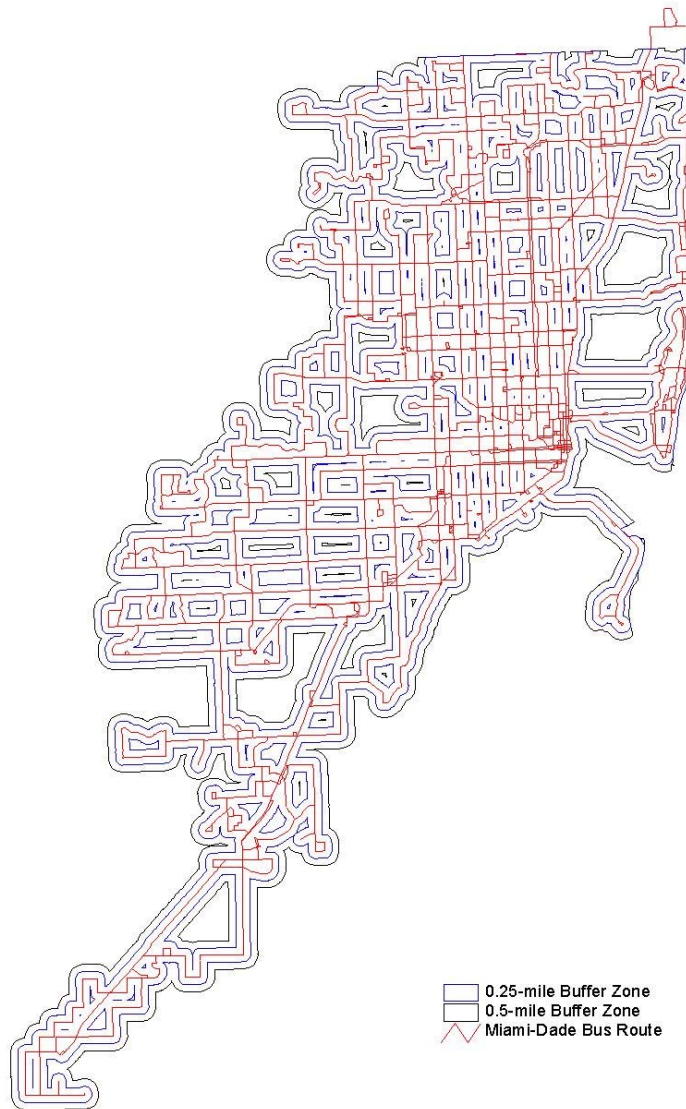


Figure 3-1. 0.25- and 0.5-Mile Buffer Zones along Miami-Dade Bus Routes

Table 3-8. Percentages of Areas Overlapped by Transit Route Buffers

Municipality Name	0.25-Mile	0.50-Mile	Municipality Name	0.25-Mile	0.50-Mile
Miami	91.4	96.6	Biscayne Park	43.9	92.0
Miami Beach	89.0	98.4	El Portal	95.6	100.0
Coral Gables	61.4	79.5	Golden Beach	99.9	100.0
Hialeah	76.8	91.7	Pinecrest	52.9	86.3
Miami Springs	66.9	92.7	Indian Creek	3.6	47.5
North Miami	75.7	95.2	Medley	13.5	29.5
N Miami Beach	88.0	99.8	N Bay Village	79.4	100.0
Opa Locka	72.1	92.7	Key Biscayne	81.3	100.0
South Miami	94.4	100.0	Sweetwater	85.7	100.0
Homestead	23.2	35.9	Virginia Gdns	100.0	100.0
Miami Shores	89.4	100.0	Hialeah Gdns	25.5	44.7
Bal Harbour	92.5	99.3	Aventura	74.0	98.0
Bay Harbor ISL	73.4	99.4	Islandia	0.0	0.0
Surfside	93.5	98.9	Unincorporated	7.67	10.9
West Miami	99.4	100.0	Sunny Isles	95.8	99.9
Florida City	61.1	94.8			

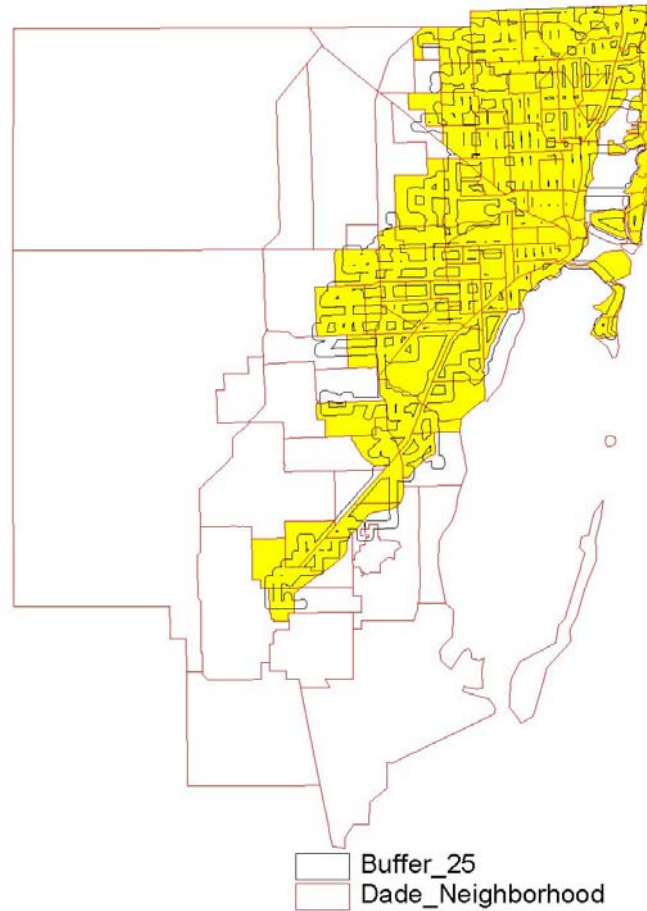


Figure 3-2. Neighborhoods Included in the Choice Set for Study Area Selection

3.5 Study Area Selection

The 128 neighborhoods obtained in Section 3.4 were first sorted by their neighborhood ID number in ascending order. Each neighborhood was then sequentially assigned a number ranging from 1 to 128, and 30 areas were randomly selected. Table 3-9 shows the neighborhoods that were selected and the percentage of their areas overlapping the 0.25-mile transit buffer. The percentages varied significantly, from a low of 36.2% to a high of 100%. The geographic locations of the selected neighborhoods are shown in Figure 3-3. Thirteen municipalities are represented by the selected neighborhoods, which include the City of Miami, Aventura, Sunny Isles Beach, North Miami Beach, Miami Shores, Opaloca, Hialeah, Miami Springs, Coral Gables, Miami Beach, Biscayne Park, Florida City, and Bay Harbor Island. The chosen areas depict a wide range of commuting tendencies and household characteristics, which reflect those in other parts of the county. The neighborhoods selected are expected to allow as much spatial and socioeconomic variation as possible to be considered in the analysis.

Table 3-9. Neighborhoods Selected for Study Areas

Neighborhood	Cover¹	Neighborhood	Cover
Central Gables	100.0%	North Miami Northwest	81.5%
West Ave	100.0%	Sunny Isles	76.9%
Civic Center	99.8%	Kendale Lakes	73.8%
Buena Vista	99.6%	Bay Harbor Island	68.5%
La Gorce	99.3%	Aventura	67.4%
Westview	99.1%	Ives Estate	65.4%
Douglas Park	96.3%	Westchester	64.8%
West Flagler	96.3%	Opalocka North	55.0%
South-Golden Glades	92.5%	Florida City	54.3%
Hialeah-Area 7	90.2%	North Miami East	51.6%
West-Miami Shores	89.5%	Perrine	50.0%
North Miami South	88.4%	Biscayne Park	42.8%
Carol City	87.2%	Doral Area	40.9%
Hialeah-Area 3	85.4%	Calusa Area or Lindgren Ac.	37.6%
Opalocka City	84.0%	Miami Springs-Area 1	36.2%

¹ Area covered by the 0.25-mile buffer.

Next, Census 2000 tract boundaries were obtained for the tracts that are inside the selected neighborhoods. These 100 tracts are used as the unit of analysis to perform detailed study on how variables such as urban form, household and trip characteristics, and transit service quality and quantity affect mode choice. A list of the tract numbers will be available when new census tract information is obtained.

The following steps summarize the procedure for selecting neighborhoods that make up the study area as stated above in detail:

- (1) A 0.25-mile buffer was created around bus routes based on the fact that most transit trips surveyed accessed transit by walking. This buffer area is assumed to be the “service area” accessible by walk mode.

- (2) Transit “service area” boundaries were overlaid on neighborhood boundaries to calculate the area of each neighborhood that fall within the service area.
- (3) Neighborhoods with less than 25 percent of their areas within the “service area” were eliminated, which resulted in 128 remaining neighborhoods.
- (4) The 128 neighborhoods were sorted by identification numbers in ascending order, and 30 were randomly selected.
- (5) 100 Census tracts falling within the selected neighborhoods were selected.

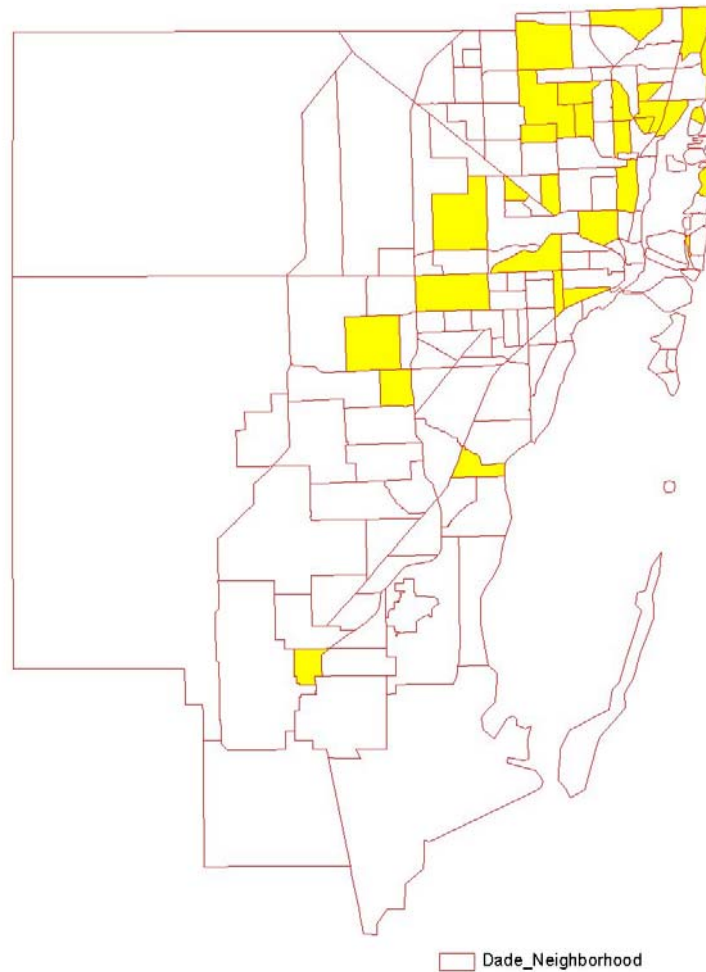


Figure 3-3. Neighborhoods Selected for Study Areas

4. DATA COLLECTION AND PROCESSING

4.1 Introduction

Multiple regression analysis has been chosen as the methodology for determining the factors affecting transit usage. Data from a transit-on-board (TOB) survey conducted in the southeast Florida area, data on transit level of service, accessibility, land use, and socioeconomic characteristics of transit users at aggregate levels were used to develop the model. More than 170 variables were compiled using GIS.

As discussed in Section 3.3, because the transit trips sampled in the SEFRTCS household survey were too few to provide adequate information on mode shares, data from the TOB survey were therefore used for the analysis. Since the number of samples at both transit trip ends were highly related to the area's transit services (in other words, usage), the dependent variable was defined as the ratio between the number of trip ends (either production or attraction or both) in the transit buffer zone in a census tract and the tract population.

For some of the variables, data of varying degrees of detail are available. For instance, land use data in Miami-Dade County are available at three levels of detail: single-family and multi-family dwelling units by TAZ, land use represented as polygons aggregated from property parcels, and individual property parcels, which offer the most detailed information about land use and the spatial distribution of different land uses. As a result, different data sets were compiled for the three kinds of land use data, which were used to develop different regression models.

While most of the analyses were performed using data compiled for the buffer areas in census tracts, tract level data were also compiled. For tract level analysis, the dependent variable was the ratio of the number of trip ends (either production or attraction or both) in a census tract to the tract population. To compare results from Miami-Dade County with those from Broward and Palm Beach counties, available data from the latter two counties were also compiled.

The independent variables may be classified generally into the following categories:

- Transit level of service (LOS);
- Land use;
- Accessibility; and
- Transit users' socioeconomic/demographic characteristics.

In the following sections, the collection and processing of data for independent variables are described. Most of the original data in GIS format or database format have been obtained from Miami-Dade County Information Technology Division (ITD).

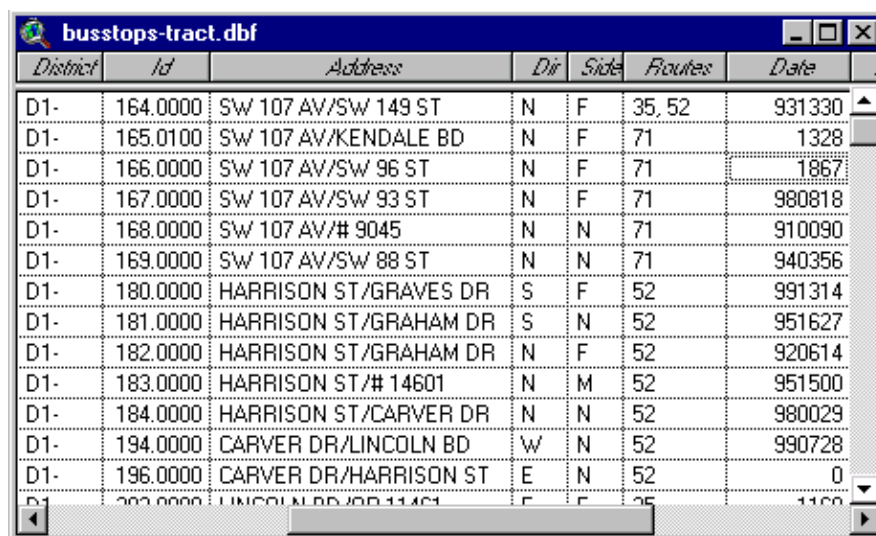
4.2 Transit Level of Service

Transit LOS is an important factor in determining transit use. However, transit LOS is also difficult to measure because an array of factors ranging from ease of access/egress, transfer and wait time, to ride comfort and security may affect whether a traveler will use transit. Transit service quality varies from one urban area to another, from mode to mode, and even from route

to route. Based on the results of a survey on transit users' satisfaction with the largest transit systems in six urban areas in Florida, hours of service, routes, and headways were identified by the transit users as the biggest concerns among the 22 factors included in the survey (Cleland *et al.* 1997). Additionally, bus ride comfort, printed schedules, safety, and vehicle cleanliness were identified as potential factors. For this project, we included headway, number of daily bus runs, and access to Metrorail stations as independent variables as they are easy to quantify and data are easily available.

The following were used to compile the data for transit LOS variables:

- Metrorail stations. This information was obtained from ITD in GIS format. Stations were represented as points.
- Bus routes. A 2000 bus route GIS theme was created based on the Miami-Dade County 2000 bus route maps.
- Bus stops. At the time of this report, the bus stop information from the ITD had not been updated since 1993. The Miami-Dade Transit Agency (MDTA) staff at the Division of Service Planning and Scheduling confirmed that a plan was being considered to re-inventory the bus stops using the Global Positioning System (GPS) technology. However, no RFP had been issued, and no completion date was estimated. Upon request, MDTA provided a printout of a list of the bus stops in the county (2000). Figure 4-1 shows the kinds of bus stop information available. Based on the address information, a bus stop GIS theme was created. Approximately 4,000 bus stops were geocoded, which were located in either the study area or the 1/4-mile buffer surrounding the study area.



District	Id	Address	Dir	Side	Routes	Date
D1-	164.0000	SW 107 AV/SW 149 ST	N	F	35, 52	931330
D1-	165.0100	SW 107 AV/KENDALE BD	N	F	71	1328
D1-	166.0000	SW 107 AV/SW 96 ST	N	F	71	1867
D1-	167.0000	SW 107 AV/SW 93 ST	N	F	71	980818
D1-	168.0000	SW 107 AV/# 9045	N	N	71	910090
D1-	169.0000	SW 107 AV/SW 88 ST	N	N	71	940356
D1-	180.0000	HARRISON ST/GRAVES DR	S	F	52	991314
D1-	181.0000	HARRISON ST/GRAHAM DR	S	N	52	951627
D1-	182.0000	HARRISON ST/GRAHAM DR	N	F	52	920614
D1-	183.0000	HARRISON ST/# 14601	N	M	52	951500
D1-	184.0000	HARRISON ST/CARVER DR	N	N	52	980029
D1-	194.0000	CARVER DR/LINCOLN BD	W	N	52	990728
D1-	196.0000	CARVER DR/HARRISON ST	E	N	52	0
D1-	202.0000	LINCOLN BD/BB 11461	E	F	35	1160

Figure 4-1. Sample of the Bus Stop Information Obtained from MDTA

- Bus schedules. To derive indicators of bus LOS, bus schedule information for each route was obtained from the MDTA published schedule and a database was constructed. In addition to headway information, the database included number of bus runs on each route

(see Figure 4-2). Headway information was coded for eight weekday time periods: 5 to 6 am, 6 to 9 am (defined as AM peak hours), 9 to 11 am, 11 am to 1 pm, 1 to 3 pm, 3 to 6 pm (defined as PM peak hours), 6 to 12 pm, and 12 to 5 am.

ROUTE	NO_TRIP	AM_5_6AM	AM6_9AM	AM9_11AM	AM11_1PM	PM1_3PM	PM3_6PM	PM6_12PM	PM12_5AM	TRIP_5_6AM	TRIP_6_9AM	TRIP_9_11A
104	17	65	57	60	60	60	60	61	0	1	3	2
9	45	22	13	40	40	40	15	37	0	2	12	3
91	17	57	60	60	60	41	54	54	0	1	3	2
A	42	11	20	20	20	20	20	19	0	1	9	6
BisMAX	30	0	15	15	0	0	15	15	0	0	12	3
BLOC	40	21	20	28	31	29	21	32	0	2	11	3
C	51	20	20	20	20	20	20	33	0	3	9	6
CORMAX	31	19	23	45	45	45	19	25	0	1	8	3
95	35	0	7	15	0	0	9	0	0	0	16	2
B	42	0	14	21	30	20	17	30	0	0	10	6
L	88	10	10	12	12	10	10	29	31	4	18	10
W	28	0	24	24	24	24	24	60	0	0	3	5

Figure 4-2. Example of the Headway Database (Partial)

From the headway database as well as the bus route GIS theme, data for the variables listed in Table 4-1 were compiled for each census tract. The methods used to compile the data are described below.

Table 4-1. Transit LOS Variables

Variable Name	Description
<i>MetroRailRatio</i>	The ratio of the area of 1/4-mile buffer surrounding the Metrorail stations to that of the bus buffers in a tract
<i>ServiceArea%</i>	Percentage of tract area served by transit based on 1/4-mile buffers around bus stops
<i>Peakhdway</i>	Composite average peak hour headway for a tract
<i>WPeakhdway</i>	Composite peak hour headway for a tract weighted by service area
<i>Avghdway</i>	Composite average daily headway in a tract
<i>WAvghdway</i>	Composite daily headway for a tract weighted by service area
<i>Dailybus</i>	Unweighted average number of bus runs per stop in a tract
<i>WDailybus</i>	Average number of bus runs per stop in a tract weighted by service area

Metrorail Service Area Ratio (*MetroRailRatio*)

The Metrorail Service Area Ratio is the ratio of the area within the 1/4-mile buffer surrounding the Metrorail stations to that of the bus buffers in the tract. It is an indicator of high-level transit services. The higher this ratio, the larger the proportion of Metrorail service to the tract, thus the better the transit service level in a tract. This variable was calculated by dividing the area of the 1/4-mile buffer of the Metrorail stations by the buffer area of the transit stations or bus stops in a tract.

Percent of Service Area (*ServiceArea%*)

Percent of Service Area is the percentage of tract area served by transit. It was calculated by dividing the area in a tract that was covered by the 1/4-mile buffers by the total area of a tract.

Peak Headways (*Peakhdway*, *WPeakhdway*)

The headway measures are indicators of transit LOS in an entire tract area. Since a tract might have multiple bus stops, at which multiple bus routes operated, these measures were computed in two steps. First, a composite headway CH_j for bus stop j was calculated, which combined the headways of different routes:

$$CH_j = \frac{1}{\sum_{i=1}^{N_j} \frac{1}{h_{ij}}} \quad [1]$$

where h_{ij} = peak hour headway for the i^{th} bus route operated at bus stop j . Peak hour was defined as morning peak between 6 am and 9 am; and
 N_j = the number of routes operated at bus stop j .

To derive the headway measurements for the entire tract area, the stop composite peak hour headways were combined again to arrive at the weighted composite peak hour headway (*WPeakhdway*) and unweighted composite peak hour headway (*Peakhdway*) as follows:

$$WPeakhdway_k = \frac{SA_k}{\sum_{j=1}^{N_k} \frac{1}{CH_{jk}} \times SA_{jk}} \quad [2]$$

$$Peakhdway_k = \frac{1}{\sum_{j=1}^{N_k} \frac{1}{CH_{jk}}} \quad [3]$$

where CH_{jk} = composite headway for bus stop j in tract k ;
 SA_{jk} = area of the 1/4-mile buffer around bus stop j in tract k ;
 SA_k = area of non-overlapping buffer zones in tract k ; and
 N_k = number of bus stops in tract k .

Average Daily Headways (*Avghdway*, *WAvghdway*)

Avghdway is a measure of the average headway during a weekday in a tract. The formula used to calculate this measure is similar to [1] and [3], except that in place of peak hour headway h_{ij} , the daily average headway is used. Daily average headway was derived by dividing the length of the service period of a route by the total number of bus runs on that route. Therefore, if a route operates 30 bus ran on a weekday and the service period was 15 hours, the daily average headway would be 30 minutes.

Similarly, *WAvghdway* is a measure that combines the daily average headway at individual bus stops for an entire tract. The difference between this measure and *Avghdway* was that the bus stop daily headways were weighted by the corresponding service areas of the stops. Formulae

[1] and [2] were used to calculate the value of the average headway variables, in which h_{ij} was replaced by the daily average headway at bus stop j for route i .

Average Bus Runs Per Stop (*Dailybus*, *WDailybus*)

For a bus stop where multiple routes operated, the numbers of bus runs of all routes were added together to arrive at a daily total bus runs for the bus stop. At the tract level, the daily bus runs at different bus stops were simply averaged to obtain the unweighted average daily bus runs (*Dailybus*), and the weighted average of the stop daily bus runs became the weighted average daily bus runs (*WDailybus*) as follows:

$$WDailybus = \frac{\sum_{i=1}^{N_s} stop\ daily\ runs(i) \times SA_i}{Total\ Service\ Area}$$

where $stop\ daily\ runs(i)$ = total buses stopping at stop i ;
 SA_i = service area of stop i ; and
 N_s = number of stops in a tract.

The above measurements are transit LOS measures at the tract level. They consider multiple bus routes at a bus stop, and the service area of each bus stop. A larger service area will improve the LOS measures at the tract level.

4.3 Land Use

Land use variables may include population density, employment density, land use mix, land use balance, and so on. Table 4-2 shows the land use variables developed. The numbers (1, 2, 3, and 4) that follow the variable names indicate the different levels of aggregation and methods for calculating the values. Generally the four methods are different in terms of (1) whether a variable is compiled for the buffer area in a tract or an entire tract and (2) whether population and employment distributions are based on property information, land use, or assumed to be evenly distributed across a tract:

- Level 1 – data were compiled for transit stop buffer areas in a tract, and population and employment were assumed to be evenly distributed in the same land uses.
- Level 2 – data were compiled for transit stop buffer areas in a tract, and population was distributed based on property information while employment was assumed to be evenly distributed in the same land uses.
- Level 3 – data were compiled for transit stop buffer areas in a tract, and both population and employment were assumed to be evenly distributed in the TAZs in the tract.
- Level 4 – data were compiled for an entire TAZ instead of just for the buffer area in the tract with the assumption that population and employment were evenly distributed in the TAZs in the tract.

The methods for calculating the values of each variable at the above levels of aggregation are referred to as the GIS Land Use Method (Level 1), Property Method (Level 2), Buffer Method (Level 3), and Tract Method (Level 4). These methods and variables given in Table 4-2 are described below.

Table 4-2. Land Use Variables

Variable Name	Area	Description
<i>JOBS HH1, 2, 3</i>	Buffer	Jobs-housing balance in buffer area of a census tract
<i>JOBS HH4</i>	Tract	Jobs-housing balance in a census tract
<i>JOBS HH^m1, 2, 3</i>	Buffer	Modified jobs-housing balance in buffer area of a census tract
<i>JOBS HH^m4</i>	Tract	Modified jobs-housing balance in a census tract
<i>AvgEntropy 1-3</i>	Buffer	Land use mix in buffer area of a census tract
<i>AvgEntropy 4</i>	Tract	Land use mix in census tract
<i>AvgParcelSFSize1-3</i>	Buffer	Average SF parcel size in buffer area of a tract
<i>AvgParcelSFSize4</i>	Tract	Average SF parcel size in a tract
<i>AvgParcelMFSize1-3</i>	Buffer	Average MF parcel size in buffer area of a tract
<i>AvgParcelMFSize4</i>	Tract	Average MF parcel size in a tract
<i>AvgParcelSize1-3</i>	Buffer	Average parcel size of single- and multi-family properties in buffer area of a tract
<i>AvgParcelSize4</i>	Tract	Average parcel size of single- and multi-family properties in a tract
<i>SFDUDensity1-2, 3</i>	Buffer	Single-family dwelling unit density in buffer area of a tract
<i>SFDUDensity4</i>	Tract	Single-family dwelling unit density in a tract
<i>MFDUDensity1-2, 3</i>	Buffer	Multi-family dwelling unit density in buffer area of a tract
<i>MFDUDensity4</i>	Tract	Multi-family dwelling unit density in a tract
<i>TotalDUDensity1-2, 3</i>	Buffer	Total dwelling unit (SF plus MF) density in buffer area of a tract
<i>TotalDUDensity4</i>	Tract	Total dwelling unit (SF plus MF) density in a tract
<i>SFPopDensity1, 2, 3</i>	Buffer	Single-family population density in buffer area of a tract
<i>SFPopDensity4</i>	Tract	Single-family population density in a tract
<i>MFPopDensity1, 2, 3</i>	Buffer	Multi-family population density in buffer area of a tract
<i>MFPopDensity4</i>	Tract	Multi-family population density in a tract
<i>TotalPopDensity1, 2, 3</i>	Buffer	Total population density in buffer area of a tract
<i>TotalPopDensity4</i>	Tract	Total population density in a tract
<i>ComEmpDensity1-2, 3</i>	Buffer	Commercial employment density in buffer area of a tract
<i>ComEmpDensity4</i>	Tract	Commercial employment density in a tract
<i>ServEmpDensity1-2, 3</i>	Buffer	Service employment density in buffer area of a tract
<i>ServEmpDensity4</i>	Tract	Service employment density in a tract
<i>IndEmpDensity1-2, 3</i>	Buffer	Industrial employment density in buffer area of a tract
<i>IndEmpDensity4</i>	Tract	Industrial employment density in a tract
<i>TotalEmpDensity1-2, 3</i>	Buffer	Total employment density in buffer area of a tract
<i>TotalEmpDensity4</i>	Tract	Total employment density in a tract
<i>TotalDensity1, 2, 3</i>	Buffer	Total employment plus population density in buffer area of a tract
<i>TotalDensity4</i>	Tract	Total employment plus population density in a tract

Jobs-Housing Balance ($JOBS_HH1$, 2, 3, 4 and $JOBS_HH^m1$, 2, 3, 4)

Ewing *et al.* (1995) developed the following indicator for jobs-housing balance:

$$JOBS - HH_i = \frac{|E_i - P_i|}{E_i + P_i} \quad [4]$$

where $JOBS_HH_i$ = jobs-housing balance index for zone i ;

E_i = employment size in zone i ; and

P_i = population size in zone i .

The value of this variable lies between 0 and 1, with 0 indicating a match between employment and population in a zone, while 1 indicating a pure residential area ($E_i = 0$) or a nonresidential area ($P_i = 0$).

To calculate $JOBS_HH1$, which is a land use mix measure for the buffer area in a tract, population and employment in the buffer area are needed. For this purpose, the land use GIS data were used. It was assumed that population and employment were uniformly distributed in the corresponding land uses in a tract (e.g., single-family population evenly distributed in single-family land use areas, and commercial employment evenly distributed in commercial land use areas). The buffer polygons in a tract were then overlaid with the land use polygons (the 1998 land use theme). The total employment and population inside the buffer areas were proportioned based on the percentages of the corresponding land use areas that fell within the buffer area. The $JOBS_HH1$ was then calculated using the employment and population in the buffer area in a census tract using the above formula. We will refer to this method of estimating population or employment in buffer areas as the GIS Land Use Method (Level 1).

The population data used for computing $JOBS_HH2$ were obtained using Property Method, which involved the use of population estimated for the buffer area based on property information. The employment data were calculated using the GIS Land Use Method (Level 1).

The third method (Level 3: Buffer Method) was used to calculate $JOBS_HH3$, also a measure of land use mix in the buffer area in a tract, does not use the land use data. Beginning with TAZs in a census tract, the employment and population of each TAZ were distributed to the buffer areas within the TAZ based on the ratio of buffer area to the total TAZ area. Uniform density throughout a TAZ was assumed. Once employment and population for the buffer area in each TAZ were obtained, they were summed up over all TAZs in the tract, and $JOBS_HH3$ was calculated using the same formula [4].

The value of $JOBS_HH4$, a land use mix measure for an entire tract, was computed using a fourth method (Tract Method), which involved the use of employment and population of the tract in Formula [4] as opposed to the first three methods.

Noting that the labor population is usually smaller than the total population, Formula [4] may overestimate the need for employment thus the jobs-housing balance index. To account for this

fact, the following formula, which is a modified version of Formula [4], was used to derive another set of jobs-housing balance indices:

$$JOBS - HH_i^m = \frac{|E_i - 1.5 \times DU_i|}{E_i + 1.5 \times DU_i}$$

where $JOBS - HH_i^m$ = modified jobs-housing balance indicator; and
 DU_i = number of dwelling units in a tract i .

Using the above formula, $JOBS_HH^m 1, 2, 3, 4$ were calculated for the four methods (Land Use, Property, Buffer, and Tract) described above.

Land Use Mix (*AvgEntropy* 1-3, 4)

Entropy is a measure of land use mix and was computed in this study using a method similar to that described in (Kockelman 1997). This method involved overlaying grid cells of 528 feet by 528 feet (see Figure 4-3) over the land use. The mean entropy value for a single cell was determined with the following formula:

$$Cell\ Entropy = - \sum_{j=1}^J \frac{p_j \ln(p_j)}{\ln(J)} \quad [5]$$

where P_j = proportion of land development of the j th type; and
 J = number of different types of land development, which include residential, commercial, public, offices and research sites, industrial, and parks ($J = 6$).

The value of entropy ranges between 0 and 1, with 0 indicating single land use and 1 indicating the highest possible mix.

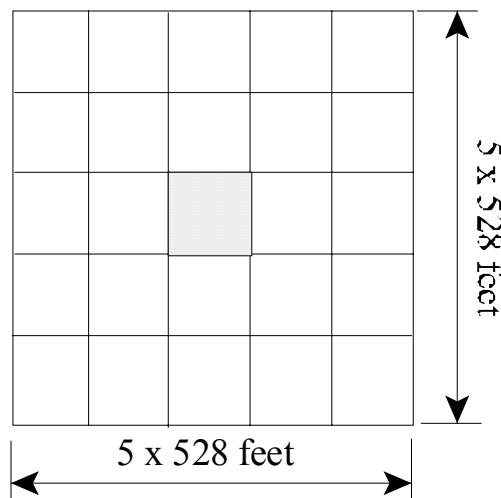


Figure 4-3. Cells Used for Calculation of Mean Entropy

To consider the effects of land use mix in near-by cells, cell mean entropy was used instead to measure the land use mix in a cell. The entropy of the 24 surrounding cells was combined with that of the center cell as follows:

$$\text{Cell mean entropy} = - \sum_k \frac{\sum_j \frac{P_{jk} \times \ln(P_{jk})}{\ln(J)}}{K} \quad [6]$$

where P_{jk} = proportion of land development of the j th type in cell k ; and
 K = number of cells used in the calculation. Here K was 25.

For any given zone, the zonal mean entropy was calculated as:

$$\text{Zonal mean entropy} = \frac{\sum_{i=1}^N (\text{cell entropy of zone } i)}{N}$$

where N is the number of cells in the zone. *AvgEntropy1-3* was calculated for the buffer area in a tract, while *AvgEntropy4* was derived for the entire tract.

Average Parcel Size (*AvgParcelSFSize1-3*, *AvgParcelSFSize4*, *AvgParcelMFSize1-3*, *AvgParcelMFSize4*, *AvgParcelSize1-3*, *AvgParcelSize4*, *AvgParcelComSize1-3*, *AvgParcelComSize4*, *AvgParcelIndSize1-3*, *AvgParcelIndSize4*, *AvgParcelSerSize1-3*, *AvgParcelSerSize4*, *AvgParcelEmpSize1-3*, and *AvgParcelEmpSize4*)

Average parcel size is an indicator of land use development intensity. In urban areas where high-density development has occurred, the parcel sizes tend to be small. Parcel sizes were computed for single-family, multi-family, and all dwelling units using the property GIS data. For this purpose, the 1999 Miami-Dade County Property Tax Appraiser's database was used because it contained information of the parcel size of each property. The parcel sizes of the single-family, multi-family, and all residential properties were averaged, respectively, to derive the values for *AvgParcelSFSize1-3* and *AvgParcelMFSize1-3*. *AvgParcelSize1-3* was calculated for the buffer areas in a tract, while *AvgParcelSFSize4*, *AvgParcelMFSize4*, and *AvgParcelSize4* were for a census tract. Variables *AvgParcelComSize1-3*, *AvgParcelComSize4*, *AvgParcelIndSize1-3*, *AvgParcelIndSize4*, *AvgParcelSerSize1-3*, *AvgParcelSerSize4*, *AvgParcelEmpSize1-3*, and *AvgParcelEmpSize4* were similarly computed for parcel size of properties that were of commercial, industrial, and service types and for all nonresidential properties.

Densities (*SFDUDensity1-2*, *3*, *4*, *MFDUDensity1-2*, *3*, *4*, *TotalDuDensity1-2*, *3*, *4*, *SFPopDensity1*, *2*, *3*, *4*, *MFPopDensity1*, *2*, *3*, *4*, *TotalPopDensity1*, *2*, *3*, *4*, *ComEmpDensity1-2*, *3*, *4*, *SerEmpDensity1-2*, *3*, *4*, *IndEmpDensity1-2*, *3*, *4*, *TotalEmpDensity1-2*, *3*, *4*, *TotalDensity1*, *2*, *3*, *4*)

The density variables were calculated for single-family dwelling units, multi-family dwelling units, total dwelling units, single-family population, multi-family population, total population,

commercial employment, service employment, industrial employment, total employment, and total employment plus population. Four methods were used to compute the densities of these statistics that measured population, dwelling units, and employment, which were available at the TAZ level. These methods were the Land Use Method, the Property Method, the Buffer Method, and the Tract Method. Variables with names followed by number "1" were computed using the Land Use Method. Variables with the names followed by a "2" were computed using the Property Method. Variables with the names followed by a "3" were computed using the Buffer Method. Variables with the names followed by a "4" were computed using the Tract Method.

4.4 Accessibility

Similar to the LOS measures of effectiveness (MOEs), accessibility variables have been shown to significantly affect transit use by some studies (Metro 2000, Parsons Brinckerhoff 2000, Sun *et al.* 1998). Numerous models are available for measuring accessibility, such as the accessibility indices proposed by Richardson and Young (1982) and Kockelman (1997).

Accessibility variables are measures of pedestrian accessibility as well as combinations of availability of opportunities and mobility that permits those opportunities to be reached. The pedestrian accessibility is a measure of walking distance to transit stops, while the accessibility measures that combine opportunities and mobility attempt to reflect how well the transit and highway systems serve residential and employment needs. These two kinds of measures are described below in detail.

Walk Accessibility (*AvewalkSFdis*, *AvewalkMFdis*, *Avewalkdis*, *AvewalkComdis*, *AvewalkINDdis*, *AvewalkSERdis* and *AvewalkEMPdis*)

This group of variables includes the average walking distances from single-family, multi-family, all dwelling units, and commercial, industrial, service and all nonresidential properties, respectively, in the buffer areas of a tract to the nearest bus stops following the street network. The locations of the properties were available from the GIS property data. To arrive at the average walking distance, the distance from each of the appropriate residential or nonresidential property was calculated using ArcView Network Analyst. The average distance was then calculated for the entire tract.

Regional Accessibility (*Accessh1sc*, *Accessh1em*, *Accessh1pop*, *Accessh2sc*, *Accessh2em*, *Accessh2pop*, *AccessTWSC1*, 2, 3, 4, *AccessTWEm1*, 2, 3, 4, *AccessTWPop1-2*, 3, 4, *AccessBSC1*, 2, 3, 4, *AccessBEm1*, 2, 3, 4, *AccessBPop1-2*, 3, 4, *With1SC1*, 2, 3, 4, *With2SC1*, 2, 3, 4, *With1Em1*, 2, 3, 4, *With2Em1*, 2, 3, 4, *With1Pop1-2*, 3, 4, *With2Pop1-2*, 3, 4, *Bth1SC1*, 2, 3, 4, *Bth2SC1*, 2, 3, 4, *Bth1Em1*, 2, 3, 4, *Bth2Em1*, 2, 3, 4, *Bth1Pop1-2*, 3, 4 and *Bth2Pop1-2*, 3, 4)

Accessibility is measured for highway and transit separately. For highway, two zone-to-zone travel skims are obtained using the 1999 Miami-Dade highway network: free-flow skim (FFSkim) and free-flow skim plus the time equivalent for toll (FFSkimT). The skims represent the zonal shortest travel times under uncongested conditions. The general form of regional accessibility by highway travel is given below:

$$highway\ accessibility(i) = \frac{\sum_{j=1}^N O_j e^{-0.0954 \times t(i,j)}}{\text{Max}_{k \in [1, N]} \{highway\ accessibility(k)\}} \quad [7]$$

where $highway\ accessibility(i)$ = regional accessibility index for TAZ i ;
 O_j = a measure of population or employment in TAZ i ;
 $t(i, j)$ = travel time determined by one of the skims; and
 N = number of TAZs.

This accessibility measure falls within 0 and 1, with 0 indicating no accessibility, and 1 the best accessibility.

To calculate the values of the transit accessibility variables, two transit skims were also obtained, i.e., AM peak with walk access (TSkimW) and AM peak with the best access (TSkimB). The difference between these two transit skims was the mode to access transit services, i.e., by walk mode or by other modes, most likely autos, with the shortest impedance. Similar to highway accessibility, the general form of regional accessibility by transit is given below:

$$transit\ accessibility(i) = \frac{\sum_{j=1}^N O_j e^{-0.0166 \times t(i,j)}}{\text{Max}_{k \in [1, N]} \{transit\ accessibility(k)\}} \quad [8]$$

The difference between formulae [7] and [8] is that the coefficients in the travel impedance term assume an exponential form. The coefficients were calculated using a method developed by Levinson and Kumar (1995).

The accessibility indices given by [7] and [8] were for TAZs. To arrive at indices for census tracts, the TAZ accessibility indices were combined as follows:

$$AI(j) = \frac{\sum_{i=1}^{N_t} accessibility(i)}{\sum_{i=1}^{N_t} O_i}$$

where $AI(j)$ = accessibility index for tract j ;
 O_i = a measure of population or employment of TAZ i ;
 $accessibility(i)$ = highway or transit accessibility measure calculated by formula [7] or [8]; and
 N_t is the number of TAZs in a tract j .

Table 4-3 describes the tract level highway accessibility variables developed using different types of opportunities and highway skims. Table 4-4 lists the transit accessibility variables developed for different areas, using different types of opportunities and transit skims. Additionally, the variables were distinguished by the area for which they were developed and the

land use data utilized in the calculations. For instance, *AccessTWSCI* represented accessibility based on transit walk access mode and using service and commercial employment data obtained by the Land Use Method. The number that immediately follows a variable name indicates which method was used to compute the opportunities: 1 for the Land Use Method, 2 for the Buffer Method, 3 for the Buffer Method, and 4 for the Tract Method.

Table 4-3. Tract Level Highway Accessibility Variables

Variable	Type of Opportunities (O_j)	Skim Used
<i>AccessH1SC</i>	Service+commercial employment	FFSkim
<i>AccessH1Em</i>	Total employment	FFSkim
<i>AccessH1Pop</i>	Total population	FFSkim
<i>AccessH2SC</i>	Service+commercial employment	FFSkimT
<i>AccessH2Em</i>	Total employment	FFSkimT
<i>AccessH2Pop</i>	Total population	FFSkimT

While the variables in Tables 4-3 and 4-4 measure the accessibility by highway travel and transit, respectively, these two sets of measures are independent and disjoint. However, it is possible that the difference between the transit and highway accessibilities will influence the mode choice decision. To evaluate the effect of the difference between transit and highway accessibilities, a third regional accessibility measure was developed using the following formula:

$$accessibility\ difference(i) = \frac{\sum_{j=1}^N O_j e^{-[0.0166tt(i,j) - 0.0954 \times th(i,j)]}}{Max_{k \in [1, N]} \{accessibility\ difference(k)\}}$$

where $tt(i,j)$ = transit travel time between zone i and zone j ; and
 $th(i,j)$ = highway travel time between zone i and zone j .

Variables that measure transit and highway accessibilities are given in Table 4-5. The area for which the variable was computed, the method used to calculate the opportunities, the type of opportunities, and the type of transit and highway travel times are also indicated for each variable.

Table 4-4. Tract Level Transit Accessibility Variables

Variable	Area Used	Method Use for O_i	Type of Opportunities	Skim Used
<i>AccessTWSC1</i>	Buffer area	Land Use	Service+comm. employment	TSkimW
<i>AccessTWSC2</i>	Buffer area	Property	Service+comm. employment	TSkimW
<i>AccessTWSC3</i>	Buffer area	Buffer	Service+comm. employment	TSkimW
<i>AccessTWSC4</i>	Tract	Tract	Service+comm. employment	TSkimW
<i>AccessTWEm1</i>	Buffer area	Land Use	Total employment	TSkimW
<i>AccessTWEm2</i>	Buffer area	Property	Total employment	TSkimW
<i>AccessTWEm3</i>	Buffer area	Buffer	Total employment	TSkimW
<i>AccessTWEm4</i>	Tract	Tract	Total employment	TSkimW
<i>AccessTWPop1-2</i>	Buffer area	Land Use	Population	TSkimW
<i>AccessTWPop3</i>	Buffer area	Buffer	Population	TSkimW
<i>AccessTWPop4</i>	Tract	Tract	Population	TSkimW
<i>AccessBSC1</i>	Buffer area	Land Use	Service+comm. employment	TSkimB
<i>AccessBSC2</i>	Buffer area	Property	Service+comm. employment	TSkimB
<i>AccessBSC3</i>	Buffer area	Buffer	Service+comm. employment	TSkimB
<i>AccessBSC4</i>	Tract	Tract	Service+comm. employment	TSkimB
<i>AccessBEm1</i>	Buffer area	Land Use	Total employment	TSkimB
<i>AccessBEm2</i>	Buffer area	Property	Total employment	TSkimB
<i>AccessBEm3</i>	Buffer area	Buffer	Total employment	TSkimB
<i>AccessBEm4</i>	Tract	Tract	Total employment	TSkimB
<i>AccessBPop1-2</i>	Buffer area	Land Use	Population	TSkimB
<i>AccessBPop3</i>	Buffer area	Buffer	Population	TSkimB
<i>AccessBPop4</i>	Tract	Tract	Population	TSkimB

Table 4-5. Tract Level Transit-Highway Accessibility Differences

Variable	Area Used	Method Use for O_i	Type of Opportunities	Skims Used
<i>Wth1SC1</i>	Buffer area	Land Use	Service+comm. Emp.	TSkimW – FFSkim
<i>Wth1SC2</i>	Buffer area	Property	Service+comm. Emp.	TSkimW – FFSkim
<i>Wth1SC3</i>	Buffer area	Buffer	Service+comm. Emp.	TSkimW – FFSkim
<i>Wth1SC4</i>	Tract	Tract	Service+comm. Emp.	TSkimW – FFSkim
<i>Wth2SC1</i>	Buffer area	Land Use	Service+comm. Emp.	TSkimW – FFSkimT
<i>Wth2SC2</i>	Buffer area	Property	Service+comm. Emp.	TSkimW – FFSkimT
<i>Wth2SC3</i>	Buffer area	Buffer	Service+comm. Emp.	TSkimW – FFSkimT
<i>Wth2SC4</i>	Tract	Tract	Service+comm. Emp.	TSkimW – FFSkimT
<i>Wth1Em1</i>	Buffer area	Land Use	Total employment	TSkimW – FFSkim
<i>Wth1Em2</i>	Buffer area	Property	Total employment	TSkimW – FFSkim
<i>Wth1Em3</i>	Buffer area	Buffer	Total employment	TSkimW – FFSkim
<i>Wth1Em4</i>	Tract	Tract	Total employment	TSkimW – FFSkim
<i>Wth2Em1</i>	Buffer area	Land Use	Total employment	TSkimW – FFSkimT
<i>Wth2Em2</i>	Buffer area	ProeprtyProperty	Total employment	TSkimW – FFSkimT
<i>Wth2Em3</i>	Buffer area	Buffer	Total employment	TSkimW – FFSkimT
<i>Wth2Em4</i>	Tract	Tract	Total employment	TSkimW – FFSkimT
<i>Wth1Pop1-2</i>	Buffer area	Land Use	Population	TSkimW – FFSkim
<i>Wth1Pop3</i>	Buffer area	Buffer	Population	TSkimW – FFSkim
<i>Wth1Pop4</i>	Tract	Tract	Population	TSkimW – FFSkim
<i>Wth2Pop1-2</i>	Buffer area	Land Use	Population	TSkimW - FFSkimT
<i>Wth2Pop3</i>	Buffer area	Buffer	Population	TSkimW - FFSkimT
<i>Wth2Pop4</i>	Tract	Tract	Population	TSkimW - FFSkimT
<i>Bth1SC1</i>	Buffer area	Land Use	Service+comm. Emp.	TSkimB – FFSkim
<i>Bth1SC2</i>	Buffer area	Property	Service+comm. Emp.	TSkimB – FFSkim
<i>Bth1SC3</i>	Buffer area	Buffer	Service+comm. Emp.	TSkimB – FFSkim
<i>Bth1SC4</i>	Tract	Tract	Service+comm. Emp.	TSkimB – FFSkim
<i>Bth2SC1</i>	Buffer area	Land Use	Service+comm. Emp.	TSkimB – FFSkimT
<i>Bth2SC2</i>	Buffer area	ProeprtyProperty	Service+comm. Emp.	TSkimB – FFSkimT
<i>Bth2SC3</i>	Buffer area	Buffer	Service+comm. Emp.	TSkimB – FFSkimT
<i>Bth2SC4</i>	Tract	Tract	Service+comm. Emp.	TSkimB – FFSkimT
<i>Bth1Em1</i>	Buffer area	Land Use	Total employment	TSkimB – FFSkim
<i>Bth1Em2</i>	Buffer area	Property	Total employment	TSkimB – FFSkim
<i>Bth1Em3</i>	Buffer area	Buffer	Total employment	TSkimB – FFSkim
<i>Bth1Em4</i>	Tract	Tract	Total employment	TSkimB – FFSkim
<i>Bth2Em1</i>	Buffer area	Land Use	Total employment	TSkimB – FFSkimT
<i>Bth2Em2</i>	Buffer area	Property	Total employment	TSkimB – FFSkimT
<i>Bth2Em3</i>	Buffer area	Buffer	Total employment	TSkimB – FFSkimT
<i>Bth2Em4</i>	Tract	Tract	Total employment	TSkimB – FFSkimT
<i>Bth1Pop1-2</i>	Buffer area	Land Use	Total population	TSkimB – FFSkim
<i>Bth1Pop3</i>	Buffer area	Buffer	Total population	TSkimB – FFSkim
<i>Bth1Pop4</i>	Tract	Tract	Total population	TSkimB – FFSkim
<i>Bth2Pop1-2</i>	Buffer area	Land Use	Total population	TSkimB - FFSkimT
<i>Bth2Pop3</i>	Buffer area	Buffer	Total population	TSkimB - FFSkimT
<i>Bth2Pop4</i>	Tract	Tract	Total population	TSkimB - FFSkimT

4.5 Transit Users' Socioeconomic Characteristics

Studies that have investigated transit users' socioeconomic/demographic characteristics including age, education, ethnicity, gender, household structure, vehicle availability, etc., have led to contradictory conclusions. For example, income was found to have insignificant impacts on transit use by Sun *et al.* (1998), Loutzenheiser (1997), Gray and Thompson (1996), Spillar and Rutherford (1990), Holtzclaw (1990), and Nelson and O'Neil (1982). However, the majority of literature concluded otherwise. Similar conflicting findings on age, gender, and vehicle availability variables can also be found in the literature.

While it is desirable to include some socioeconomic variables in the model calibration, the data from the TOB survey of individual transit user are not utilized in this study. Due to the survey's relatively small sample size in a given tract and its focus on transit users only, the socioeconomic information collected from TOB survey was considered unrepresentative and thus was excluded. Only the socioeconomic variables that can reveal the true characteristics of the entire tract population were included in the analysis since the regression model was calibrated with data aggregated at the tract level. Census data are commonly employed to generate the socioeconomic and demographic data at the tract level. Numerous variables were therefore prepared from available census data. Income data from the 2000 census were not available at the time of this writing. Because of the additional efforts required to obtain the 1990 income data at the tract level from the Census Bureau, income information from the Census was not included in the model calibration.

Table 4-6 lists the variables included in the analyses as socioeconomic indicators. The average assessed value of properties in a tract was obtained using GIS property data and the tax appraiser's database. This variable was used as a measure of income level in a tract. The advantage of this information over income data from census is that they are always up-to-date while census is available only every ten years. For multi-family properties, assessed values for apartments were not available on a unit basis and were not included in the analyses.

Table 4-6. Socioeconomic Variables

Variable Name	Description
<i>AvgSFValue</i>	Average single-family housing value
<i>AvgMFValue</i>	Average multi-family housing value
<i>AvgValue</i>	Average housing value
<i>AvgCOMValue</i>	Average commercial property value
<i>AvgINDValue</i>	Average industrial property value
<i>AvgSERValue</i>	Average service property value
<i>AvgEMPValue</i>	Average nonresidential property value
<i>AvgAuto0Child</i>	Average number of cars owned by households without children
<i>AvgAutoChild</i>	Average number of cars owned by households with children
<i>SFDUNoAuto%</i>	Percentage of single-family households with no automobiles in a tract
<i>MFDUNoAuto%</i>	Percentage of multi-family households with no automobiles in a tract
<i>DUNoAuto%</i>	Percentage of households with no automobiles in a tract

The average values of properties that were of commercial, service, and industrial type as well as that for all nonresidential properties were also calculated. Auto ownership was another possible proxy for income. Auto ownership data came from two sources: the 1999 ZDATA1 file and the 1999 ZDATA1a file, both prepared by the Miami-Dade County Planning Department. ZDATA1a file contains data for the lifestyle groups that are used in the trip generation step in the FSUTMS model, while ZDATA1 file contains data for the standard trip generation of FSUTMS. Based on data from ZDATA1a, data for the following variables were prepared: average number of cars in households without children (*AvgAuto0Child*) and average number of cars in households with children (*AvgAutoChild*). Data on percentage of single-family households with no cars (*SFDUNoAuto%*), percentage of multi-family households with no cars (*MFDUNoAuto%*), and percentage of all households with no cars (*DUNoAuto%*) were aggregated from TAZ level to tract level from the ZDATA1 file. Because auto ownership data were computed for the entire tracts, they were only used in Regression Model 3, in which the dependent variable was the ratio of the number of transit trip samples in a tract to the tract population.

4.6 Demographics

The sources of demographic data were the 1999 ZDATA1a file for Miami-Dade County and 2000 census. The variables based on data from the ZDATA1a file included percentages of households without children (*%HH0Child*), Average number of workers in households without children (*AvgWrkr0Child*), average number of workers in households with children (*AvgWrkrChild*), average number of persons in households without children (*AvgP0Child*), and average number of persons in households with children (*AvgPChild*). The data were given at TAZ level. Aggregating the data to the tract level was straight forward by simple summations. Race groups from the 2000 Census data were considered to be potential explanatory variables. Table 4-7 lists the variables that concern the race. Only single races were included (about 3.8% of population declared they are of mixed race). Since the Census data are provided at the tract level, no additional processing was necessary. Variable *ForeignBorn* is the percentage of immigrant population in a census tract. Since this information was not available in Census 2000 data at the time of this writing, Census 1990 data were used.

Table 4-7. Demographic Variables

Variable	Description
<i>%HH0Child</i>	Percentage of households without children
<i>AvgWrkr0Child</i>	Average number of workers in households without children
<i>AvgWrkrChild</i>	Average number of workers in households with children
<i>AvgP0Child</i>	Average number of persons in households without children
<i>AvgPChild</i>	Average number of persons in households with children
<i>White</i>	Percentage of population that are White in a tract
<i>Black</i>	Percentage of population that are Black in a tract
<i>Native</i>	Percentage of population that are American Indian or Alaska Native in a tract
<i>Asian</i>	Percentage of population that are Asian and Pacific Islander in a tract
<i>Hispanic</i>	Percentage of population that are Hispanic origin in a tract
<i>ForeignBorn</i>	Percentage of population that are immigrants in a tract

4.7 Data for Broward County

For Broward County, the GIS data on bus stops were available from Broward County Transit Agency (BCTA). Therefore, the study area included the entire county. However, since information for land use and property was not available, only buffer method and tract level analysis could be performed. The available independent variables may be classified into the following four categories: transit LOS (shown in Table 4-8), land use (Table 4-9), accessibility (Tables 4-10 through 4-12), and socioeconomic (Tables 4-13 and 4-14). The headway information provided by Broward County Transit Agency did not include number of bus or peak hour information. Therefore, only average headway (*AVGHDWAY*) and weighted average headway (*WAVGHDWAY*) were compiled.

Table 4-8. Transit LOS Variables for Broward County

Variable Name	Description
<i>ServiceArea%</i>	Percentage of tract area served by transit based on ¼-mile buffers around
<i>Avghdway</i>	Composite average daily headway in a tract
<i>WAvghdway</i>	Composite peak hour headway for a tract weighted by service area

Table 4-9. Land Use Variables for Broward County

Variable Name	Source	Description
<i>TotalDUDensity3</i>	Census 2000	Total dwelling unit density in buffer area of a census tract
<i>TotalDUDensity4</i>	Census 2000	Total dwelling unit density in a census tract
<i>TotalPopDensity3</i>	Census 2000	Total population density in buffer area of a census tract
<i>TotalPopDensity4</i>	Census 2000	Total population density in a census tract
<i>ComEmpDensity3</i>	1999 ZDATA2	Commercial employment density in buffer area of a tract
<i>ComEmpDensity4</i>	1999 ZDATA2	Commercial employment density in a census tract
<i>ServEmpDensity3</i>	1999 ZDATA2	Service employment density in buffer area of a tract
<i>ServEmpDensity4</i>	1999 ZDATA2	Service employment density in a census tract
<i>IndEmpDensity3</i>	1999 ZDATA2	Industrial employment density in buffer area of a tract
<i>IndEmpDensity4</i>	1999 ZDATA2	Industrial employment density in a census tract
<i>TotalEmpDensity3</i>	1999 ZDATA2	Total employment density in buffer area of a census tract
<i>TotalEmpDensity4</i>	1999 ZDATA2	Total employment density in a census tract
<i>TotalDensity3</i>		Total employment plus population density in buffer area
<i>TotalDensity4</i>		Total employment plus population density in a tract
<i>JOBS_HH3</i>		Jobs-housing balance in buffer area of a census tract
<i>JOBS_HH4</i>		Jobs-housing balance in a census tract
<i>JOBS_HH3^m</i>		Modified jobs-housing balance in buffer area of a tract
<i>JOBS_HH4^m</i>		Modified jobs-housing balance in a census tract

Table 4-10. Tract Level Highway Accessibility Variables for Broward County

Variable	Type of Opportunities (O_i)	Skim Used
<i>AccessH1SC</i>	Service+comm. employment	FFSkim
<i>AccessH1Em</i>	Total employment	FFSkim
<i>AccessH1Pop</i>	Total population	FFSkim
<i>AccessH2SC</i>	Service+comm. employment	FFSkimT
<i>AccessH2Em</i>	Total employment	FFSkimT
<i>AccessH2Pop</i>	Total population	FFSkimT

Table 4-11. Tract Level Transit Accessibility for Broward County

Variable	Method Use for Oj	Type of Opportunities	Skim Used
<i>AccessTWSC3</i>	Buffer	Service+comm. employment	TskimW
<i>AccessTWSC4</i>	Tract	Service+comm. employment	TskimW
<i>AccessTWEm3</i>	Buffer	Total employment	TskimW
<i>AccessTWEm4</i>	Tract	Total employment	TskimW
<i>AccessTWPop3</i>	Buffer	Population	TskimW
<i>AccessTWPop4</i>	Tract	Population	TskimW
<i>AccessBSC3</i>	Buffer	Service+comm. employment	TSkimB
<i>AccessBSC4</i>	Tract	Service+comm. employment	TSkimB
<i>AccessBEm3</i>	Buffer	Total employment	TSkimB
<i>AccessBEm4</i>	Tract	Total employment	TSkimB
<i>AccessBPop3</i>	Buffer	Population	TSkimB
<i>AccessBPop4</i>	Tract	Population	TSkimB

Table 4-12. Tract Level Transit-Highway Accessibility Differences for Broward County

Variable	Method Use for Oj	Type of Opportunities	Skim Used
<i>Wth1SC3</i>	Buffer	Service+comm. Emp.	TSkimW - FFSkim
<i>Wth1SC4</i>	Tract	Service+comm. Emp.	TSkimW - FFSkim
<i>Wth2SC3</i>	Buffer	Service+comm. Emp.	TSkimW - FFSkimT
<i>Wth2SC4</i>	Tract	Service+comm. Emp.	TSkimW - FFSkimT
<i>Wth1Em3</i>	Buffer	Total employment	TSkimW - FFSkim
<i>Wth1Em4</i>	Tract	Total employment	TSkimW - FFSkim
<i>Wth2Em3</i>	Buffer	Total employment	TSkimW - FFSkimT
<i>Wth2Em4</i>	Tract	Total employment	TSkimW - FFSkimT
<i>Wth1Pop3</i>	Buffer	Population	TSkimW - FFSkim
<i>Wth1Pop4</i>	Tract	Population	TSkimW - FFSkim
<i>Wth2Pop3</i>	Buffer	Population	TSkimW - FFSkimT
<i>Wth2Pop4</i>	Tract	Population	TSkimW - FFSkimT
<i>Bth1SC3</i>	Buffer	Service+comm. Emp.	TSkimB - FFSkim
<i>Bth1SC4</i>	Tract	Service+comm. Emp.	TSkimB - FFSkim
<i>Bth2SC3</i>	Buffer	Service+comm. Emp.	TSkimB - FFSkimT
<i>Bth2SC4</i>	Tract	Service+comm. Emp.	TSkimB - FFSkimT
<i>Bth1Em3</i>	Buffer	Total employment	TSkimB - FFSkim
<i>Bth1Em4</i>	Tract	Total employment	TSkimB - FFSkim
<i>Bth2Em3</i>	Buffer	Total employment	TSkimB - FFSkimT
<i>Bth2Em4</i>	Tract	Total employment	TSkimB - FFSkimT
<i>Bth1Pop3</i>	Buffer	Total population	TSkimB - FFSkim
<i>Bth1Pop4</i>	Tract	Total population	TSkimB - FFSkim
<i>Bth2Pop3</i>	Buffer	Total population	TSkimB - FFSkimT
<i>Bth2Pop4</i>	Tract	Total population	TSkimB - FFSkimT

Table 4-13. Socioeconomic Variables for Broward County

Variable Name	Description
<i>AvgAuto0Child</i>	Average number of cars owned by households without children
<i>AvgAutoChild</i>	Average number of cars owned by households with children

Table 4-14. Demographic Variables for Broward County

Variable Name	Description
<i>%HH0Child</i>	Percentage of households without children
<i>AvgWrkr0Child</i>	Average number of workers in households without children
<i>AvgWrkrChild</i>	Average number of workers in households with children
<i>AvgP0Child</i>	Average number of persons in households without children
<i>AvgPChild</i>	Average number of persons in households with children
<i>White</i>	Percentage of population that are White in a tract
<i>Black</i>	Percentage of population that are Black in a tract
<i>Native</i>	Percentage of population that are American Indian or Alaska Native in a tract
<i>Asian</i>	Percentage of population that are Asian and Pacific Islander in a tract
<i>Hispanic</i>	Percentage of population that are Hispanic origin in a tract
<i>ForeignBorn</i>	Percentage of population that are immigrants in a tract

4.8 Data for Palm Beach County

For Palm Beach County, GIS bus stop data were unavailable; therefore the buffer method could not be applied. Consequently, only tract level analysis was performed. The dependent variable was the ratio between the number of trip ends (either production or attraction or both) in a census tract and the tract population.

The independent variables may be classified into the following three categories: land use, accessibility, and transit users' socioeconomic/demographic characteristics. They were compiled using the Tract Method. These variables are listed in Tables 4-15 through 4-20.

Table 4-15. Land Use Variables for Palm Beach County

Variable Name	Source	Description
<i>TotalDUDensity4</i>	Census 2000	Total dwelling unit density in a census tract
<i>TotalPopDensity4</i>	Census 2000	Total population density in a census tract
<i>ComEmpDensity4</i>	1999 ZDATA2	Commercial employment density in a census tract
<i>ServEmpDensity4</i>	1999 ZDATA2	Service employment density in a census tract
<i>IndEmpDensity4</i>	1999 ZDATA2	Industrial employment density in a census tract
<i>TotalEmpDensity4</i>	1999 ZDATA2	Total employment density in a census tract
<i>TotalDensity4</i>		Total employment plus population density in a tract
<i>JOBS_HH4</i>		Jobs-housing balance in a census tract
<i>JOBS_HH^m4</i>		Modified jobs-housing balance in a census tract

Table 4-16. Tract Level Highway Accessibility Variables for Palm Beach County

Variable	Type of Opportunities (O_i)	Skim Used
<i>AccessH1SC</i>	Service+comm. employment	FFSkim
<i>AccessH1Em</i>	Total employment	FFSkim
<i>AccessH1Pop</i>	Total population	FFSkim
<i>AccessH2SC</i>	Service+comm. employment	FFSkimT
<i>AccessH2Em</i>	Total employment	FFSkimT
<i>AccessH2Pop</i>	Total population	FFSkimT

Table 4-17. Tract Level Transit Accessibility Variables for Palm Beach County

Variable	Type of Opportunities	Skim Used
<i>AccessTWSC4</i>	Service+comm. employment	TSkimW
<i>AccessTWEm4</i>	Total employment	TSkimW
<i>AccessTWPop4</i>	Population	TSkimW
<i>AccessBSC4</i>	Service+comm. employment	TSkimB
<i>AccessBEm4</i>	Total employment	TSkimB
<i>AccessBPop4</i>	Population	TSkimB

Table 4-18. Palm Beach Tract Level Transit-Highway Accessibility Differences

Variable	Type of Opportunities	Skim Used
<i>Wth1SC4</i>	Service+comm. Emp.	TSkimW - FFSkim
<i>Wth2SC4</i>	Service+comm. Emp.	TSkimW - FFSkimT
<i>Wth1Em4</i>	Total employment	TSkimW - FFSkim
<i>Wth2Em4</i>	Total employment	TSkimW - FFSkimT
<i>Wth1Pop4</i>	Population	TSkimW - FFSkim
<i>Wth2Pop4</i>	Population	TSkimW - FFSkimT
<i>Bth1SC4</i>	Service+comm. Emp.	TSkimB - FFSkim
<i>Bth2SC4</i>	Service+comm. Emp.	TSkimB - FFSkimT
<i>Bth1Em4</i>	Total employment	TSkimB - FFSkim
<i>Bth2Em4</i>	Total employment	TSkimB - FFSkimT
<i>Bth1Pop4</i>	Total population	TSkimB - FFSkim
<i>Bth2Pop4</i>	Total population	TSkimB - FFSkimT

Table 4-19. Socioeconomic Variables for Palm Beach County

Variable	Description
<i>AvgAuto0Child</i>	Average number of cars owned by households without children
<i>AvgAutoChild</i>	Average number of cars owned by households with children

Table 4-20. Demographic Variables for Palm Beach County

Variable	Description
<i>%HH0Child</i>	Percentage of households without children
<i>AvgWrkr0Child</i>	Average number of workers in households without children
<i>AvgWrkrChild</i>	Average number of workers in households with children
<i>AvgP0Child</i>	Average number of persons in households without children
<i>AvgPChild</i>	Average number of persons in households with children
<i>White</i>	Percentage of population that are White in a tract
<i>Black</i>	Percentage of population that are Black in a tract
<i>Native</i>	Percentage of population that are American Indian or Alaska Native in a tract
<i>Asian</i>	Percentage of population that are Asian and Pacific Islander in a tract
<i>Hispanic</i>	Percentage of population that are Hispanic origin in a tract
<i>ForeignBorn</i>	Percentage of population that are immigrants in a tract

5. TRANSIT ACCESSIBILITY

5.1 Introduction

One of the objectives of this research is to develop a method more accurate than the traditional buffer method for estimation of population that have access to transit services. Also of interest is if any new measures developed will have predictive power on transit use. The development of methods for transit service population and employment will be presented in this chapter, and the testing of the new measures.

A critical factor in transit planning transit mode share analysis is the level of accessibility to transit services by population and employment. “Transit accessibility” here refers to the ability of residents and workers to reach transit facilities, including bus stops and/or rail stations. Transit accessibility is affected by many factors including safe, pleasant, and comfortable streets for walking to transit facilities, parking facilities for cars and bicycles, handicap access, and so on. The majority of transit users access transit systems by walking. Walking distance is an important factor in the choice of transit use. The TOB survey results indicated that individuals walk for over 75% of all trips from point of origin to transit station and from station to destination.⁷ It is commonly accepted that most people are willing to walk up to 0.25 mile to use transit, and the farther away from the transit stops/stations, the less likely it is for people to use transit. In this context, easy access to transit means proximity.

Easy access to transit services also depends, to a large degree, on the design of a community. Traditional communities are typically laid out in a grid system, in which streets form the grid and residential and commercial developments occur along the streets. In such communities, street blocks are normally small and roads are well connected, allowing easy access to major roads where transit stops are often located. In recent years, the traditional neighborhood design (TND) approach has been used less often in suburban areas. New developments in the American suburbs tend to be “enclosed” by a local street system design that limits access from a development to major roads. Additionally, dead-end streets, cul-de-sacs, and community walls are popular means to seal a community off from the noisy traffic or to provide residents with a sense of security. However, such an approach to neighborhood design has great implications to the transportation system since it tends to reduce public transit use and increase roadway congestion.

Typically, transit accessibility is estimated using the geographic information system (GIS) buffer method to calculate the proportion of population or employees that are close to transit facilities such as bus stops or rail stations. The buffer method assumes that population and employment are evenly distributed across the spatial unit of analysis, usually in terms of traffic analysis zones (TAZs), census tracts, or block groups. Buffers around transit stops or stations are then created with a given size and are defined as the “service area.” The percentage of population and employment that have access to transit facilities in a zone is assumed to be the same as the ratio of the buffer area falling within the zone to the total area of the zone. Therefore, if a zone has a population of 1,000 and buffers created around transit stops/stations cover 20% of its total area,

⁷ The TOB survey results showed that 79.8% of the 4,152 surveyed trips involved walking to transit stops/stations from an origin, while 75.2% of the 4,159 trips involved walking to destinations after leaving the transit systems

the proportion of the population that is served by transit, or the service population, is assumed to be $1,000 \times 20\% = 200$. Everything else being equal, the larger the proportion of population and employees that fall in the service area, transit is more accessible, and it is more likely that transit ridership will be higher.

In FUSTMS, the population and employment within 1/3- and 1-mile walking distances (referred to as short walk and long walk distances) from transit facilities are one type of the information used in estimating the transit mode share. Population and employment figures are estimated for each traffic analysis zone (TAZ) using the buffer method in which a 1/3-mile or 1-mile buffer are created around transit routes or stops/stations. Estimating potential ridership through the use of buffers around transit stops/stations will reduce errors caused by overestimation should buffers be utilized around transit routes with bus stop spacing larger than 0.25 mile.

The buffer method, while used in FSUTMS as well as widely used by transit properties in their planning applications, is flawed in its fundamental assumption that population or employment in a zone is evenly distributed across the zone. In reality, this assumption only holds occasionally when land use in a zone is uniform with the same density. In most cases, a zone may have the same land use but density varies, or it may have different land uses with significant variations in density. Another problem with the buffer method is the assumption that the walking distance for a transit user accessing a transit stop or a station is the same as Euclidian distance (also referred to as straight line or air distance). The actual walking distance is, in fact, usually longer due to the “crookedness” of streets. A person may live near a transit stop; however, if no streets or walking paths connect his or her residence to the transit stop, the person does not have access to that transit stop. Other problems that cannot be handled by the buffer zone method include natural or man-made barriers such as highways with limited access, canals, and community walls or fences that surround a development that prevent people from accessing transit facilities in a direct manner.

Recognizing the problems underlying the buffer method, various researchers have looked for ways to improve the estimation of the transit service population. O'Neill *et al.* (1995) developed the network ratio method, which assumes pedestrian travel occurs on streets, therefore lines of equal travel time or distance were constructed around a transit line defining its service area. Additionally, it is assumed that population is evenly distributed along streets. Therefore, the proportion of population within the transit service area was calculated as the ratio of total length of streets that are within the 1/4-mile walking distance to that of all streets. Hsiao *et al.* (1997) also employed this approach in the analysis of links between transit usage and pedestrian accessibility and demographics of transit users and found a strong relationship between bus riders and pedestrian access: higher pedestrian access areas correspond to higher transit usage.

The network ratio method would perform well in an area with a single density residential development (e.g. single family or multi-family housing), but could not account for land uses with different densities and could not handle barriers. In order to improve the network ratio method, Zhao (1998) considered the population distribution in areas that include single- and multi-family land use. The effect of barriers was also investigated. Land use data were helpful in better depicting the population distribution, especially in cases when multi-family housing was concentrated in areas close to transit stops and that barriers could have significant negative impact on transit accessibility. Zhao's study points to the difficulty in collecting barrier

information in a large urban area. Initial data collection will involve some significant effort, and the information will need to be updated from time to time as barriers change when, for example, a community wall is constructed or when pedestrian access is created. In addition to density variations and barrier problems, the estimation of transit walk accessibility for a forecast year in FSUTMS when no street data or detailed land use data are available is an additional challenge if a method such as the network ratio method is to be used to estimate transit accessibility.

The goal of the analysis described in this chapter is to develop a set of methodologies that can overcome problems associated with barriers and uneven distribution of population in the estimation of transit walk accessibility for the base year, and to forecast transit accessibility for future years given forecast population and employment data. The objectives may be stated as follows:

- (1) Investigate existing data that provide better information on the spatial distribution of population and employment;
- (2) Develop a methodology to utilize existing population and employment data to improve the existing methods for estimation of transit service population such as the buffer method and the network ratio method;
- (3) Develop a methodology that is capable of handling barriers, which does not require extensive field data collection; and
- (4) Develop a methodology to forecast transit accessibility given future land uses.

The next section (Section 5.2) methodologies for estimating transit accessibility in terms of population and employees served by transit are described. The analysis results are compared with those from the traditional buffer method as well as the network ratio method. Section 5.3 discusses the models developed for transit accessibility forecasting. Section 5.4 presents analysis on auto access to transit.

5.2 Transit Walk Accessibility

In this section, methodologies for estimating transit walk accessibility are described.

5.2.1 Estimation of Population Accessibility to Transit

Zhao (1998) described the development of a modified network ratio method that considered walking distance to transit stops and dwelling unit distribution based on street network and land use data. One problem associated with this method was the assumption that properties were evenly distributed along all streets. However, some streets do not have properties on them, such as when it is merely an access road into a community or when there is a barrier like a community wall along one side or both sides of the street as illustrated in Figure 5-1. The modified network ratio method has the potential for improvement if better knowledge of the spatial distribution of population and the existence of barriers can be obtained.

To address the issue of population distribution, more detailed information on the spatial distribution of dwelling units than land uses is desired. Such information can be obtained if parcel level GIS data are available. Parcel GIS data are becoming increasingly available in Florida counties. For the purpose of calculating the walking distance from a property to a transit

stop, parcel boundary data are not necessary. Only a point representing the location of a property is needed. The location of this point may be simply the centroid location of the polygon representing the property or the location of the front door of the property, reflecting the setback of the property therefore the additional walking distance on the property. In our analysis, setbacks are not considered to be a significant component of the walking distance for residential properties. They may be significant, however, for nonresidential establishments. Due to a lack of accurate information on employment distribution, the method for estimating workers' walk accessibility is unable to take setbacks into consideration.

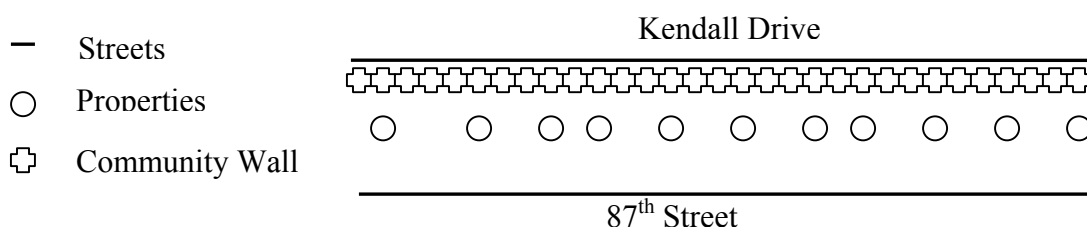


Figure 5-1. Streets That Do Not Allow Access to Transit from Properties

The Miami-Dade County's property tax database provides detailed information on each property, including the number of bedrooms. While we do not have information on the household size for each residential property, the number of bedrooms in a dwelling unit may be considered a proxy indicator of the household size, which allows for a better estimation of population distribution.

Determination of man-made barriers such as community walls has been identified a difficulty in (Zhao 1998). Collecting such data will be expensive since field observation of every community will be necessary and field data will require significant effort to process to create a GIS database. Such a database will also need to be updated periodically to reflect any possible changes in existing barriers or addition of new barriers. The solution to this problem lies in the fact that, if there is a barrier along a street, it is likely that there will be no properties having addresses on that street. The occupants of properties located not on but nearby a street on which transit services are operated will need to access transit services via streets connecting to that street. This understanding leads to the assumption underlying the methodology described in this section, i.e. a property's occupants can access transit services on a given street only if (1) it is located on that street; (2) it is connected by other streets to the street where transit services are available; and (3) it is within walking distance of a transit stop. Making this assumption eliminates the need for hard coding barrier information.

To determine on which street a property is located requires matching that property's address with a street using GIS. This, in turn, requires that addresses in the property tax database and the street attribute database must be formatted in a consistent manner. Some processing was necessary to modify the Miami-Dade County property addresses to be consistent with the addresses in the street database. The following steps describe the methodology for estimating population with walk accessibility to transit services.

Step 1: Based on the transit stop locations, a transit service catchment area was determined by identifying street segments that were connected to a bus stop and were within a distance of 0.5-mile of the bus stop. Here, "within a distance of 0.5-mile" means the walking

distance along streets starting from a bus stop. An example of the transit catchment area is shown in Figure 5-2, together with the street network, bus stop locations, and location of properties. The size of a point representing properties indicates the estimated number of residents at that location. Expressways, freeways, and ramps have been excluded from the street network since pedestrians are not able to use them.

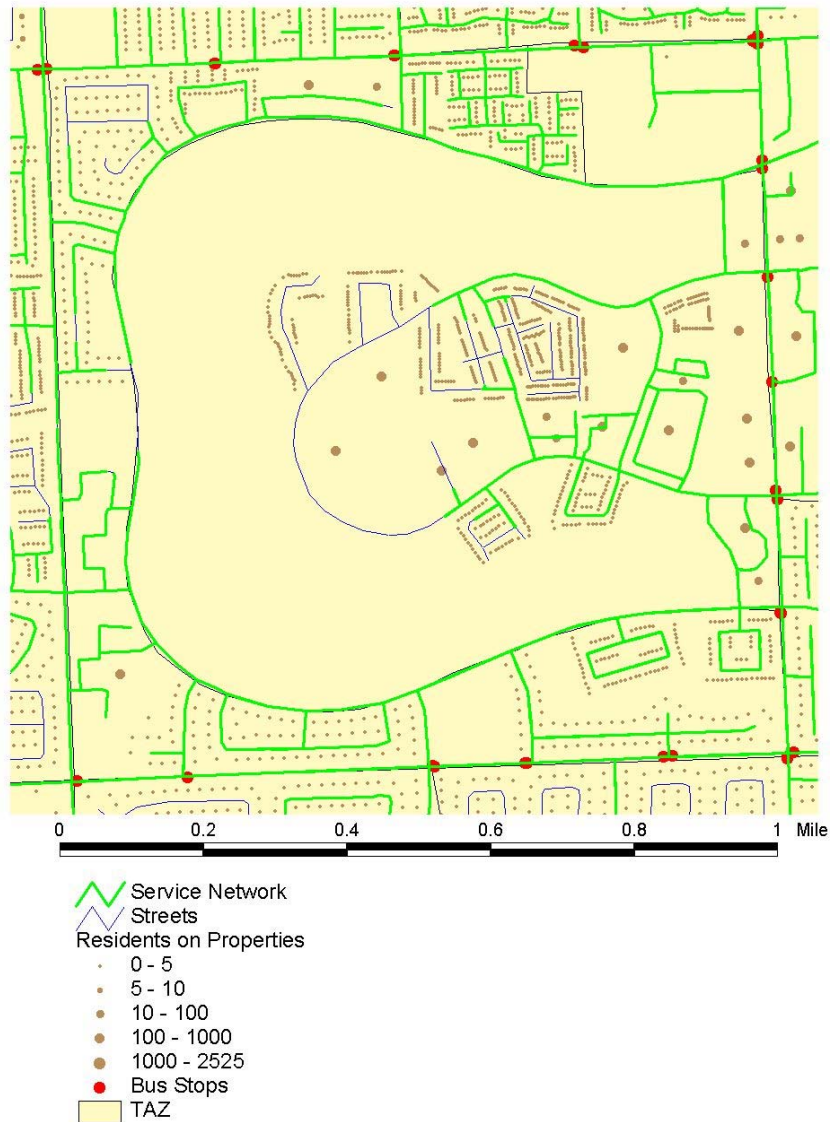


Figure 5-2. Transit Catchment Area, Streets, and Property Distribution

Step 2: The street network in the catchment area was intersected with TAZ to assign street segments to each TAZ.

Step 3: Using the TAZ population information from ZDATA1 file, which included single-family and multi-family population, the average household sizes for the two types of residential properties (single- and multi-family) were calculated.

Step 4: The average household sizes within a TAZ, calculated in Step 3, were then used to estimate how many people might live on each property.

- (1) Decide household type (single-family or multi-family). The type of residential properties was determined based on the CLU code in the property database. After verifying the code description and land use type, including a field survey in some cases, the relationship between the CLU code and the type of dwelling units was established:
 - CLU code 0001 and 0004 – single family
 - Most (6518) of CLU code 0005 – single family
 - CLU code 0002, 0003, 0006-0010 – multi-family
 - A few (590) properties with CLUC 0005 – multi-family
- (2) Determine the number of bedrooms for each property.
 - (a) Join the property label point (LPROP) attribute table with the property tax file to obtain the number of bedrooms for each property. Some multi-family properties with the same street address (for example, condominiums) were represented by the same label point in LPROP and shared one record in the LPROP (CLUC=0000) attribute table. For these properties, the number of bedrooms had to be summed up for the reference record in LPROP.
 - (b) Deal with missing bedroom information. In some property records, the number of bedrooms was missing. For such single-family dwelling units, the number of bedrooms was assumed to be the average number of bedrooms in the TAZ calculated from property database records that had the number of bedroom information. For multi-family dwelling units, the average bedroom number per unit was calculated by averaging the known numbers of bedrooms in properties of a particular multi-family dwelling type as defined by CLUC types, for example, 2.68 bedrooms per townhouse.
 - (c) Manually check and adjust inconsistencies and errors. For example, TAZ 47 (mobile home park) had four commercial properties but a multi-family population of 3500.
- (3) Determine the household size for each property. From 1999 ZDATA1, the population for the two types (single- and multi-family) of residential properties was found for each TAZ, which was then divided by the total number of bedrooms to get the average number of persons per bedroom. These factors were applied to each property to find the household size.

Step 5: Match residential properties in the LPROP file with street segments based on their addresses. To do so, address formats in both street network and property database were standardized. For example, 103 Street was changed to 103rd St, 105th St Rd to 105th St, or 22nd Street to 22nd St., etc. Properties were then snapped to the closest streets that matched their addresses.

Step 6: Determine the negative impact of walking distance on transit accessibility. The TOB survey data have indicated that the number of transit users decreases with the increase

in walking distance to transit stops. Figure 5-3 shows the distribution of survey samples from the transit-onboard survey based on walking distance from or to home. It is apparent from this data that the origin of most trips were within 1,800 feet of transit stops. Few trips were more than 2,700 feet away from transit stops. This is consistent with conclusions from other studies that transit use will sharply drop after the first 0.06 mile (300 feet), and will diminish beyond 0.36 mile (1900 feet) from a transit stop or station (Lam and Morrall, 1982; Levinson and Brown-West, 1984).

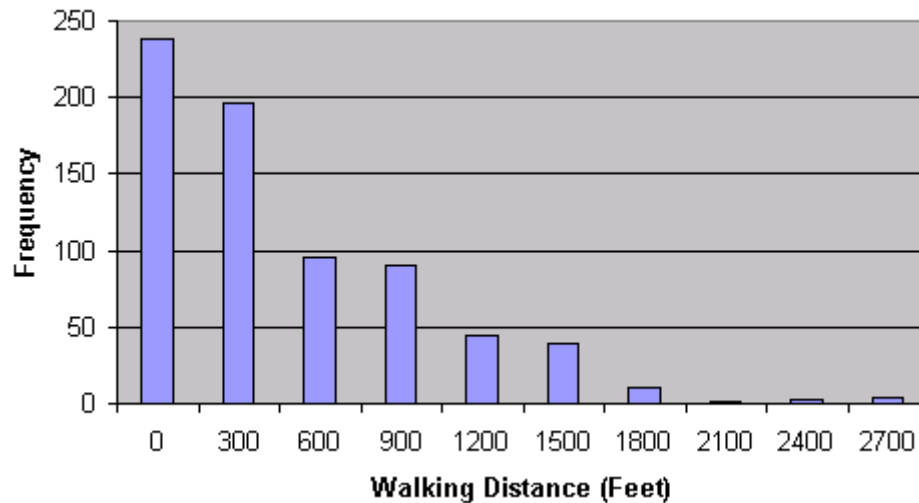


Figure 5-3. Frequency Distribution of Transit Trips versus Walking Distance

To reflect the deterioration of transit use due to increasing walking distance to transit services, a decay function was estimated based on the data from the transit onboard (TOB) survey (applied in the next step) when the TAZ's population within walking distance was calculated. The procedure for estimating the decay function is described below.

- (1) Select samples of transit users walking to transit stops from their homes or walking to their homes from transit stops from TOB survey data. There were a total of 744 such samples. Twenty-two of these 744 samples were excluded because of possible geocoding errors (e.g. some claimed that they walked less than three blocks, ranging from less than 1000 ft to 2000 ft on street network, but their homes were located more than 2000 feet away from the closest transit stops); 722 samples were used.
- (2) Calculate the walking distances from the homes of those 722 transit users to the closest transit stops based on shortest path algorithms.
- (3) Calculate the frequencies of the samples based on walking distances based on equal intervals, a minimum five samples in each interval, and 10 intervals. Intervals of 200, 300, and 400 feet were tested.
- (4) Normalize the frequencies by the population living within each interval contour. To estimate the population living inside each interval contour, the walking distance between each property and the closest transit stop was calculated based on network analysis. Since the household size of each property has been estimated based on the number of bedrooms, as described in Step 4, the population within each interval contour (in increment of 300 feet) was calculated for the entire study area. For instance, there were

196 samples of which the residence was located between 300 and 600 feet of a transit stop. The population living within this contour interval was found to be 113,587. Thus the normalized frequency was $196/113,587 = 0.00172556$. The normalized frequencies were then multiplied by a constant such that the maximum frequency, which occurred in interval $[0, 300]$, was 1.0. This constant did not affect the power of the exponential function estimated, only its coefficient, which was not of interest.

- (5) Fit the weighted frequency with an exponential curve. The 400-foot interval grouping had one interval with no samples and was not used. The 200-foot intervals resulted in a R^2 of 0.6897, and the 300-foot intervals produced a R^2 of 0.7703. The 300-foot interval was selected because of its higher R^2 . The decay function has the following form: Decay function = $\exp(-0.0013 x)$, where x is the walking distance from a transit stop. The fitted curve is shown in Figure 5-4. The above decay function indicates the rate at which transit use will decrease as walking distance grows relative to the transit use in areas within 300 feet of a bus stop. Farther than 0.5 mile away from a transit stop, transit use diminishes to three percent of that within 300 feet of a transit stop. Therefore, in calculating percent of population served by transit and considering the likelihood of their use of transit based on walking distance, use of 0.5 mile as the limit for walk access for the purpose of evaluating transit accessibility by foot is reasonable.

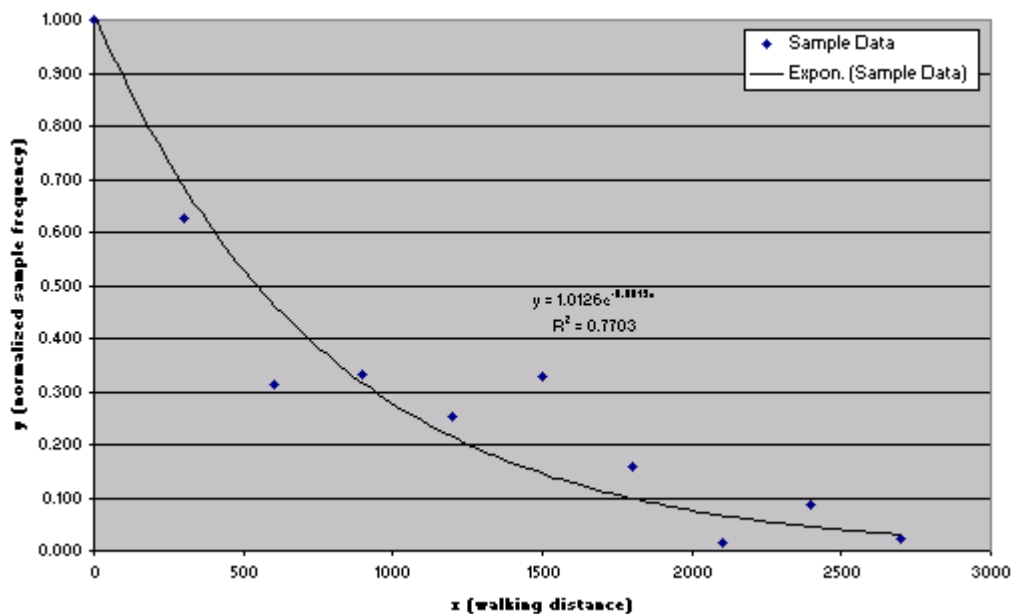


Figure 5-4. Estimation of the Decay Function

Step 7: The total population served by transit in a TAZ was obtained by summing the estimated household size along the street segments in the catchment area within 0.5 mile of walking distance, weighted by the decay function.

Step 8: The percentage of population with walk access to transit services in a TAZ was the ratio of the total population served by transit in the TAZ to the total population in the TAZ.

To compare the result from the above procedure with those of the traditional buffer method and the network ratio method, a 0.25-mile buffer size and walking distance were assumed and a selected area was examined (see Figure 5-5). For the area shown in Figure 5-5, the populations served by transit based on each of the three methods are compared in Table 5-1.

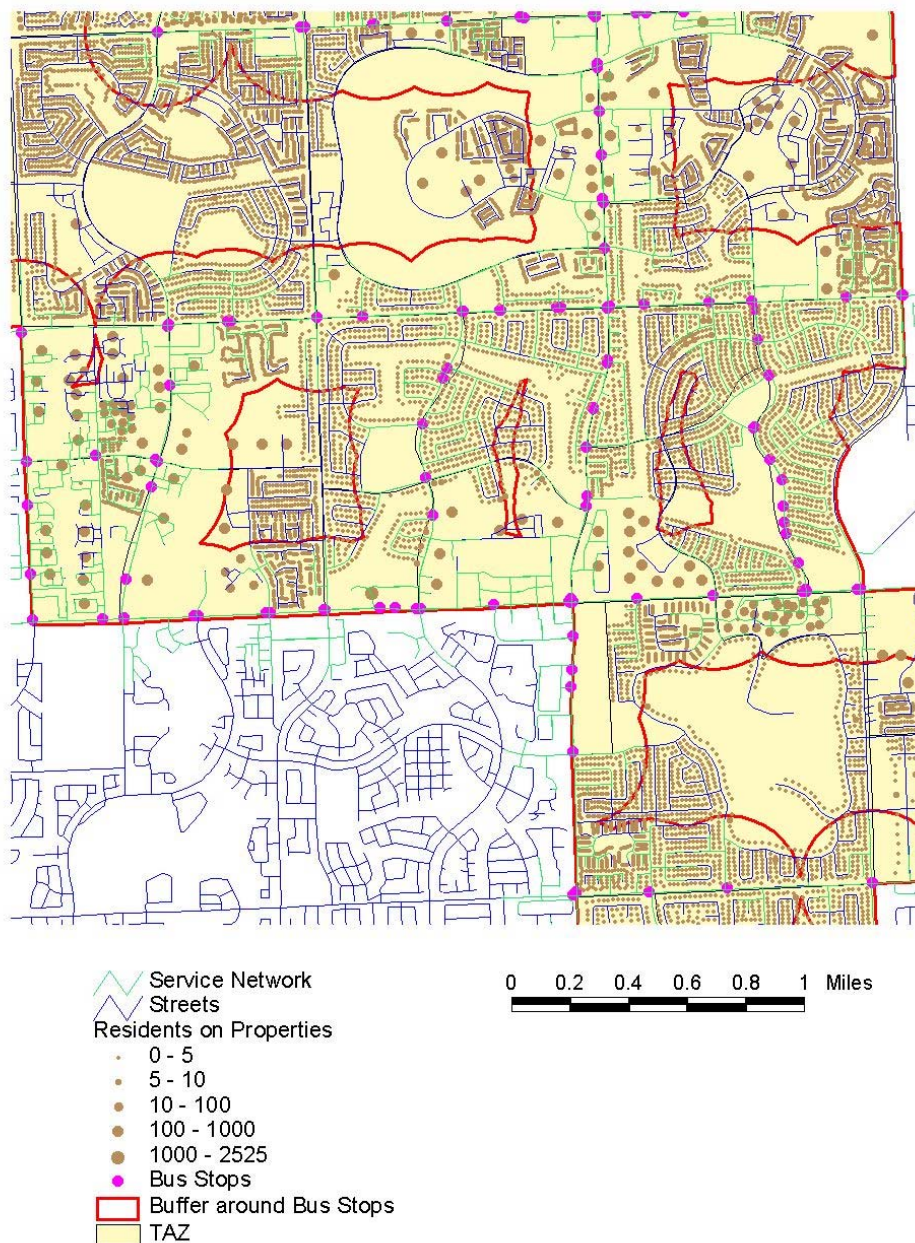


Figure 5-5. Transit Catchment Area, 0.25-mile Buffer, Streets, and Property Distribution

The first column identifies the TAZs used in the comparison. The second column provides the zonal population as given in the 1999 ZDATA1 file. The third and fourth columns are the population served and percentage of population served by transit, respectively, based on the buffer method, with a buffer size of 0.25 mile. The fifth and sixth columns provide the same

information obtained using the network ratio method, which assumes that population is evenly distributed along streets within 0.25-mile walking distance (and walking is assumed to happen along streets). The seventh and eighth columns show the results from the distance decayed method, in which walking distance is assumed to be up to 0.25 mile, the decay function is applied, and actual property locations are used to estimate population distribution along streets. The last two columns measure the differences between the distance decayed method and the two other methods in terms of percentage of difference. It may be seen that while the network ratio method in most cases results in a reduction in the service population when compared to the buffer method, the distance decayed method reduces the service population significantly and consistently, a result due mainly to the application of the decay function.

Table 5-1. Comparison of Transit Service Population (1/4-Mile Distance)

TAZ	Population	Buffer Method (1)		Network Ratio Method (2)		Distance Decayed Method (3)		Difference (%)	
		Pop. Served	%	Pop. Served	%	Pop. Served	%	(3)-(1)	(3)-(2)
867	4550	4550	100.0	3655	80.3	1490	32.8	-67.2	-47.6
870	5045	4918	97.5	2643	52.4	1210	24.0	-73.5	-28.4
873	2935	2034	69.3	1424	48.5	634	21.6	-47.7	-26.9
874	3070	2257	73.5	1302	42.4	724	23.6	-49.9	-18.8
875	1524	1296	85.0	1385	90.9	552	36.2	-48.8	-54.7
876	2527	2409	95.3	2033	80.4	677	26.8	-68.5	-53.6
877	3741	3537	94.6	1387	37.1	806	21.5	-73.0	-15.5
878	1211	1026	84.8	624	51.5	199	16.4	-68.4	-35.1
879	1938	1689	87.1	1282	66.1	596	30.8	-56.4	-35.4
880	1696	1604	94.6	785	46.3	326	19.2	-75.4	-27.1
881	1367	1312	95.9	703	51.4	245	17.9	-78.0	-33.5
882	1119	942	84.2	299	26.7	147	13.1	-71.0	-13.6
Total	30723	27574	89.8	17520	57.0	7605	24.8	-65.0	-32.3

5.2.2 Estimation of Workers' Accessibility to Transit

Currently in FSUTMS, workers' accessibility to transit services is estimated by the buffer method with some modification based on the consideration that most commercial developments are located near the main streets where transit services are provided. For example, the percentage of workers in a TAZ served by transit calculated using the buffer method (which assumes even distribution of workers throughout a TAZ) may be doubled or even assumed to be 100% (FDOT 1997). Improving the estimation of workers' access to transit services requires that location of employers and size of employment at each establishment be known. Estimation of location, type, and size of employment has remained a challenge. PBSJ (1998) discussed employment data sources and problems associated with them. Comparison between the employment data in the Miami-Dade County 1999 ZDATA2 file and the American Sales Leads (formerly InfoUSA) data also revealed large discrepancies. While it is possible to use a commercial employment database in employment transit accessibility analysis if the database has been verified and errors corrected, it has been decided for this study not to use the American Sales Leads' data since the data have not been validated.

One possible improvement to the estimation of employment accessibility to transit may be by using the land use GIS data, which provide information of spatial extent of employment. Miami-Dade County has land use GIS data created based on parcel data. The 1998 land use layer used in this study contains 18 different land uses. For this study, these land uses were matched with those in FSUTMS to estimate trips produced by different land uses. Table 5-2 establishes the equivalency between the 18 county land uses and the FSUTMS land uses.

Table 5-2. Conversion between Dade County and FSUTMS Land Use Categories

Land Use Description	FSUTMS Land Use Category
Agriculture, Industrial, Industrial Extraction	Industrial
Office, Airports/Ports, Communications, Utilities, Terminals, Plants, Institutions	Service
Shopping Centers, Commercial, Stadiums, Tracks	Commercial
Single-Family	Single-Family
Mobile Home Parks, Multi-Family (Including Migrant Camps), Townhouses, Two-Family (Duplex)	Multi-Family
Transient-Residential (Hotels/Motels)	Hotel/Motel
Water, Water Conservation Areas, Cemeteries, Parks (Including Preserves & Conservation), Streets/Roads, Expressways, Ramps, Vacant, Government Owned, Vacant Unprotected, Vacant, Protected, Privately Owned	-

The percentage of a given type of employment (industrial, commercial, and service) with access to transit service in a TAZ was estimated as the ratio between the area of the corresponding land use that fall within the transit buffer and the total area of that land use in the TAZ. In this calculation, workers of each type were assumed to be evenly distributed in areas of the corresponding land use. Figure 5-6 illustrates that for TAZs that do not fall into the 0.25-mile buffer of the bus stops, it is possible that some of the workers will not have transit access.

The choice of 0.25-mile buffer size was based on the TOB survey data, which showed that 97.5 of the transit trips were with 0.25 mile of a transit stop/station for non-home-based trips. Table 5-3 gives the cumulative percentages of transit trips from the TOB survey by walking distance.

Table 5-3. Cumulative Percentages of Transit Trips by Walking Distance

Walking Distance (ft)	Number of Samples	%	Cumulative %
0 – 100	71	10.36	10.36
100 – 200	309	45.11	55.47
200 – 300	100	14.60	70.07
300 – 400	38	5.55	75.62
400 – 500	41	5.99	81.61
500 – 1000	92	13.43	95.04
1000 – 1320	17	2.48	97.52
> 1320	17	2.48	100.00
Total	685	100.00	100.00

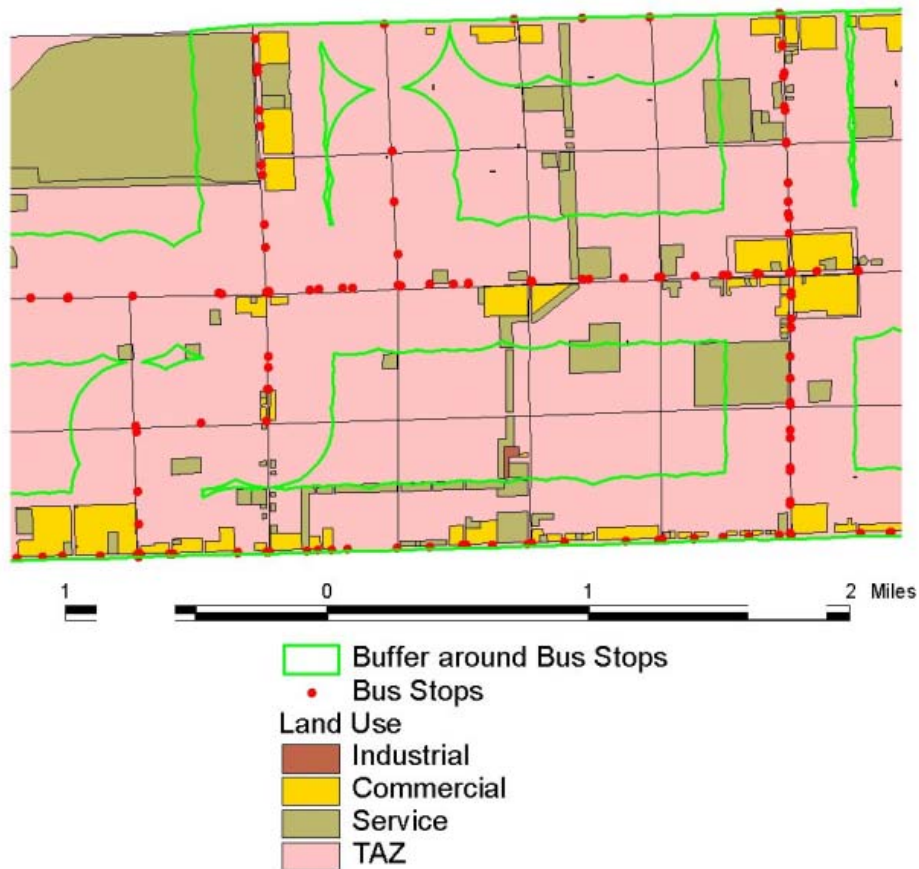


Figure 5-6. Transit Buffer and Distribution of Nonresidential Land Use

5.2.3 Estimation of Percentages of Trips with Transit Accessibility

In Section 5.2.1, transit service population, defined as people that are likely to use transit within the transit catchment area, has been estimated using the detailed street network to calculate walking distance and property information to estimate population distribution in a zone. In this section, the number of trips in a zone that may be made by transit mode is considered.

In FSUTMS, the transit walk accessibility in a zone is estimated as the percentages of trips within a short walk distance of 1/3 of a mile and a long walk distance of one mile. The buffer method is applied to create buffers of 1/3 of a mile and one mile around transit routes. With the assumption that trips are evenly distributed throughout a zone, the percentage of the area of a zone covered by the buffer area becomes the percentage of trips that are candidates for transit mode. In the absence of detailed information at the level of household types as defined in the trip production models (standard or lifestyle), even distribution of household, population, or trips is perhaps the best that can be achieved. Should some information be available about household locations and types, it would be interesting to see how the spatial distribution of trips could be estimated.

It would be possible to estimate the spatial distribution of trips if individual household location, household size, number of workers in household, number of children in household, and auto ownership were known. For our study, we had information at zonal level on population, the

number and locations of single-family dwelling units, the location and number of bedrooms of multi-family dwelling units, and the number of trips by purpose. Information on workers, children, and auto ownership, which are variables in the trips generation model of the Miami-Dade County FSUTMS model, was missing. Households, therefore, must be aggregated over these variables, and trips can be assigned to these generalized households based only on household size.

To convert the percentage of population with walk access to transit stops to the percentage of trips that may have a walk link to transit, the production trips must be first split between single-family and multi-family households. Trips by household type can then be assigned to each household according to the household size. The five-step procedure is described below.

Step 1. Split production trips between single-family and multi-family households

Since the 1999 FSUTMS trip generation model uses a lifestyle structure, separate production trips for single-family and multi-family households cannot be derived directly using the FSUTMS model. Therefore, the following procedure was applied to obtain single-family and multi-family production trips for each TAZ:

- (1) Using Travel 2000 survey data, sums of single-family households, single-family trips, multi-family households, and multi-family trips for each survey district were calculated. Trip rates were calculated for single-family households and multi-family households separately.
- (2) The ratio of trip rates between single-family households and multi-family households in each survey district was calculated.
- (3) Since the TAZ structure used in the survey was different from the 1999 TAZ structure, TAZs in the 1999 ZDATA1 file were assigned to the corresponding survey district using GIS.
- (4) The 1999 FSUTMS model was run to obtain production trips for HBW (Home-Based Work) and HBNW (Home-Based Non-Work) trip purposes (since HBW and HBNW have different lifestyle trip rate structures) for each TAZ. (Note: hotels from ZDATA1A were not included in the analysis because we only considered households.) The production trips were split between single-family and multi-family households based on the number of households from ZDATA1 and the ratio of trip rates from Step (2).

Step 2. Assign non-HBW trips to single-family households by household size

- (1) From the 1999 FSUTMS model, trip rates of HBSHOP (Home-Based Shop), HBSOC (Home-Based Social-Recreation), HBSCH (Home-Based School), and HBO (Home-Based Other) were added together for HBNW trip rates as shown below:

		Persons			
Children	Vehicle	1	2	3	4+
Without Children	0	0.32	0.58	1.09	3.19
	1	0.51	0.83	1.33	3.57
	2	0.75	1.19	2.27	4.37
	3+	2.68	3.34	4.66	6.85
With Children	0	0.00	1.80	3.14	5.31
	1	0.00	2.17	3.47	5.67
	2	0.00	3.01	4.32	6.42
	3+	0.00	5.57	6.89	8.99

- (2) The above HBNW trip rates were aggregated by household size since the information of auto ownership and presence of children was not available in the property file.

(A) From STP 283, the number of households in each cell was found:

		Persons			
Children	Vehicle	1	2	3	4+
Without Children	0	57667	21137	3737	1455
	1	102246	67613	11672	3928
	2	9435	86540	20041	7838
	3+	1320	10637	18200	18316
With Children	0	0	3527	6702	13613
	1	0	9593	21964	38891
	2	0	1435	32934	74523
	3+	0	196	5677	41400

- (B) The number of households was multiplied with trip rates to obtain the number of trips in each cell.
- (C) The number of households and the number of trips by household size were summed.
- (D) Trip rates by household size was calculated with results shown below:

Household Size	Persons			
	1	2	3	4+
Trip Rates	0.48	1.19	3.54	6.61

- (3) HBNW trips were assigned to each single-family household based on the ratio of trip rates for different household sizes.

Step 3. Assign non-HBW trips to multi-family households by household size

In the property parcel GIS layer, a point may represent a property of an apartment with 300 bedrooms or a group of condominium units with 100 bedrooms; both were considered multi-family dwelling units. To assign trips to these records, the number of households and the average household size in these complexes must be estimated. From the survey data, the average household size was derived for different complex types:

MF Housing Type	Average Household Size
Duplex	2.66
Apartment	2.32
Condo	1.81
Mobile Home	2.51

The number of units of each complex was determined by dividing the number of residents by the average household size. At this point, non-HBW trips were assigned to each multi-family household based on the ratio of trip rates for different household sizes.

Step 4. Assign HBW trips to single-family households by household size

- (1) The following HBW trip rates (from the 1999 FSUTMS model) were aggregated by number of workers in each household using a similar method as aggregating HBNW trips in Step 2.

Children	Vehicle	Workers		
		0	1	2+
Without Children	0	0	1.37	2.74
	1	0	1.48	3.07
	2	0	1.60	3.21
	3+	0	1.71	3.98
With Children	0	0	1.43	2.79
	1	0	1.54	3.15
	2	0	1.65	3.28
	3+	0	1.77	4.46

The aggregate trip rates by number of workers were:

Workers	0	1	2+
Trip Rates	0	1.54	3.48

- (2) The number of workers in each household was not available in the property file. ZDATA1A was used to obtain the relationship between household size and number of workers for each TAZ. For example, for TAZ 200, the ratio of number of workers to number of persons was 0.5312. Therefore, for households with three people, the estimated number of workers was 1.59.
- (3) HBW trips was assigned to each single-family household based on the number of workers weighted by the ratio of trip rates for different number of workers (interpolate trip rates for decimal number of workers).

Step 5. Assign HBW trips to multi-family households using the same method as for single-family households.

A comparison of the results from assigning trips to households with the transit service population estimated in Section 4.2 revealed that they were similar, with a maximum 8% differences. Table 5-4 illustrates for the same TAZs as in Table 5-1 the differences in percentage of population served and percentages of trips served by transit. For the trip purposes considered, including HBW and HB-Non-Work purposes, these percentages were virtually the same. These small differences may be explained by the fact that household size is the main variable based on which the spatial distribution of trips is determined, as in the case of population. Therefore, we may conclude that without detailed information of the spatial distribution of households of different types as specified in the FSUTMS trip generation model, the population estimate may suffice and an estimation of percentage of trips with transit service area may be unnecessary.

Table 5-4. Comparison of Percentage of Population and Percentage of Trips Served by Transit within a 0.5-Mile Distance

TAZ	Population	Pop. Served	%	HBW Trips	HBW Trips Served	%	HBNW Trips	HBNW Trips Served	%
867	4550	1514	33.3	4033	1308	32.4	2052	6384	32.1
870	5045	1426	28.3	4296	1192	27.7	1838	6672	27.5
873	2935	795	27.1	2478	643	25.9	994	3929	25.3
874	3070	816	26.6	2619	717	27.4	1150	4202	27.4
875	1524	568	37.3	879	323	36.7	603	1645	36.6
876	2527	735	29.1	1469	414	28.2	781	2776	28.1
877	3741	1128	30.1	2603	781	30.0	1437	4782	30.0
878	1211	244	20.2	1209	245	20.3	509	2503	20.3
879	1938	659	34.0	1933	664	34.4	1385	4023	34.4
880	1696	412	24.3	1678	405	24.1	853	3508	24.3
881	1367	317	23.2	1482	342	23.1	681	2939	23.2
882	1119	205	18.3	1210	222	18.3	444	2418	18.4
Total	30723	8819	28.7	25887	7257	28.0	12724	45781	27.8

5.3 Forecast of Transit Accessibility

Transit accessibility can be calculated for the base year model using the procedures outlined in the previous two subsections. While most built out TAZs may have little change in land use density or street configuration for a future forecast year, some currently underdeveloped TAZs may see significant changes in its land use and street configuration. The question thus arises: how can the transit accessibility for such TAZs be estimated for the future? Forecast of “transit friendliness” has been described by John *et al.* (1997), where transit friendliness was defined by a set of criteria including pedestrian use of sidewalks (presence of sidewalk, presence of shoulder, width of side walk, paved or unpaved sidewalk), street crossing rating (roadway width, traffic speed, traffic control devices), transit amenities rating, which considered both the physical environment and services provided at transit stop/station areas. The equation used for forecasting transit friendliness factor (*TTF*) was given as

$$TFF_{future} = TFF_{base} + [(TFF_{max} - TFF_{base}) \times (1 - e^{f(x)})]$$

where $f(x)$ is a function that governs *TFF* policy, which dictates how much change will occur in the transit environment between a base year and a future year.

Cervero (2001) studied the pedestrian access to Bay Area Rapid Transit (BART) stations in San Francisco; his regression model (shown in Table 5-5) used 34 cases to predict the percentage of access trips to BART stations by walking for all trip purposes.

Table 5-5. Regression Model For Predicting Percentage of Access trips to BART Stations by Walking, All Trip Purposes, 1992 (Cervero 2001)

Variables	Coefficients	Standard Error	Probability
Density			
Employment density (workers per acre in 0.5-mile buffer)	0.330	0.057	0.000
Residential density (households per acre in 0.5-mile buffer)	1.130	0.314	0.001
Land Use Type and Diversity			
Percent of residential land use in 0.5-mile buffer	0.532	0.312	0.100
Normalized entropy index of land use mix in 0.5-mile buffer	55.746	35.308	0.127
Transit Provisions			
Number of park-n-ride spaces at station	-0.020	0.004	0.000
Route miles per 1,000 households in 0.5-mile buffer	-3.121	1.099	0.009
Terminal or near terminal station: 0 = no, 1 = yes	19.569	6.886	0.009
Constant	-18.664	42.474	0.664

The model had a R^2 of 0.887; however, the land use type and diversity variables were difficult to forecast for a future year. Additionally, the BART-related variables did not apply in most cases to a dominantly bus transit network.

We attempt to forecast the walk accessibility as defined in Section 4 by identifying a set of variables that may be indicative of land use and street configuration and by establishing the relationship between transit accessibility with these variables. The variables were selected based on the criteria that they could be forecast for the future year and their compilation required minimum data processing.

Multiple regression models were developed for both production trips and attraction trips. Since the percentage of transit service population and percentage of trips served by transit in a zone are similar, the former will be used as the measure of transit accessibility for production trips. The transit accessibility for the attraction trips will be measured by the percentage of employees served by transit in a zone.

Section 5.3.1 describes the regression model for forecasting transit accessibility for production trips. Section 5.3.2 discusses the regression model used for forecasting transit accessibility for attraction trips. Section 5.3.3 addresses transit access via automobile.

5.3.1 Forecast of Transit Accessibility for Production Trips

The dependent variable, or the variable to be forecast, is the percentage of population served by transit calculated using the distance decayed method described in Section 4.1. The service population is estimated based on a 0.5-mile walk network and the application of the decay

function. For forecast purposes, the independent variables to be used to predict the dependent variable must be able to describe future land use, street configuration, and transit services. The data also need to be available and relatively easy to process. Based on these considerations, the following variables are considered as possible independent variables:

<i>MHH_RATIO</i>	—	ratio of multi-family households to total zonal households
<i>TOTHH_DEN</i>	—	number of households per acre in a zone
<i>MPOP_RATIO</i>	—	ratio of multi-family population to total zonal population
<i>TOTPOP_DEN</i>	—	number of residents per acre in a zone
<i>COM_RATIO</i>	—	ratio of commercial employees to total zonal employees
<i>SER_RATIO</i>	—	ratio of service employees to total zonal employees
<i>EMP_DEN</i>	—	number of employees per acre in a zone
<i>EMPPOP_DEN</i>	—	number of employees plus population per acre in a zone
<i>BUSRT_DEN</i>	—	bus route density in feet per acre
<i>ST_INT_DEN</i>	—	number of internal streets intersecting the boundary (per 1000 feet length of the TAZ perimeter) of a TAZ

For all the variables except the last one, *ST_INT_DEN*, data are readily available or can be easily compiled. The last variable, *ST_INT_DEN*, is an indicator of street configuration. For instance, an area with grid street network and small blocks will have a larger number of intersections between the local streets and the TAZ boundary. A community that has walls surrounding it with limited access roads and curvilinear streets will have a low number of intersections. Figures 5-7 and 5-8 illustrate the value ranges of *ST_INT_DEN* and the street network structure. Figure 5-9 shows the *ST_INT_DEN* values for TAZ for the study area. It may be seen that there is a general trend that *ST_INT_DEN* value is higher in areas that are closer to downtown and that are older parts of the county. *ST_INT_DEN* value deteriorates in the Miami Beach area and in suburban communities farther from downtown.



Figure 5-7. Value Ranges for *ST_INT_DEN* and Corresponding Mostly Grid Street Structures



Figure 5-8. Ranges of Value of *ST_INT_DEN* and Corresponding Mostly Curvilinear Street Structures

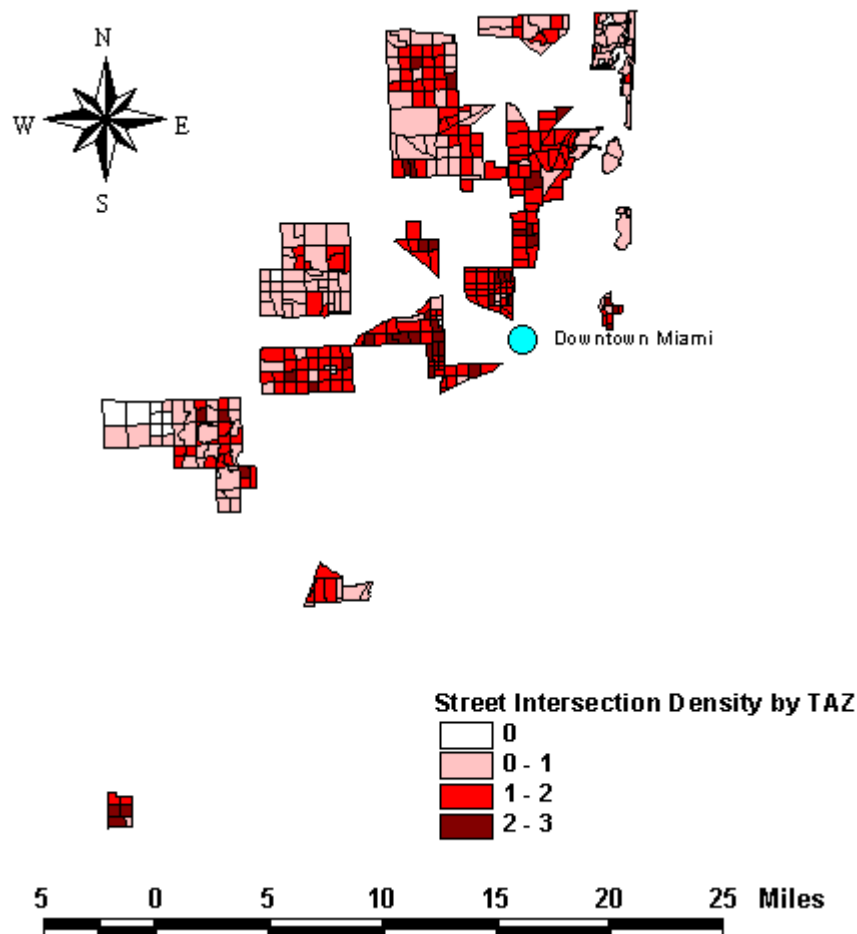


Figure 5-9. Street Intersection Density Distribution by TAZ in Study Area

For model development, 312 TAZs with nonzero HBW trips from the study area were selected. These TAZs were divided into two groups of equally size. One group (156 TAZs) was used for model estimation and the other (156 TAZs) for model validation. By applying the stepwise method to the first group of samples, the following variables were selected for linear regression.

Root MSE	0.12721	R-Square	0.9000
Dependent Mean	0.35641	Adj R-Sq	0.8974
Coeff Var	35.69176		

Parameter Estimates							
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
MHH_RATIO	MHH_RATIO	1	0.17309	0.02599	6.66	<.0001	2.79945
EMP_DEN	EMP_DEN	1	-0.00174	0.00058241	-3.00	0.0032	1.60381
BUSRUT_DEN	BUSRUT_DEN	1	0.00349	0.00031891	10.95	<.0001	3.94884
INT_DEN	ST_INT_DEN	1	0.06481	0.01268	5.11	<.0001	3.23715

The signs of the parameters estimated were reasonable and the variance inflation factors (VIF) did not indicate high correlations between independent variables. Additionally, the correlation matrix did not show high correlation between selected explanatory variables. However, since the partial R^2 for *EMP_DEN* to enter the model was not significant, it was excluded for the next stepwise procedure. The second candidate model was:

Root MSE	0.12759	R-Square	0.8994
Dependent Mean	0.35641	Adj R-Sq	0.8967
Coeff Var	35.79955		

Parameter Estimates							
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
MHH_RATIO	MHH_RATIO	1	0.20109	0.02835	7.09	<.0001	3.31199
EMPPPOP_DEN	EMPPPOP_DEN	1	-0.00154	0.00054337	-2.83	0.0053	3.06407
BUSRUT_DEN	BUSRUT_DEN	1	0.00339	0.00030890	10.99	<.0001	3.68253
INT_DEN	ST_INT_DEN	1	0.07171	0.01263	5.68	<.0001	3.19111

The signs of the estimated parameters, VIF, and correlation matrix were inspected again. The variable *EMPPPOP_DEN* was removed because it was the last variable entered into the model with lowest partial R^2 . The stepwise procedure was repeated and a new candidate model was obtained. This procedure was repeated and three more candidate models were estimated:

Model 3: ($R^2 = 0.8920$)
 $0.00306 \times BUSRUT_DEN + 0.06978 \times INT_DEN + 0.16933 \times MHH_RATIO$

Model 4: ($R^2 = 0.8841$)
 $0.00343 \times BUSRUT_DEN + 0.11471 \times SER_RATIO + 0.17777 \times MHH_RATIO$

Model 5: ($R^2 = 0.8722$)
 $0.00387 \times BUSRUT_DEN + 0.22723 \times MHH_RATIO$

For each candidate model, the standard influence on predicted value was checked for the presence of outliers. After excluding outliers, the new models were:

Model 1: ($R^2 = 0.9206$)
 $0.00306 \times BUSRUT_DEN + 0.12808 \times MHH_RATIO - 0.00138 \times EMP_DEN + 0.09092 \times ST_INT_DEN$

Model 2: ($R^2 = 0.9193$)
 $0.00293 \times BUSRUT_DEN + 0.14590 \times MHH_RATIO - 0.00097273 \times EMPPPOP_DEN + 0.09584 \times ST_INT_RUT$

Model 3: ($R^2 = 0.9174$)
 $0.00271 \times BUSRUT_DEN + 0.12409 \times MHH_RATIO + 0.09550 \times ST_INT_DEN$

Model 4: ($R^2 = 0.8942$)
 $0.00335 \times BUSRUT_DEN + 0.15367 \times MHH_RATIO + 0.12528 \times SER_RATIO$

Model 5: ($R^2 = 0.8778$)
 $0.00384 \times BUSRUT_DEN + 0.20832 \times MHH_RATIO$

To test the models, the second group of sample TAZs was used to calculate the errors in model predictions. Table 5-6 measures the model performance in terms of their adjusted R^2 and the sum

of square errors in their predictions. It may be seen that Model 2 had a higher adjusted R^2 but a larger prediction error than Model 3. Although Model 1 had the highest adjusted R^2 with the smallest prediction error, the result was not significantly better than Model 3, which had one less explanatory variable. Therefore, the structures of both Model 1 and 3 were carried forward for the final model calibration, for which the parameters were estimated using both groups of sample TAZs.

Table 5-6. Comparison of Five Accessibility Models

	Model 1	Model 2	Model 3	Model 4	Model 5
R^2	0.9206	0.9193	0.9174	0.8942	0.8778
Sum of Square Prediction Errors	2.5975	2.6837	2.6192	3.1188	0.3680

Using the same independent variables and both groups of TAZ, the model parameters were estimated again after examining the data for outliers and excluding them:

Model 1

Root MSE	0.10849	R-Square	0.9223
Dependent Mean	0.35472	Adj R-Sq	0.9213
Coeff Var	30.58513		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
MHH_RATIO	MHH_RATIO	1	0.10752	0.01636	6.57	<.0001	2.76168
ST_INT_DEN	ST_INT_DEN	1	0.10303	0.00791	13.03	<.0001	3.43572
BUSRUT_DEN	BUSRUT_DEN	1	0.00307	0.00018911	16.26	<.0001	3.59523
EMP_DEN	EMP_DEN	1	-0.00122	0.00041498	-2.93	0.0036	1.48441

Model 3

Root MSE	0.10870	R-Square	0.9223
Dependent Mean	0.35583	Adj R-Sq	0.9215
Coeff Var	30.54963		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
MHH_RATIO	MHH_RATIO	1	0.09586	0.01633	5.87	<.0001	2.74673
ST_INT_DEN	ST_INT_DEN	1	0.10730	0.00777	13.82	<.0001	3.32039
BUSRUT_DEN	BUSRUT_DEN	1	0.00289	0.00017305	16.69	<.0001	3.06967

Model 1 has one more variable (employment density) than Model 3 and a lower adjusted R^2 . Therefore Model 3 was chosen to be the final model, which had the following form:

$$\text{Percentage of Service Population} = 0.09586 \times \text{MHH_RATIO} + 0.10730 \times \text{ST_INT_DEN} + 0.00289 \times \text{BUSRUT_DEN}$$

The model states that the percentage of population in a zone served by transit increases with multi-family dwelling units, the number of intersections between internal streets and TAZ boundary, and bus route density.

All variables can be easily computed given future year dwelling units, employment, and bus route data. The number of intersections between internal streets (*ST_INT_DEN*) and TAZ boundary will not always be available as in cases of TAZs that have not yet been developed. Because this variable is devised to reflect the philosophy of urban design, it may be considered a policy variable and must be determined based on the planning guidelines. To facilitate the choice of a value for this variable, Table 5-7 is provided to select a value for *ST_INT_DEN*.

Table 5-7. Recommended Values for Intersection Density for Different Types of Street Structures

Street Structure Type	Intersection Density Interval (<i>ST_INT_DEN</i>)
No internal streets	0.0
Mostly curvilinear, irregular, or cal-de-sac streets	0.0 – 1.0
Majority grid, some discontinued, curvilinear, or irregular streets	1.0 – 2.0
Traditional/Neotraditional, mostly regular grid	2.0 – 3.0

5.3.2 Forecast of Transit Accessibility for Attraction Trips

To forecast transit accessibility for attraction trips, a regression model has been developed similar to that of production trips. Recall that employment accessibility was estimated using the buffer method and land use information (see Section 4.2). Because a large percentage of businesses are typically located near major arterials where buses operate, meaning that commercial land uses are mostly located near bus routes, the percentage of employment served by transit is high. In fact, out of the 324 sample TAZs served by transit in the category “nonzero employees,” over 95% of the employees were served by transit in 190 TAZs (58.64%). Therefore, for forecasting purposes, small TAZs that fall completely within the transit buffer areas can be considered 100% accessible. Such TAZs can be easily identified by calculating the percentage of the TAZs covered by buffers created around existing or planned transit routes. For the TAZs that are not entirely served, other factors may come into play, such as limited transit service coverage and large TAZ size with nonresidential use (resulting in a larger number of employees that are too far away from transit stops). Forecasting of employee transit accessibility for these TAZs is the focus of this section.

The forecasting model is a regression model similar to that of the production trips. The dependent variable is the percentage of employees in a zone served by transit. The independent variables investigated include the following:

<i>COM_RATIO</i>	—	the ratio of commercial employees to total employees in a zone
<i>SER_RATIO</i>	—	the ratio of service employees to total employees in zone
<i>SER_COM</i>	—	the ratio of service and commercial employees to total employees in a zone
<i>EMP_DEN</i>	—	the number of employees per acre
<i>EMPPOP_DEN</i>	—	the number of employees and residents per acre
<i>C_AREA_RATIO</i>	—	the ratio of commercial area to total area of a zone
<i>S_AREA_RATIO</i>	—	the ratio of service area to total area of a zone

I_AREA_RATIO — the ratio of industry area to total area of a zone
BUSRT_DEN — length (in feet) of bus routes per acre

There are 128 sample TAZs with nonzero employees that had transit services within 0.25 mile and for which accessibility was lower than 95% (i.e., less than 95% of the employees had access to transit services). The samples were divided into two equal-size groups (64 cases each) for model estimation and validation.

Applying the step-wise selection procedure resulted in three variables being selected: *SER_COM*, *EMPPOP_DEN*, and *BUSRT_DEN* as shown in the following SAS output.

Root MSE			0.22103	R_Square	0.9264		
Dependent Mean			0.75328	Adj_R_Sq	0.9228		
Coeff Var			29.34275				
Parameter Estimates							
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
SER_COM	SER_COM	1	0.62747	0.06334	9.91	<.0001	4.59971
EMPPOP_DEN	EMPPOP_DEN	1	−0.00617	0.00169	−3.65	0.0006	2.19185
BUSRUT_DEN	BUSRUT_DEN	1	0.00901	0.00184	4.90	<.0001	5.04448

There was no correlation between explanatory variables in this model. However, the sign for *EMPPOP_DEN* did not seem reasonable. Since this variable had the lowest partial R^2 , it was excluded for the next stepwise process. The first candidate model was:

Root MSE			0.24195	R_Square	0.9104		
Dependent Mean			0.75328	Adj_R_Sq	0.9075		
Coeff Var			32.11991				
Parameter Estimates							
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
SER_COM	SER_COM	1	0.58846	0.06833	8.61	<.0001	4.46844
BUSRUT DEN	BUSRUT DEN	1	0.00675	0.00189	3.56	0.0007	4.46844

Another candidate model was obtained by removing *BUSRT_DEN* because of its lower partial R^2 . We then reapplied the stepwise selection:

Root MSE		0.26347	R_Square	0.8920			
Dependent Mean		0.75328	Adj_R_Sq	0.8903			
Coeff Var		34.97638					
Parameter Estimates							
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
SER_COM	SER_COM	1	0.80304	0.03520	22.81	<.0001	1.00000

For both candidate models, the standard influence on predicted value was checked for presence of outliers, which were not found. The sums of square errors were calculated for the two models using the second group of TAZs. Table 5-9 compares the performance of two models.

Table 5-8. Comparison of Four Models for HBW Trips

	Model 1	Model 2
R^2	0.9104	0.8903
Sum of Square Prediction Errors	3.4466	3.8867

Model 1 had a higher R^2 and a smaller sum of square error, thus was selected as the final model structure. The parameters were estimated again using data from both groups of sample TAZs.

Root MSE	0.23530	R_Square	0.9185
Dependent Mean	0.77672	Adj_R_Sq	0.9172
Coeff Var	30.29440		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
BUSRUT_DEN	BUSRUT_DEN	1	0.00663	0.00141	4.69	<.0001	4.92441
SER_COM	SER_COM	1	0.62169	0.04910	12.66	<.0001	4.92441

There was one outlier, which was removed. After re-estimating, SAS produced the following model:

Root MSE	0.23369	R_Square	0.9193
Dependent Mean	0.77496	Adj_R_Sq	0.9180
Coeff Var	30.15564		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
BUSRUT_DEN	BUSRUT_DEN	1	0.00707	0.00143	4.95	<.0001	5.10309
SER_COM	SER_COM	1	0.60447	0.04987	12.12	<.0001	5.10309

The model may be written as

$$\text{Percentage of employees served} = 0.00707 \times \text{BUSRUT_DEN} + 0.60447 \times \text{SER_COM}$$

The model indicates that the percentage of employees served by transit in a zone increases with the ratio of service and commercial employees to total employees in a zone and with the bus route density in a zone.

5.3.3 Auto Accessibility

Some transit users access transit stations/stops by automobiles. In FSUTMS, a certain distance (e.g., 10 miles) is assumed to be the distance limit for auto access. Here we examine the auto access mode in terms of distance traveled by auto and the distance traveled by transit. For this purpose, the TOB survey data collected in Miami-Dade County were used. Survey samples were selected based on the following criteria:

- The sample involved a tour that began or ended at home
- The sample involved the use of automobile to access transit station/stop at the home end of the tour. The transit user could be a driver or a passenger during the trip to a transit stop/station by car.
- The origin and destination of the tour were geocoded.
- The non-home end of the tour was accessed by walk mode.

Here a transit tour means a series of trips that begin at the home and end at a non-home destination, or vice versa, with one or more trips accomplished using transit. The last condition was to ensure that the transit trip length could be estimated. It was assumed that a transit user would not switch to a non-transit mode before completing all trip segments that involved the use of transit. (Any walking occurring during transfers was assumed to be short and negligible). If the user did not access the non-home-based end of a tour by walk mode, it would be impossible most of the time to determine where the transit user completed the transit portion of the tour. On the other hand, if the non-home end of the tour was accessed by walk mode, the walking distance was reported in the survey, and the transit tour length could be estimated. There were a total of 54 samples that met the criteria.

To estimate the auto access distance, park-n-ride locations were geocoded. In addition to the 17 parking-n-ride facilities at Metrorail stations, seven additional park-n-ride facilities existed in Miami-Dade County. They were located at SW 152nd Street and U.S. 1, Cutler Ridge Mall on U.S. 1 in south Miami-Dade, SW 117th Avenue and 152nd Street, SW 104th Street and Hammocks Boulevard, SW 104th Street and SW 107th Avenue in Miami-Dade Community College, SW 107th Avenue and SW 72nd Street, and Golden Glades interchange that serve the 95 express bus route. Figure 5-10 illustrates the Metrorail alignment and stations, bus 95 routes and stops, and park-n-ride locations.

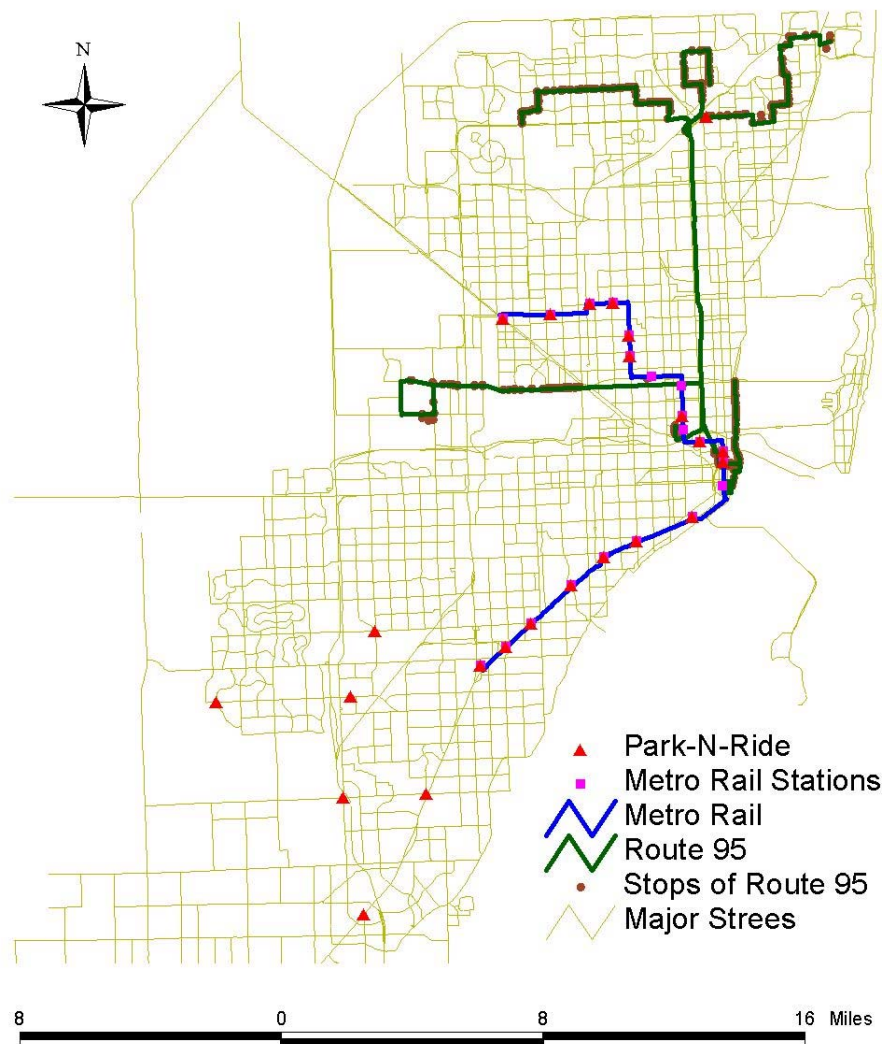


Figure 5-10. Metrorail Alignment, Bus 95 Routes, an Park-n-Ride Sites

The length of the auto access trip and the length of the transit portion of the tour (D1 as shown in Figure 5-11) were estimated as follows. Since the survey samples did not provide information as to which park-n-ride site the transit user drove to or was dropped off, it was assumed that the closest park-n-ride site to a user's home was the one that was used. Therefore, the auto access trip length was the shortest network distance between home and the closest park-n-ride location. The walking distance was estimated as the shortest distance from a Metrorail station or a 95 express bus stop in downtown to the non-home tour end (D3 in Figure 5-11). The transit tour length (D2 in Figure 5-11), which could include several linked transit trips if transfers were involved, was calculated as the route distance between the park-n-ride site and the transit tour non-home end. The tour might involve the use of multiple transit routes.

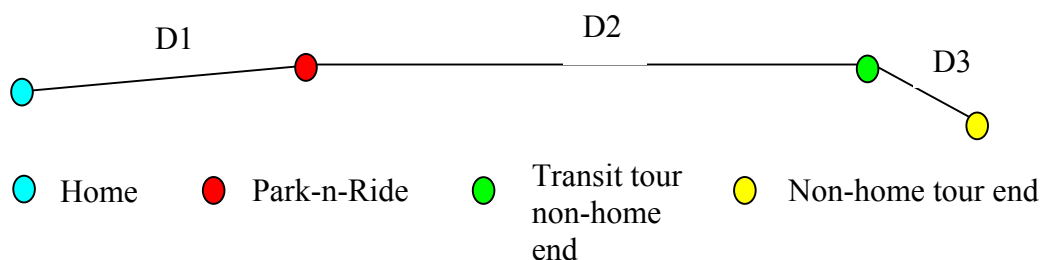


Figure 5-11. Transit Tours

Table 5-9 provides the auto access distance, the transit trip distance, the walk access distance, and the ratio between the transit trip distance to auto access distance for each of the samples. In Figure 5-12, the distribution of trip samples by auto access distance is plotted. The average driving distance was 4.8 miles. Most samples (92%) involved an auto access distance of less than 10 miles. Therefore, the 10-mile limit assumption used in FSUTMS appears to be reasonable.

Table 5-9. Home-Based Transit Linked Trips

ID	D1 (mile)	D2 (mile)	D3 (mile)	D2/D1
51385	2.548	24.114	0.053	9.464
50512	1.057	19.455	0.055	18.410
51351	0.564	18.866	0.539	33.455
51395	7.730	18.589	0.187	2.405
50310	1.349	14.673	0.401	10.879
50478	4.653	13.072	0.068	2.809
50663	12.940	12.648	0.048	0.977
50553	10.571	12.648	0.055	1.196
50516	9.122	12.648	0.055	1.387
50521	9.037	12.648	0.371	1.400
50515	8.512	12.648	0.340	1.486
51398	8.004	12.841	0.509	1.604
50513	7.324	12.648	0.035	1.727
50669	6.929	12.648	0.048	1.825
50511	6.492	12.648	0.055	1.948
50524	5.759	12.648	0.340	2.196
50499	5.216	12.648	0.055	2.425
50468	4.909	12.648	0.055	2.576
50469	4.481	12.648	0.055	2.822
50543	1.398	12.648	0.048	9.049
50425	0.924	12.648	0.472	13.683
50473	0.922	12.648	0.457	13.714
50505	0.642	12.648	0.055	19.711
50452	0.152	12.648	0.055	83.360
50508	1.263	11.711	0.340	9.274
50642	0.894	11.711	0.145	13.102
50572	3.992	10.893	0.697	2.729
50501	14.566	10.773	0.311	0.740

ID	D1 (mile)	D2 (mile)	D3 (mile)	D2/D1
50546	9.158	10.773	0.311	1.176
50559	4.598	10.773	0.372	2.343
58144	1.490	10.540	0.060	7.074
58157	5.160	10.510	0.120	2.037
50455	9.178	10.414	0.257	1.135
50539	7.598	10.414	0.180	1.371
50518	3.948	10.414	0.280	2.638
50533	3.297	10.414	0.180	3.159
50635	3.045	10.414	0.292	3.420
58147	8.410	10.350	0.110	1.231
58151	5.260	10.350	0.110	1.968
58142	2.070	10.350	0.110	5.000
58139	7.800	10.340	0.030	1.326
50608	0.377	10.007	0.311	26.531
50520	5.665	9.648	0.521	1.703
50646	1.566	9.648	0.121	6.159
51352	8.657	8.941	0.724	1.033
50595	2.250	8.512	0.260	3.784
51374	3.798	8.175	0.074	2.152
50380	1.995	8.175	0.553	4.098
50739	1.992	8.175	0.553	4.103
50424	10.977	5.941	0.055	0.541
51354	1.662	3.898	0.053	2.345
50369	1.890	2.658	0.551	1.407
50625	1.768	2.299	0.311	1.301
58141	3.220	1.200	0.310	0.373

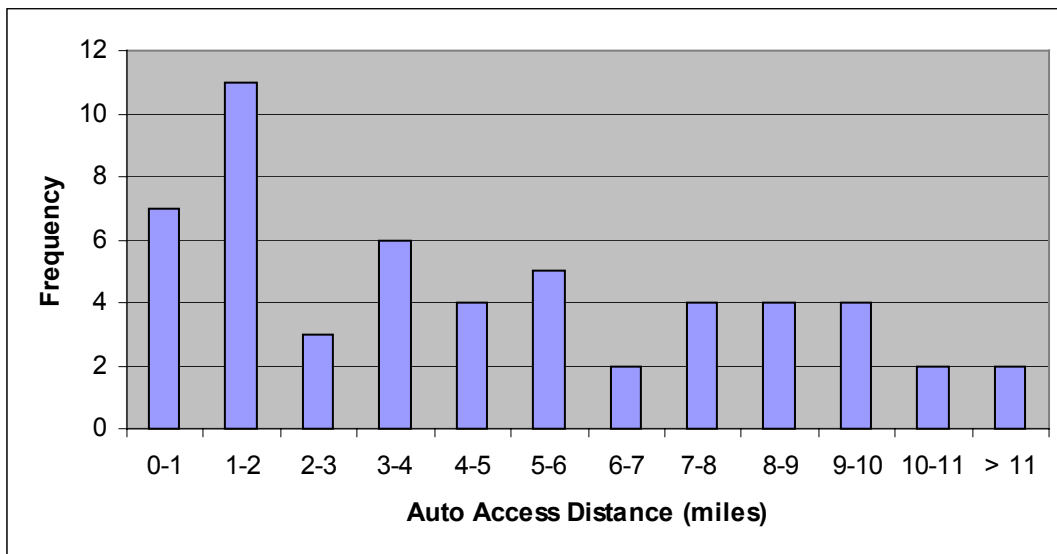


Figure 5-12. Distribution of Transit Tours by Auto Access Distance

Figure 5-13 plots the data in Table 5-9, including the auto access distance, linked transit trip distance, and the ratio between the two. While the auto access trip distance ($D1$) is plotted in an ascending order, the transit trip length ($D2$) does not reveal any trend, suggesting that transit trip length is not related to the auto access distance in Miami-Dade County. From data provided in Table 5-9, it may be observed that ratio between transit trip length and auto access trip distance ($D2/D1$) rarely (less than 5% of the time) falls under 1.0. The distribution of linked transit trips by trip length is shown in Figure 5-14. It may be seen that most trips (72%) are between eight and 14 miles. The average transit trip length is 11.2 miles, and the shortest transit trip is about 1.3 miles. This latter case was further investigated because of the short transit trip length compared to the auto trip length. It is believed that the transit user in that particular case was taking advantage of a Park-n-Ride lot at the Golden Galdes interchange to park the vehicle and use transit as a shuttle service between the Park-n-Ride lot and the workplace.

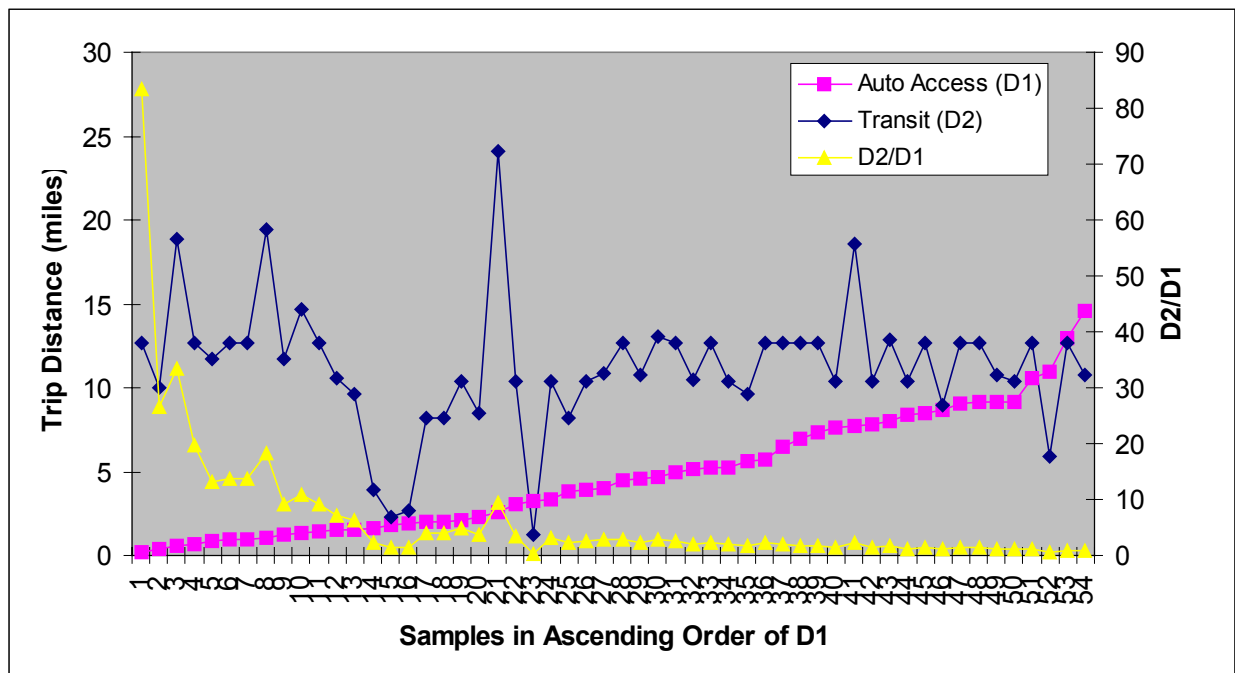


Figure 5-13. Comparison of Auto Access Distance and Linked Transit Trip Distance

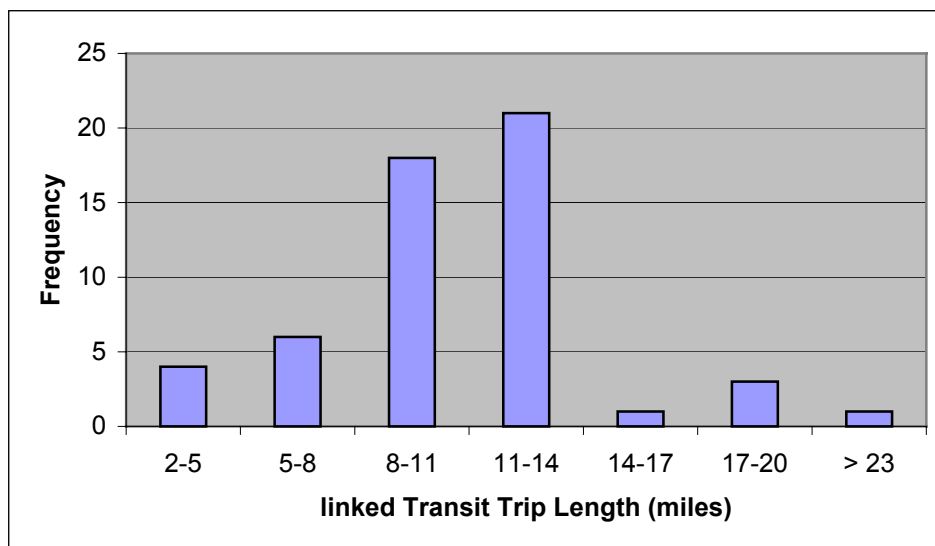


Figure 5-14. Distribution of Linked Transit Trips by Length (miles)

6. REGRESSION ANALYSIS OF FACTORS CONTRIBUTING TO TRANSIT USE

This chapter describes the multiple regression analyses performed to determine the most significant factors that affect transit use and that may be potentially incorporated into the FSUTMS modal split model. The data used in the regression analyses are described in Chapter 4, which are compiled using transit LOS, land use, demographics, and socioeconomic data at census tract level. The newly developed transit LOS variable, *DECAY_POP*, i.e., transit service population as a percentage of total zonal population described in Section 5, is also included in the regression analysis for Miami-Dade County. In the next section, the regression analysis procedure is briefly discussed. Sections 6.2 and 6.3 describe the regression models obtained for Miami-Dade, Broward, and Palm Beach counties. Section 6.4 further investigates variables that determine transit demand by controlling transit LOS variables. Section 6.5 provides a summary of the results of the regression analyses.

6.1 Multiple Linear Regression Analysis Procedure

Regression is a commonly used statistical tool for a variety of problems due to its simplicity. The regression analysis may be used for a number of purposes, including prediction, variable screening, model specification (system explanation), and parameter estimation (Myers 1990). In this study, the regression analysis was used for variable screening. This means that only the importance of the variables to transit usage was determined and that the models were not to be used for prediction purposes.

Linear models are merely empirical approximations and several models can have nearly equal performance in effectiveness. As a result, a model selection procedure, aimed at reducing the number of candidate models to a relatively small set, was developed to identify variables that had a significant linear relationship with transit use. These candidate models were then further investigated with more information on residuals, outlier and leverage computation, etc.

The model selection procedure applied in the study is as follows:

1. Create a pool of all the variables that may have effects on transit use;
2. Specify zero intercept (see below for the reasons);
3. Employ the stepwise procedure to select variables by specifying the significance levels to enter and stay in the model at 0.01;
4. Run regression;
5. Stop if no variables in the factor pool are entered in the model; otherwise, verify coefficient sign and exam variables' VIF for multicollinearity and Durbin-Watson statistics for autocorrelation; and
6. Eliminate those variables that do not meet the criteria in Step 5 from the variable pool and go to Step 1. If no variable are eliminated, go to Step 7.
7. Log the model and remove variables in the resulting model from the variable pool. Go to Step 2 if there are variables left in the variable pool.

Fixing the intercept at zero allows slope estimation with retention of the virtues of least squares, for example, performing least squares with extra information, namely the known value of intercept introduced into the model. Although one may argue that model intercept may not be

zero since some variations in the data may not be explained by the model variables, it is unimportant since we are attempting to identify the variables that may help better estimate the portion of person trips that are utilizing transit modes instead of determining the exact functional form of a transit ridership model. Additionally, region-wide accessibility variables have been included in the analysis and it is only reasonable to assume that zero accessibility results in zero transit trips. Consequently, intercept value was set to zero during the model selection.

Stepwise regression was applied in the variable selection procedure due to the large number of regressor variables, making all possible regressions prohibitive. One weakness of such a sequential algorithm is that it is designed to give one answer without providing information on a large number of subset models. The truly best model, if it exists, may not survive in such a procedure. Significance levels were set at 0.01 for regressor variables to enter and stay in the model in stepwise regression to assure only those with significant effects on transit use were included in the model.

The variance inflation factors (VIF) for the i th regression coefficient can be written as:

$$VIF = \frac{1}{1 - R_i^2}$$

where R_i^2 is the coefficient of multiple determination of the regression produced by regressing the variable x_i against the other regressor variables. The stronger the multiple correlation in this artificial regression, the lower the precision in the estimate of the coefficient b_i . The VIFs represent a considerably more productive approach for detection of multicollinearity than do the simple correlation values. They supply the user with an indication of which variables are causing multicollinearity and to what degree. Although there are no fixed rules regarding what values of VIF are sure signs of multicollinearity, it is generally believed that if any VIF exceeds 10, there is a reason for at least some concern, warranting considerations of variable deletion or alternatives to least squares estimation to combat the problem.

Note that once the regressor variables were eliminated, they were not moved back to the independent variables pool for the remaining regression runs. The procedure would reveal variables that met the minimum R^2 criteria but were prevented from entering the model by variables with larger partial F values. The procedure, however, was not designed to produce all probable models, which required every possible combination of variables to be tested. Since variables were removed due to violation of multicollinearity, autocorrelation, and, in some occasions, judgmental criteria, it was expected that most variables eliminated from the pool were not the factors that had great impacts on transit use.

6.2 Regression Results for Miami-Dade County

Tables 6-1 to 6-8 show the candidate models for production and attraction transit trips for Miami-Dade County. The first column indicates the sequence in which a model was calibrated. There might have been more models calibrated that are not included in the lists in the tables due to violation of the model selection criteria.

Table 6-1. Transit Production Trip Models for Miami-Dade (GIS Land Use Method)*

Order	Model**	Adjusted R^2	MSE
1	$4.08337 \times AccessTWPop1 - 2.273877E-7 \times AvgCOMValue$	0.7872	1.04107
2	$3.46374 \times AccessBPop1 + 1.90323 \times AvgEntropy$	0.7429	1.25766
3	$3.93377 \times AccessTWEm1 + 0.01213 \times Dailybus$	0.7174	1.38244
4	$3.79545 \times AccessTWSC1 + 0.87198 \times Bth1Pop1$	0.7153	1.39300
5	$0.02645 \times DECAY_POP + 3.40651 \times AccessBEm1$	0.7106	1.41602
6	$1.42110 \times AccessHISC + 0.00178 \times WDailybus$	0.7045	1.44569
7	$2.65313 \times AccessH2SC$	0.6632	1.64778
8	$2.62374 \times AccessH1Em$	0.6628	1.64980
9	$2.63267 \times AccessH2Em$	0.6612	1.65746
10	$2.96676 \times AccessH1Pop - 0.16291 \times WAvghdway$	0.6875	1.52874
11	$2.32094 \times AccessH2Pop$	0.6536	1.69485
12	$0.32101 \times SFDUDensity1$	0.6520	1.70242
13	$2.35974 \times Wth1Pop1 - 1.07394 \times WAvghdway$	0.6801	1.56489
14	$1.89892 \times Wth2Pop1$	0.6506	1.70953
15	$2.29572 \times Wth1Em1 - 0.16524 \times WPeakhdway$	0.6716	1.60659
16	$1.85250 \times Wth2Em1$	0.6455	1.73456
17	$1.84901 \times Wth1SC1$	0.6448	1.73777
18	$1.84111 \times Bth2Pop1$	0.6415	1.75413
19	$0.09369 \times SFPopDensity1$	0.6408	1.75732
20	$1.81802 \times Bth1Em1$	0.6387	1.76736
21	$1.81770 \times Bth2Em1$	0.6387	1.76754
22	$1.81741 \times Bth1SC1$	0.6387	1.76762
23	$1.81712 \times Bth2SC1$	0.6387	1.76779
24	$2.11868 \times ServiceArea\%$	0.6330	1.79551

Notes:

* The list does not include the models with the adjusted R^2 's that are less than 0.6.

** Variable descriptions in alphabetic order:

AccessBEm1: transit best skim accessibility weighted by total employment.*AccessBPop1*: transit best skim accessibility weighted by population opportunities.*AccessH1Em*: highway Time1 accessibility weighted by total employment.*AccessH1Pop*: highway Time1 accessibility weighted by total population.*AccessHISC*: highway Time1 accessibility weighted by service plus commercial employment.*AccessH2Em*: highway Time2 accessibility weighted by total employment.*AccessH2Pop*: highway Time2 accessibility weighted by total population.*AccessH2SC*: highway Time2 accessibility weighted by service plus commercial employment.*AccessTWEm1*: transit walk skim accessibility weighted by total employment.*AccessTWPop1*: transit walk skim accessibility weighted by population opportunities.*AccessTWSC1*: transit walk skim accessibility weighted by service plus commercial employment.*AvgCOMValue*: Average commercial property value.*AvgEntropy*: land use mix in buffer area of a census tract.*Avghdway*: composite average daily headway in a tract.*Bth1Em1*: transit best skim over highway FFSkim weighted by total employment.*Bth1Pop1*: transit best skim over highway FFSkim weighted by total population.*Bth1SC1*: transit best skim over highway FFSkim weighted by service plus commercial employment.*Bth2Em1*: transit best skim over highway FFSkimT weighted by total employment.*Bth2Pop1*: transit best skim over highway FFSkimT weighted by total population.*Bth2SC1*: transit best skim over highway FFSkimT weighted by service plus commercial employment.*Dailybus*: unweighted average number of bus runs per stop in a tract.*DECAY_POP*: percentage of service population.*ServiceArea%*: percentage of tract area served by transit based on ¼-mile buffers around.*SFDUDensity1*: single-family dwelling unit density in buffer area of a tract.

SFPopDensity1: single-family population density in buffer area of a tract.
WAvghdway: composite daily headway for a tract weighted by service area.
WDailybus: average number of bus runs per stop in a tract weighted by service area.
WPeakhdway: composite peak hour headway for a tract weighted by service area.
Wth1Pop1: transit walk skim over highway FFSkim weighted by total population.
Wth1Em1: transit walk skim over highway FFSkim weighted by total employment.
Wth1SC1: transit walk skim over highway FFSkim weighted by service plus commercial employment.
Wth2Pop1: transit walk skim over highway FFSkimT weighted by total population.
Wth2Em1: transit walk skim over highway FFSkimT weighted by total employment.

Table 6-2. Transit Attraction Trip Models for Miami-Dade (GIS Land Use Method)*

Order	Model**	Adjusted R^2	MSE
1	$4.34419 \times JOBS_HH1 + 3.14845 \times ServiceArea\%$	0.5769	17.78523
2	$7.41456 \times AccessTWPop1$	0.5310	19.71456
3	$7.74808 \times AccessBPop1$	0.5272	19.87342
4	$6.20752 \times AccessH1Pop$	0.5249	19.97282
5	$7.69955 \times AccessTWEm1$	0.5241	20.00388
6	$6.23680 \times AccessH2Pop$	0.5228	20.05838
7	$7.73858 \times AccessTWSC1$	0.5227	20.06394
8	$7.96710 \times AccessBEm1$	0.5213	20.12173
9	$8.00816 \times AccessBSC1$	0.5207	20.14942
10	$7.07078 \times AccessH1Em$	0.5010	20.97579
11	$4.20759 \times JOBS_HH^m1 + 4.68718 \times AccessH2Em$	0.5317	19.68582
12	$7.09590 \times AccessH1SC$	0.4978	21.11049
13	$7.12104 \times AccessH2SC$	0.4953	21.21720
14	$0.06812 \times Dailybus$	0.4865	21.58746
15	$0.79084 \times SFDUDensity1$	0.4761	22.02462
16	$0.23093 \times SFPopDensity1$	0.4703	22.26703
17	$0.06905 \times MFPopDensity1$	0.4485	23.18487
18	$0.00491 \times WDailybus + 0.00027380 \times AvgParcelSize$	0.4950	21.22858

Notes:

* The list does not include the models obtained after the 18th model.

** Variable descriptions in alphabetic order:

AccessBEm1: transit best skim accessibility weighted by total employment.
AccessBPop1: transit best skim accessibility weighted by population opportunities.
AccessBSC1: transit best skim accessibility weighted by service plus commercial employment.
AccessH1Em: highway Time1 accessibility weighted by total employment.
AccessH1Pop: highway Time1 accessibility weighted by total population.
AccessH1SC: highway Time1 accessibility weighted by service plus commercial employment.
AccessH2Em: highway Time2 accessibility weighted by total employment.
AccessH2Pop: highway Time2 accessibility weighted by total population.
AccessH2SC: highway Time2 accessibility weighted by service plus commercial employment.
AccessTWEm1: transit walk skim accessibility weighted by total employment.
AccessTWSC1: transit walk skim accessibility weighted by service plus commercial employment.
AccessTWPop1: transit walk skim accessibility weighted by population opportunities.
AvgParcelSize: average parcel size of single- and multi-family properties in buffer area of a tract.
Dailybus: unweighted average number of bus runs per stop in a tract.
JOBS_HH1: job house balance.
JOBS_HH^m1: modified jobs-housing balance in buffer area of a census tract.
MFPopDensity1: multi-family population density in buffer area of a tract.
ServiceArea%: percentage of tract area served by transit based on ¼-mile buffers around.
SFDUDensity1: single-family dwelling unit density in buffer area of a tract.
SFPopDensity1: single-family population density in buffer area of a tract.
WDailybus: average number of bus runs per stop in a tract weighted by service area.

Table 6-3. Transit Production Trip Models for Miami-Dade (Property Method)*

Order	Model**	Adjusted R^2	MSE
1	$3.64651 \times \text{AccessTWPop2} + 1.70315 \times \text{AvgEntropy}$	0.7688	1.46294
2	$3.99117 \times \text{AccessBPop2} + 0.97263 \times \text{Bth1SC2}$	0.7478	1.59539
3	$1.91722 \times \text{AccessH1SC} + 0.00165 \times \text{WDailybus}$	0.7114	1.82629
4	$2.68863 \times \text{AccessH1Pop}$	0.6855	1.99016
5	$2.80528 \times \text{Wth1Pop2} - 0.17126 \times \text{WAvghdway}$	0.7105	1.83146
6	$3.07629 \times \text{Wth2Pop2} - 1.61882 \times \text{JOBS_HH2}$	0.7075	1.85038
7	$3.06468 \times \text{AccessH2SC}$	0.6846	1.99568
8	$2.69966 \times \text{AccessH2Pop}$	0.6843	1.99777
9	$3.03072 \times \text{AccessH1Em}$	0.6842	1.99840
10	$3.04198 \times \text{AccessH2Em}$	0.6830	2.00590
11	$2.16272 \times \text{Wth1Em2}$	0.6801	2.02391
12	$2.16232 \times \text{Wth2Em2}$	0.6801	2.02424
13	$2.15833 \times \text{Wth1SC2}$	0.6795	2.02809
14	$2.45539 \times \text{Bth1Pop2} - 0.00000458 \times \text{AvgParcelCOMSize2}$	0.6994	1.90164
15	$2.63279 \times \text{Bth2Pop2} - 0.17573 \times \text{WPeakhdway}$	0.6983	1.90913
16	$2.12266 \times \text{Bth1Em2}$	0.6737	2.06444
17	$2.12225 \times \text{Bth2Em2}$	0.6737	2.06481
18	$2.12157 \times \text{Bth2SC2}$	0.6736	2.06513
19	$0.20935 \times \text{SFDUDensity2} + 0.01523 \times \text{Dailybus}$	0.6861	1.98585
20	$2.43348 \times \text{ServiceArea\%}$	0.6459	2.24023
21	$0.05325 \times \text{DECAY_POP}$	0.6456	2.24253
22	$0.04262 \times \text{TotalDensity2} + 0.00013569 \times \text{AvgParcelSize2}$	0.6437	2.25444

Notes:

* Models with the adjusted R^2 's that are less than 0.6 are not included.

** Variable descriptions in alphabetic order:

AccessBPop2: transit best skim accessibility weighted by population opportunities.*AccessH1Em*: highway Time1 accessibility weighted by total employment.*AccessH1Pop*: highway Time1 accessibility weighted by total population.*AccessH1SC*: highway Time1 accessibility weighted by service plus commercial employment.*AccessH2Em*: highway Time2 accessibility weighted by total employment.*AccessH2Pop*: highway Time2 accessibility weighted by total population.*AccessH2SC*: highway Time2 accessibility weighted by service plus commercial employment.*AccessTWPop2*: transit walk skim accessibility weighted by population opportunities.*AvgParcelCOMSize2*: average parcel size for commercial properties.*AvgParcelSize2*: average parcel size for all residential properties.*AvgEntropy*: land use mix in buffer area of a census tract.*Bth1Em2*: transit best skim over highway FFSkim weighted by total employment.*Bth1Pop2*: transit best skim over highway FFSkim weighted by total population.*Bth1SC2*: transit best skim over highway FFSkim weighted by service plus commercial employment.*Bth2Em2*: transit best skim over highway FFSkimT weighted by total employment.*Bth2Pop2*: transit best skim over highway FFSkimT weighted by total population.*Bth2SC2*: transit best skim over highway FFSkimT weighted by service plus commercial employment.*Dailybus*: unweighted average number of bus runs per stop in a tract.*DECAY_POP*: percentage of service population.*JOBS_HH2*: job house balance.*ServiceArea%*: percentage of tract area served by transit based on ¼-mile buffers around.*SFDUDensity2*: single-family dwelling unit density in buffer area of a tract.*TotalDensity2*: total employment plus population density in buffer area.*WAvghdway*: composite daily headway for a tract weighted by service area.*WDailybus*: average number of bus runs per stop in a tract weighted by service area.*WPeakhdway*: composite peak hour headway for a tract weighted by service area.*Wth1Em2*: transit walk skim over highway FFSkim weighted by total employment.

Wth1Pop2: transit walk skim over highway FFSkim weighted by total population.
Wth1SC2: transit walk skim over highway FFSkim weighted by service plus commercial employment.
Wth2Em2: transit walk skim over highway FFSkimT weighted by total employment.
Wth2Pop2: transit walk skim over highway FFSkimT weighted by total population.

Table 6-4. Transit Attraction Trip Models for Miami-Dade (Property Method)*

Order	Model**	Adjusted R^2	MSE
1	$0.54304 \times SFPopDensity2 + 4.75169 \times JOBS_HH2$	0.6180	16.05603
2	$5.87543 \times ServiceArea\%$	0.5439	19.17108
3	$7.41456 \times AccessTWPpop2$	0.5310	19.71456
4	$7.74808 \times AccessBPop2$	0.5272	19.87342
5	$6.20752 \times AccessH1Pop$	0.5249	19.97282
6	$7.69036 \times AccessTWEm2$	0.5234	20.03564
7	$6.23680 \times AccessH2Pop$	0.5228	20.05838
8	$7.72936 \times AccessTWSC2$	0.5220	20.09466
9	$7.95846 \times AccessBEm2$	0.5207	20.14991
10	$7.99978 \times AccessBSC2$	0.5200	20.17735
11	$7.07078 \times AccessH1Em$	0.5010	20.97579
12	$4.20759 \times JOBS_HH^m2 + 4.68718 \times AccessH2Em$	0.5317	19.68582
13	$7.09590 \times AccessHISC$	0.4978	21.11049
14	$7.12104 \times AccessH2SC$	0.4953	21.21720
15	$0.06812 \times Dailybus$	0.4865	21.58746
16	$0.79084 \times SFDUDensity2$	0.4761	22.02462
17	$0.00491 \times WDailybus + 0.00027380 \times AvgParcelSize$	0.4950	21.22858
18	$0.17142 \times TotalPopDensity2 + 0.00029256 \times AvgParcelSFSize$	0.4795	21.87810
19	$9.45346 \times AvgEntropy$	0.4099	24.80696

Notes:

* Models with the adjusted R^2 's that are less than 0.4 are not included.

** Variable descriptions in alphabetic order:

AccessBEm2: transit best skim accessibility weighted by total employment.
AccessBPop2: transit best skim accessibility weighted by population opportunities.
AccessBSC2: transit best skim accessibility weighted by service plus commercial employment.
AccessH1Em: highway Time1 accessibility weighted by total employment.
AccessH1Pop: highway Time1 accessibility weighted by total population.
AccessHISC: highway Time1 accessibility weighted by service plus commercial employment.
AccessH2Em: highway Time2 accessibility weighted by total employment.
AccessH2Pop: highway Time2 accessibility weighted by total population.
AccessH2SC: highway Time2 accessibility weighted by service plus commercial employment.
AccessTWEm2: transit walk skim accessibility weighted by total employment.
AccessTWPpop2: transit walk skim accessibility weighted by population opportunities.
AccessTWSC2: transit walk skim accessibility weighted by service plus commercial employment.
AvgEntropy: land use mix in buffer area of a census tract.
AvgParcelSFSize: average parcel size of single-family properties in buffer area of a tract.
AvgParcelSize: average parcel size of single- and multi-family properties in buffer area of a tract.
Dailybus: unweighted average number of bus runs per stop in a tract.
JOBS_HH2: job house balance.
JOBS_HH^m2: modified jobs-housing balance in buffer area of a census tract.
ServiceArea%: percentage of tract area served by transit based on ¼-mile buffers around.
SFPopDensity2: single-family population density in buffer area of a tract.
SFDUDensity2: single-family dwelling unit density in buffer area of a tract.
TotalPopDensity2: total population density in buffer area of a census tract.
WDailybus: average number of bus runs per stop in a tract weighted by service area.

Table 6-5. Transit Production Trip Models for Miami-Dade (Buffer Method)*

Order	Model**	Adjusted R^2	MSE
1	$2.78918 \times \text{AccessTWPop3} + 2.23290 \times \text{AvgEntropy3}$	0.7349	1.56248
2	$4.05242 \times \text{AccessBPop3} + 0.01181 \times \text{Dailybus}$	0.6867	1.84663
3	$1.78761 \times \text{AccessHISC} + 0.00156 \times \text{WDailybus}$	0.6657	1.97009
4	$2.83203 \times \text{AccessH1Em}$	0.6408	2.11708
5	$2.51575 \times \text{AccessH2Pop}$	0.6382	2.13253
6	$2.85733 \times \text{AccessH2SC}$	0.6379	2.13435
7	$2.83641 \times \text{AccessH2Em}$	0.6369	2.14007
8	$0.05051 \times \text{DECAY_POP}$	0.6165	2.26010

Notes:

* The list does not include the models with the adjusted R^2 's that are less than 0.6.

** Variable descriptions in alphabetic order:

AccessBPop3: transit best skim accessibility weighted by population opportunities.*AccessH1Em*: highway Time1 accessibility weighted by total employment.*AccessHISC*: highway Time1 accessibility weighted by service plus commercial employment*AccessH2Em*: highway Time2 accessibility weighted by total employment*AccessH2Pop*: highway Time2 accessibility weighted by total population.*AccessH2SC*: highway Time2 accessibility weighted by service plus commercial employment.*AccessTWPop3*: transit walk skim accessibility weighted by population opportunities.*AvgEntropy3*: land use mix in buffer area of a census tract.*Dailybus*: unweighted average number of bus runs per stop in a tract.*DECAY_POP*: percentage of service population.*Wdailybus*: average number of bus runs per stop in a tract weighted by service area.**Table 6-6. Transit Attraction Trip Models for Miami-Dade (Buffer Method)***

Order	Model**	Adjusted R^2	MSE
1	$5.53718 \times \text{JOBS_HH3} + 0.04240 \times \text{Dailybus}$	0.5779	25.84608
2	$6.06660 \times \text{AccessTWPop3} + 0.00001618 \times \text{AvgValue}$	0.5508	27.50937
3	$6.60408 \times \text{AccessBPop3} + 0.00001505 \times \text{AvgSFValue}$	0.5448	27.87717
4	$9.13106 \times \text{AccessTWEm3}$	0.5041	30.36674
5	$9.18368 \times \text{AccessTWSC3}$	0.5036	30.39884
6	$9.49700 \times \text{AccessBSC3}$	0.5010	30.55620
7	$9.44338 \times \text{AccessBEm3}$	0.5009	30.56152
8	$7.31302 \times \text{AccessH1Pop}$	0.4934	31.02256
9	$7.30918 \times \text{AccessH2Pop}$	0.4862	31.46228
10	$0.00477 \times \text{WDailybus} + 0.00041309 \times \text{AvgParcelSFSize}$	0.5139	29.76686
11	$5.75022 \times \text{JOBS_HH}^m3 + 0.00037010 \times \text{AvgParcelSize}$	0.5008	30.56641
12	$6.55752 \times \text{ServiceArea\%}$	0.4598	33.08095
13	$15.83873 \times \text{Wth2Pop3} + 6.52689 \times \text{AccessH1Em}$	0.4975	30.76869
14	$17.46226 \times \text{Bth2Pop3} + 6.64814 \times \text{AccessHISC}$	0.4952	30.90945
15	$8.15706 \times \text{AccessH2Em}$	0.4440	34.04829
16	$8.19789 \times \text{AccessH2SC}$	0.4422	34.15851
17	$2.86335 \times \text{SFDUDensity3}$	0.4264	35.12224
18	$0.83178 \times \text{SFPopDensity3}$	0.4171	35.69524

Notes:

* The list does not include the models with the adjusted R^2 's that are less than 0.4.

** Variable descriptions in alphabetic order:

AccessBEm3: transit best skim accessibility weighted by total employment.*AccessBPop3*: transit best skim accessibility weighted by population opportunities.*AccessBSC3*: transit best skim accessibility weighted by service plus commercial employment.*AccessH1Em*: highway Time1 accessibility weighted by total employment.

AccessH1Pop: highway Time1 accessibility weighted by total population.
AccessH1SC: highway Time1 accessibility weighted by service plus commercial employment
AccessH2Em: highway Time2 accessibility weighted by total employment
AccessH2Pop: highway Time2 accessibility weighted by total population.
AccessH2SC: highway Time2 accessibility weighted by service plus commercial employment.
AccessTWEm3: transit walk skim accessibility weighted by total employment.
AccessTWP3: transit walk skim accessibility weighted by population opportunities.
AccessTWSC3: transit walk skim accessibility weighted by service plus commercial employment.
AvgSFValue: average single-family housing value.
AvgParcelSFSize: average parcel size of single-family properties in buffer area of a tract.
AvgParcelSize: average parcel size of single- and multi-family properties in buffer area of a tract.
AvgValue: average housing value.
Bth2Pop3: transit best skim over highway Time2 accessibility weighted by population opportunities.
Dailybus: unweighted average number of bus runs per stop in a tract.
JOBS_HH3: job house balance.
JOBS_HH^m3: modified jobs-housing balance in buffer area of a census tract.
ServiceArea%: percentage of tract area served by transit based on ¼-mile buffers around.
SFDUDensity3: single-family dwelling unit density in buffer area of a tract.
SFPopDensity3: single-family population density in buffer area of a tract.
WDailybus: average number of bus runs per stop in a tract weighted by service area.
Wth2Pop3: transit walk skim over highway Time2 accessibility weighted by population opportunities.

Table 6-7. Transit Production Trip Models for Miami-Dade (Tract Method)*

Order	Model**	Adjusted R^2	MSE
1	$4.38241 \times AccessTWP3$	0.8087	0.76548
2	$4.60293 \times AccessTBP3 + 0.00980 \times Black$	0.7985	0.80667
3	$3.76624 \times AccessTWEm3 + 1.63266 \times AvgEntropy3$	0.7733	0.90719
4	$4.13504 \times AccessTWSC3 + 0.15332 \times AvgPChild$	0.7679	0.92879
5	$6.32606 \times AccessTBEm3$	0.7278	1.08938
6	$4.69664 \times AccessTBSC3 + 0.92305 \times \%HH0Child$	0.7613	0.95529
7	$1.72310 \times AccessH1SC + 3.05882 \times MFDUNoAuto\%$	0.7539	0.98505
8	$1.74977 \times AccessH2SC + 3.09875 \times DUNoAuto\%$	0.7537	0.98574
9	$1.79217 \times AccessH2Em + 3.21592 \times SFDUNoAuto\%$	0.7529	0.98909
10	$1.10775 \times ServiceArea\% + 0.00161 \times Dailybus$	0.7301	1.08024
11	$0.01579 \times Dailybus + 0.02637 \times TotalDensity3$	0.6503	1.39965
12	$0.97623 \times AvgP0Child - 0.16558 \times AvgHdway$	0.6618	1.35353
13	$0.94494 \times AvgWrkrChild - 1.02232 \times AvgHdway + 0.03667 \times MFPopDensity3$	0.6178	1.52956

Notes:

* The list does not include the models with the adjusted R^2 's that are less than 0.6.

** Variable descriptions in alphabetic order:

%HH0Child: percentage of households without children.

AccessH1SC: highway Time1 accessibility weighted by service plus commercial employment

AccessH2SC: highway Time2 accessibility weighted by service plus commercial employment.

AccessH2Em: highway Time2 accessibility weighted by total employment.

AccessTBEm3: transit best skim accessibility weighted by total employment.

AccessTBP3: transit best skim accessibility weighted by population opportunities.

AccessTBSC3: transit best skim accessibility weighted by service plus commercial employment.

AccessTWEm3: transit walk skim accessibility weighted by total employment.

AccessTWP3: transit walk skim accessibility weighted by population opportunities.

AccessTWSC3: transit walk skim accessibility weighted by service plus commercial employment.

AvgEntropy3: land use mix in a census tract.

AvgHdway: composite average daily headway in a tract.

AvgP0Child: average number of persons in households without children.

AvgPChild: average number of persons in households with children.

AvgWrkrChild: average number of workers in households with children.
Black: percentage of black population in a census tract.
Dailybus: unweighted average number of bus runs per stop in a tract.
DUNoAuto%: percentage of households without auto.
TotalDensity4: total residential property density in a census tract.
MFDUNoAuto%: percentage of multi-family households without auto.
MFPopDensity4: multi-family population density in a census tract.
ServiceArea%: percentage of tract area served by transit in a census tract.
SFDUNoAuto%: density of single-family households without auto in a census tract.

Table 6-8. Transit Attraction Trip Models for Miami-Dade (Tract Method)*

Order	Model**	Adjusted R^2	MSE
1	$1.18955 \times SFDUDensity4 + 4.99499 \times JOBS_HH4$	0.5838	17.39299
2	$5.86252 \times ServiceArea\%$	0.5365	19.36766
3	$0.38240 \times SFPopDensity4 + 4.46452 \times AccessTWPpop4$	0.5706	17.94212
4	$7.79699 \times AccessBPop4$	0.5273	19.75343
5	$7.76121 \times AccessTWEm4$	0.5243	19.87962
6	$7.80060 \times AccessTWSC4$	0.5230	19.93177
7	$6.23331 \times AccessH1Pop$	0.5226	19.94784
8	$8.01836 \times AccessBEm4$	0.5203	20.04413
9	$1.13679 \times AvgPChild4$	0.5203	20.04562
10	$8.05977 \times AccessBSC4$	0.5199	20.06391
11	$6.25771 \times AccessH2Pop$	0.5197	20.07132
12	$0.06993 \times Dailybus$	0.5142	20.29973
13	$2.65693 \times AvgWrkrChild4$	0.5077	20.57177
14	$6.88689 \times \%HH0Child$	0.4996	20.91250
15	$4.63213 \times JOBS_HH^m4 + 4.17396 \times AccessH1Em$	0.5295	19.65952
16	$7.04111 \times AccessHISC$	0.4810	21.68659
17	$7.02991 \times AccessH2Em$	0.4804	21.71401
18	$1.11579 \times AvgAuto0Child4 + 0.00428 \times WDailybus$	0.5353	19.41892
19	$7.05800 \times AccessH2SC$	0.4774	21.83842

Notes:

* The list does not include the models obtained after the 19th model.

** Variable descriptions in alphabetic order:

%HH0Child: percentage of households without children.

AccessBEm4: transit best skim accessibility weighted by total employment.

AccessBPop4: transit best skim accessibility weighted by population opportunities.

AccessBSC4: transit best skim accessibility weighted by service plus commercial employment.

AccessH1Em: highway Time1 accessibility weighted by total employment.

AccessH1Pop: highway Time1 accessibility weighted by total population.

AccessHISC: highway Time1 accessibility weighted by service plus commercial employment

AccessH2Em: highway Time2 accessibility weighted by total employment

AccessH2Pop: highway Time2 accessibility weighted by total population.

AccessH2SC: highway Time2 accessibility weighted by service plus commercial employment.

AccessTWEm4: transit walk skim accessibility weighted by total employment.

AccessTWPpop4: transit walk skim accessibility weighted by population opportunities.

AccessTWSC4: transit walk skim accessibility weighted by service plus commercial employment.

AvgAuto0Child4: average number of cars owned by households without children.

AvgPChild4: average number of persons in households with children.

AvgWrkrChild4: average number of workers in households with children.

Dailybus: unweighted average number of bus runs per stop in a tract.

JOBS_HH4: job house balance.

JOBS_HH^m4: modified jobs-housing balance in a census tract.

ServiceArea%: percentage of tract area served by transit.

SFDUDensity4: single-family dwelling unit density in a tract.
SFPopDensity4: single-family population density in a tract.
WDailybus: average number of bus runs per stop in a tract weighted by service area.

From the results, it may be observed that regardless of which method was used to compile the variables for the Miami-Dade models, accessibility variables, especially those weighted by population, were always an important factor for both production and attraction transit trips. Models with a single accessibility variable explained more than 50%, some even up to 70%, of variation in the response data. The production models calibrated using the census tract level data (compiled with Tract Method), however, showed that accessibility variables weighted by total and service/commercial employment to be more significant. These production models may be misleading since the data were not detail enough as those using the Land Use Method, Buffer Method, or Property Method.

DECAY_POP, which is the transit service population in a tract as a percentage of the total tract population, was significant in three out of the four models, which are those based on data compiled with the GIS Land Use Method, Property Method, and Buffer Method. A similar and more crude measure of pedestrian accessibility, the percentage of transit service area in a tract defined by a 1/4-mile buffer also appeared in three models: those based on data compiled using the GIS Land Use Method, Property Method, and Tract Method. It is interesting to note that while the Buffer Method is the most popular method in transit demand analysis, *ServiceArea%* did not appear in the related models.

ServiceArea% is much easier to calculate than most of the accessibility and LOS variables, including *DECAY_POP*. Since it performed nearly as well as *DECAY_POP* for the zero-intercept linear models calibrated with the data compiled by the GIS Land Use Method and Property Method, one may question the need for *DECAY_POP*, which requires much more effort to compute. In fact, the correlation coefficient between these two variables is around 0.7.

Figure 6-1 illustrates the relationship between *DECAY_POP* and *ServiceArea%* by plotting the two values for each census tract included in the study, sorted by *ServiceArea%*. It may be seen that even when service area percentage reaches 100 percent for some tracts, the distance decayed transit service population percentage is much lower, remaining below 50 percent in most cases. In fact, service area coverage reaches 95% to 100% for nearly 40% of the census tracts in the study. For these tracts, *ServiceArea%* is no longer sensitive to the differences in transit use. Table 6.9 shows the adjusted R^2 's of regression models before and after eliminating nearly 40% of the sampled tracts with more than 95% service coverage. The results showed a relatively large drop in the adjusted R^2 's for the *ServiceArea%* models. Table 6-9 suggests that *ServiceArea%* may not be a good indicator to model transit use for areas with lower transit service coverage. Consequently, it is preferable to implement *DECAY_POP* to estimate transit use.

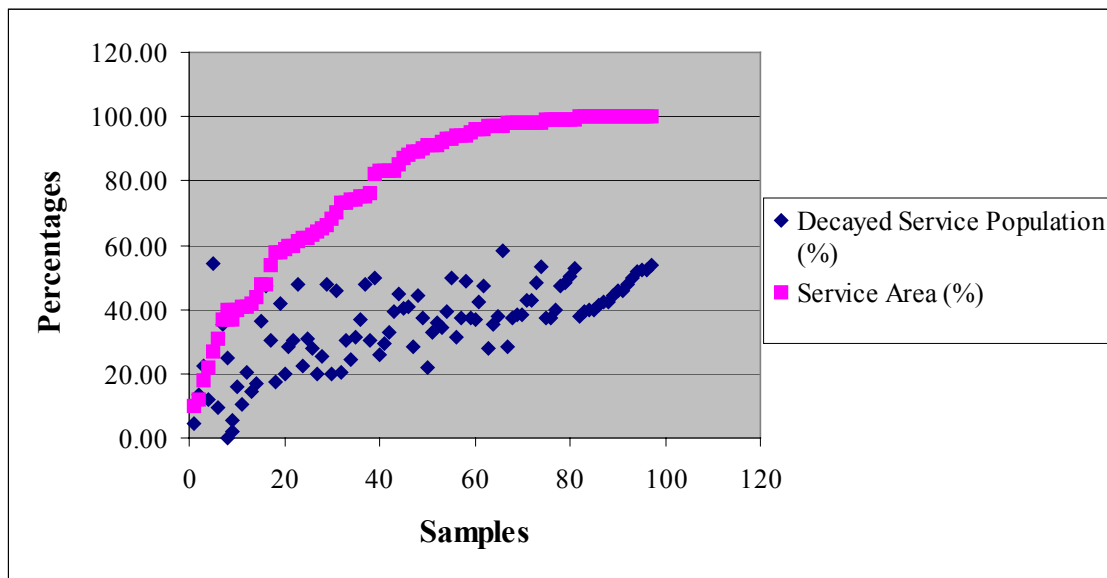


Figure 6-1. Distance decayed Transit Service Population and Percentage of Service Area in a Tract

Table 6-9. Adjusted R^2 's for the *DECAY POP* and *ServiceArea%* Models

	GIS Land Use Method		Property Method		Buffer Method	
	Before	After	Before	After	Before	After
<i>DECAY POP</i>	0.6662	0.5408	0.6456	0.5256	0.6165	0.4834
<i>ServiceArea%</i>	0.6330	0.4625	0.6459	0.5049	0.5750	0.4032

Among the variables other than accessibility variables that were included in the pool of candidate models, *Dailybus*, a transit LOS variable representing the average number of bus runs per stop in a tract, appeared to be a better indicator for transit use since the variable alone explained nearly 60% of the variation for transit production models. Consequently, since accessibility-based variables are relatively more complicated to calculate, *Dailybus* may be utilized to quantify transit use. This also implies that data compiled using different methods as described in Chapter 4 will produce similar results, since the majority of the special characteristics based on smaller spatial units (e.g., TAZs) are lost in the process of aggregation to the tract level. An analysis of correlation between the *Dailybus* variable and other headway variables revealed that they were strongly correlated. This is expected since the shorter the headway the higher number of daily bus runs for a given service period.

The models explained more variations in transit use at the production end than at the attraction end, probably due to the fact that data at the production sides were more accurate and frequently updated and, in comparison, employment data were less reliable. Consequently, models for transit production trips could be better estimated.

For attraction models, all of the four methods revealed that *ServiceArea%* was a critical factor and the variable alone could explain nearly 50% of the variation in the data. Although the jobs-housing balance variable was significant, most of the variations in the data had been explained by other variable(s) included in the same model. Therefore, the partial R^2 for this variable was much lower.

6.3 Regression Results for Broward and Palm Beach Counties

The same model selection procedure was applied to examine the Broward and Palm Beach data. As mentioned in Sections 4-7 and 4-8, data were compiled using the Buffer Method and the Tract Method for Broward County and the Tract Method for Palm Beach County. Tables 6-10 to 6-15 show the resulting candidate models for these two counties for transit production and attraction trips. The results suggested the following:

1. Although accessibility variables were consistently significant for the production trip models estimated for both counties, the models' adjusted R^2 's were relatively lower than those for Miami-Dade models. One of the probable causes is that a small number of samples were used to calibrate the transit impedance functions for both counties (i.e., 64 for Broward County and 36 for Palm Beach County). On the other hand, highway impedance functions for Broward and Palm Beach counties are calibrated using 1,628 and 1,618 samples, respectively. The ability to make meaningful interpretations of the models based on the impedance functions calibrated from a relatively small number of samples may be questionable.
2. Different from the Miami-Dade models, transit levels of service and demographic variables did not appear to be relevant to transit production trips.
3. Lower than expected adjusted R^2 's were obtained for the Broward and Palm Beach transit attraction models. As a result, attraction models were inconclusive.

Table 6-10. Transit Production Trip Models for Broward (Buffer Method)*

Order	Model**	Adjusted R^2	MSE
1	$7.58103 \times \text{AccessTWP}3$	0.5644	6.58990
2	$6.86415 \times \text{Wth2Pop3}$	0.5543	6.74322
3	$6.47177 \times \text{Wth1Em3}$	0.5504	6.80128
4	$6.45828 \times \text{Wth2Em3}$	0.5504	6.80191
5	$6.38996 \times \text{Wth1SC3}$	0.5499	6.81001
6	$6.37684 \times \text{Wth2SC3}$	0.5498	6.81091
7	$7.33734 \times \text{AccessTWSC3}$	0.5426	6.91955
8	$7.47374 \times \text{AccessTWEm3}$	0.5425	6.92159
9	$7.49584 \times \text{AccessBSC3}$	0.5341	7.04771
10	$7.57539 \times \text{AccessBEm3}$	0.5333	7.06003
11	$5.52147 \times \text{Bth1Pop3}$	0.5285	7.13327
12	$5.43488 \times \text{Bth1SC3}$	0.5252	7.18233
13	$5.42692 \times \text{Bth2SC3}$	0.5252	7.18293
14	$5.50305 \times \text{Bth1Em3}$	0.5246	7.19199
15	$5.49478 \times \text{Bth2Em3}$	0.5246	7.19250
16	$4.36714 \times \text{AccessHISC}$	0.5077	7.44824
17	$4.39861 \times \text{AccessH2SC}$	0.5069	7.45978

Notes:

* The list does not include the models with the adjusted R^2 's that are less than 0.5.

** Variable descriptions in alphabetic order:

AccessBEm3: transit best skim accessibility weighted by total employment.

AccessBSC3: transit best skim accessibility weighted by service plus commercial employment.

AccessHISC: highway Time1 accessibility weighted by service plus commercial employment.

<i>AccessH2SC</i> :	highway Time2 accessibility weighted by service plus commercial employment.
<i>AccessTWEm3</i> :	transit walk skim accessibility weighted by total employment.
<i>AccessTWPop3</i> :	transit walk skim accessibility weighted by population opportunities.
<i>AccessTWSC3</i> :	transit walk skim accessibility weighted by service plus commercial employment.
<i>Bth1Em3</i> :	transit best skim over highway Time1 accessibility weighted by total employment.
<i>Bth1Pop3</i> :	transit best skim over highway Time1 accessibility weighted by population opportunities.
<i>Bth1SC3</i> :	transit best skim over highway Time1 accessibility weighted by social plus commercial employment.
<i>Bth2Em3</i> :	transit best skim over highway Time2 accessibility weighted by total employment.
<i>Bth2SC3</i> :	transit best skim over highway Time2 accessibility weighted by social plus commercial employment.
<i>Wth2Pop3</i> :	transit walk skim over highway Time2 accessibility weighted by population opportunities.
<i>Wth1Em3</i> :	transit walk skim over highway Time1 accessibility weighted by total employment.
<i>Wth1SC3</i> :	transit walk skim over highway Time1 accessibility weighted by service plus commercial employment.
<i>Wth2Em3</i> :	transit walk skim over highway Time2 accessibility weighted by total employment.
<i>Wth2SC3</i> :	transit walk skim over highway Time2 accessibility weighted by service plus commercial employment.

Table 6-11. Transit Attraction Trip Models for Broward (Buffer Method)*

Order	Model**	Adjusted R^2	MSE
1	$31.82185 \times JOBS_{HH^m}3$	0.3796	382.52197
2	$26.58495 \times JOBS_{HH}3$	0.3749	385.41922
3	$1.44683 \times TotalPopDensity3$	0.3174	420.87990

Notes:

* The list does not include the models with the adjusted R^2 's that are less than 0.3.

** Variable descriptions in alphabetic order:

JOBS_{HH}3: job house balance.

JOBS_{HH^m}3: modified jobs-housing balance in buffer area of a census tract.

TotalPopDensity3: total population density in buffer area of a census tract.

Table 6-12. Transit Production Trip Models for Broward (Tract Method)*

Order	Model**	Adjusted R^2	MSE
1	$6.64969 \times AccessTWPop4$	0.5521	4.60767
2	$6.69978 \times AccessTWEm4$	0.5485	4.64541
3	$6.57190 \times AccessTWSC4$	0.5474	4.65656
4	$6.24768 \times AccessBPop4$	0.5425	4.70712
5	$6.79136 \times AccessBEm4$	0.5408	4.72434
6	$6.71358 \times AccessBSC4$	0.5404	4.72845
7	$5.54739 \times Wth1Em4$	0.5249	4.88822
8	$5.53406 \times Wth2Em4$	0.5245	4.89190
9	$5.47016 \times Wth1SC4$	0.5227	4.91107
10	$5.45752 \times Wth2SC4$	0.5223	4.91458
11	$5.79719 \times Wth1Pop4$	0.5143	4.99693
12	$4.66474 \times Bth1Pop4$	0.4977	5.16795
13	$4.65256 \times Bth2Pop4$	0.4970	5.17456
14	$4.66667 \times Bth1Em4$	0.4952	5.19319
15	$4.65838 \times Bth2Em4$	0.4949	5.19608
16	$4.60368 \times Bth1SC4$	0.4944	5.20143
17	$4.59581 \times Bth2SC4$	0.4942	5.20423
18	$3.55934 \times AccessH2SC$	0.4650	5.50371
19	$3.52313 \times AccessH1SC$	0.4635	5.51928

Notes:

*	The list does not include the models obtained after the 19 th model.
**	Variable descriptions in alphabetic order:
<i>AccessBEm4:</i>	transit best skim accessibility weighted by total employment.
<i>AccessBPop4:</i>	transit best skim accessibility weighted by population opportunities.
<i>AccessBSC4:</i>	transit best skim accessibility weighted by service plus commercial employment.
<i>AccessHISC:</i>	highway Time1 accessibility weighted by service plus commercial employment
<i>AccessH2SC:</i>	highway Time2 accessibility weighted by service plus commercial employment.
<i>AccessTWEm4:</i>	transit walk skim accessibility weighted by total employment.
<i>AccessTWPop4:</i>	transit walk skim accessibility weighted by population opportunities.
<i>AccessTWSC4:</i>	transit walk skim accessibility weighted by service plus commercial employment.
<i>Bth1Em4:</i>	transit best skim over highway Time1 accessibility weighted by total employment.
<i>Bth1Pop4:</i>	transit best skim over highway Time1 accessibility weighted by population opportunities.
<i>Bth1SC4:</i>	transit best skim over highway Time1 accessibility weighted by social plus commercial employment.
<i>Bth2Em4:</i>	transit best skim over highway Time2 accessibility weighted by total employment.
<i>Bth2Pop4:</i>	transit best skim over highway Time2 accessibility weighted by population opportunities.
<i>Bth2SC4:</i>	transit best skim over highway Time2 accessibility weighted by social plus commercial employment.
<i>ForeignBorn:</i>	percentage of population that are immigrants in a tract.
<i>ServiceArea%:</i>	percentage of tract area served by transit.
<i>Wth1Em4:</i>	transit walk skim over highway Time1 accessibility weighted by total employment.
<i>Wth1Pop4:</i>	transit walk skim over highway Time1 accessibility weighted by population opportunities.
<i>Wth1SC4:</i>	transit walk skim over highway Time1 accessibility weighted by service plus commercial employment.
<i>Wth2Em4:</i>	transit walk skim over highway Time2 accessibility weighted by total employment.
<i>Wth2SC4:</i>	transit walk skim over highway Time2 accessibility weighted by service plus commercial employment.

Table 6-13. Transit Attraction Trip Models for Broward (Buffer Method)*

Order	Model**	Adjusted R^2	MSE
1	$1.15340 \times TotalPopDensity4$	0.3825	176.26305
2	$21.05298 \times JOBS_{HH4} + 21.26576 \times AccessTWSC4 - 4.64754 \times AvgAuto0Child$	0.4324	162.01245
3	$18.67696 \times Wth2SC4 + 17.96289 \times JOBS_{HH^m4} - 3.81654 \times AvgAuto0Child$	0.4126	167.66791
4	$23.19681 \times Wth1SC4$	0.3391	188.63588
5	$23.35029 \times Wth2Em4$	0.3371	189.21638
6	$23.39105 \times Wth1Em4$	0.3369	189.25673
7	$27.22611 \times AccessTWEm4$	0.3283	191.72787
8	$24.31474 \times Wth1Pop4$	0.3280	191.80408
9	$24.25466 \times Wth2Pop4$	0.3277	191.88477
10	$26.53106 \times AccessTWPop4$	0.3195	194.24378
11	$13.77653 \times AccessH2Pop$	0.3123	196.28119
12	$13.60269 \times AccessH1Pop$	0.3118	196.42679
13	$15.23312 \times AccessHISC$	0.3097	197.02841
14	$15.34730 \times AccessH2SC$	0.3090	197.22229

Notes:

*	The list does not include the models with the adjusted R^2 's that are less than 0.3.
**	Variable descriptions in alphabetic order:
<i>AccessH1Pop:</i>	highway Time1 accessibility weighted by total population.
<i>AccessH1SC:</i>	highway Time1 accessibility weighted by service plus commercial employment
<i>AccessH2Pop:</i>	highway Time2 accessibility weighted by total population.
<i>AccessH2SC:</i>	highway Time2 accessibility weighted by service plus commercial employment.
<i>AccessTWEm4:</i>	transit walk skim accessibility weighted by total employment.
<i>AccessTWPop4:</i>	transit walk skim accessibility weighted by population opportunities.
<i>AccessTWSC4:</i>	transit walk skim accessibility weighted by service plus commercial employment.

AvgAuto0Child: average number of cars owned by households without children.
JOBS_HH4: job house balance.
JOBS_HH^m4: modified jobs-housing balance in a census tract.
TotalPopDensity4: total population density in a census tract.
Wth1Em4: transit walk skim over highway Time1 accessibility weighted by total employment.
Wth1Pop4: transit walk skim over highway Time1 accessibility weighted by population opportunities.
Wth1SC4: transit walk skim over highway Time1 accessibility weighted by service plus commercial employment.
Wth2Em4: transit walk skim over highway Time2 accessibility weighted by total employment.
Wth2Pop4: transit walk skim over highway Time2 accessibility weighted by population opportunities.
Wth2SC4: transit walk skim over highway Time2 accessibility weighted by service plus commercial employment.

Table 6-14. Transit Production Trip Models for Palm Beach (Tract Method)*

Order	Model**	Adjusted R^2	MSE
1	$24.48368 \times AccessTWEm4$	0.6797	2.33270
2	$25.38213 \times AccessTWSC4$	0.6737	2.37654
3	$11.24775 \times AccessTWPop4$	0.5382	3.36306
4	$7.65490 \times Wth1SC4 - 0.80089 \times AvgAuto0Child$	0.5421	3.33468
5	$7.44418 \times Wth2SC4 - 0.94698 \times AvgAuto0Child$	0.5336	3.39645
6	$6.57754 \times Wth1Em4 - 1.41710 \times White$	0.5231	3.47291
7	$4.38788 \times Wth2Em4$	0.4679	3.87536
8	$5.26366 \times Bth1Em4$	0.4552	3.96804
9	$5.27067 \times Bth2Em4$	0.4547	3.97149
10	$5.13728 \times Bth1SC4$	0.4530	3.98367
11	$5.14425 \times Bth2SC4$	0.4526	3.98702
12	$3.80155 \times Bth1Pop4 + 9.47064 \times AccessBEm4$	0.4717	3.84724
13	$3.85980 \times Bth2Pop4 + 9.36043 \times AccessBSC4$	0.4696	3.86325
14	$3.91680 \times Wth1Pop4$	0.4294	4.15548
15	$8.17180 \times AccessBPop4 + 0.12668 \times TotalDensity4$	0.4031	4.34692

Notes:

* The list does not include the models with the adjusted R^2 's that are less than 0.3.

** Variable descriptions in ascending order:

AccessBEm4: transit best skim accessibility weighted by total employment.
AccessBPop4: transit best skim accessibility weighted by population opportunities.
AccessBSC4: transit best skim accessibility weighted by service plus commercial employment.
AccessTWEm4: transit walk skim accessibility weighted by total employment.
AccessTWPop4: transit walk skim accessibility weighted by population opportunities.
AccessTWSC4: transit walk skim accessibility weighted by service plus commercial employment.
AvgAuto0Child: average number of cars owned by households without children.
Bth1Em4: transit best skim over highway Time1 accessibility weighted by total employment.
Bth1Pop4: transit best skim over highway Time1 accessibility weighted by population opportunities.
Bth1SC4: transit best skim over highway Time1 accessibility weighted by social plus commercial employment.
Bth2Em4: transit best skim over highway Time2 accessibility weighted by total employment.
Bth2Pop4: transit best skim over highway Time2 accessibility weighted by population opportunities.
Bth2SC4: transit best skim over highway Time2 accessibility weighted by social plus commercial employment.
TotalDensity4: total employment plus population density in a census tract.
White: percentage of population that are White in a tract.
Wth1Em4: transit walk skim over highway Time1 accessibility weighted by total employment.
Wth1Pop4: transit walk skim over highway Time1 accessibility weighted by population opportunities.
Wth1SC4: transit walk skim over highway Time1 accessibility weighted by service plus commercial employment.

WithEm4: transit walk skim over highway Time2 accessibility weighted by total employment.
WithSC4: transit walk skim over highway Time2 accessibility weighted by service plus commercial employment.

Table 6-15. Transit Attraction Trip Model for Palm Beach (Tract Method)*

Order	Model**	Adjusted R^2	MSE
1	$11.10076 \times JOBS_HH4$	0.2762	104.43712

Notes:

* The list does not include the models with the adjusted R^2 's that are less than 0.3.

** Variable description:

JOBS_HH4: job house balance.

6.4 Regressions with Demand Variables

The regression results described in the proceeding three subsections showed that only supply variables, such as transit accessibility, total number of bus runs, and percentage of area of a census tract covered by bus stop buffer zones, were significant and consistently included in the possible models that were identified. It is well recognized that transit supply and demand are interdependent, i.e., supply is determined by demand and demand is stimulated by supply. As supply variables are much easier to measure than demand variables, it is not surprising that supply variables are the leading indicators of transit use.

While supply variables may be employed to determine the mode share where transit services have been established, they do not help predict mode share where transit services are being considered in the future. For future year forecasts, planning of transit projects will require the determination of demand. As a result, additional regressions are performed to further explore the relationship between transit use and probable influential factors by excluding the dominant transit supply variables. The data set was first categorized into three groups of low, medium, and high accessibility, each with nearly equal number of observations, i.e., around 30. The purpose of data regrouping was to utilize transit accessibility as a control variable to examine the variation of demand variables in each of the three new data subsets. Accessibility was chosen as the control variable because it had the highest R^2 thus the best explanatory power. The following sections will describe the results of the regression analyses for each of the four data aggregation methods, namely the GIS Land Use Method, the Property Method, the Buffer Method, and the Tract Method.

6.4.1 Data Compiled Using GIS Land Use Method

Tables 6-16, 6-17, and 6-18 respectively list the regression models with adjusted R^2 's greater than 0.5 for the low, medium, and high accessibility groups. These models were produced by the procedure described in Section 6 with land use/socio-economic/demographic data compiled using GIS land use method, that is data compiled for transit stop buffer areas in a tract assuming population and employment to be evenly distributed in the same land uses. The *AccessTWPpop1* variable, i.e., transit walk skim accessibility weighted by population opportunities, was used to classify the data into three accessibility groups. The accessibility indices for the low, medium, and high accessibility groups ranged from 0.0441 to 0.2413, 0.2440 to 0.3990, and 0.4228 to 0.8784, respectively. The results in Tables 6-16 to 6-18 show that the *SFDUDensity1* variable, i.e., single-family dwelling unit (DU) density in the buffer area of a tract, appears to be a

significantly factor on transit use at all accessibility levels. Additionally, the effect of single-family dwelling unit density on transit use increases with the accessibility level as indicated by the increase in the magnitude of the coefficients of the variable in the models.

Table 6-16. Models for Miami-Dade (GIS Land Use Method, Low Accessibility)

Order	Model*	Adjusted R^2	MSE
1	$0.20772 \times SFDUDensityI$	0.5340	0.90874
2	$0.05316 \times TotalDensityI$	0.4847	1.00481

* Variable descriptions in alphabetic order:

SFDUDensityI: single-family dwelling unit density in the buffer area of a census tract.

TotalDensityI: total employment plus population density.

Table 6-17. Models for Miami-Dade (GIS Land Use Method, Medium Accessibility)

Order	Model*	Adjusted R^2	MSE
1	$2.36922 \times AvgEntropyI + 0.00014312 \times AvgParcelMFSzI$	0.7487	1.16151
2	$0.04498 \times DECAY_POP$	0.6119	1.79371
3	$0.00021529 \times AvgParcelSzI$	0.6063	1.81957
4	$0.00019464 \times SFPopDensityI$	0.5736	1.97075
5	$0.29410 \times SFDUDensityI$	0.5627	2.02112
6	$0.02187 \times MFPopDensityI$	0.5436	2.10957
7	$0.05884 \times TotalDensityI$	0.5078	2.27472

* Variable descriptions in alphabetic order:

AvgEntropyI: land use mix in the buffer area of a census tract.

AvgParcelMFSzI: average parcel size for multi-family dwelling units in the buffer area of a census tract.

AvgParcelSzI: average parcel size for residential properties in the buffer area of a census tract.

DECAY_POP: percentage of service population.

MFPopDensityI: multi-family population density in the buffer area of a census tract.

SFDUDensityI: single-family dwelling unit density in buffer area of a census tract.

SFPopDensityI: single family population density in the buffer area of a census tract.

TotalDensityI: total employment plus population density.

Table 6-18. Models for Miami-Dade (GIS Land Use Method, High Accessibility)

Order	Model	Adjusted R^2	MSE
1	$5.33338 \times AvgEntropyI$	0.8515	1.21974
2	$0.12239 \times SFPopDensityI$	0.8354	1.35194
3	$0.05830 \times DECAY_POP$	0.8108	1.55404
4	$0.42001 \times SFDUDensityI$	0.8098	1.56172
5	$0.00020751 \times AvgParcelSzI + 0.03109 \times TotalPopDensityI$	0.8403	1.31122
6	$0.00034931 \times AvgParcelSFSzI$	0.8011	1.63323
7	$0.04374 \times MFPopDensityI$	0.7975	1.66259
8	$0.11160 \times MFDUDensityI$	0.7352	2.17463
9	$0.00018959 \times AvgParcelMFSzI + 0.08546 \times TotalDUDenitI$	0.7794	1.81142
10	$0.00002452 \times AvgMFValue$	0.6812	2.61798
11	$5.19856 \times JOBS_HH^mI$	0.6725	2.68920
12	$0.00002703 \times AvgValue$	0.6617	2.77778
13	$0.00002871 \times AvgSFValue$	0.6586	2.80372
14	$2.35597 \times JOBS_HHI + 0.04014 \times TotalDensityI$	0.6930	2.52128

* Variable descriptions in alphabetic order:

AvgEntropyI: land use mix in the buffer area of a census tract.

AvgMFValue: average multi-family property value in the buffer area of a census tract.

AvgParcelMFSzI: average parcel size for multi-families in the buffer area of a census tract.

<i>AvgParcelSFSize1</i> :	average parcel size for single families in the buffer area of a census tract.
<i>AvgParcelSize1</i> :	average parcel size for residential properties in the buffer area of a census tract.
<i>AvgSFValue</i> :	average value for single families in the buffer area of a census tract.
<i>AvgValue</i> :	average value for residential properties in the buffer area of a census tract.
<i>DECAY_POP</i> :	percentage of service population.
<i>JOBS_HH1</i> :	job house balance.
<i>JOBS_HH^m1</i> :	modified jobs-housing balance in buffer area of a census tract.
<i>MFDUDensity1</i> :	multi-family dwelling unit density in buffer area of a tract.
<i>MFPopDensity1</i> :	multi-family population density in buffer area of a tract.
<i>SFDUDensity1</i> :	single-family dwelling unit density in buffer area of a census tract.
<i>SFPopDensity1</i> :	single-family population density in buffer area of a tract.
<i>TotalDensity1</i> :	total employment plus population density.
<i>TotalDUDensity1</i> :	total single- plus multi-family dwelling unit density in buffer area of a census tract.
<i>TotalPopDensity1</i> :	total single- plus multi-family population density in buffer area of a census tract.

Figure 6-2 illustrates the relationships between single-family DU density and transit productions by assuming zero model intercept as previously described. When the model intercept is fixed at zero, the effects of single-family DU density on transit use at different accessibility levels can be directly estimated by comparing the coefficients from the models. When non-zero intercept is specified, however, the variable appears to have different effects on transit productions, i.e., it is negatively related to transit production at the high accessibility level while positively related at the other levels (see Figure 6-3). Note that the variable itself has a positive relationship with transit productions for both no-intercept and intercept models when the data are not grouped by accessibility level as shown in Figure 6-4.

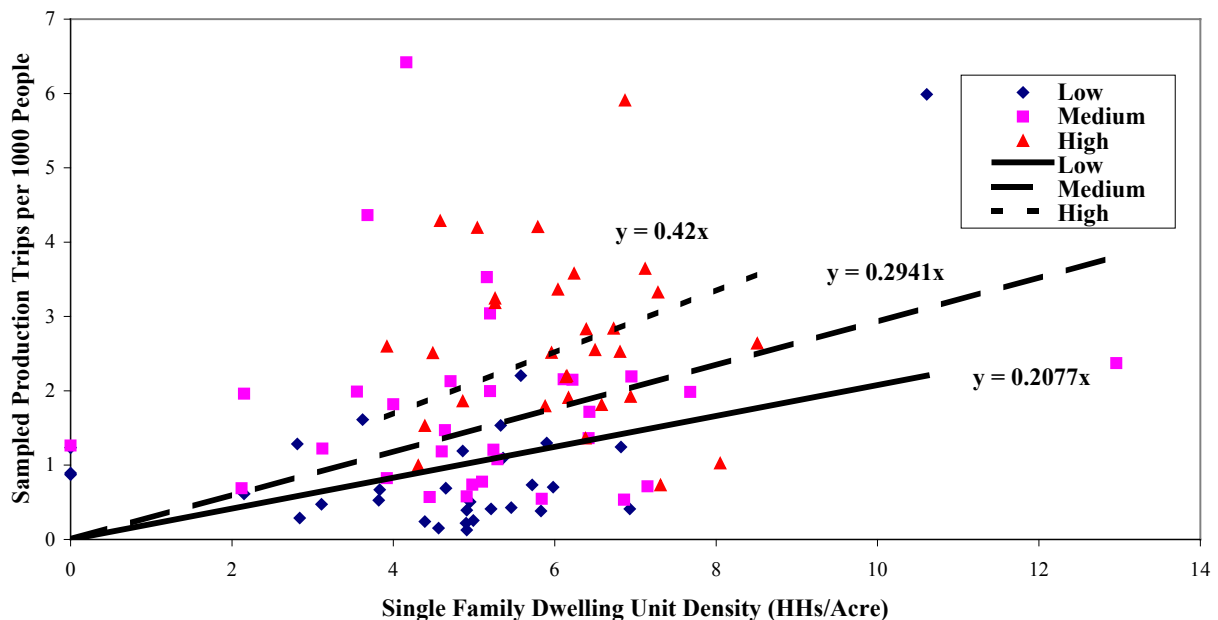


Figure 6-2. Through Origin Relationships between *SFDUDensity1* and Transit Production by Transit Accessibility Level (GIS Land Use Method)

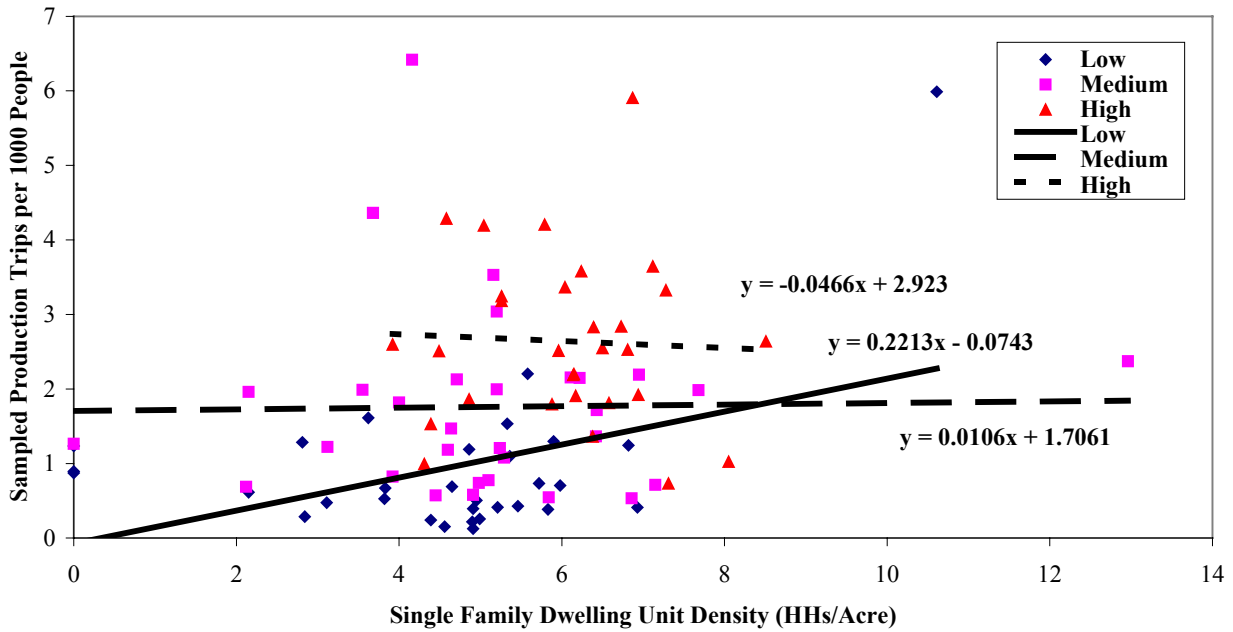


Figure 6-3. Non-Zero Intercept Relationships between *SFDUDensity1* and Transit Production by Transit Accessibility Level (GIS Land Use Method)

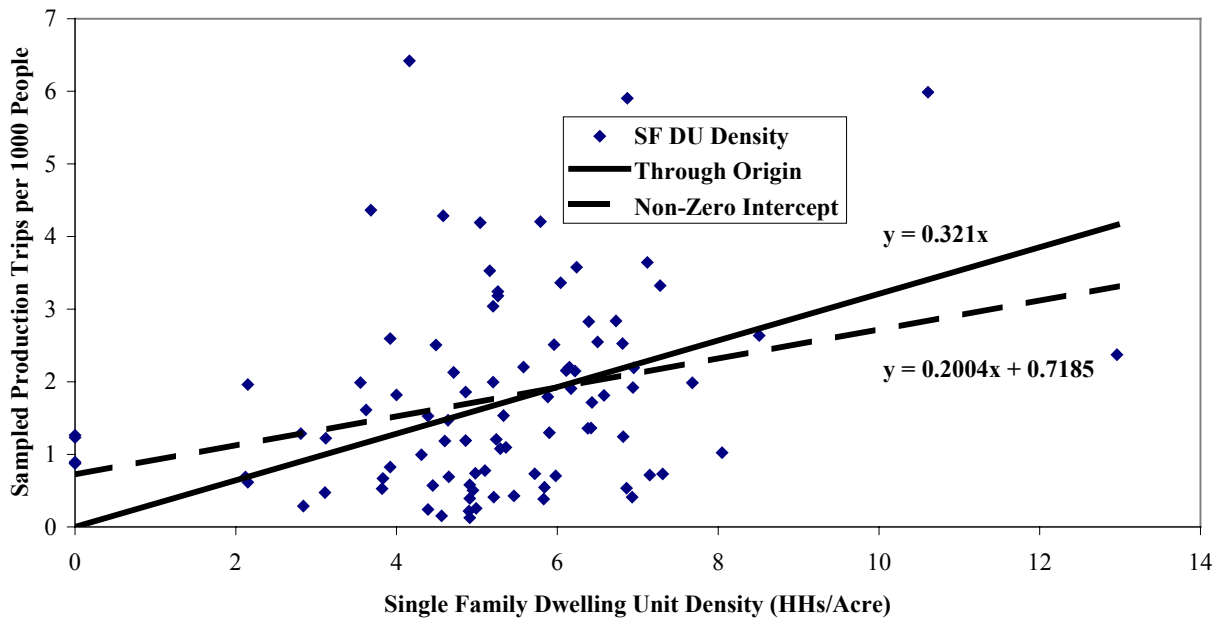


Figure 6-4. Intercept and No-Intercept *SFDUDensity1* Models for Transit Productions (GIS Land Use Method)

Following single-family dwelling unit density, the next variable that seemed to have an effect on transit use is *TotalDensity1*. The *TotalDensity1* variable, i.e., employment plus population density in the buffer area of a tract, has a R^2 of 0.4847, which is close to the 0.5 threshold value. Figures 6-5 and 6-6 illustrate the zero-intercept and intercept models at different accessibility levels, respectively. Both figures show that *TotalDensity1* is positively related to transit use

regardless of the accessibility level. Similar positive relationship can be observed when the data are not grouped according to the accessibility levels (see Figure 6-7).

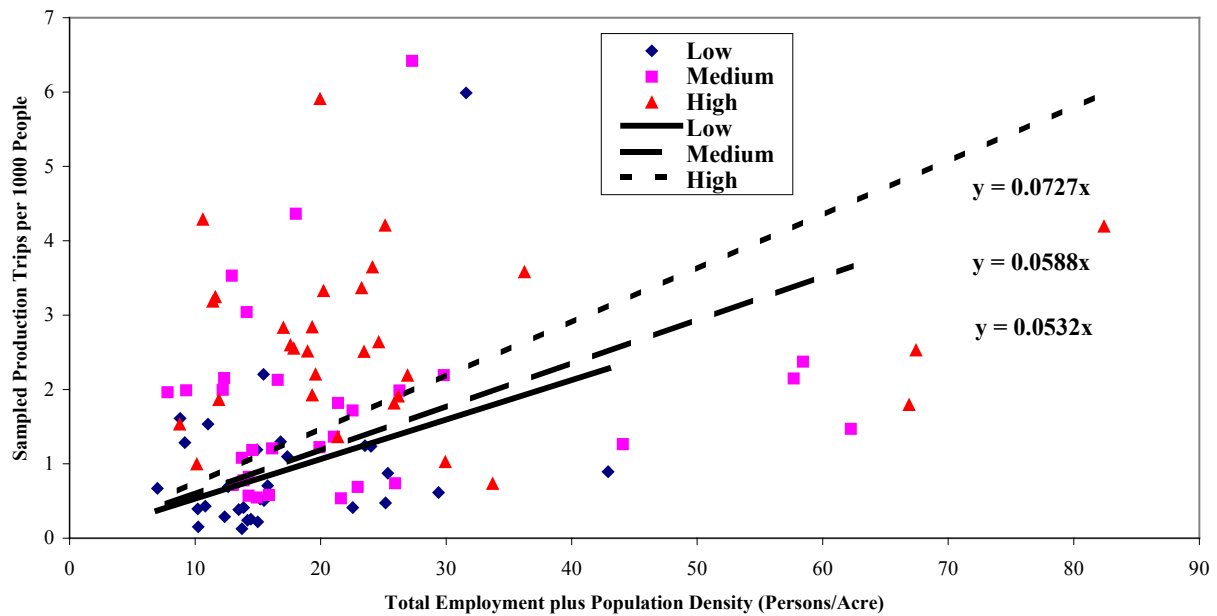


Figure 6-5. Through Origin Relationships between *TotalDensity1* and Transit Production by Transit Accessibility Level (GIS Land Use Method)

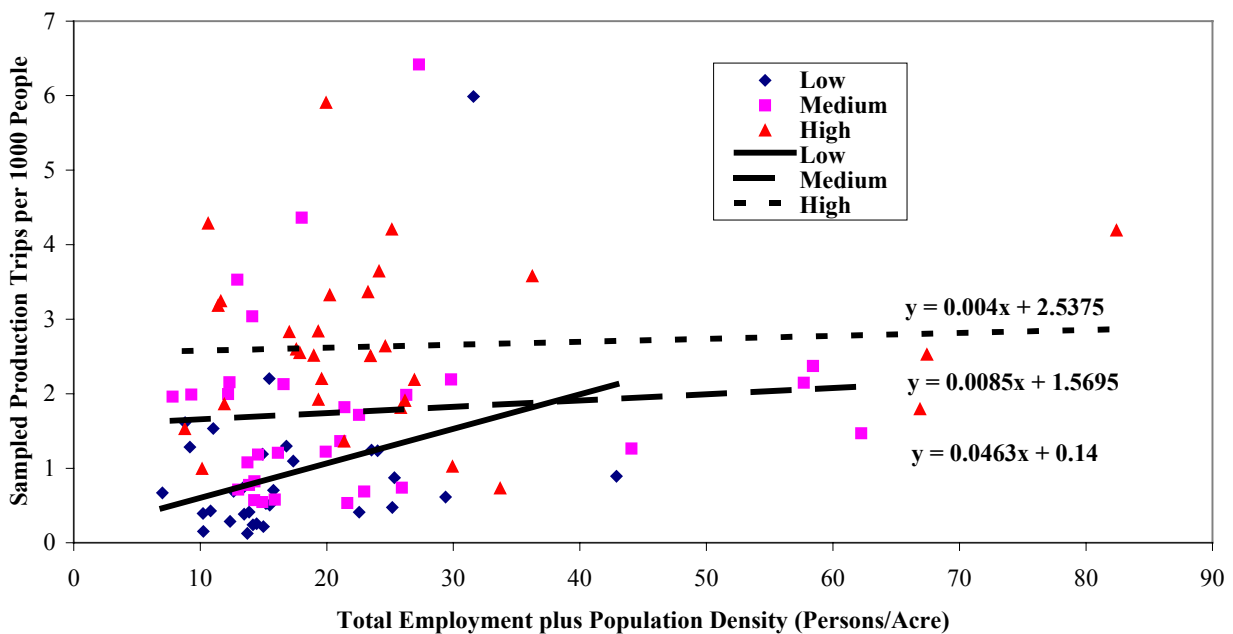


Figure 6-6. Non-Zero Intercept Relationships between *TotalDensity1* and Transit Production by Transit Accessibility Level (GIS Land Use Method)

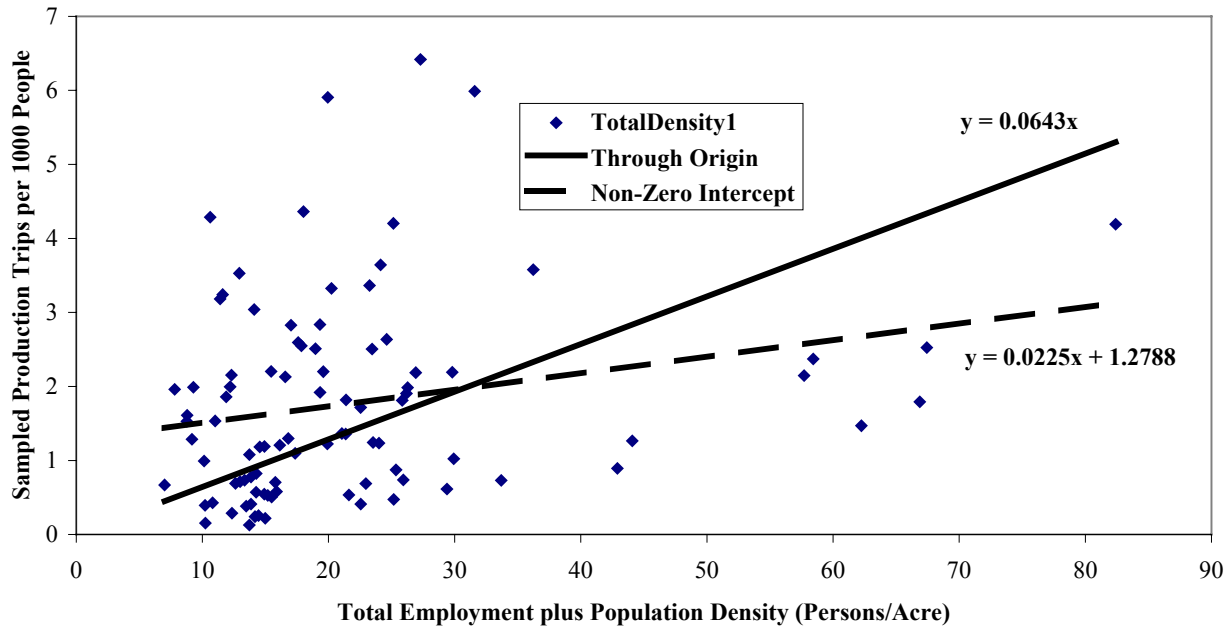


Figure 6-7. Intercept and No-Intercept *TotalDensity1* Models for Transit Productions (GIS Land Use Method)

6.4.2 Data Compiled Using Property Method

Tables 6-19, 6-20, and 6-21 respectively list the regression models with adjusted R^2 's that are greater than 0.5 for the low, medium, and high accessibility groups. The data were compiled by the Property Method procedure described in Section 6. The difference between Methods 1 and 2 is that population is distributed based on property information while employment is assumed to be evenly distributed in the same land uses. The *AccessTWPpop2* variable, i.e., the transit walk skim accessibility weighted by population, was used to classify the data into three accessibility groups. The accessibility indices for the low, medium, and high accessibility groups range from 0.0425 to 0.2406, 0.2413 to 0.3931, and 0.4229 to 0.8870, respectively.

Table 6-19. Models for Miami-Dade (Property Method, Low Accessibility)

Order	Model	Adjusted R^2	MSE
1	$2.80975 \times AvgEntropy2$	0.6491	0.63998
2	$0.05667 \times TotalDensity2$	0.5712	0.78193
3	$0.02796 \times MFDUDensity2 + 0.12393 \times SFDUDensity2$	0.6218	0.68968
4	$0.00007318 \times AvgParcelSize2 + 0.02999 \times TotalDUDensity2$	0.5770	0.77138

* Variable descriptions in alphabetic order:
AvgEntropy2: land use mix in the buffer area of a census tract.
AvgParcelSize2: average parcel size for residential properties in the buffer area of a census tract.
MFDUDensity2: multi-family dwelling unit density in buffer area of a tract.
SFDUDensity2: single-family dwelling unit density in buffer area of a census tract.
TotalDensity2: total employment plus population density.
TotalDUDensity2: total single- plus multi-family dwelling unit density in buffer area of a census tract.

Table 6-20. Models for Miami-Dade (Property Method, Medium Accessibility)

Order	Model	Adjusted R^2	MSE
1	$4.96075 \times AvgEntropy2$	0.6920	1.99476
2	$0.00021369 \times AvgParcelMFSIZE2 + 0.03501 \times TotalDensity2$	0.7189	1.82029
3	$0.34974 \times SFDUDensity2$	0.6368	2.35228
4	$0.05269 \times DECAY_POP$	0.6299	2.39670
5	$0.00026489 \times AvgParcelSize2$	0.6172	2.47889
6	$0.00023809 \times AvgParcelSFSize2$	0.5808	2.71469

* Variable descriptions in alphabetic order:
AvgEntropy2: land use mix in the buffer area of a census tract.
AvgParcelMFSIZE2: average parcel size for multi-families in the buffer area of a census tract.
AvgParcelSFSize2: average parcel size for single-families in the buffer area of a census tract.
AvgParcelSize2: average parcel size for residential properties in the buffer area of a census tract.
DECAY_POP: percentage of service population.
SFDUDensity2: single-family dwelling unit density in buffer area of a census tract.
TotalDensity2: total employment plus population density.

Table 6-21. Models for Miami-Dade (Property Method, High Accessibility)

Order	Model	Adjusted R^2	MSE
1	$0.00039897 \times AvgParcelSFSize2$	0.7924	2.24814
2	$0.00039490 \times AvgParcelSize2$	0.7911	2.26113
3	$5.88032 \times AvgEntropy2$	0.7824	2.35575
4	$0.47025 \times SFDUDensity2$	0.7683	2.50849
5	$0.06428 \times DECAY_POP$	0.7448	2.76254
6	$0.00034835 \times AvgParcelMFSIZE2$	0.7137	3.09966
7	$0.12263 \times MFDUDensity2$	0.6705	3.56763
8	$0.00002778 \times AvgMFValue$	0.6621	3.65811
9	$0.00003274 \times AvgSFValue$	0.6492	3.79762
10	$0.00003073 \times AvgValue$	0.6480	3.81070
11	$5.75953 \times JOBS_HH^m2$	0.6238	4.07327
12	$0.12776 \times TotalDensity2 + 0.23175 \times SFPopDensity2$	0.7051	3.19270

* Variable descriptions in alphabetic order:
AvgEntropy2: land use mix in the buffer area of a census tract.
AvgMFValue: average multi-family property value in the buffer area of a census tract.
AvgParcelMFSIZE2: average parcel size for multi-families in the buffer area of a census tract.
AvgParcelSFSize2: average parcel size for single families in the buffer area of a census tract.
AvgParcelSize2: average parcel size for residential properties in the buffer area of a census tract.
AvgSFValue: average value for single families in the buffer area of a census tract.
AvgValue: average value for residential properties in the buffer area of a census tract.
DECAY_POP: percentage of service population.
JOBS_HH^m2: modified jobs-housing balance in buffer area of a census tract.
MFDUDensity2: multi-family dwelling unit density in buffer area of a tract.
SFDUDensity2: single-family dwelling unit density in buffer area of a census tract.
SFPopDensity2: single-family population density in buffer area of a tract.
TotalDensity2: total employment plus population density.

The results from Tables 6-19 to 6-21 identify the following five variables as possible influential factors on transit use at various accessibility levels:

<i>AvgEntropy</i>	-	Average entropy (a land use mix measure)
<i>TotalDensity</i>	-	Total employment plus population density
<i>SFDUDensity</i>	-	Single-family dwelling units density

AvgParcelSize - Average single-family plus multi-family parcel size
AvgParcelSFSize - Average single-family parcel size

The *AvgEntropy* and *TotalDensity* variables were further examined to verify their effects on transit productions at different accessibility levels. Figure 6-8 illustrates the no-intercept models for transit productions when *AvgEntropy* was the only explanatory variable included in the model. As expected, the effect of *AvgEntropy* on transit use increases with transit accessibility. As illustrated in Figure 6-9, the similar ascending pattern was also observed at all accessibility levels when non-zero intercept was specified. For ungrouped data, the no-intercept and intercept models again revealed similar ascending pattern between *AvgEntropy* and transit production (see Figure 6-10).

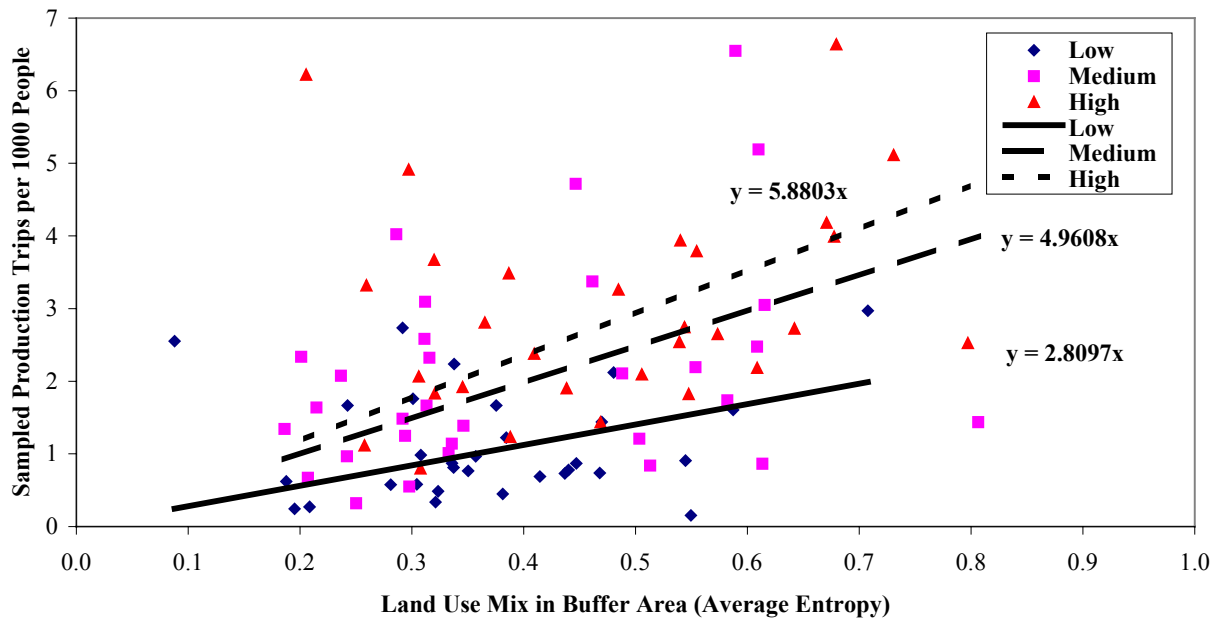


Figure 6-8. Through Origin Relationships between *AvgEntropy* and Transit Production by Transit Accessibility Level (Property Method)

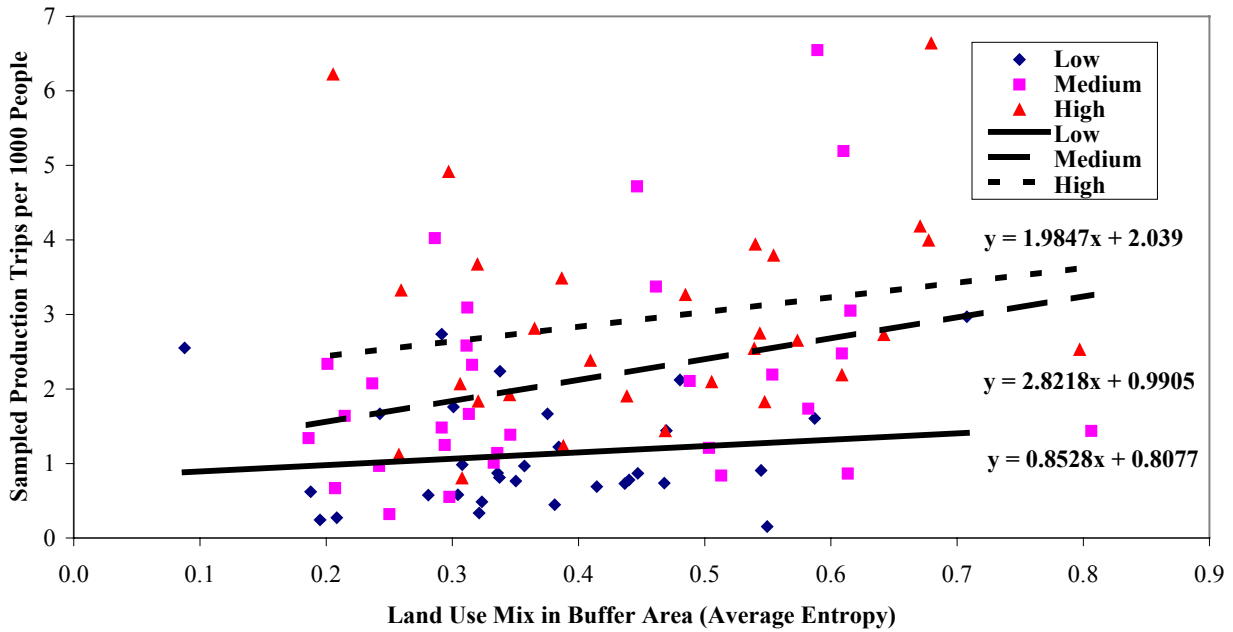


Figure 6-9. Non-Zero Intercept Relationships between *AvgEntropy* and Transit Production by Transit Accessibility Level (Property Method)

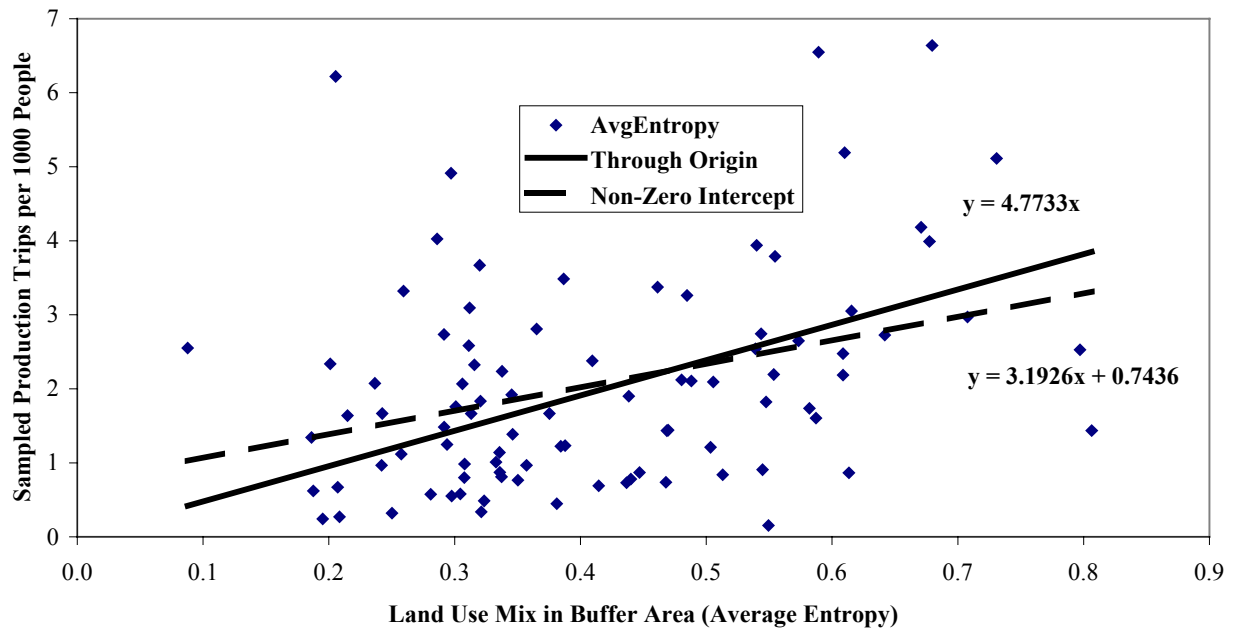


Figure 6-10. Intercept and No-Intercept *AvgEntropy* Models for Transit Production (Property Method)

Figure 6-11 illustrates the no-intercept models for transit productions when *TotalDensity* was the only explanatory variable included in the model. As expected, the effect of *TotalDensity* on transit use again increases with transit accessibility. As illustrated in Figure 6-12, the similar ascending pattern was also observed at all accessibility levels when non-zero intercept was specified. The *TotalDensity* variable, however, does not appear to have significant effects on

transit productions at the low and high accessibility levels when non-zero intercept is specified since the model trend lines nearly parallel the horizontal axis. For ungrouped data, the no-intercept and intercept models again revealed similar ascending pattern between *TotalDensity* and transit production (see Figure 6-13).

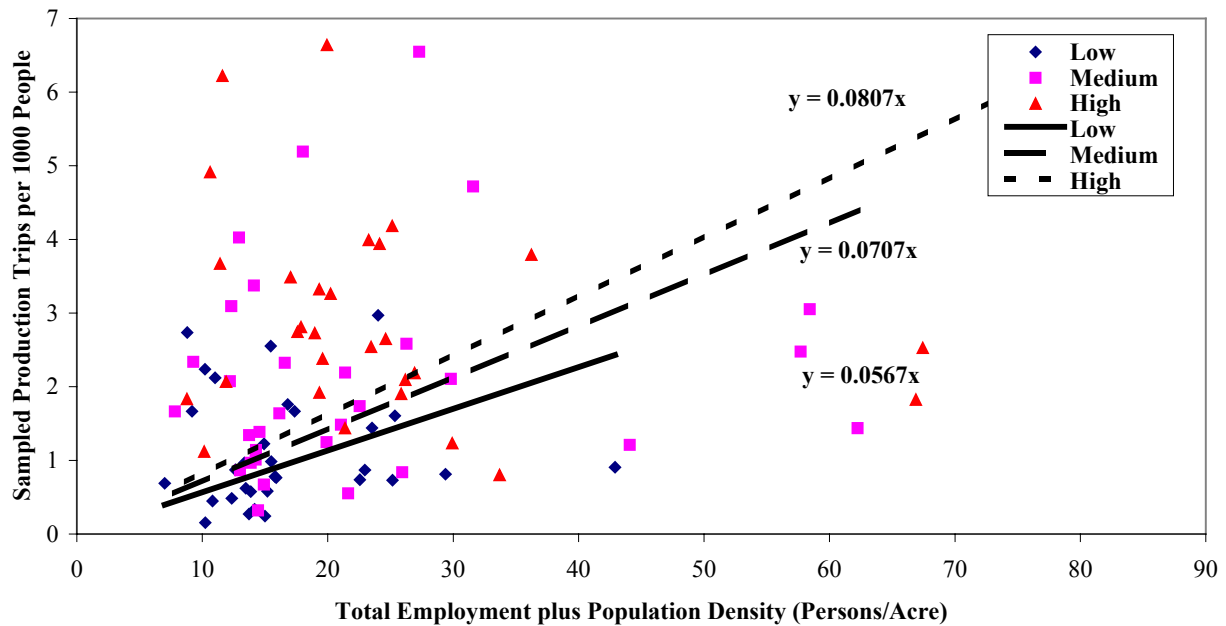


Figure 6-11. Through Origin Relationships between *TotalDensity* and Transit Production by Transit Accessibility Level (Property Method)

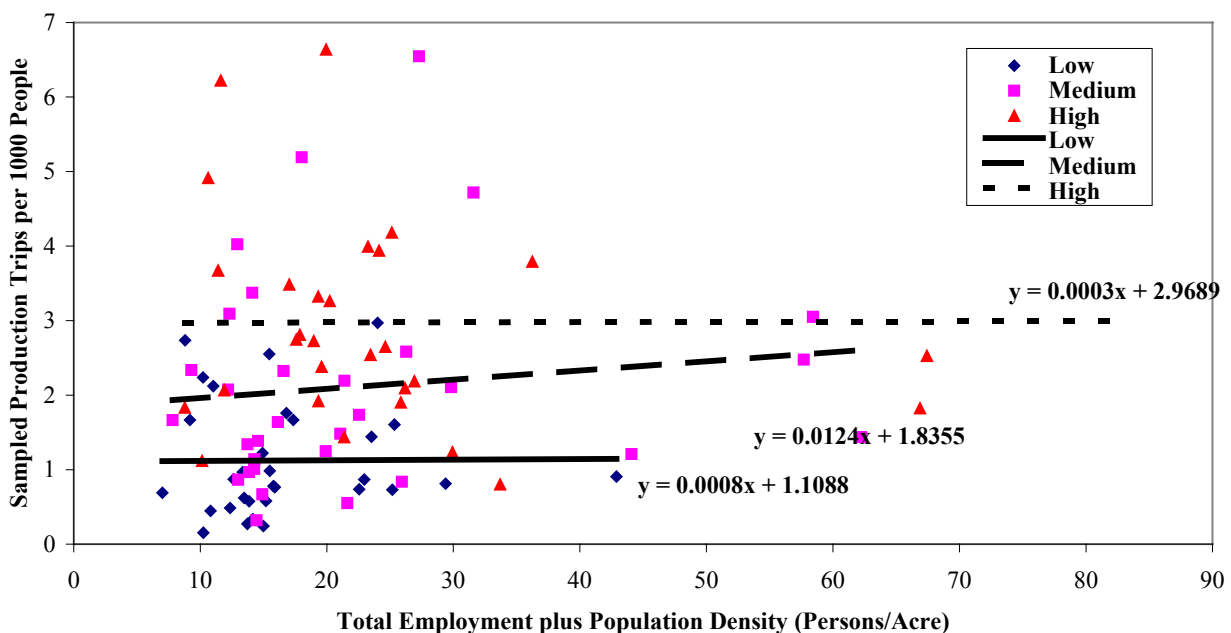


Figure 6-12. Non-Zero Intercept Relationships between *TotalDensity* and Transit Production by Transit Accessibility Level (Property Method)

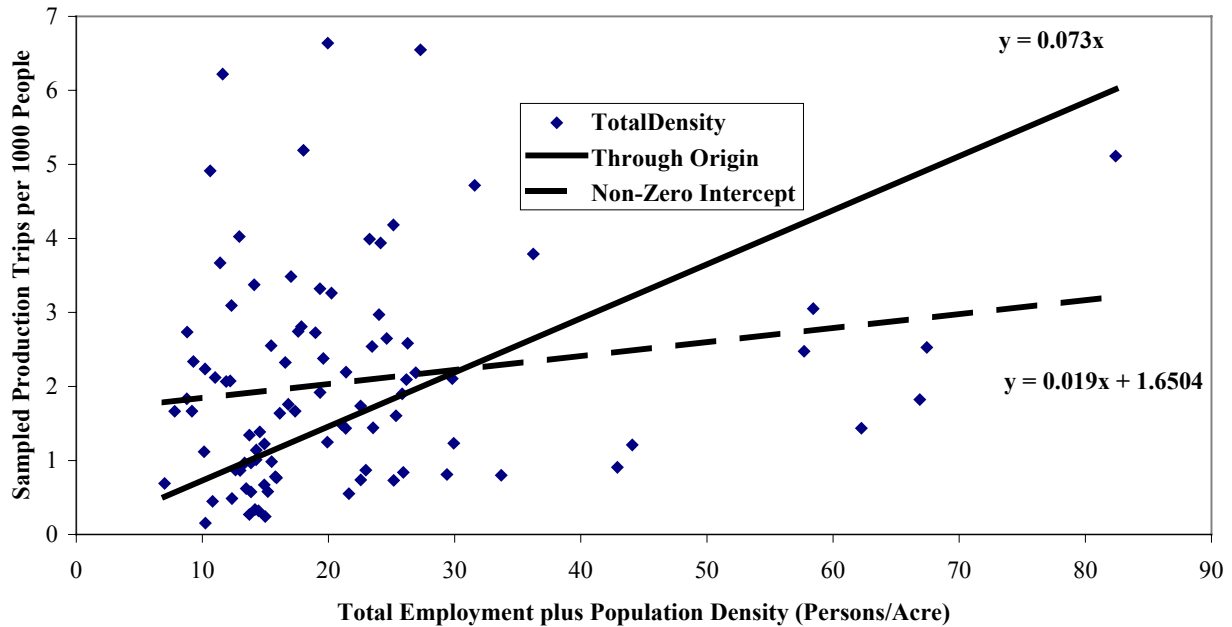


Figure 6-13. Intercept and No-Intercept *TotalDensity* Models for Transit Production (Property Method)

6.4.3 Data Compiled Using Buffer Method

Tables 6-22, 6-23, and 6-24 present regression models obtained for Miami-Dade County for the low, medium, and high transit accessibility groups, respectively. The data for land use variables were compiled using the Buffer Method by assuming even distribution of population and employment and were calculated for buffer zones in a tract. The *AccessTWPop3* variable (transit walk skim accessibility weighted by population opportunities) was used to classify the data into the desired three accessibility groups. The accessibility index values for the low, medium, and high accessibility groups range from 0.0442 to 0.2441, 0.2448 to 0.3989, and 0.4229 to 0.8870, respectively.

Table 6-22. Models for Miami-Dade (Buffer Method, Low Accessibility)

Order	Model	Adjusted R^2	MSE
1	$1.72017 \times JOBS_HH^m3$	0.4855	1.02291
2	$2.56300 \times AvgEntropy3$	0.4702	1.05329
3	$1.39009 \times JOBS_HH3$	0.4553	1.08290

* Variable descriptions in alphabetic order:
AvgEntropy3: land use mix in the buffer area of a census tract.
JOBS_HH3: job house balance.
JOBS_HH^m3: modified jobs-housing balance in buffer area of a census tract.

Table 6-23. Models for Miami-Dade (Buffer Method, Medium Accessibility)

Order	Model	Adjusted R^2	MSE
1	$5.05978 \times AvgEntropy3$	0.6589	2.46326
2	$0.05475 \times DECAY_POP$	0.5542	3.21934
3	$0.00022794 \times AvgParcelSFSize3$	0.5297	3.39657
4	$0.00024459 \times AvgParcelSize3$	0.5255	3.42695
5	$0.00029550 \times AvgParcelMFSIZE3$	0.5104	3.53548

* Variable descriptions in alphabetic order:
AvgEntropy3: land use mix in the buffer area of a census tract.
AvgParcelMFSIZE3: average parcel size for multi-families in the buffer area of a census tract.
AvgParcelSFSize3: average parcel size for single families in the buffer area of a census tract.
AvgParcelSize3: average parcel size for residential properties in the buffer area of a census tract.
DECAY_POP: percentage of service population.

Table 6-24. Models for Miami-Dade (Buffer Method, High Accessibility)

Order	Model	Adjusted R^2	MSE
1	$5.44422 \times AvgEntropy3$	0.8566	1.21435
2	$0.05941 \times DECAY_POP$	0.8212	1.51486
3	$0.00035157 \times AvgParcelSize3$	0.8060	1.64301
4	$0.00035490 \times AvgParcelSFSize3$	0.8052	1.65034
5	$0.00030793 \times AvgParcelMFSIZE3$	0.7437	2.17109
6	$3.17832 \times JOBS_HH^m3 + 0.08397 \times TotalPopDensity3$	0.7825	1.84221
7	$0.00002509 \times AvgMFValue$	0.6800	2.71033
8	$0.00002761 \times AvgValue$	0.6557	2.91670
9	$0.00002933 \times AvgSFValue$	0.6533	2.93662
10	$2.48323 \times JOBS_HH3 + 0.03973 \times TotalDensity3$	0.7011	2.53159
11	$0.34436 \times SFPopDensity3$	0.5599	3.72781

* Variable descriptions in alphabetic order:
AvgEntropy3: land use mix in the buffer area of a census tract.
AvgMFValue: average multi-family property value in the buffer area of a census tract.
AvgSFValue: average value for single families in the buffer area of a census tract.
AvgValue: average value for residential properties in the buffer area of a census tract.
AvgParcelMFSIZE3: average parcel size for multi-families in the buffer area of a census tract.
AvgParcelSFSize3: average parcel size for single families in the buffer area of a census tract.
AvgParcelSize3: average parcel size for residential properties in the buffer area of a census tract.
DECAY_POP: percentage of service population.
JOBS_HH3: job house balance.
JOBS_HH^m3: modified jobs-housing balance in buffer area of a census tract.
SFPopDensity3: single-family population density in buffer area of a tract.
TotalDensity3: total employment plus population density.
TotalPopDensity3: total single- plus multi-family population density in buffer area of a census tract.

The results in Tables 6-22 to 6-24 show that only *AvgEntropy3* (land use mix in buffer area of a census tract) appears to be a consistently significant factor on transit use at various accessibility levels. Figure 6-14 illustrates the relationship between *AvgEntropy3* and transit productions by assuming zero model intercept as previously described. The resulted models show that the effects of *AvgEntropy3* on transit use at different accessibility levels increases with accessibility. In addition, a similar pattern can be observed when non-zero intercept is specified since the model coefficients at the medium and high accessibility levels are nearly equal (see Figure 6-15). Note that the variable itself also has a positive relationship with transit productions for both no-intercept and intercept models (see Figure 6-16) when the data are not grouped by accessibility level.

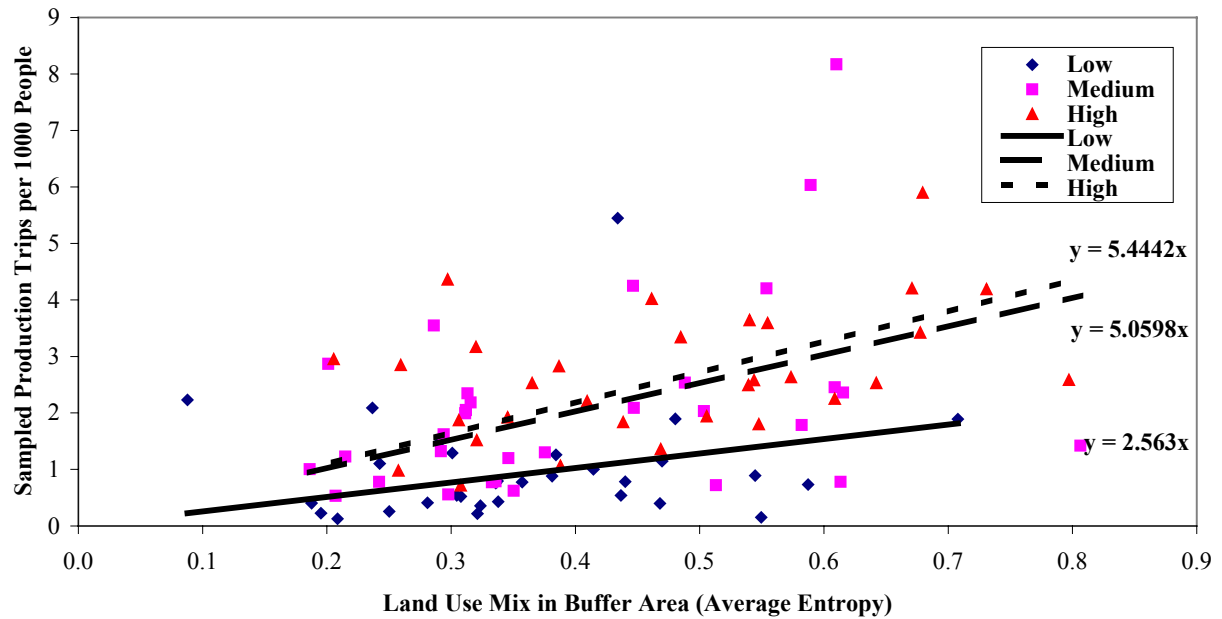


Figure 6-14. Through Origin Relationships between *AvgEntropy3* and Transit Production by Transit Accessibility Level (Buffer Method)

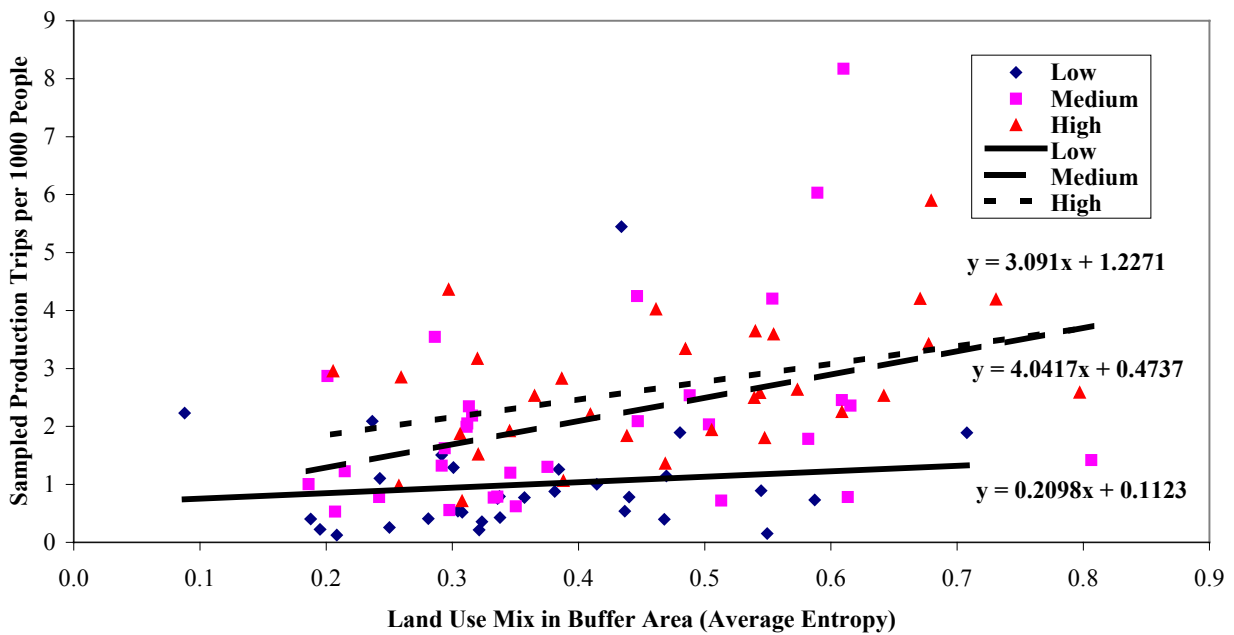


Figure 6-15. Non-Zero Intercept Relationships between *AvgEntropy3* and Transit Production by Transit Accessibility Level (Buffer Method)

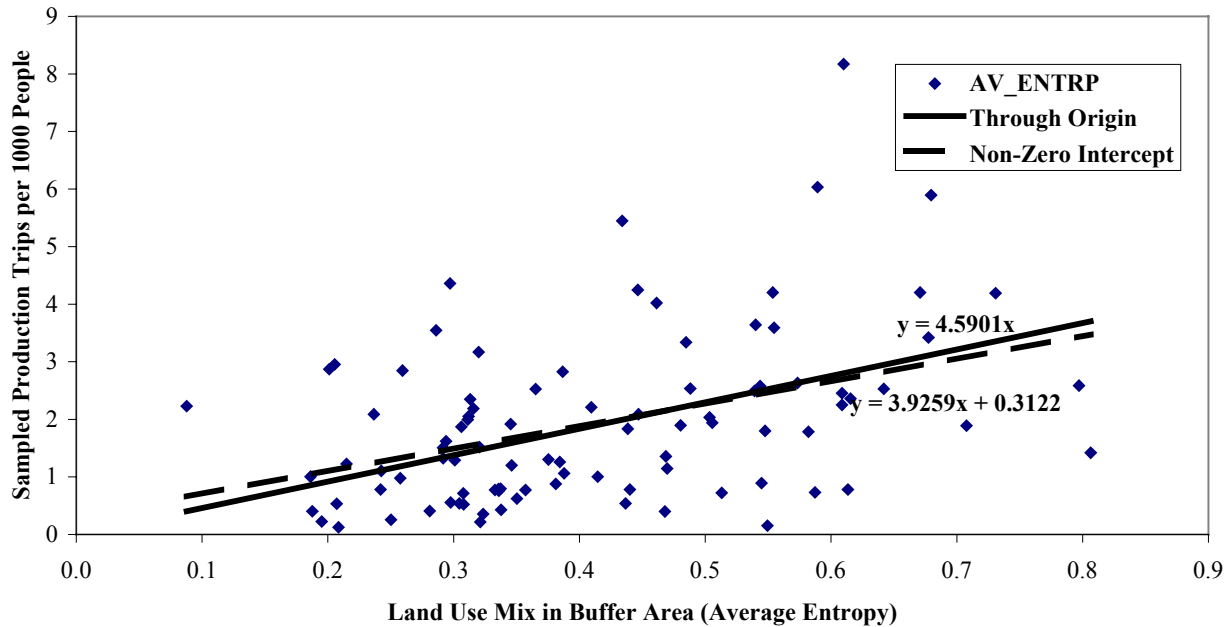


Figure 6-16. Intercept and No-Intercept *AvgEntropy3* Models for Transit Productions (Buffer Method)

6.4.4 Data Compiled Using Tract Method

Tables 6-25, 6-26, and 6-27 present regression models for Miami-Dade County for the low, medium, and high transit accessibility groups, respectively. The data for land use variables were compiled using Level 4 data by assuming population and employment as evenly distributed in a tract. The *AccessTWPpop4* variable, i.e., transit walk skim accessibility weighted by population opportunities, was used to classify the data into the low, medium, and high accessibility groups, ranging from 0.0425 to 0.2374, 0.2406 to 0.3905, and 0.3931 to 0.8870, respectively.

Table 6-25. Models for Miami-Dade (Tract Method, Low Accessibility)

Order	Model	Adjusted R^2	MSE
1	$0.25965 \times \text{AvgAutoChild}$	0.5836	0.38297
2	$0.38603 \times \text{AvgWrkrChild}$	0.5626	0.40232
3	$1.12190 \times \%HH0Child$	0.5610	0.40376
4	$0.17138 \times \text{AvgPChild}$	0.5382	0.42476
5	$0.34812 \times \text{AvgAuto0Child}$	0.5347	0.42798
6	$0.31503 \times \text{AvgP0Child}$	0.5339	0.42876
7	$0.56067 \times \text{AvgWrkr0Child}$	0.5142	0.44686
8	$1.01909 \times \text{JOBS_HH4}$	0.5060	0.45437

* Variable descriptions in alphabetic order:
%HH0Child: percentage of households without children.
AvgAuto0Child: average number of autos in households without children.
AvgAutoChild: average number of autos in households with children.
AvgP0Child: average number of persons in households without children.
AvgPChild: average number of persons in households with children.
AvgWrkr0Child: average number of workers in households without children.
AvgWrkrChild: average number of workers in households with children.
JOBS_HH4: job house balance.

Table 6-26. Models for Miami-Dade (Tract Method, Medium Accessibility)

Order	Model	Adjusted R^2	MSE
1	$0.36487 \times \text{AvgPChild}$	0.7437	0.76152
2	$1.20737 \times \%HH0Child + 0.00011800 \times \text{AvgParcelMFSIZE4}$	0.8085	0.56897
3	$3.58684 \times \text{AvgEntropy4}$	0.7370	0.78149
4	$0.59011 \times \text{AvgWrkrChild} + 2.73390 \times \text{MFDUNoAuto}\%$	0.7602	0.71249
5	$0.47044 \times \text{AvgP0Child} + 2.85744 \times \text{DUNoAuto}\%$	0.7538	0.73160
6	$0.79378 \times \text{AvgWrkr0Child} + 0.03056 \times \text{TotalDensity4}$	0.7251	0.81673
7	$0.00010802 \times \text{AvgParcelSFSize4} + 3.54081 \times \text{SFDUNoAuto}\%$	0.6961	0.90288
8	$0.00013584 \times \text{AvgParcelSize4} + 0.03832 \times \text{MFPopDensity4}$	0.7017	0.88641
9	$3.00923 \times \text{JOBS_HH}^m4$	0.6090	1.16181
10	$2.34044 \times \text{JOBS_HH4}$	0.6082	1.16425
11	$0.00000823 \times \text{AvgMFValue} + 0.01719 \times \text{Black}$	0.7440	0.76054
12	$2.62393 \times \text{ForeignBorn}$	0.5696	1.27895
13	$0.07133 \times \text{TotalPopDensity4}$	0.5678	1.28403

*

Variable descriptions in alphabetic order:
%HH0Child: percentage of households without children.
AvgEntropy4: land use mix in a census tract.
AvgMFValue: average multi-family property value in a census tract.
AvgP0Child: average number of persons in households without children.
AvgParcelMFSIZE4: average parcel size for multi-families in a census tract.
AvgParcelSFSize4: average parcel size for single families in a census tract.
AvgParcelSize4: average parcel size for residential properties in a census tract.
AvgPChild: average number of persons in households with children.
AvgWrkr0Child: average number of workers in households without children.
AvgWrkrChild: average number of workers in households with children.
Black: percentage of black population in a census tract.
DUNoAuto%: percentage of households without auto.
ForeignBorn: percentage of population that are immigrants in a tract.
JOBS_HH4: job house balance.
JOBS_HH^m4: modified jobs-housing balance in a census tract.
MFDUNoAuto%: percentage of multi-family households without auto.
MFPopDensity4: multi-family population density in a census tract.
SFDUNoAuto%: percentage of single-family households without auto.
TotalDensity4: total employment plus population density.
TotalPopDensity4: total single- plus multi-family population density in a census tract.

Table 6-27. Models for Miami-Dade (Tract Method, High Accessibility)

Order	Model	Adjusted R^2	MSE
1	$0.63133 \times AvgPChild$	0.8580	1.17172
2	$5.30550 \times AvgEntropy4$	0.8527	1.21531
3	$1.21018 \times AvgP0Child$	0.8342	1.36739
4	$1.54220 \times AvgWrkr0Child + 4.37589 \times MFDUNoAuto\%$	0.8535	1.20873
5	$0.00021483 \times AvgParcelSize4 + 4.4047 \times DUNoAuto\%$	0.8540	1.20442
6	$0.00022813 \times AvgParcelSFSize4 + 4.37339 \times SFDUNoAuto\%$	0.8516	1.22407
7	$2.95817 \times \%HH0Child + 0.01952 \times Black$	0.8330	1.37725
8	$1.52631 \times AvgWrkrChild$	0.7488	2.07212
9	$2.73285 \times ForeignBorn + 2.69109 \times JOBS_HH^m4$	0.7844	1.77868
10	$1.08698 \times AvgAutoChild + 0.03456 \times TotalDensity4$	0.7624	1.95961
11	$0.00029918 \times AvgParcelMFSize4$	0.7017	2.46092
12	$0.00002478 \times AvgMFValue$	0.6878	2.57562
13	$0.45504 \times TotalDUDensity4$	0.6769	2.66501
14	$0.00002538 \times AvgValue$	0.6533	2.85978
15	$0.00002612 \times AvgSFValue$	0.6386	2.98150
16	$4.00601 \times JOBS_HH4$	0.6271	3.07628
17	$0.22862 \times SFPopDensity4 + 0.12419 \times MFPopDensity4$	0.7144	2.35625
18	$1.14223 \times SFDUDensity4$	0.5637	3.59934

* Variable descriptions in alphabetic order:
%HH0Child: percentage of households without children.
AvgEntropy4: land use mix in a census tract.
AvgMFValue: average multi-family property value in a census tract.
AvgP0Child: average number of persons in households without children.
AvgParcelMFSize4: average parcel size for multi-families in a census tract.
AvgParcelSFSize4: average parcel size for single-families in a census tract.
AvgParcelSize4: average parcel size for residential properties in a census tract.
AvgPChild: average number of persons in households with children.
AvgWrkr0Child: average number of workers in households without children.
AvgWrkrChild: average number of workers in households with children.
Black: percentage of black population in a census tract.
DUNoAuto%: percentage of households without auto.
ForeignBorn: percentage of population that are immigrants in a tract.
JOBS_HH4: job house balance.
JOBS_HH^m4: modified jobs-housing balance in a census tract.
MFDUNoAuto%: percentage of multi-family households without auto.
MFPopDensity4: multi-family population density in a census tract.
SFDUNoAuto%: density for single-family households without auto in a census tract.
TotalDenity4: total residential property density in a census tract.
TotalPopDensity4: total single- plus multi-family population density in a census tract.

The results in Tables 6-25 to 6-27 show that *%HH0Child*, *AvgWrkrChild*, *AvgWrkr0Child*, *AvgPChild*, *AvgP0Child*, and *JOBS_HH4* appear to be significantly factors on transit use at various accessibility levels. Figure 6-17 illustrates the zero-intercept relationships between percentage of households without children (*%HH0Child*) and transit productions. These models show that the effects of *%HH0Child* on transit use at different accessibility levels increases with accessibility. However, the same pattern holds only at the low and medium accessibility levels for intercept models (see figure 6-18). In addition, no-intercept and intercept models show different relationships between *%HH0Child* and transit productions for unclassified data (see Figure 6-19).

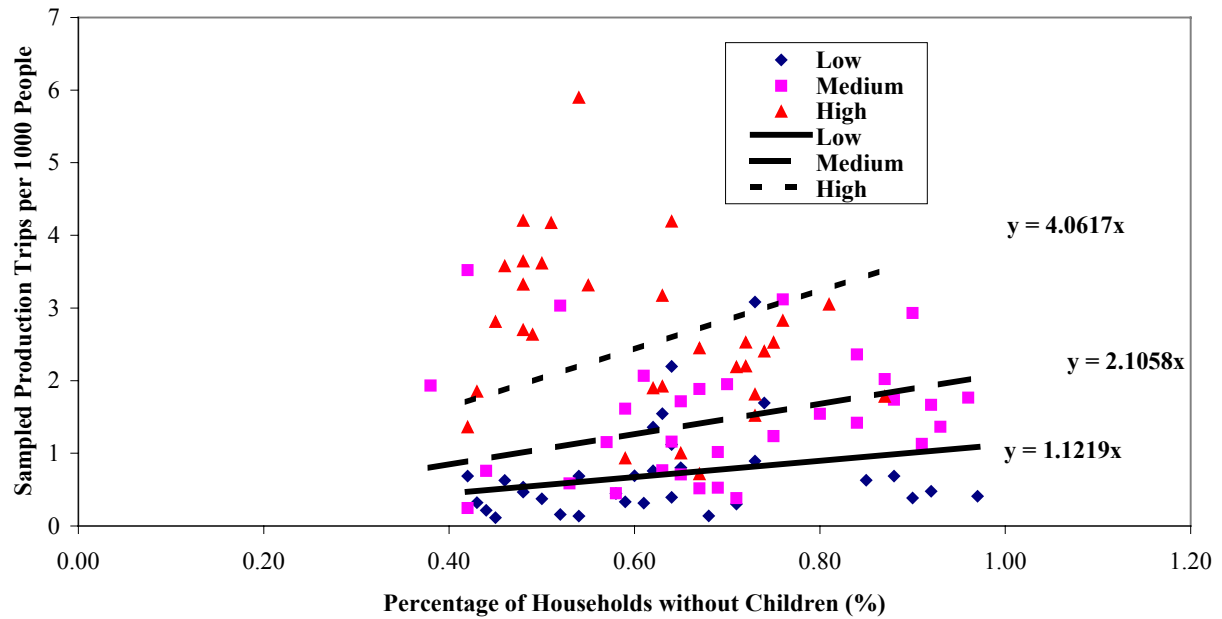


Figure 6-17. Through Origin Relationships between %HH0Child and Transit Production by Transit Accessibility Level (Tract Method)

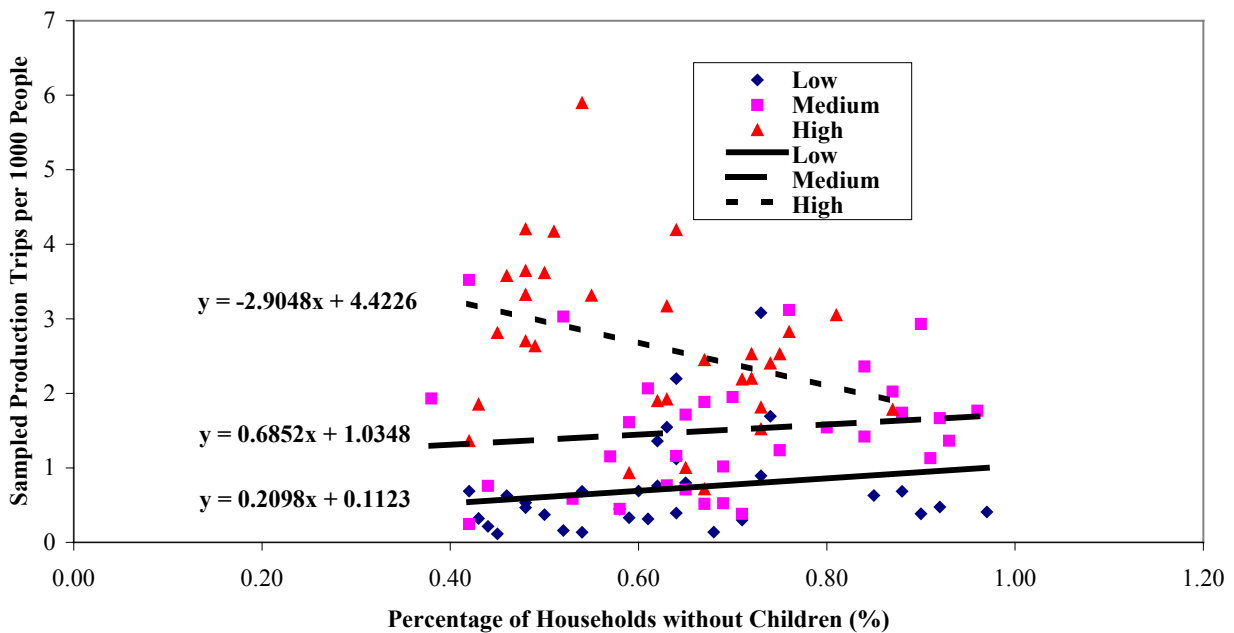


Figure 6-18. Non-Zero Relationships between %HH0Child and Transit Production by Transit Accessibility Level (Tract Method)

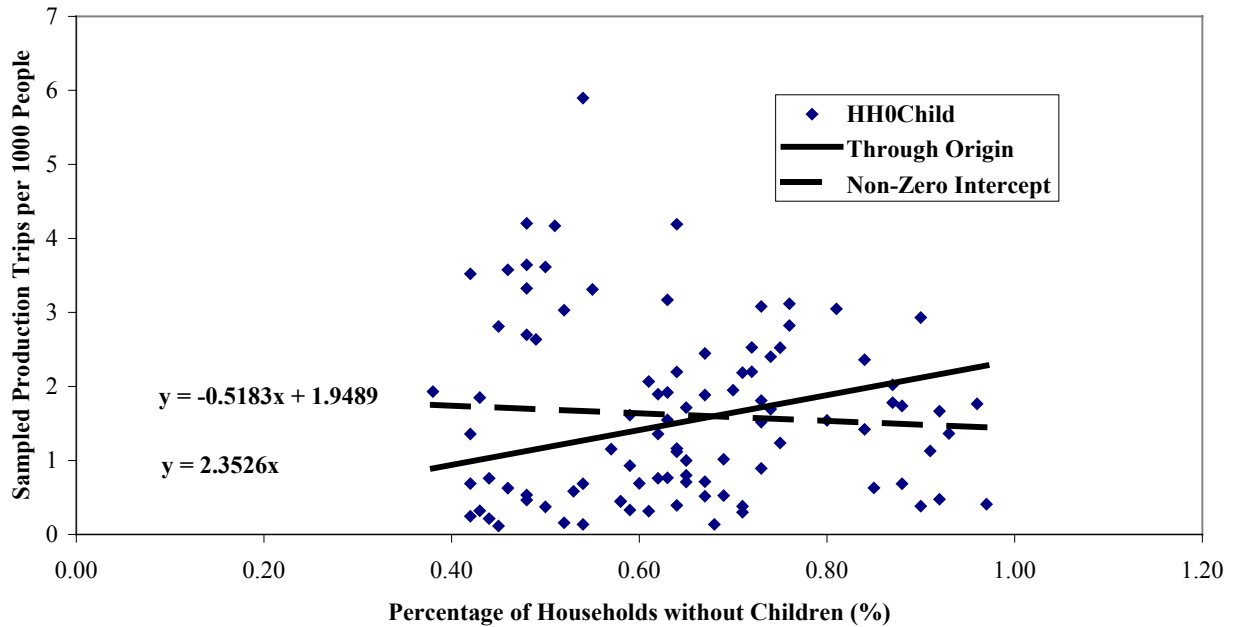


Figure 6-19. Intercept and No-Intercept %HH0Child Models for Transit Production (Tract Method)

Figure 6-20 illustrates the no-intercept models for the average number of workers in households with children (*AvgWrkrChild*). The models show that the effect of *AvgWrkrChild* on transit use increases with the accessibility level. Similar pattern can only be observed at the low accessibility level for intercept models (Figure 6-21). The variable itself had a negative relationship with transit productions for no-intercept models (see Figure 6-22) when the data were not categorized by accessibility level.

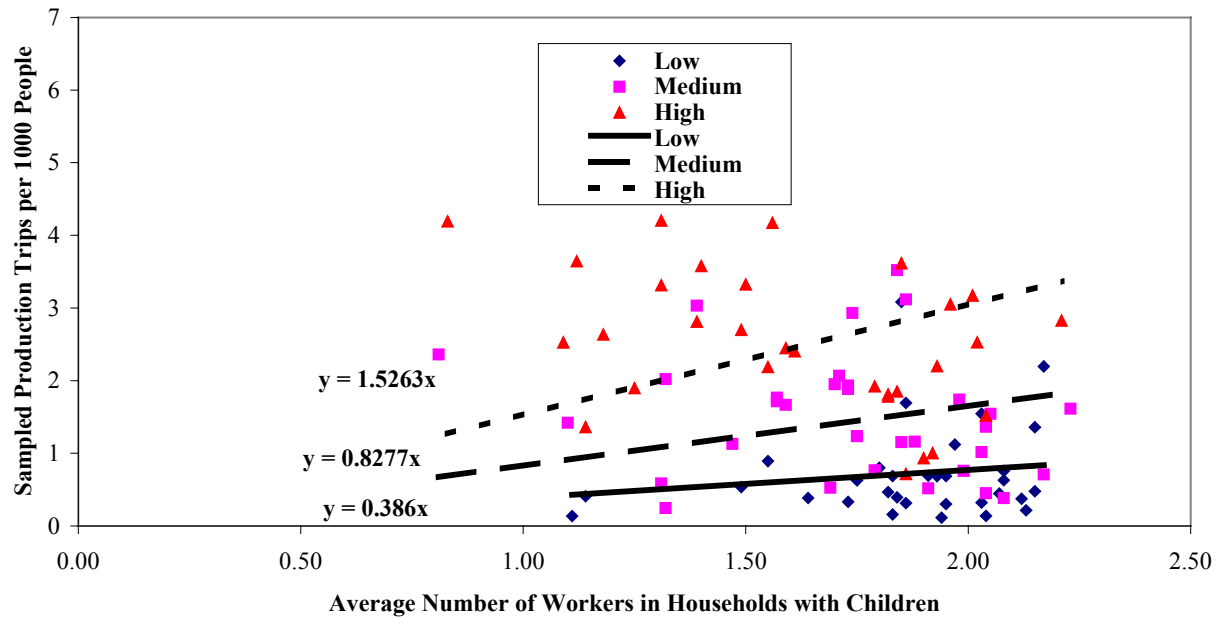


Figure 6-20. Through Origin Relationships between *AvgWrkrChild* and Transit Production by Transit Accessibility Level (Tract Method)

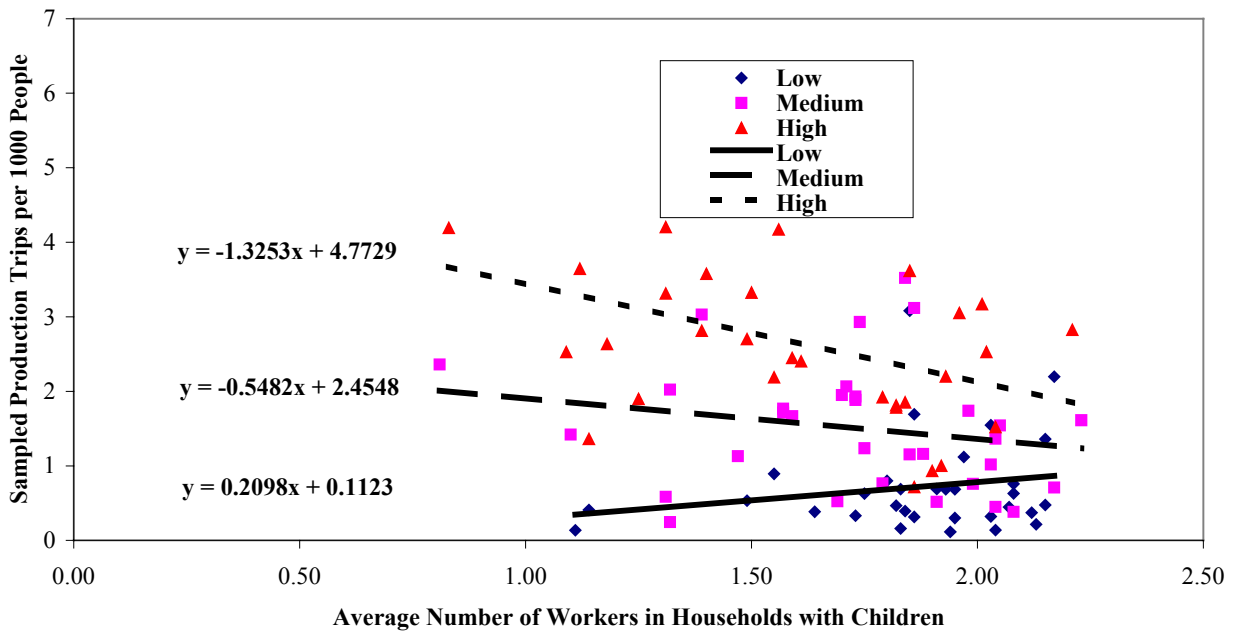


Figure 6-21. Non-Zero Intercept Relationships between *AvgWrkrChild* and Transit Production by Transit Accessibility Level (Tract Method)

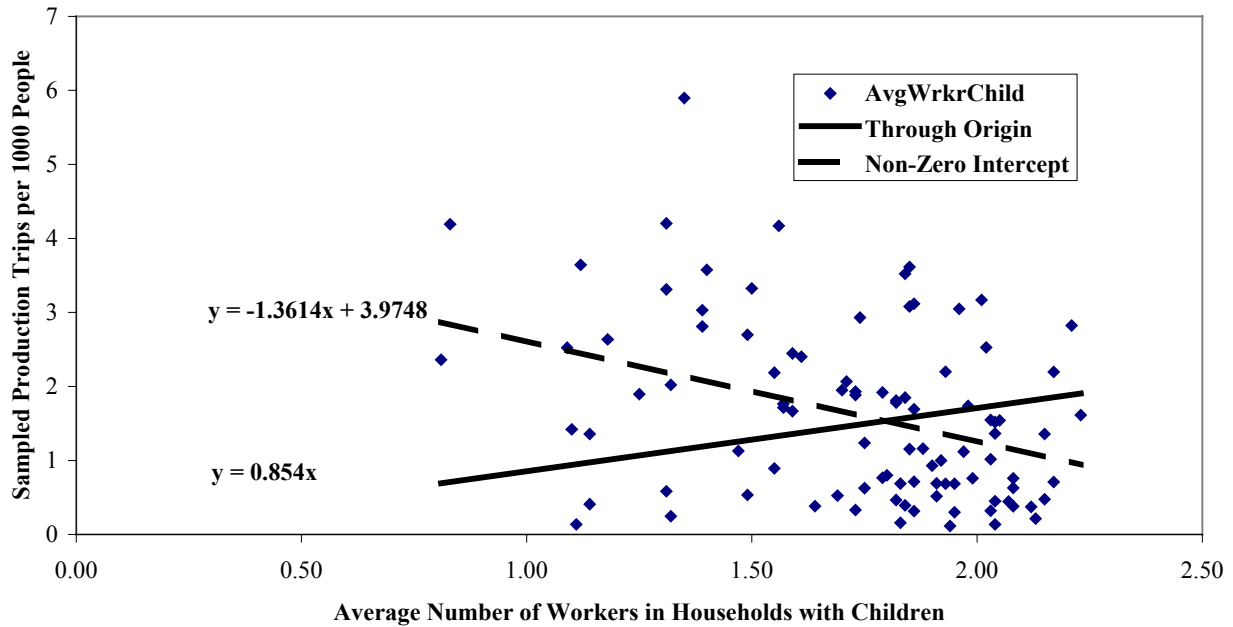


Figure 6-22. Intercept and Non-Intercept *AvgWrkrChild* Models for Transit Productions (Tract Method)

Figure 6-23 illustrates the no-intercept models for the average number of workers in households without children (*AvgWrkr0Child*). The models show that the effect of *AvgWrkr0Child* on transit use at different accessibility levels increases with accessibility. Same as *AvgWrkrChild*, similar pattern can only be observed at the low accessibility level for intercept models (Figure 6-24). The variable itself also has a negative relationship with transit productions for no-intercept models (see Figure 6-25) when the data were not categorized by accessibility level.

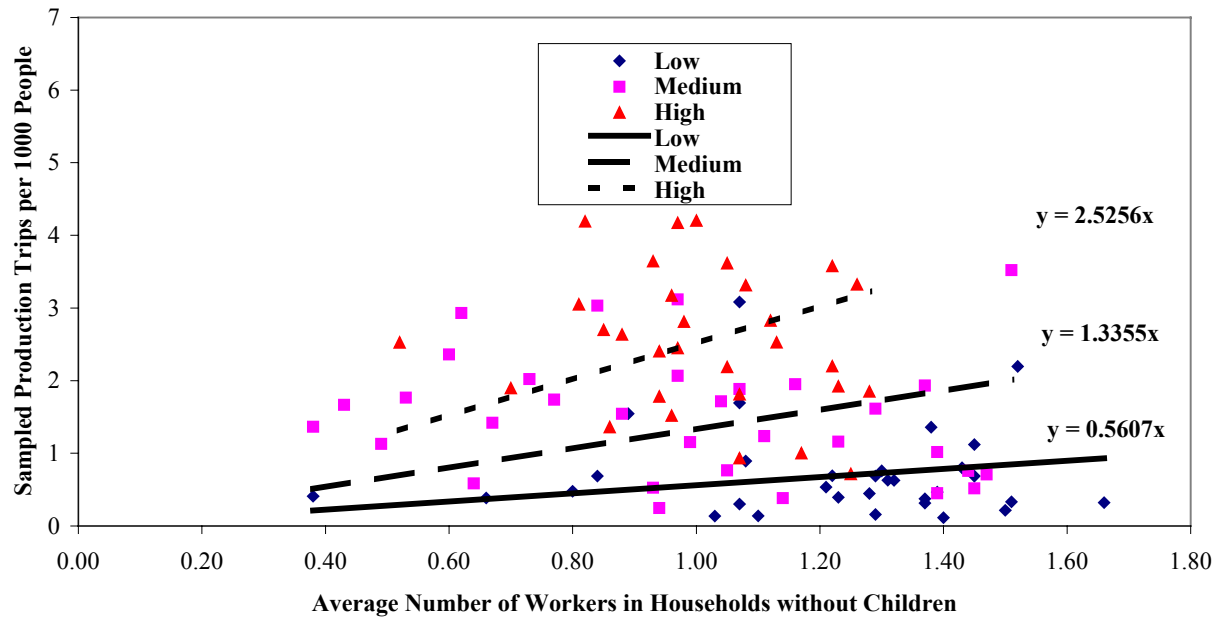


Figure 6-23. Through Origin Relationships between *AvgWrkr0Child* and Transit Production by Transit Accessibility Level (Tract Method)

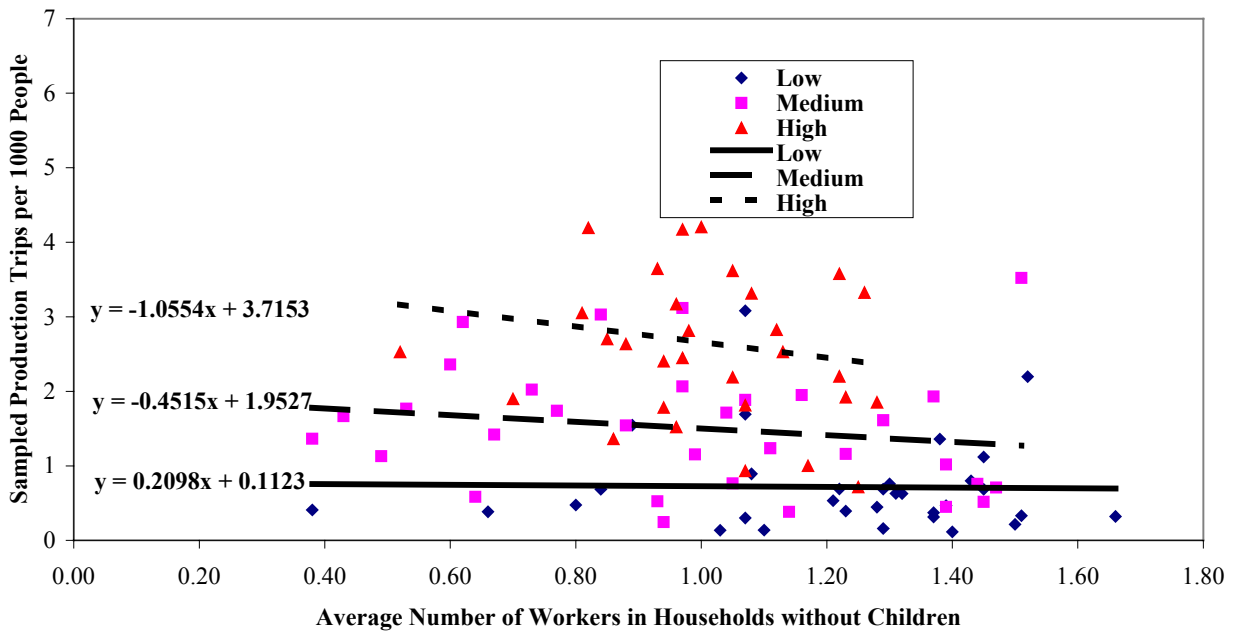


Figure 6-24. Non-Zero Intercept Relationships between *AvgWrkr0Child* and Transit Production by Transit Accessibility Level (Tract Method)

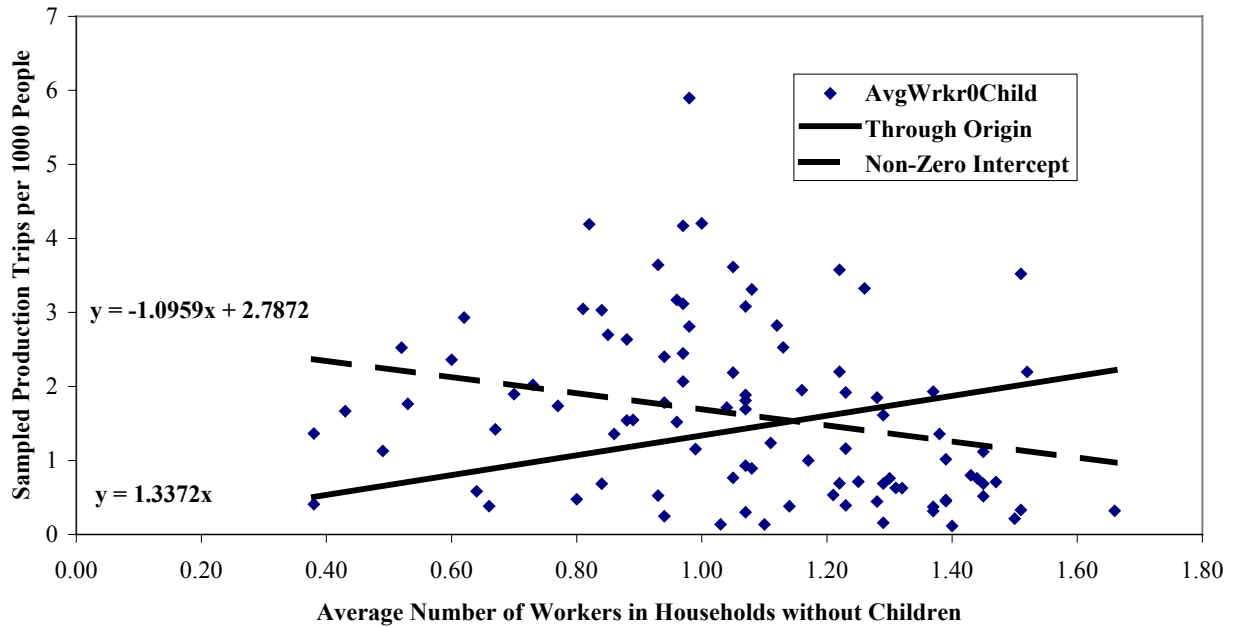


Figure 6-25. Intercept and Non-Intercept *AvgWrkr0Child* Models for Transit Productions (Tract Method)

Figure 6-26 illustrates the zero-intercept models for *AvgPChild*, i.e., average number of persons in households with children, and transit productions at different accessibility levels. These models reveal a positive relationship between the regressor and the dependent variable. In addition, the effect of *AvgPChild* on transit production increases with accessibility. As illustrated in Figure 6-27, the similar positive relationship can only be observed at the high accessibility level when the intercept was specified as non-zero. However, *AvgPChild* itself has a positive relationship with transit productions for both no-intercept and intercept models (see Figure 6-28) when the data were not grouped by accessibility level.

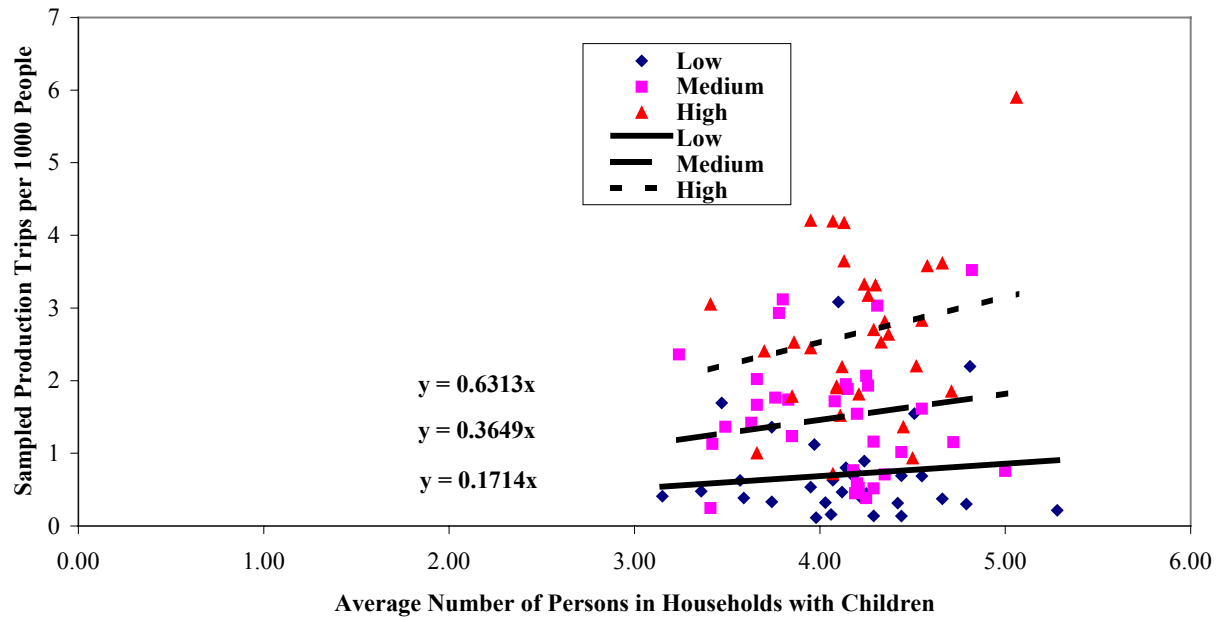


Figure 6-26. Through Origin Relationships between *AvgPChild* and Transit Production by Transit Accessibility Level (Tract Method)

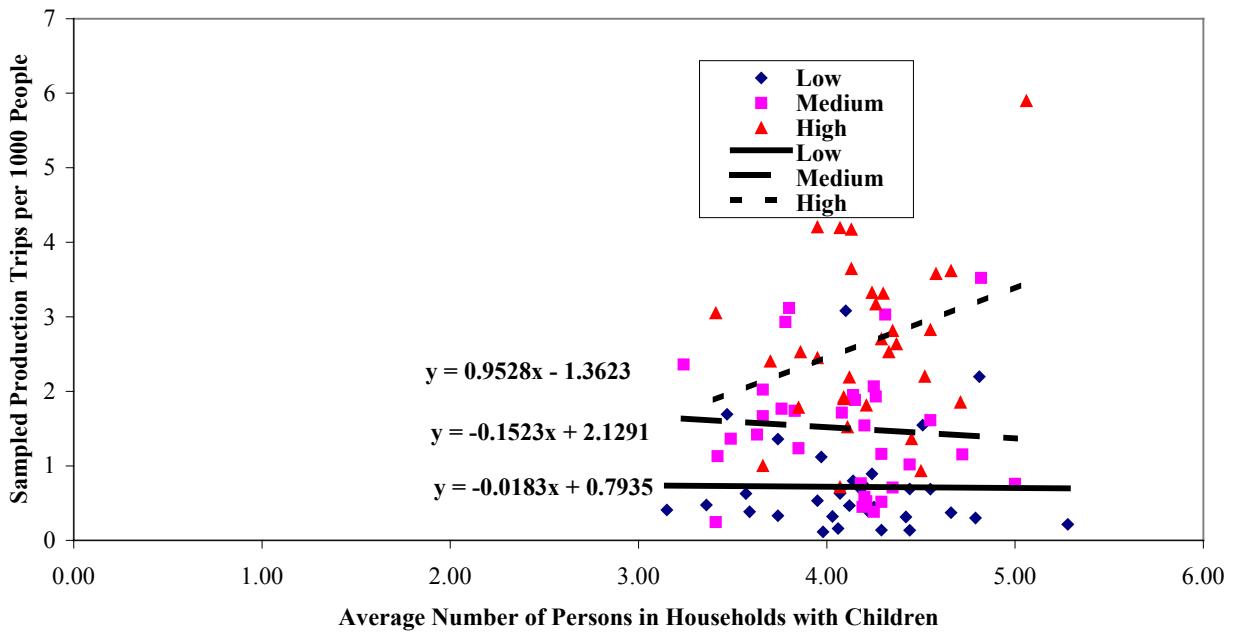


Figure 6-27. Non-Zero Intercept Relationships between *AvgPChild* and Transit Production by Transit Accessibility Level (Tract Method)

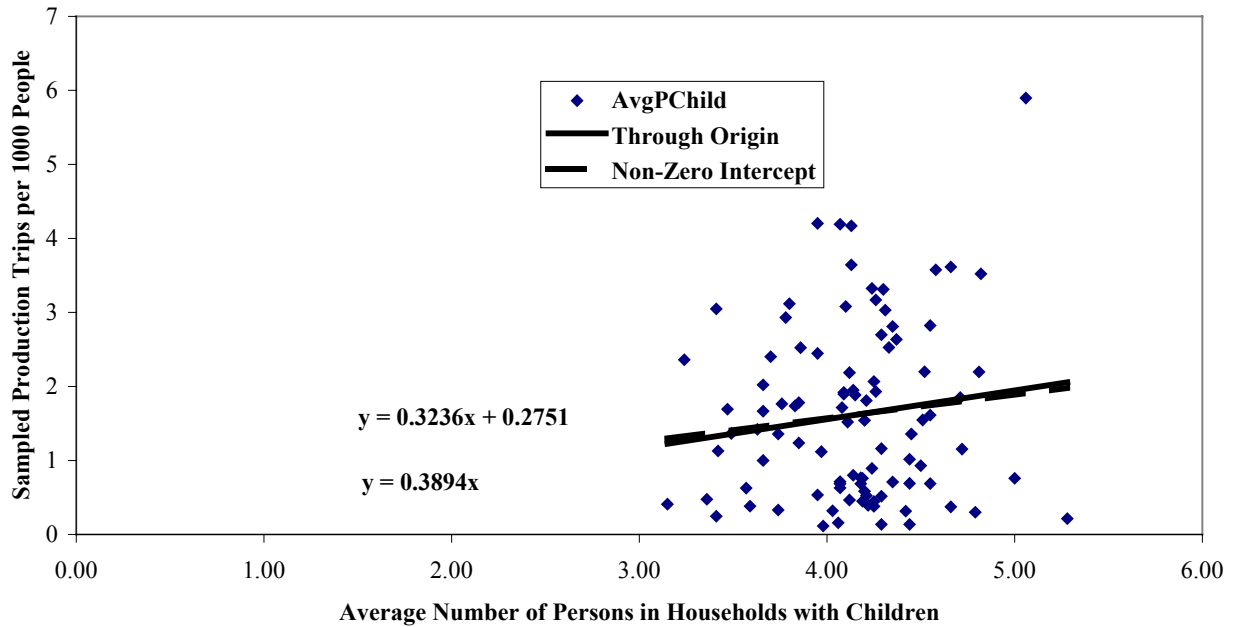


Figure 6-28. Intercept and No-Intercept *AvgPChild* Models for Transit Productions

Figure 6-29 illustrates the zero-intercept models for *AvgP0Child*, i.e., average number of persons in households without children, and transit productions at different accessibility levels. The resulted models show that a positive relationship exists between these two variables and the effect of *AvgP0Child* on transit production increases with accessibility. As illustrated in Figure 6-30, the similar positive relationship can be observed for the low and high accessibility groups when the intercept was specified as non-zero. However, *AvgP0Child* is negatively related with transit productions for intercept model (see Figure 6-31) when the data were not grouped by accessibility level.

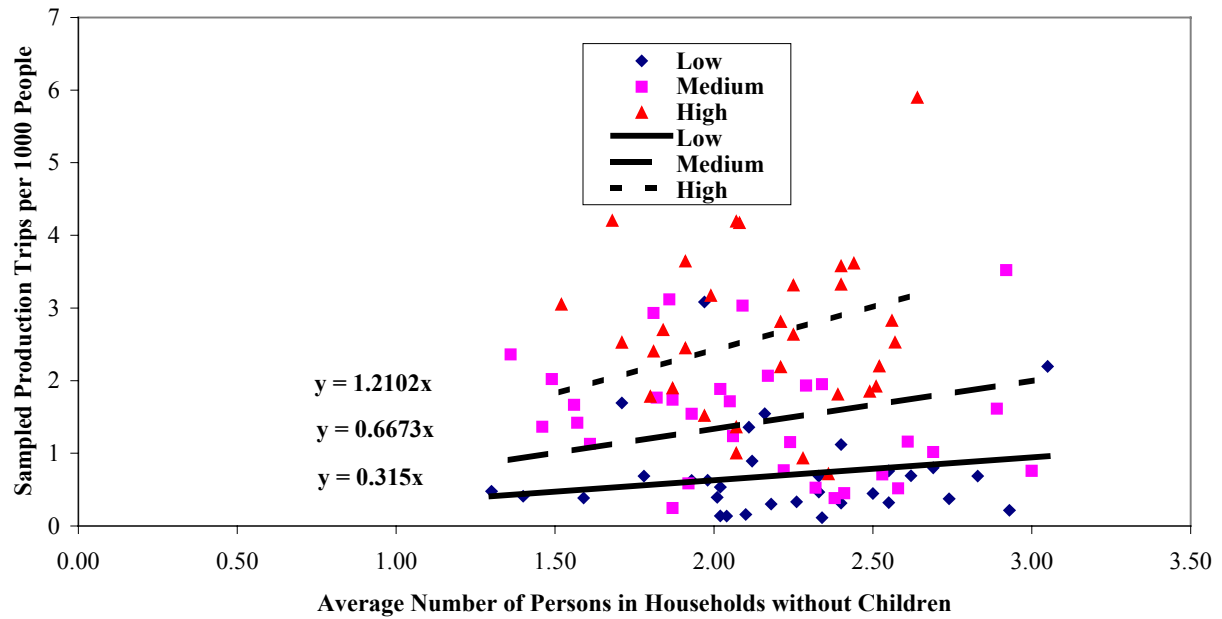


Figure 6-29. Through Origin Relationships between *AvgP0Child* and Transit Production by Transit Accessibility Level (Tract Method)

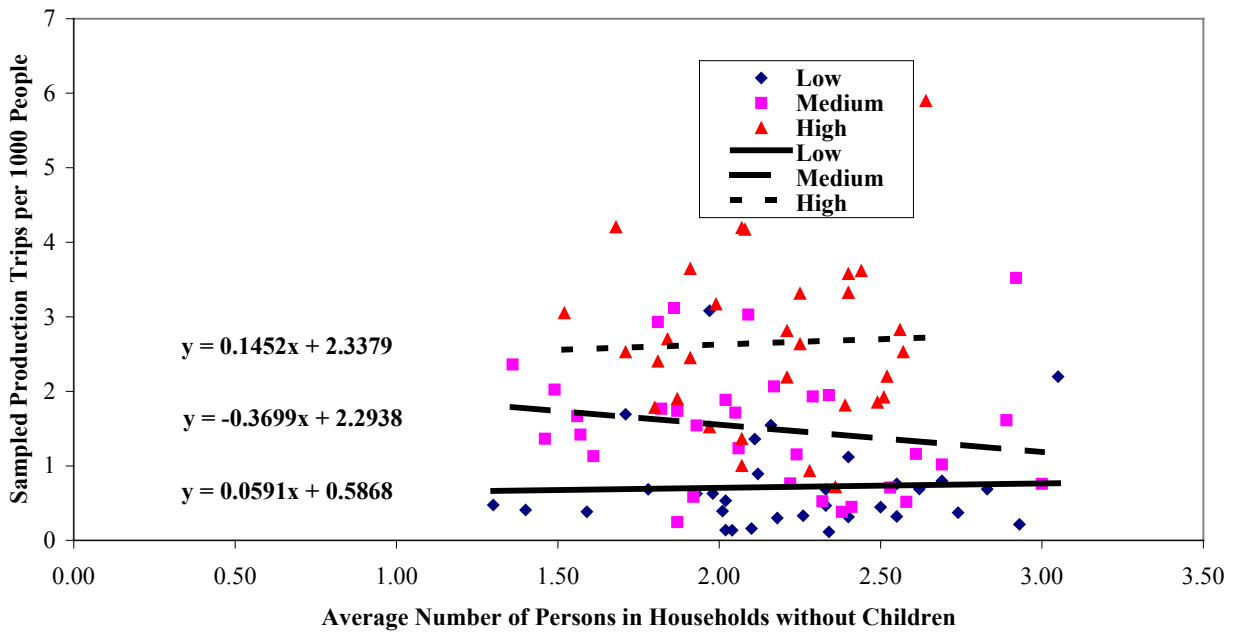


Figure 6-30. Non-Zero Intercept Relationships between *AvgP0Child* and Transit Production by Transit Accessibility Level (Tract Method)

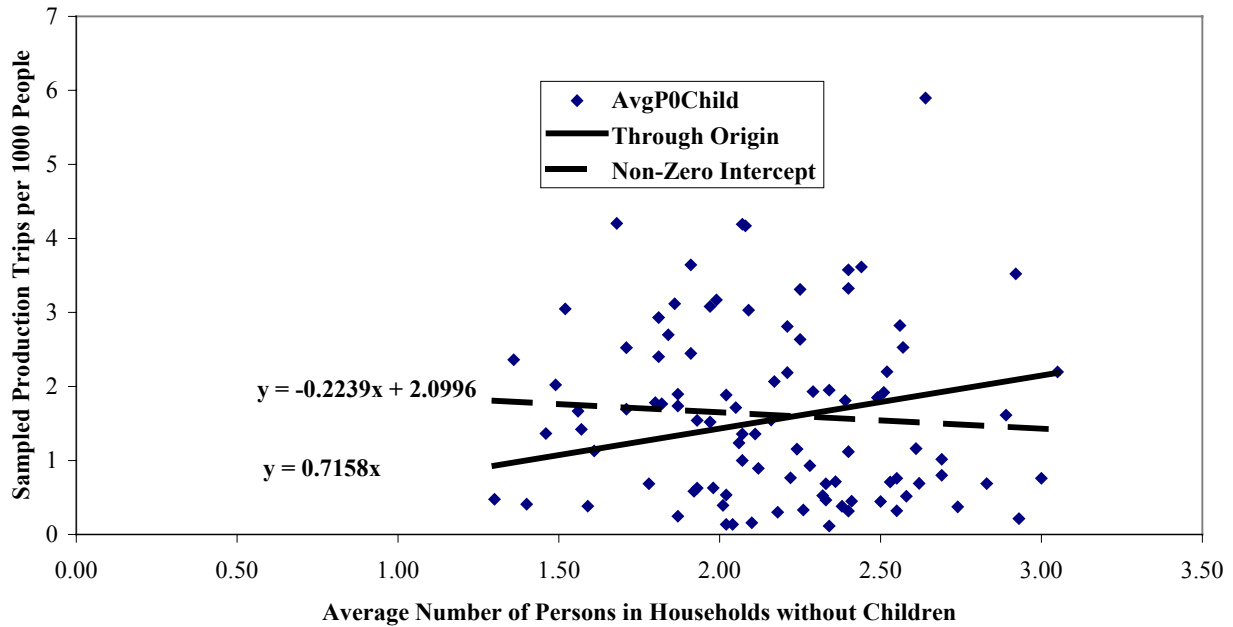


Figure 6-31. Intercept and No-Intercept *AvgP0Child* Models for Transit Productions

Figure 6-32 illustrates the zero-intercept models for the job house balance variable, i.e., *JOBS_HH4*, and transit productions at different accessibility levels. These models show that a positive relationship exists between these two variables and the effect of *JOBS_HH4* on transit production increases with accessibility. As illustrated in Figure 6-33, the similar positive relationship can only be observed at the low accessibility level for non-zero intercept models. In addition, *AvgP0Child* is negatively related with transit productions for intercept model (see Figure 6-34) when the data were not grouped by accessibility level.

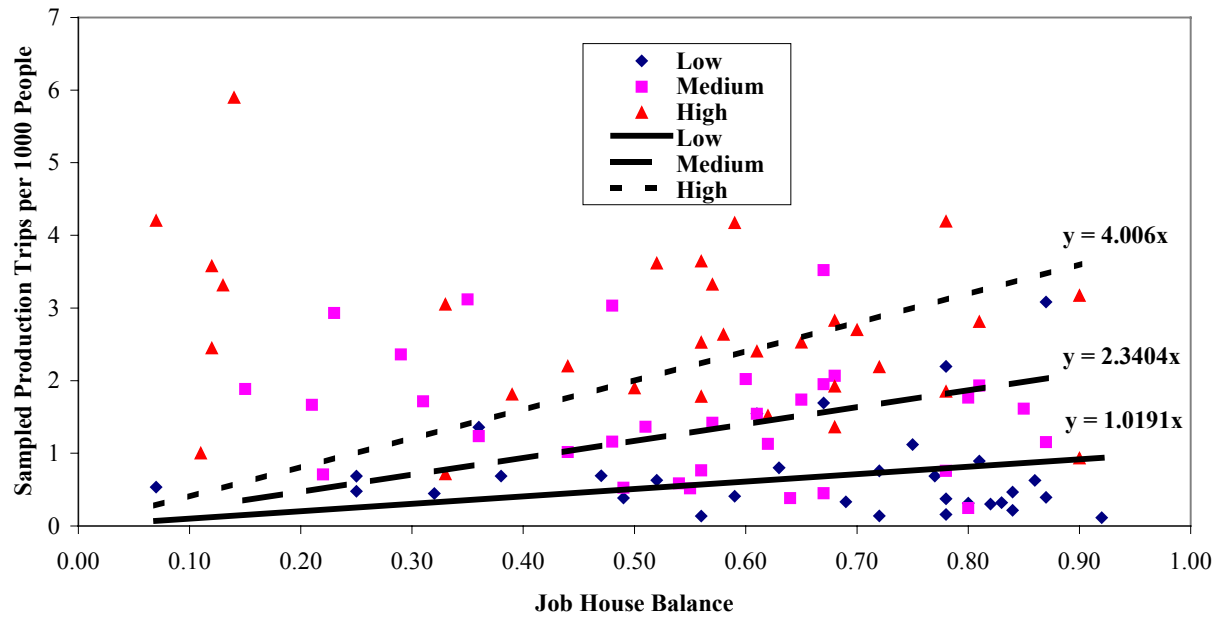


Figure 6-32. Through Origin Relationships between *JOBS_HH4* and Transit Production by Transit Accessibility Level (Tract Method)

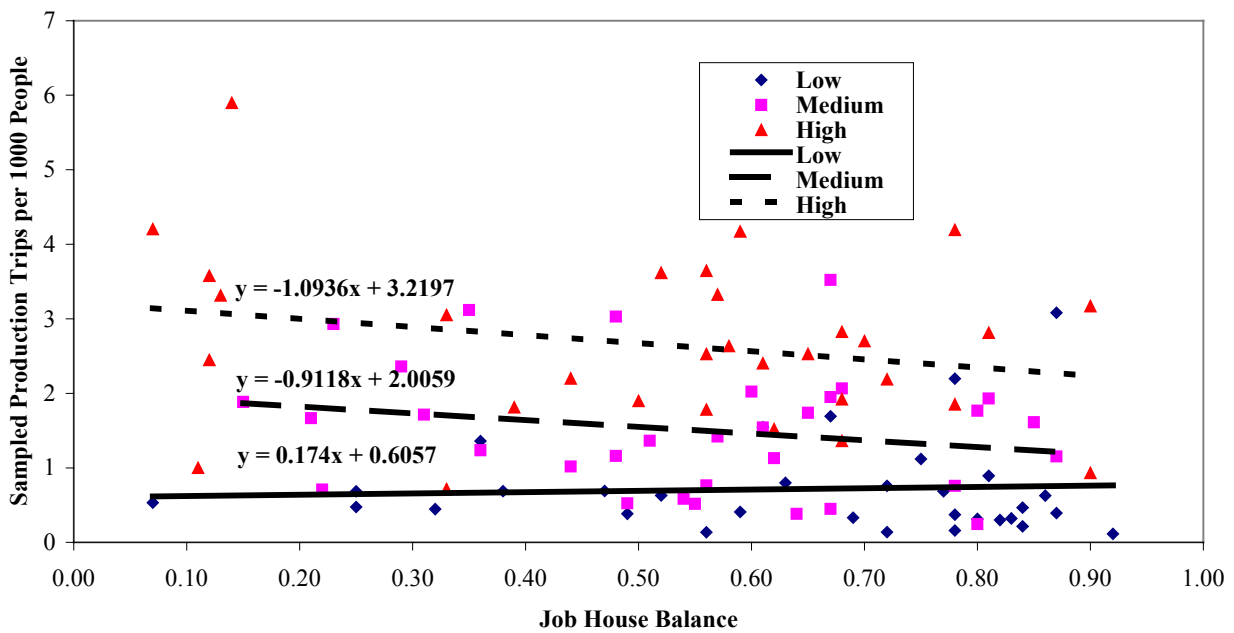


Figure 6-33. Non-Zero Intercept Relationships between *JOBS_HH4* and Transit Production by Transit Accessibility Level (Tract Method)

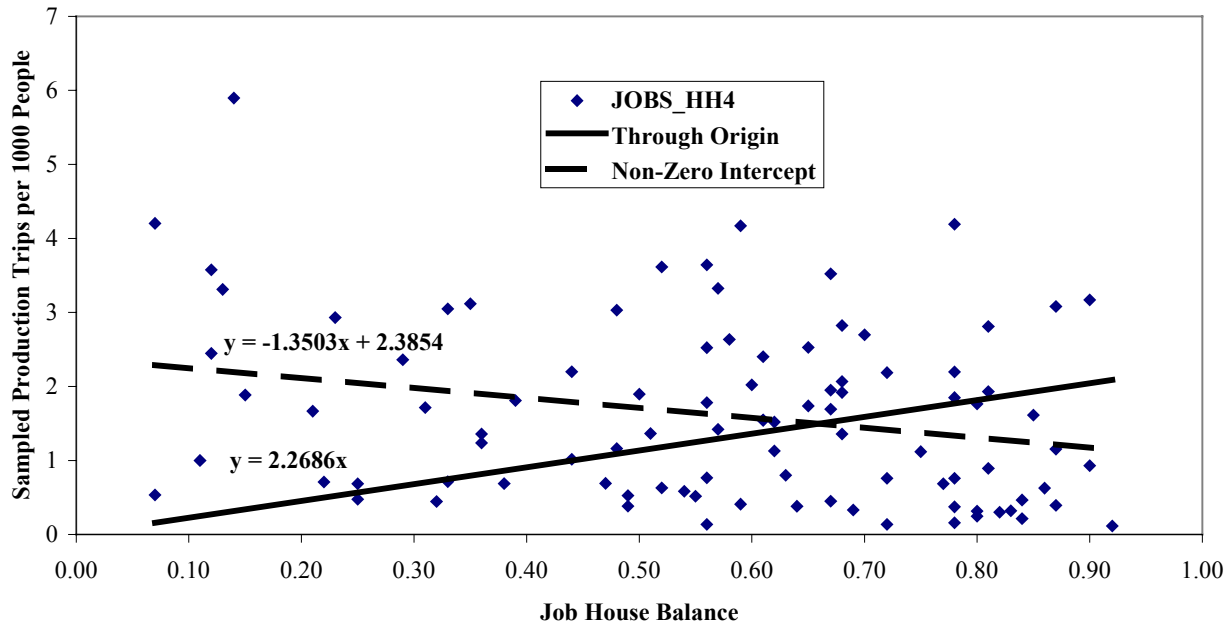


Figure 6-34. Intercept and No-Intercept *JOBS_HH4* Models for Transit Productions

6.5 Summary

In this section, linear regression models developed based on data compiled using different methods and with all variables and demand variables only have been presented. The models presented in Section 6.2 were calibrated using data compiled by four different methods for Miami-Dade County: GIS Land Use Method, Property Method, Buffer Method, and Tract Method. Models based on these four data compilation methods consistently revealed *AccessTWPop*, i.e., transit walk skim accessibility weighted by population, as the significant variable that had a linear relationship with transit productions. Other accessibility variables such as those weighted by number of workers were found to be significant as well. The variables of transit LOS, e.g., *Dailybus*, were also significant. In addition, the results from the GIS Land Use and Buffer methods indicate that *DECAY_POP*, i.e., transit service population percentage described in Section 5, had a linear relationship with transit productions. The R^2 s for the models with this new variable were even higher than those with accessibility and LOS variables obtained from the GIS Land Use Method. Other than the accessibility and transit LOS variables, the socioeconomic and demographic variables compiled at the tract level by different methods did not show such promising relationship with transit productions. *JOBS_HH*, however, appeared to be a significant factor in the attraction models for Miami-Dade County. However, the explanatory power of this variable was not as significant as the accessibility/level of service variables for transit productions.

The models calibrated based on the available data from Broward and Palm Beach counties have been described in Section 6.3. Data for Broward County were compiled using both the Buffer Method and Tract Method, while the data for Palm Beach County were compiled using only the Tract Method. Similar to those for Miami-Dade County, the results for Broward and Palm Beach counties suggest that *AccessTWPop* is a significant factor for transit productions and

JOBS_HH is a significant factor for attractions. The R^2 's for transit attraction models were also much lower than those for production models.

Table 6-28 summarizes the results presented in Section 6.4, which were obtained after excluding the transit supply variables and controlling for transit LOS as measured by transit accessibility. Models for three groups of data with low, medium, and high accessibility were developed. Table 6-28 shows the effects of the demand variables on transit use at different accessibility levels, with the +/- symbols indicating a positive or negative effect and *Null* representing a nearly neutral effect. Those variables that appear in all models and have consistent positive (or negative) effects on transit uses among various transit accessibility levels are better candidates for serving as predictors of transit use. Such variables include total density (*TotalDensity*) for data compiled with the GIS Land Use Method, and entropy measures (*AvgEntropy*) for data from the Property Method and the Buffer Method.

Table 6-28. Variables for Intercept and No-Intercept Models

Method	Variable	Zero Intercept				Intercept			
		Low	Medium	High	Total	Low	Medium	High	Total
GIS Land Use Method	<i>SFDUDensity</i>	+	+	+	+	+	+	-	+
	<i>TotalDensity</i>	+	+	+	+	+	+	+	+
Property Method	<i>AvgEntropy</i>	+	+	+	+	+	+	+	+
	<i>TotalDensity</i>	+	+	+	+	Null	+	Null	+
Buffer Method	<i>AvgEntropy</i>	+	+	+	+	+	+	+	+
Tract Method	<i>%HH0Child</i>	+	+	+	+	+	+	-	-
	<i>AvgWrkrChild</i>	+	+	+	+	+	-	-	-
	<i>AvgWrkr0Child</i>	+	+	+	+	+	-	-	-
	<i>AvgPChild</i>	+	+	+	+	-	-	+	+
	<i>AvgP0Child</i>	+	+	+	+	+	-	+	-
	<i>JOBS_HH4</i>	+	+	+	+	+	-	-	-

The results in Table 6-27 indicate that none of the variables from the Tract Method models had consistent effect on transit use after the non-zero intercept was specified. The results also imply that the effects of density variables such as *TotalDensity* can only be modeled when more detailed GIS data are utilized in the model calibration process since it was significant only for models based on data compiled using the GIS Land Use data and the Property Method. In other words, data aggregated based on simplified assumptions, such as those in the Tract Method, do not provide adequate information for estimating transit use. Therefore, factors appropriate to estimate transit use depend on the level of detail of the data and the method used to compile the data.

A word of caution is in order regarding entropy as a transit use predictor. Entropy is an indicator of land mix. A large entropy value indicates dissimilarity of land uses in an area. Entropy has been identified as having a positive effect on increasing walk trips by some researchers (Frank and Pivo, 1994). It has also been positively associated with transit use (Frank and Pivo, 1994; Seskin, 1996). While it is relatively easy to understand that a good land use mix may mean close proximity of opportunities reachable by walk, thus more walk trips, it is less certain that good

land use mix is the cause for higher transit use. One possible explanation may be that good land use mix was a common characteristic for most areas when transit services were initially introduced in the early part of the century and this characteristic has remained for transit-oriented neighborhoods. In comparison, newer suburban developments have been designed for automobiles and typically lack mixed land uses. Therefore, entropy itself may not be a cause of low or high transit use, rather it is a characteristic associated neighborhoods, with transit oriented neighborhoods having high entropy while auto-oriented ones having low entropy. In other words, the relationship between entropy and transit use is not causal, but association. Therefore, if a community has been designed with mixed land use, walk trips may be encouraged but unless good transit services are provided, it is unlikely that transit usage will be high. As a result, the predictive power of land use mix (average entropy) may be limited.

The same is suspected to be true for jobs-housing balance (*JOBS_HH*), another land use mix measurement. Good jobs-housing balance means that employment opportunities are nearby and there may be fewer long work commutes and more walk trips. However, similar to entropy, good jobs-housing balance itself may not actually encourage transit use but is a characteristic of transit-oriented neighborhoods. As such, its power to explain transit use will be limited.

7. SUMMARY AND CONCLUSIONS

This report presented research methodologies and results on transit accessibility and factors that impact transit use and that have the potential of being considered in the FSUTMS modal split model. This research utilized the TOB survey data from the SEFRTCS, demographic and socioeconomic data from FSUTMS models, and extensive and detailed land use information such as land use GIS data and property tax data.

A new methodology was developed for estimating transit accessibility for production trips. This method took into consideration walking distance to transit stops, population distribution, and existence of barriers to pedestrians. Data included the detailed street network, bus stop locations, bus routes, population and dwelling unit information by TAZ, property locations as represented by their label points, and property tax database, which provided information on number of bedrooms, used as a proxy for household size, at each property location. Additionally, the transit onboard survey data from the 1999 Southeast Florida Travel Characteristics Study were used to determine the effect of walking distance on transit use. The results of this research suggested that transit use deteriorated exponentially with walking distance to transit stops. A decay function was determined based on the survey data to reflect this deteriorating trend in transit use with respect to walk distance, and transit accessibility was measured by the percentage of the population, weighted by the decay function, in a zone within 0.5 mile of walking distance from transit stops. Increasing the limit of walking distance longer than 0.5 mile produced no noticeable increase in accessibility based on the survey data. Because of the decay in transit use due to increases in walking distance, transit accessibility was much lower than the traditional buffer method or the network ratio method would estimate.

Results from analysis also showed that there was no significant difference between the percentage of population with transit access and the percentage of production trips with transit access in a given zone. This means that the percentage of population with transit access can be used directly in place of percentage production trips with transit access as required as input by FSUTMS modal split model.

To allow the use of the transit walk accessibility measure for forecast purposes, where detailed information on street configuration and population distribution may be lacking for new TAZs, regression models were developed using easily obtainable data to predict transit accessibility. One of the variables used in predicting transit accessibility for production trips was the number of streets in a TAZ intersecting TAZ boundary per 1000 feet. This variable in fact measures how well the street configuration provides walk accessibility, for example, in traditional neighborhoods with grid street patterns and small blocks versus suburban neighborhoods where streets are often curvilinear with cul-de-sacs. The value of this variable must be determined based on the anticipated type of community in new TAZs. This may be used as a planning tool to examine the impact of different alternatives of development in a new area in terms of its transit accessibility.

Employment accessibility to transit is defined as the percentage of employees in a zone within 0.25 mile of transit stops. Here the 0.25-mile distance is air distance rather than actual walking distance. This choice was because most commercial developments are located along arterials and thus are rather accessible to transit, and because the spatial distribution of employees in a

zone is difficult to determine. Employing land use data improves the information on the spatial distribution of employees in a zone as opposed to assuming employees are evenly distributed across the entire zone or they are evenly distributed along all arterials, although on a few occasions land use data have not been consistent with employment data from the ZDATA file.

Forecasting future employment accessibility to transit services may be accomplished using the regression model developed in this study. Application of the model only requires the calculation of bus route density in a zone and the service and commercial employment ratio; both are easily obtainable.

Auto access of transit was analyzed using the TOB data of transit trips that involved accessing a transit station/stop either by park-and-ride or kiss-and-ride. Analysis results showed that while there was no apparent relationship between auto access trip distance and the transit trip length, most transit trips were longer than the auto access trips. The auto access distance for most (91.5%) sampled transit trips was shorter than 10 miles; the longest auto access trip was 14.6 miles. Therefore, we concluded that auto access distance in a zone may be assumed to be up to the longest transit trips likely from that zone (by considering premium transit modes and major activity centers) up to 14 miles. This result, obtained based on data from Miami-Dade County, however, may not be applicable to areas with a different urban structure, premier transit alignment, and parking availability. This upper limit also depends on the route length of the rapid transit services and may change if the route configuration or length changes.

In addition to transit walk and auto accessibility, transit factors that are most significant in impacting transit use have also been investigated. Multiple regression analysis was employed for this purpose using data compiled for Miami-Dade, Broward, and Palm Beach counties. Four methods were used to compile the model data, which range in their levels of detail of the land use and demographic information. Regional accessibility variables, especially those weighted by population, appeared to have the most significant impact on both transit productions and attractions, regardless of which method was used to compile the data. Regional accessibility measures were calculated based on zonal population, employment, zonal trip time, and impedance functions calibrated based on travel survey data for both transit and highway networks. However, the impedance functions are unlikely to be transferable and may need to be calibrated for individual urban areas, which will require a moderate effort.

For Miami-Dade County, the average number of daily bus runs per bus stop in a given tract and percentage of tract area falling into a 0.25-mile transit buffer were also relatively good indicators for predicting transit use at both production and attraction ends. Because these two kinds of data are easier to compile and are more reliable than regional accessibility, they may be used as alternatives to regional accessibility measures. However, based on the results from the GIS Land Use Method, Property Method, and Buffer Method, the newly developed *DECAY_POP* variable, i.e., percentage of distance decayed transit service population is a better indicator for transit production than percentage of service area in a tract.

No significant linear relationships were found consistently between transit use and the demographic, socio-economic, and other transit LOS variables for either production or attraction trips beside the regional accessibility measures, the number of bus runs, and transit service coverage, which are all related to transit supply. This may be due to the loss of spatial variation

in these variables when large spatial analysis units (such as census tracts) are used to compile the data. While smaller spatial units will preserve the characteristics of an area or population better, this research has determined that the TOB survey cannot be used due to the limited number of samples available resulting in a significant number of spatial units with no samples.

Compared with transit production trip models, the goodness-of-fit for attraction models were relatively poor. The R^2 's for these models were consistently lower than those for the production trip models. This may be the result of less reliable data on zonal employment. Furthermore, the models for Broward and Palm Beach counties showed similar effects of the accessibility variables on transit productions. Again, no other variables can be identified as the supplement/substitute of the accessibility variables.

The results suggest that transit supply dominates other factors in contributing to transit use, which in turn is the result of transit service supply being determined based on demand. The results from excluding the primary transit supply factors, such as the accessibility and transit LOS variables, showed that total density (population plus employment) in the buffer area of transit routes performed better than other demand variables based on the data compiled by the GIS Land Use Method. The land use mix variables in buffer areas also performed better than other variables based on the data compiled using the Property Method and Buffer Method. No similar conclusions can be drawn from the models calibrated for the Tract Method. Consequently, mode split models that utilize data of different details to estimate the proportion of trips using transit may require different explanatory variables. Additional research efforts are necessary to estimate the coefficients of the potential variables identified in this research for each travel mode's utility function utilized in the modal split process.

This research has used GIS extensively to compile data for various variables created for the accessibility analyses and the transit use analyses. Transit accessibility analyses require bus stop locations and property locations. As it is becoming more common for counties to possess GIS parcel data, the availability of property location data is not foreseen as a problem for the application of the methodology developed in this research. Property tax records should include information on number of bedrooms for each residential property and information on the type of the property to indicate if it is a single- or multi-family dwelling. If number of bedroom information is unavailable, then an average household size will have to be assigned to dwelling units that are of the same type (single- or multi-family).

Finally, the transit accessibility analysis can be automated with a specially designed GIS program. This program will automate the process of matching properties to streets, assign household size, create transit service network, calculate the percentage of population with transit access in a TAZ, estimate the percentage of workers with transit access in a zone using land use information, and so on.

8. RECOMMENDATIONS

Based on the research results described in this report, the following recommendations are provided for consideration for future effort in improving FSUTMS modal split models:

- (1) The long transit walk (assuming an one-mile walking distance) file is unnecessary since transit use will be extremely low beyond one-half mile from transit stops. Instead, a single transit walk file can be used that is based on one-half mile walking distance and the decay function calibrated in this research. The data may also be used for transit service planning as it can provide much more accurate information regarding service population at transit stop level.
- (2) The percentage of population served by transit in a zone may be estimated using the transit service population forecast model developed in this research for any reasons such as lack of data or skilled GIS staff. The forecast models can also be used for forecasting transit service population for future year models.
- (3) A GIS application should to be implemented that will calculate percentage of transit service population in a zone based on street network data, ZDATA1 and ZDATA2 files, property tax database, and parcel GIS data. In the case that GIS parcel data are unavailable, the regression model will be applied to estimate the percentage of distance decayed service population.
- (4) If a buffer method is to be used to estimate transit service population, cautions should be used when the calculated service population percentage from the buffer method exceeds 50%, a level that the distance decayed transit service population rarely exceeds. Only when population is actually clustered around transit stops, a percentage higher than 50% may be justified.
- (5) Regional transit accessibility may be considered for the inclusion in the modal split model. This requires that accessibility to be estimated for future, which is possible given forecast population, employment, and transportation improvements.
- (6) Density as measured by total employment plus population per acre may serve as predictors of demand.
- (7) To ensure adequate data are available to support a more conclusive study, carefully designed surveys are desired in the future to obtain adequate number of observations in areas that reflect the different spectrum of socioeconomic and land use characteristics. Future household surveys should target transit households so that studies can be carried out to directly determine the modal split instead of using an indirect approach as in this study. If a community leadership sees transit development as necessary, then the development of information, data, and forecasting methods to better place investments is worthwhile.

REFERENCES

- Allen, W.B., Liu, D., and Singer, S. (1993). Accessibility Measures of U.S. Metropolitan Areas. *Transportation Research Part B*, 27B, Transportation Research Board, National Research Council, Washington, D.C., pp. 439-449.
- Altinoglu, I. and Smith, N. (1992). *Approaching a Dynamic Urban Transit Demand Model for Sydney*. Working Paper ITS-WP-92-8, Institute of Transport Studies, Graduate School of Business, The University of Sydney, Sydney, Australia, 513-528.
- Azar, K.T., Ferrira, Jr. J., and Wiggins, L. (1994). Using GIS Tools to Improve Transit Ridership on Routes Serving Large Employment Centers: The Boston South End Medical Area Case Study. *Computers, Environmental and Urban Systems*, 18, pp. 205-231.
- Batchelder, J. et al. (1983). *Bus Route Demand Models: Final Report*. Contract/Grant No. DOT-TSC-1756, U.S. Department of Transportation, Mass Transportation Administration and Transportation System Centers, Washington, D.C.
- Black, W.R. (1993). Development of Ideal Model for Identification of Rural Public Transit Needs. *Transportation Research Record 1402*, Transportation Research Board, National Research Council, Washington, D.C., pp. 107-109.
- Box, G.E.P. and Jenkins, G.M. (1976). *Time-Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, California.
- Box, G.E.P. and Tiao, G.C. (1975). Intervention Analysis with Application to Economic and Environmental Problems. *Journal of the American Statistical Association*, 70, pp.70-79.
- Cambridge Systematics (1994). *The Effects of Land Use and Travel Demand Management Strategies on Commuting Behavior*. Final report DOT-T-95-06, Washington, D.C., U.S. Department of Transportation.
- Cervero, R. (1994). Transit-Based Housing in California: Evidence on Ridership Impacts. *Transport Policy*, 1, pp. 174-183.
- Cervero, R. and Gorham, R. (1995). Commuting in Transit Versus Automobile Neighborhoods: *American Planning Association Journal*, 61, pp. 210-225.
- Cleland, F., Hinebaugh, D., and Rey, J.R. (1997). *Transit Customer Satisfaction Index for Florida Transit Properties. Technical Memorandum No. 3: Results and Analysis of Florida Transit Properties*. Center for Urban Transportation Research, University of South Florida, Tampa, Florida.
- Clifton, K, Fisher, J. and Handy, S.L. (1998). *The Effectiveness of Land Use Policies as a Strategy for Reducing Automobile Dependence: A Study of Austin Neighborhoods*. Research Report 465650-1, SWUTC/98/465650-1, Energy Office, Office of the Governor of the State of Texas, Austin, Texas.

Dehghani, Y., and Harvey, R. (1994). Fully Incremental Model for Transit Ridership Forecasting: Seattle Experience, *Transportation Research Record 1452*, Transportation Research Board, National Research Council, Washington, D.C., pp. 52-61.

DKS Associates (1994). *Model Development and User Reference Report*, Sacramento Area Council of Governments, Sacramento, California.

FDOT (1997). *FSUTMS Transit Network Model (TNET). Technical Report No. 11 in the Series of Documentation and Procedural Updates to the Florida Standard Urban Transportation Model Structure (FSUTMS)*. Technical Report, Florida Department of Transportation, Tallahassee, Florida.

Frank, L.D. and Pivo, G. (1994). Impacts of Mixed Use and Density on Utilization of Three Modes of Travel: Single-Occupant Vehicle, Transit, and Walking. *Transportation Research Record 1466*, Transportation Research Board, National Research Council, Washington, D.C., pp. 44-52.

Gannett Fleming, Inc. (1998). *Broward Urban Study Area Travel Forecast Model Validation, Technical Report No. 2*, prepared for Florida Department of Transportation.

Gray, C. and Thompson, G. (1996). *Determinants of Transit Usage and Mode Choice. Final Report for Task 1b: Transit Patronage as a Product of Land Use Potential and Connectivity: The Sacramento Case*. NUTI-2-FSU3. Department of Urban and Regional Planning, Florida State University, Tallahassee, Florida.

Handy, S. (1992). Regional Versus Local Accessibility: Neotraditional Development and Its Implications for Non-Work Travel. *Built Environment*, 18, pp. 253-267.

Harrington, I.E. and Carakatsane, M.J. (1984). *Forest Hills-Needham Corridor: Ridership Forecasts for Transit-Service Alternatives*. CTPS Technical Report 45. Central Transportation Planning Staff, MPO, Boston, Massachusetts.

Holtzclaw, J. (1990). *Explaining Urban Density and Transit Impacts on Auto Use*. Docket No. 89-CR-90. State of California Energy Resources Conservation and Development Commission, California.

Holtzclaw, J. (1990). *Manhattanization vs. Sprawl: How Urban Density Impacts Auto Use - Comparing Five Bay Area Communities*. Paper presented at the 11th International Pedestrian Conference, San Francisco, California.

Holtzclaw, J. (1994). *Using Residential Patterns and Transit to Decrease Automobile Dependence and Cost*. Report Prepared for California Home Energy Efficiency Rating Systems. Natural Resources Defense Council, San Francisco, California.

Horowitz, A.J. (1985). *Transit Ridership Forecasting - Reference Manual*. Technical Report, Center for Urban Transportation Studies, University of Wisconsin, Milwaukee, Wisconsin.

- Hsiao, S., Lu, J., Sterling, J. and Weatherford, M. (1997). *Using GIS for Transit Pedestrian Access Analysis*. Preprint of the Transportation Research Board 76th Annual Meeting, Washington, D.C., January 12-16.
- Hsiao, S., Lu, J., Sterling, J., and Weatherford, M. (1997). Use of Geographic Information System for Analysis of Transit Pedestrian Access. *Transportation Research Record 1604*, Transportation Research Board, National Research Council, Washington, D.C., pp. 50-59.
- Hunt, D.T., Still, S.E., Carroll, J.D., and Kruse, A.O. (1986). A Geodemographic Model for Bus Service Planning and Marketing. *Transportation Research Record 1051*, Transportation Research Board, National Research Council, Washington, D.C., pp. 1-12.
- Kitamura, R., Laidet, L., Mokhtarian, P.L., Buckinger, C., and Gianelli, F. (1994). *Land Use and Travel Behavior*. Technical Report LUCD-ITS-RR-94-27, report prepared for the State of California Air Resources Board. Institute of Transportation Studies, University of California. Davis, California.
- 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, pp. 125-158.
- Kittelsohn & Associates, Inc. (1999a). *Transit Capacity and Quality of Service Manual*. TCRP A-15 Final Report. Transportation Research Board, National Research Council, Washington, D.C.
- Kittelsohn & Associates, Inc. (1999b). *Development of a Transit Level of Service (TLOS) Indicator*. Final Report, Florida Department of Transportation, Tallahassee, Florida.
- Kockelman, K.M. (1995). Which Matters More in Mode Choice: Density or Income? *Compendium of Technical Papers: the 65th Annual Meeting and 1995 District 6 Annual Meeting of the Institute of Transportation Engineers*. The Institute of Transportation Engineers, Washington, D.C., pp. 844-867.
- Kockelman, K.M. (1997). Travel Behavior as a Function of Accessibility, Land Use Mixing, and Land Use Balance: Evidence from Francisco Bay Area. *Transportation Research Record 1607*, Transportation Research Board, National Research Council, Washington, D.C., pp. 116-125.
- Koppelman, F.S. (1983). Predicting Transit Ridership in Response to Transit Service Changes. *Journal of Transportation Engineering*, 109, pp. 548-564.
- Krechmer, D., Lantos, G., and Golengerg, M. (1983). *Bus Route Demand Models: The Cleveland Prototype Study*. Technical Report OH-09-7002-83, Office of Planning Assistance, UMTA, U.S. Department of Transportation, Washington, D.C.
- Kyte, M., Stoner, J., and Cryer, J. (1985). Development and Application of Time-Series Transit Ridership Models for Portland, Oregon, *Transportation Research Record 1036*, Transportation Research Board, National Research Council, Washington, D.C., pp. 9-18.
- Lam, W. and Morrall, J. (1982). Bus Passenger Walking Distances and Waiting Times: A Summer-Winter Comparison. *Transportation Quarterly*, 36, pp. 407-421.

- Levinson, D., and Kumar, A. (1995). A Multimodal Trip Distribution Model: Structure and Application. *Transportation Research Record 1446*, Transportation Research Board, National Research Council, Washington, D.C.
- Levinson, H.S. (1985). Forecasting Future Transit Route Ridership. *Transportation Research Record 1036*, Transportation Research Board, National Research Council, Washington, D.C., pp. 19-28.
- Levinson, H.S., Brown-West, O. (1984). Estimating Bus Ridership, *Transportation Research Record 994*, Transportation Research Board, National Research Council, Washington, D.C., pp. 8-12.
- Loutzenheiser, D.R. (1997). Pedestrian Access to Transit: Model of Walk Trips and Their Design and Urban Form Determinants around Bay Area Rapid Transit Stations, *Transportation Research Record 1604*, Transportation Research Board, National Research Council, Washington, D.C., pp. 40-49.
- McLeod Jr., M.S., Flannelly, K.J., and Flannelly, L. (1991). Multivariate Time-Series Model of Transit Ridership Based on Historical, Aggregate Data: The Past, Present and Future of Honolulu, *Transportation Research Record 1297*, Transportation Research Board, National Research Council, Washington, D.C., pp. 76-84.
- Messenger, T. and Ewing, R. (1996). Transit-Oriented Development in the Sun Belt, *Transportation Research Record 1552*, Transportation Research Board, National Research Council, Washington, D.C., pp. 145-153.
- Metro (2000). *Metro Estimated Mode Choice Model*. Draft Report, Metro, Transportation Department, Travel Forecasting Division, Portland, Oregon, August 2000.
- Nelson, A.C., Sanchez, T.L., Ross, C.L., and Meyer, M.D. (1997). Rail Transit in the Suburbs: Case Study of Transit Use in Atlanta's Affluent Northern Tier. *Transportation Research Record 1571*, Transportation Research Board, National Research Council, Washington, D.C., pp. 142-150.
- Nelson, D.O. and O'Neil, K.K. (1982). Home-Origin Transit Travel Analysis Model, *Transportation Research Record 915*, Transportation Research Board, National Research Council, Washington, D.C., pp. 24-30.
- Nelson\Nygaard Consulting Associates (1997). *Tri-Met Primary Transit Network Phase II Report*, Portland, Oregon.
- Newman, P. and Kenworth, J. (1989). *Cities and Automobile Dependence: An International Sourcebook*, Gower Publishing, Aldershot, England.
- Nickesen, A.H., Meybug, A.H., and Turnquist, M.A. (1983). Ridership Estimation for Short-Range Transit Planning. *Transportation Research. Part B: Methodological*, **17B**, pp. 233-244.
- Myers, Raymond H. (1990). *Classical and Modern Regression with Applications*, 2nd Edition. PWS-Kent Publishing Company, Boston, Massachusetts.

O'Neill, W., Ramsey, D., and Chou, J. (1995). "Analysis of Transit Service Areas Using Geographic Information Systems," *Transportation Research Record No. 1364*, Transportation Research Board, National Research Council, Washington, D.C., pp. 131-138.

Parsons Brinckerhoff (2000) Tour and Trip Mode Choice Models. Draft Technical Report, Prepared for Sacramento Area Council of Governments, Sacramento, California.

Pendyala, R.M. (1996). *An Assessment of Transit System Modeling in the State of Florida*. Final Report, Contract No. B-9815. Public Transit Office, Florida Department of Transportation, Tallahassee, Florida.

Pendyala, R.M. (1999). *Integrated Transit Demand and Supply Model User Manual*. ITSUP Version 0.50. Public Transit Office, Florida Department of Transportation, Tallahassee, Florida.

Peng, Z., Dueker, K.J., Strathman, J., and Hopper, J. (1997). A Simultaneous Route-Level Transit Patronage Model: Demand, Supply, and Inter-Route Relationship, *Transportation*, **24**, pp. 159-181.

Preslar, D.A. (1998). Transit Ridership Forecasting Using a GIS. *Proceedings of the ASCE Conference on Transportation, Land Use, and Air Quality - Making the Connection*, pp. 595-605.

Pushkarev, B.S. and Zupan, J.M. (1977). *Public Transportation and Land Use Policy*. Indiana University Press, Bloomington, Indiana.

Richardson, A.J. and Young, W. (1982). A Measure of Linked-Trip Accessibility. *Transportation Planning and Technology*, **7**, pp. 73-82.

Rood, T. (1997). *Local Index of Transit Availability, Riverside County, California Case Study Report*. Local Government Commission, Sacramento, California.

Ryus, P. (1998). *Working Definition of the Transit Level of Service Indicator*. Memorandum. Public Transit Office, Florida Department of Transportation, Tallahassee, Florida.

Seskin, S. (1996). *Transit and Urban Form*, Transit Cooperative Research Program Report 16. Transportation Research Board, National Research Council, Washington, DC, 1996.

Spillar, R.J. and Rutherford, G.S. (1990). The Effects of Population Density and Income on Per Capita Transit Ridership in Western American Cities. *Compendium of Technical Papers: the 60th Annual Meeting of the Institute of Transportation Engineering*, ITE Publication; No. PP-020, the Institute of Transportation Engineers, Washington, D.C., pp. 327-331.

Sun, X., Wilmont, C.G., and Kasturi, T. (1998). Household Travel, Household Characteristics, and Land Use: An Empirical Study from the 1994 Portland Travel Survey. *Transportation Research Record 1617*, Transportation Research Board, National Research Council, Washington, D.C.

Thompson, G. and Frank, J.E. (1995). *Evaluating Land Use Methods for Alternating Travel Behavior*. Final Report for Task 1b: Transit Patronage as a Product of Land Use Potential and

Connectivity: The Sacramento Case. Department of Urban and Regional Planning, Florida State University, Tallahassee, Florida.

Ulberg, C. (1982) Short-Term Ridership-Projection Model, *Transportation Research Record* 854, Transportation Research Board, National Research Council, Washington, D.C., pp. 12-16.

Vaga, K. and Shortreed, J.H. (1982). *Impact of Transfers on Transit Ridership*, RTAC Forum 5, pp. 38-44.

Vaziri, M., Hutchinson, J., and Kermanshah, M. (1990). Short-Term Demand for Specialized Transportation: Time-Series Model, *Journal of Transportation Engineering*, **116**, pp. 105-121.

Zhao, F. (1998). GIS Analysis of the Impact of Community Design on Transit Accessibility. *Proceedings of the ASCE South Florida Section Annual Meeting*. Sanibel Islands, Florida.